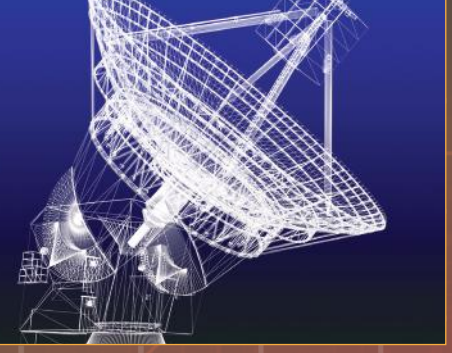


SECOND EDITION

Measurement, Instrumentation, and Sensors Handbook

Electromagnetic, Optical, Radiation, Chemical, and Biomedical Measurement

EDITED BY
John G. Webster
Halit Eren

CRC Press
Taylor & Francis Group

SECOND EDITION

Measurement, Instrumentation, and Sensors Handbook

Electromagnetic, Optical, Radiation, Chemical, and Biomedical Measurement

SECOND EDITION

Measurement, Instrumentation, and Sensors Handbook

Electromagnetic, Optical, Radiation, Chemical, and Biomedical Measurement

EDITED BY
John G. Webster
Halit Eren


CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20140113

International Standard Book Number-13: 978-1-4398-4893-7 (eBook - PDF)

Contents

PART II Electrical Variables

PART V Optical

PART VI Chemical Variables

PART VII Medical, Biomedical, and Health

PART VIII Environmental

PART X Displays and Recorders

Preface

Introduction

The second edition of *The Measurement, Instrumentation, and Sensors Handbook* comes in two volumes. This edition reflects the state of the art in measurement, instrumentation, and sensors. In this two-volume set, all chapters have been updated and 40 new chapters have been included to provide the finest possible reference that is both concise and useful for engineers practicing in industry, scientists and engineers involved in R&D, designers, college and university personnel and students, as well as managers together with many others who are involved in instrumentation and measurement design and applications.

The handbook covers an extensive range of topics that comprise the subject of measurement, instrumentation, and sensors. It describes the use of instruments and techniques for practical measurements required in engineering, physics, chemistry, environmental science, and the life sciences. It also explains sensors, techniques, hardware, and the associated software. The handbook includes information processing systems, automatic data acquisition, reduction and analysis, operation characteristics, accuracy, errors, calibrations, standards, and their incorporation for control purposes. Emphasis is given on modern intelligent instruments and techniques, wireless network operations, human factors, and modern display methods as well as virtual instruments.

The chapters include descriptive information for professionals, students, and workers interested in measurement. They consist of equations to assist engineers and scientists who seek to discover applications and solve problems that arise in fields not in their specialty. They also include specialized information needed by informed specialists who seek to learn advanced applications of the subject, evaluative opinions, and possible areas for future study. Thus, the handbook serves the reference needs of the broadest group of users—from the advanced high school science student to industrial and university professionals.

Organization

In this edition, the first volume has 10 parts, each consisting of several chapters, for a total of 96 chapters written by experts in their areas. It concentrates on concepts in instrumentation and measurements, spatial variable measurement, displacement measurement, mechanical variable measurement, acoustics, flow and spot velocity, thermal and temperature measurement, and radiation. It reflects recent trends in instrumentation and measurements with the addition of a new part on wireless instrumentation. Concepts in control systems and human factors are given as a separate part.

The second volume has 10 parts, each having several chapters, for a total of 98 chapters written by experts in their areas as in volume 1. It concentrates on sensors and sensor technology, electric variable measurement, electromagnetic variables, time and frequency, optical measurements, chemical variables, and medical, biomedical and health, and environmental measurement. Signal processing, and displays and recorders constitute the last two parts of this volume.

Locating Your Topic

To find out how to measure a given variable, skim the table of contents, turn to that part, and find the chapters that describe the different methods of making the measurement. Consider the alternative methods of making the measurement and each of their advantages and disadvantages. Select a method, sensor, and signal processing method. Many chapters list a number of vendors to contact for more information. You can also visit http://www.globalspec.com/ to obtain a list of vendors.

For more detailed information, consult the index, since certain principles of measurement may appear in more than one chapter.

John G. Webster and Halit Eren
Editors

Acknowledgments

We would like to thank all the authors for their valuable contribution toward this two-volume-set book. We appreciate the time and effort devoted by all our new authors and those authors who went an extra mile to revise and update their chapters. We are grateful to the CRC Press team for their encouragement to prepare this second edition. The publication of this book would not have been possible without their tireless dedication in putting it together. Last, but not least, we would like to thank all our readers for selecting this book in advancing their knowledge and technical skills.

John G. Webster
Halit Eren
Editors

Editors

John G. Webster received his BEE from Cornell University, Ithaca, New York, in 1953 and his MSEE and PhD from the University of Rochester, Rochester, New York, in 1965 and 1967, respectively.

He currently serves as professor emeritus of biomedical engineering at the University of Wisconsin-Madison. He serves as a highly cited researcher at King Abdulaziz University, Jeddah, Saudi Arabia. He also teaches undergraduate and graduate courses in the field of medical instrumentation and conducts research on intracranial pressure monitoring, ECG dry electrodes, and tactile vibrators.

Dr. Webster is author of *Transducers and Sensors*, an IEEE/EAB Individual Learning Program (Piscataway, NJ: IEEE, 1989). He is coauthor, with B. Jacobson, of *Medicine and Clinical Engineering* (Englewood Cliffs, NJ: Prentice-Hall, 1977); with R. Pallas-Areny, of *Sensors and Signal Conditioning, Second Edition* (New York: Wiley, 2001) and *Analog Signal Conditioning* (New York: Wiley, 1999). He is the editor of the *Encyclopedia of Medical Devices and Instrumentation, Second Edition* (New York: Wiley, 2006), *Tactile Sensors for Robotics and Medicine* (New York: Wiley, 1988), *Electrical Impedance Tomography* (Bristol, U.K.: Adam Hilger, 1990), *Teaching Design in Electrical Engineering* (Piscataway, NJ: Educational Activities Board, IEEE, 1990), *Prevention of Pressure Sores: Engineering and Clinical Aspects* (Bristol, U.K.: Adam Hilger, 1991), *Design of Cardiac Pacemakers* (Piscataway, NJ: IEEE Press, 1995), *Design of Pulse Oximeters* (Bristol, U.K.: IOP Publishing, 1997), *Medical Instrumentation: Application and Design, Fourth Edition* (Hoboken NJ: Wiley, 2010), *Encyclopedia of Electrical and Electronics Engineering* (New York, Wiley, 1999), *Minimally Invasive Medical Technology* (Bristol, U.K.: IOP Publishing, 2001), and *Bioinstrumentation* (Hoboken NJ: Wiley, 2004). He is also the coeditor, with A. M. Cook, of *Clinical Engineering: Principles and Practices* (Englewood Cliffs, NJ: Prentice-Hall, 1979) and *Therapeutic Medical Devices: Application and Design* (Englewood Cliffs, NJ: Prentice-Hall, 1982); with W. J. Tompkins, of *Design of Microcomputer-Based Medical Instrumentation* (Englewood Cliffs, NJ: Prentice-Hall, 1981) and *Interfacing Sensors to the IBM PC* (Englewood Cliffs, NJ: Prentice Hall, 1988); and with A. M. Cook, W J. Tompkins, and G. C. Vanderheiden, of *Electronic Devices for Rehabilitation* (London: Chapman & Hall, 1985).

Dr. Webster has been a member of the IEEE-EMBS Administrative Committee and the NIH Surgery and Bioengineering Study Section. He is a fellow of the Institute of Electrical and Electronics Engineers, the Instrument Society of America, the American Institute of Medical and Biological Engineering, the Biomedical Engineering Society, and the Institute of Physics. He is the recipient of the IEEE EMBS Career Achievement Award.

Halit Eren received his BEng in 1973, his MEng in 1975, and his PhD in 1978 from the University of Sheffield, United Kingdom. He obtained an MBA from Curtin University in 1999.

After his graduation, Dr. Eren worked in Etibank (a mining and metallurgy company in Turkey) as an instrumentation engineer for two years. He held the post of an assistant professor at Hacettepe University in 1980–1981 and at Middle East Technical University in 1982. He has been at Curtin University

since 1983, researching and teaching primarily in the areas of control systems, instrumentation, and engineering management.

Dr. Eren was appointed as a visiting associate professor at Polytechnic University, Hong Kong, in 2004 and served as a visiting professor at the University of Wisconsin, USA, in 2013. He is a senior member of IEEE and participates in Region 10 activities, in various committees for organizing conferences, and member of editorship in several transactions. Dr. Eren has over 180 publications in conference proceedings, books, and transactions. He is the author of *Electronic Portable Instruments—Design and Applications* (Boca Raton, FL: CRC Press, 2004) and *Wireless Sensors and Instruments—Networks, Design and Applications* (Boca Raton, FL: CRC Press, 2006). He has coedited, with Bela Liptak, *Instruments Engineers' handbook—Process Software and Digital Networks*, Vol. 3, edn.4 (Boca Raton, FL: CRC Press, 2011). He is involved in writing a number of books in the field of instrumentation and measurement. Dr. Eren researches and publishes on intelligent sensors, wireless instrumentation, wireless sensor networks, wireless biomedical devices, automation and control, and very large control systems.

Contributors

Mushtaq Ali
The National Grid Company
Sweden

Joseph H. Altman
Pittsford, New York

Roberto Ambrosini
Institute of Radioastronomy, INAF
Bologna, Italy

Jeff P. Anderson
LTV Steel Corporation
Independence, Ohio

Siamak Aram
Department of Electronics and Telecommunications
Politecnico di Torino
Turin, Italy

Pasquale Arpaia
Dipartimento di Ingegneria
Università del Sannio
CERN-European Laboratory for Nuclear Research
Geneva, Switzerland

Per Ask
Department of Biomedical Engineering
Linkoping University
Linkoping, Sweden

Francesco Avallone
Università di Napoli Federico II
Naples, Italy

Aldo Baccigalupi
Università di Napoli Federico II
Naples, Italy

William A. Barrow
Planar Systems, Inc.
Beaverton, Oregon

Cipriano Bartoletti
Sapienza University of Rome
Rome, Italy

L. Basano
Department of Physics
Università di Genova
Genova, Italy

M.W. Bautz
The Pennsylvania State University
University Park, Pennsylvania

David M. Beams
Department of Electrical Engineering
The University of Texas at Tyler
Tyler, Texas

K. Beilenhoff
Institut für Hochfiequenztechnik
Technische Universität Darmstadt
Munich, Germany

Michael Bennett
Centre for Aviation, Transport and the Environment
Manchester Metropolitan University
Manchester, United Kingdom

Tushar Kanti Bera
Department of Instrumentation and Applied Physics
Indian Institute of Science, Bangalore
Bangalore, India

Paul L. Bishop
College of Engineering
University of Rhode Island
Kingston, Rhode Island

Richard J. Blotzer
LTV Steel Corporation
Independence, Ohio

L. Boon-Brett
European Commission-JRC Institute for Energy
Petten, the Netherlands

Claudio Bortolotti
Institute of Radioastronomy, INAF
Bologna, Italy

Arnaldo Brandolini
Dipartimento di Elettrotecnica
Politecnico di Milano
Milan, Italy

Saps Buchman
Stanford University
Stanford, California

B.E. Burke
The Pennsylvania State University
University Park, Pennsylvania

Andrea Cadei
Dipartimento di Ingegneria dell'Informazione
University of Brescia
Brescia, Italy

Jarvis A. Caffrey
Department of Nuclear Engineering and Radiation Health Physics
Oregon State University
Corvallis, Oregon

Robert B. Campbell
Sandia National Laboratories
Livermore, California

Woo-Hyuck Choi
Globalfoundries Incorporation
Malta, New York

James F. Christian
Radiation Monitoring Devices, Inc.
Watertown, Massachusetts

Robert F. Cromp
Goddard Space Flight Center
National Aeronautics Space and Administration
Greenbelt, Maryland

Robert M. Crovella
NVIDIA Corporation
Plano, Texas

Timothy R. DeGrado
Duke University Medical Center
Durham, North Carolina

Alfons Dehé
Institut für Hochfiequenztechnik
Technische Universität Darmstadt
Munich, Germany

Alessandro Dionisi
Dipartimento di Ingegneria dell'Informazione
University of Brescia
Brescia, Italy

James T. Dobbins III
Duke University Medical Center
Durham, North Carolina

F. Domingue
Laboratoire de microsystèmes et télécommunications
LMST-Université du Québec à Trois-Rivières
Trois-Rivières, Québec, Canada

Enrique Dorronzoro
Department of Electronic Technology
University of Sevilla
Seville, Spain

Achim Dreher
German Aerospace Center
Wessling, Germany

Maria Eklund
Nynas AB
Stockholm, Sweden

H. El Matbouly
Laboratoire de microsystèmes et télécommunications
LMST-Université du Québec à Trois-Rivières
Trois-Rivières, Québec, Canada

Halit Eren
Department of Electrical and Computer Engineering
Curtin University
Bentley, Western Australia, Australia

Robert L. Fagaly
Quasar Federal Systems
San Diego, California

Luis Fernández-Luque
Department of Architecture and Computer Technology
University of Sevilla
Seville, Spain

and

ICT Norut
Northem Research Institute
Tromsø, Norway

Alessandro Ferrero
Dipartimento di Elettrotecnica
Politecnico di Milano
Milan, Italy

Michael Fisch
Department of Physics
John Carroll University
University Heights, Ohio

Gary R. Fones
School of Earth and Environmental Sciences
University of Portsmouth
Portsmouth, United Kingdom

Jacob Fraden
Fraden Corporation
San Diego, California

Jan Francke
International Groundradar Consulting, Inc.
Vancouver, British Columbia, Canada

Richard Frayne
University of Wisconsin, Madison
Madison, Wisconsin

K. Fricke
Institut für Hochfiequenztechnik
Technische Universität Darmstadt
Munich, Germany

Alessandro Gandelli
Dipartimento di Elettrotecnica
Politecnico di Milano
Milan, Italy

Julian W. Gardner
School of Engineering
University of Warwick
Coventry, United Kingdom

John D. Garrison
San Diego State University
San Diego, California

Manel Gasulla
Universitat Politècnica de Catalunya
Barcelona, Catalonia, Spain

Daryl Gerke
Kimmel Gerke Associates, Ltd.
Mesa, Arizona

Stéphane Gervais-Ducouret
Freescale Semiconductor
Tempe, Arizona

W.A. Gillespie
University of Abertay Dundee
Dundee, Scotland

Olaf Glück
Institut für Schicht-und Ionentechnik
Forschungszentrum Jülich GmbH
Jülich, Germany

Isabel Gómez
Department of Electronic Technology
University of Sevilla
Seville, Spain

Jose Antonio Gómez
Department of Electronic Technology
University of Sevilla
Seville, Spain

J.A. Gregory
The Pennsylvania State University
University Park, Pennsylvania

R.E. Griffiths
Pennsylvania State University
University Park, Pennsylvania

Donghai Guan
Harbin Engineering University
Harbin, China

Prasanta K. Guha
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur
Kharagpur, India

J.Y. Gui
General Electric Research and Development Center
General Electric Company Schenectady
New York, New York

Anthony Guiseppi-Elie
Virginia Commonwealth University
and
Abtech Scientific, Inc.
Richmond, Virginia

Mehdi Habibi
Department of Electrical Engineering
University of Isfahan
Isfahan, Iran

David M. Hamby
Department of Nuclear Engineering and Radiation Health Physics
Oregon State University
Corvallis, Oregon

Sean M. Hames
Duke University Medical Center
Durham, North Carolina

Daniel Harrison
Department of Physics
John Carroll University
University Heights, Ohio

H.L. Hartnagel
Institut für Hochfiequenztechnik
Technische Universität Darmstadt
Munich, Germany

Bruce H. Hasegawa
University of California, San Francisco
San Francisco, California

Alireza Hassanzadeh
Department of Electrical and Computer Engineering
Shahid Beheshti University
Tehran, Iran

Michael B. Heaney
Huladyne Research and Consulting
Palo Alto, California

Albert D. Helfrick
Embry–Riddle Aeronautical University
Dayton Beach, Florida

David A. Hill
U.S. Department of Commerce
National Institute of Standards and Technology
Boulder, Colorado

Stanley S. Ipson
Department of Electronic and Electrical Engineering
University of Bradford
West Yorkshire, United Kingdom

Rahman Jamal
National Instruments Germany
Munich, Germany

Roger Jones
Primary Children's Medical Center
Salt Lake City, Utah

Petr Kadlec
Department of Process Technology
Evonik Industries
Marl, Germany

Akihiro Kajiwara
University of Kitakyushu
Kitakyushu, Japan

Motohisa Kanda
U.S. Department of Commerce
National Institute of Standards and Technology
Boulder, Colorado

Chari V. Kandala
Agricultural Research Service
National Peanut Research Laboratory
United States Department of Agriculture
Dawson, Georgia

Mohammad A. Karim
The University of Tennessee, Knoxville
Knoxville, Tennessee

William Kimmel
Kimmel Gerke Associates, Ltd.
South St. Paul, Minnesota

H. Klingbeil
Institut für Hochfiequenztechnik
Technische Universität Darmstadt
Munich, Germany

R.L. Kraft
Pennsylvania State University
University Park, Pennsylvania

Silke Kröger
Lowestoft Laboratory
Centre for Environment, Fisheries and
Aquaculture Science
Suffolk, United Kingdom

V. Krozer
Institut für Hochfiequenztechnik
Technische Universität Darmstadt
Munich, Germany

H.L. Kwok
The Pennsylvania State University
University Park, Pennsylvania

C.K. Laird
School of Applied Chemistry
Kingston University
Surrey, United Kingdom

Carmine Landi
Università de L'Aquila
L'Aquila, Italy

W. Marshall Leach, Jr.
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia

Jin-Hwan Lee
Intel Corporation
Rancho, New Mexico

Sungyoung Lee
Department of Computer Engineering
Kyung Hee University
Seoul, South Korea

Woo Hyoung Lee
Department of Civil, Environmental, and
Construction Engineering
University of Central Florida
Orlando, Florida

Kathleen M. Leonard
Department of Civil and Environmental
Engineering
The University of Alabama in Huntsville
Huntsville, Alabama

Yufeng Li
HDD R&D Center
Samsung Information Systems America
San Jose, California

Robert G. Lindquist
Department of Electrical and Computer
Engineering
The University of Alabama in Huntsville
Huntsville, Alabama

E.B. Loewenstein
National Instruments
Austin, Texas

Michael A. Lombardi
Time and Frequency Division
National Institute of Standards and Technology
Boulder, Colorado

Oscar Lopez-Lapeña
Universitat Politècnica de Catalunya
Barcelona, Catalonia, Spain

Michael Z. Lowenstein
Harmonics Limited
Mequou, Wisconsin

Vincenzo Luciano
Dipartimento di Ingegneria dell'Informazione
University of Brescia
Brescia, Italy

D.H. Lumb
The Pennsylvania State University
University Park, Pennsylvania

Steven A. Macintyre
Macintyre Electronic Design
Herndon, Virginia

A.M. MacLeod
School of Engineering
University of Abertay Dundee
Dundee, Scotland

Martin Maier
Institut National de la Recherche Scientifique
Montréal, Québec, Canada

Robert T. Marcus
Datacolor International
Middletown, New Jersey

P.F. Martin
University of Abertay Dundee
Dundee, Scotland

Edward McConnell
National Instruments
Austin, Texas

Robert T. McGrath
Department of Engineering Science and Mechanics
The Pennsylvania State University
University Park, Pennsylvania

Douglas P. McNutt
The MacNauchtan Laboratory
Colorado Springs, Colorado

Ana Verónica Medina
Department of Electronic Technology
University of Sevilla
Seville, Spain

Gerry H. Meeten
Schlumberger Gould Research
Cambridge, United Kingdom

John T. Mester
W.W. Hansen Experimental Physics Laboratory
Stanford University
Stanford, California

Graham A. Mills
School of Pharmacy and Biomedical Sciences
University of Portsmouth
Portsmouth, United Kingdom

Jeffrey P. Mills
Illinois Institute of Technology
Chicago, Illinois

Devendra K. Misra
Department of Electrical Engineering and Computer Science
University of Wisconsin-Milwaukee
Milwaukee, Wisconsin

William C. Moffatt
Sandia National Laboratories
Livermore, California

Stelio Montebugnoli
Institute of Radioastronomy, INAF
Bologna, Italy

Jerry Murphy
Electronic Measurements Division
Hewlett-Packard
Colorado Springs, Colorado

J. Nagaraju
Department of Instrumentation and Applied Physics
Indian Institute of Science, Bangalore
Bangalore, India

Ryohei Nakamura
University of Kitakyushu
Kitakyushu, Japan

Soe-Mie F. Nee
Research and Technology Division
U.S. Naval Air Warfare Center
China Lake, California

Nam-Trung Nguyen
Berkeley Sensor & Actuator Center
The University of California, Berkeley
Berkeley, California

John A. Nousek
Department of Astronomy and Astrophysics
The Pennsylvania State University
University Park, Pennsylvania

P. Åke Öberg
Department of Biomedical Engineering
Linkoping University Hospital
Linkoping, Sweden

Chima Okereke
Department of Electronic and Electrical Engineering
University of Bradford
West Yorkshire, United Kingdom

Cesar Ortega-Sanchez
Curtin University
Bentley, Western Australia, Australia

Peter O'Shea
Department of Computer and Electrical Engineering
Royal Melbourne Institute of Technology
Melbourne, Victoria, Australia

P. Ottonello
Department of Physics
Università di Genova
Genova, Italy

Behrooz Pahlavanpour
Nynas AB
Nynashamn, Sweden

Ramón Pallás-Areny
Escola d'Enginyeria de Telecomunicació i Aerospacial de Castelldefels
Universitat Politècnica de Catalunya
Barcelona, Spain

V. Palmisano
European Commission-JRC Institute for Energy
Petten, the Netherlands

Ronney B. Panerai
Leicester Royal Infirmary
University of Leicester
Leicester, United Kingdom

Ian Papautsky
BioMicroSystems Laboratory
School of Electronic and Computing Systems
University of Cincinnati
Cincinnati, Ohio

E. Pasero
Department of Electronics and Telecommunications
Politecnico di Torino
Turin, Italy

Maria Teresa Penella
Urbiotica S.L.
Barcelona, Catalonia, Spain

B.W. Petley
Centre for Basic Metrology
National Physical Laboratory
Middlesex, United Kingdom

Thad Pickenpaugh
Air Force Research Laboratory
Wright-Patterson Air Force Base, Ohio

Luca Podestà
Sapienza University of Rome
Rome, Italy

Arshak Poghossian
Institute of Nano- and Biotechnologies
Aachen University of Applied Sciences
Aachen, Germany

Rodney Pratt
University of South Australia
Adelaide, South Australia, Australia

Naveen Puppala
Agricultural Science Center at Clovis
New Mexico State University
Las Cruces, New Mexico

M.M. Rad
Ciena
Ottawa, Ontario, Canada

Gerasimos Rigatos
Department of Engineering
Harper Adams University College
Shropshire, United Kingdom

and

Unit of Industrial Automation
Industrial Systems Institute
Rion Patras, Greece

Pavel Ripka
Faculty of Electrical Engineering
Czech Technical University in Prague
Prague, Czech Republic

Shyam Rithalia
Department of Rehabilitation
University of Salford
Salford, United Kingdom

Gordon W. Roberts
Department of Electrical Engineering
McGill University
Montréal, Québec, Canada

Stephen B.W. Roeder
Department of Chemistry
San Diego State University
San Diego, California

Mauro Roma
Institute of Radioastronomy, INAF
Bologna, Italy

F. Rugiano
Department of Electronics and
Telecommunications
Politecnico di Torino
Turin, Italy

Terry L. Rusch
Douglas Scientific
Alexandria, Minnesota

Giancarlo Sacerdoti
Sapienza University of Rome
Rome, Italy

Ravi Sankar
Department of Electrical Engineering
University of South Florida
Tampa, Florida

Sumita Santra
Department of Physics and Meteorology
Indian Institute of Technology, Kharagpur
Kharagpur, India

Emilio Sardini
Dipartimento di Ingegneria dell'Informazione
University of Brescia
Brescia, Italy

Kalluri R. Sarma
Department of Advanced Displays
Honeywell, Inc.
Phoenix, Arizona

Michael J. Schöning
Institute of Nano- and Biotechnologies
Aachen University of Applied Sciences
Aachen, Germany

Fritz Schuermeyer
United States Air Force Wright Laboratory
Wright-Patterson Air Force Base, Ohio

Gourab Sen Gupta
School of Engineering and Advanced Technology
Massey University
Palmerston North, New Zealand

Mauro Serpelloni
Dipartimento di Ingegneria dell'Informazione
University of Brescia
Brescia, Italy

Kanai S. Shah
Radiation Monitoring Devices, Inc.
Watertown, Massachusetts

Norman F. Sheppard, Jr.
Gamera Bioscience Corporation
Medford, Massachusetts

Christopher J. Sherman
Merrimack
New Hampshire

Pierluigi Siano
Department of Industrial Engineering
University of Salerno
Fisciano, Italy

Michael R. Squillante
Radiation Monitoring Devices, Inc.
Watertown, Massachusetts

Michael E. Stanley
Freescale Semiconductor
Tempe, Arizona

Robert Steer
Frequency Devices
Haverhill, Massachusetts

T.J. Sumner
Imperial College
London, United Kingdom

Haiyin Sun
Coherent Auburn Group
Auburn, California

Mark Sun
NeoPath, Inc.
Redmond, Washington

Peter H. Sydenham
University of South Australia
Mawsons Lakes, South Australia, Australia

Michal Szyper
AGH University of Science and Technology
and
University of Mining and Metallurgy
Kracow, Poland

Nitish V. Thakor
Department of Biomedical Engineering
School of Medicine
Johns Hopkins University
Baltimore, Maryland

Marion Thust
Institut für Schicht-und Ionentechnik
Forschungszentrum Jülich GmbH
Jülich, Germany

Michael F. Toner
Nortel Networks
Nepean, Ontario, Canada

A. Toriono
Department of Electronics and
Telecommunications
Politecnico di Torino
Turin, Italy

E.E. Uzgiris
General Electric Research and Development
Center
General Electric Company
Schenectady, New York

Ramanapathy Veerasingam
The Pennsylviana State University
University Park, Pennsylvania

Gert J.W. Visscher
Institute of Agricultural and Environmental
Engineering
Wageningen, the Netherlands

Manfred Wendisch
Leipzig Institute for Meteorology
University of Leipzig
Leipzig, Germany

Denise Wilson
Department of Electrical Engineering
University of Washington
Seattle, Washington

Weiwei Yuan
College of Computer Science & Technology
Harbin Engineering University
Harbin, China

James A. Zagzebski
Department of Medical Physics
University of Wisconsin, Madison
Madison, Wisconsin

I

Sensors and Sensor Technology

1

Soft Sensors: Software-Based Sensors

Petr Kadlec
Evonik Industries

1.1 Introduction

One of the major trends driving and driven by technological advancement is increasing use of the sensors and instruments in the world around us. Examples of sensors enabling the proliferated use of instruments around the globe are in great number presented in this chapter. A consequence of this effect is that increasing amount of digital data is available for further processing and exploitation. In fact, data are developing into a precious commodity, which is often traded at very high price. These data are the essence of soft sensors discussed in this work, and as such there is a strong pressure on the quality of the data. With some exceptions (e.g., [1]), algorithms for soft sensors development require a substantial amount of high-quality historical data in order to be able to develop useful soft sensors, and yet at the same time the quality of real-life industrial data is often very low. By analyzing the data obtained from their sources, one very often finds data impurities like outliers, missing value, and measurement noise [2]. The causes of these issues are numerous: it can be the physical limits of the hardware sensors, for example, in case of noisy data, and hardware sensor failures or maintenance, and in cases of missing data. For these reasons, the historical data often have to be manually treated to remove the impurities mentioned previously.

Once the data are pretreated, they should be ready to be processed in order to exploit their full potential by creating the useful information out of them. This process is called information extraction, and it is achieved by applying machine learning [3] and data mining [4] techniques to the raw data. These methods can either be univariate with a single variable, that is, transforming a single measurement stream (input variable) into another measurement stream (output or target variable). An example of such a technique is forecasting, where based on past values, the soft sensor is trying to forecast future values of the same variable. Other, more common, techniques are multivariate methods, which use

several input measurement streams and transform them into the output measurement stream. To name a few practical examples, soft sensors can provide predictions of the values of critical variables, which are not practical to be measured by hardware sensors or they can predict the expected trends and developments of quality variables, process states, etc.

In particular in the process industry, soft sensors have been established as useful tools, which can provide additional information about the underlying processes. One can distinguish two fundamentally different soft sensor types, namely, model-driven soft sensors and data-driven soft sensors [2]. There are several other terms that can be used for these two types of models; these include parametric vs. nonparametric models, white box vs. black box models [5], and phenomenological vs. empirical models [6]. The model-driven soft sensors use the chemical and physical principles underlying the process. In general if such knowledge is available and exploitable for the required purposes, it should be used. The problem is that, in chemical and many other industries, often this is not the case because the processes are too complex to be described by rigorous models in the form of mathematical or chemical equations. A prominent example of such process type is the biochemical processes such as penicillin production [7].

For this reason, the focus of this work is on multivariate data-driven soft sensors. These models rely on data modeling techniques and are trained on data collected during the operation of the process. The most common data-driven techniques applied for soft sensor modeling are the principal component regression (PCR), principal component analysis (PCA), and partial least squares (PLS) method, which will be explained later in the section dedicated to data-driven techniques. These techniques gained popularity due to their statistical background and ease of interpretability of the model and because they deal efficiently with data colinearity, which is common among industrial data sets. Examples of soft sensor applications based on PCA/PLS have been discussed in [8–11]. Multilayer Perceptron (MLP) [12] is another predictive technique widely applied as a soft sensor [13]. The popularity of the Layer Perceptrons (LPs) originates in their ability to model nonlinear functions. Furthermore, one can find a range of other techniques like support vector machines (SVM) [14] and neuro-fuzzy systems [15] applied for soft sensors [16,17].

While the field of application of soft sensors is broad, the most common application type is what is referred to as online prediction [2]. In the case of this application type, the soft sensor is trained using historical data before it is launched in the real-life environment where its task is to provide an online prediction on the basis of the incoming input data stream. Another frequent application type of soft sensors is their use for process monitoring and process fault detection. The systems can be either applied to describe/analyze the normal operating state or to recognize possible process faults before these states occur. Commonly, process monitoring techniques are based on multivariate statistical techniques like PCA or more precisely on Hotelling's T2 [18] and *Q*-statistics [19]. Compared to traditional univariate control charts, these measures have, on the one hand, the advantage of considering all input features and, on the other hand, providing information about the contribution of the particular features to a possible violation of the monitoring statistics [20]. Soft sensors providing a certain degree of transparency may also be very useful for the interpretation of the underlying process. Particularly useful are soft sensors based on PCA and PLS as these two methods deliver a measure of how much a given input variable contributes (1) to explaining the input data variance (in case of PCA) or (2) the output variable of interest (in case of the PLS). Soft sensors can also be developed to replicate hardware sensors. In such a case, the soft sensor predictions are used in case of failure or unavailability of the original hardware sensor, which may occur due to its maintenance, replacement, etc. [21].

A particular drawback of many of current soft sensors is their static nature [2]. Traditionally, the models are not adaptive, and once deployed into the real-life operation, the model does not change (see, e.g., [21,22]). In contrast to this, the environment, in which the soft sensors are applied, is often changing [23]. The combination of the static model and the changing data leads to performance deterioration because the model usually represents the out-of-date state of the process as it was observed during the training phase. In order to cope with this problem, one of the following two conditions has to be fulfilled: (1) the historical data have to contain all of the states, which may occur in the future, in this case the soft sensor can be trained to cope with any of the states to be expected in the future; and (2) if the

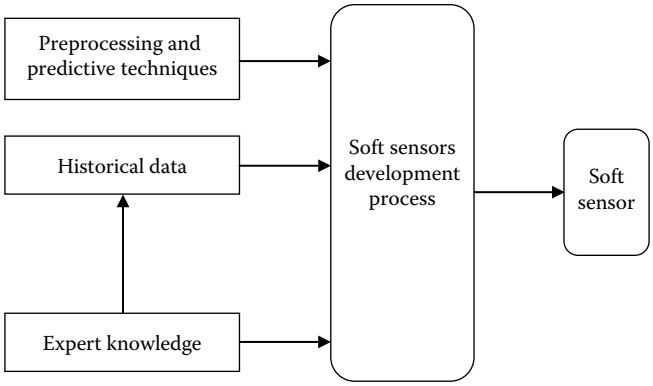


FIGURE 1.1 Soft sensor development process.

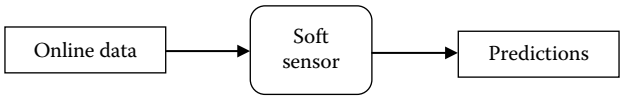


FIGURE 1.2 Soft sensor operation.

first condition cannot be met, it is necessary to equip the soft sensor with a self-adapting capability [24]. Examples of adaptive techniques and soft sensors based on these are as follows:

- Preprocessing and predictive techniques: The statistical and machine learning techniques for the data preprocessing and for the actual soft sensor. One preprocessing and two soft sensor techniques are presented in the section on data-driven technique.
- Historical data: The historical data for the training and parameterization of the preprocessing and the predictive methods. The characteristics and role of the historical data are in more detail discussed in the section on data for soft sensor development.
- Expert knowledge: The role of expert knowledge was briefly discussed in the introduction of this work. The golden rule for expert knowledge is that it has to be used for the soft sensor development when available.

All these three parts are input for the actual soft sensor development, as shown in Figure 1.1. A more detailed methodology for soft sensor development can be found in [2]. The outcome of this process is a trained predictive model—a soft sensor.

1.1.1 Soft Sensor Operations

Once developed, the soft sensor can be deployed into the real-life operation where its task is to deliver the predictions calculated measurement of the output variables as shown in Figure 1.2.

In the next section, the characteristics of the historical and online data required for the soft sensor development and operation are discussed in detail.

1.2 Data for Soft Sensor Development

1.2.1 Historical Data

Usually, when dealing with real-life industrial modeling tasks, there is a set of historical recordings available. This forms the basis for the development of the soft sensor. These data have a number of properties relevant for the soft sensor development process. The historical data consist of a number

of input variables, for example, hardware sensor measurements, and one or more target variables. Sometimes there can be a relevant delay between taking the measurement, that is, sampling the input variables, and providing the corresponding target value, that is, sampling the target variable. This is caused by the fact that the target variables may have to be obtained in a time-consuming manner (e.g., manual evaluation of quality measurement in a laboratory). Nevertheless, for the historical data, it can be assumed that the delays are compensated by entering the target values at the time point of taking the sample for the historical data.

Furthermore, although the sampling rate of the input data is usually higher than the one of the target variable, in most cases only the labeled samples (i.e., data points containing the target values) are recorded within the historical data. Therefore, the sampling rate of the input variables and the target variable can be assumed as equal.

Another relevant and possibly harming property of the data is the delays between the input variables themselves. Without prior process knowledge, these types of delays are difficult to compensate for and have to be dealt with by the application of an appropriate feature selection algorithm.

1.2.2 Online or Real-Time Data

Once the soft sensor building phase is finished, the model is applied in the online operation and needs to deal with the online data stream. In comparison to historical data, the online data have slightly different characteristics.

Data are arriving in an incremental way, that is, one sample (or a small batch of samples) after another. In general, the sampling rate between the input and the target data can differ.

The real values of the target measurement, which usually arrive at a lower sampling rate than the input samples and often with certain delays, can be used to evaluate the model performance during the online prediction phase. If there is a notable deterioration of the model performance, an adaptation of the model needs be performed using the target data. The characteristics of historical data and real-time data are summarized in Table 1.1.

1.2.3 Process Data Issues

Figure 1.3 shows examples of variables affected by common issues of industrial data that will be discussed further in this section.

1. *Missing values*: Missing data are single samples or consequent sets of them, where one or more variables (i.e., measurements) have a value that does not reflect the real state of the physical measured quantity. The affected variables usually have values like $\pm\infty$, 0.

   Missing values in industrial context have various causes. The most common ones are the failure of a hardware sensor, its maintenance or removal. As already mentioned, industrial plants are heavily instrumented for process control purposes; therefore, the recorded process data also consist of a large number of diverse variables. In such a scenario, there is a certain probability

TABLE 1.1 Characteristics of the Historical and Online Data

	Historical Data	Real-Time Data
Purpose	Model training and parameter optimization	Soft sensor operation and adaptation (if required)
Mode	Batch	Incremental
Input variable delays	Possibly present	Possibly present
Target variable delays	Compensated	Possibly present
Input vs. target sampling rate	Equal	Possibly lower

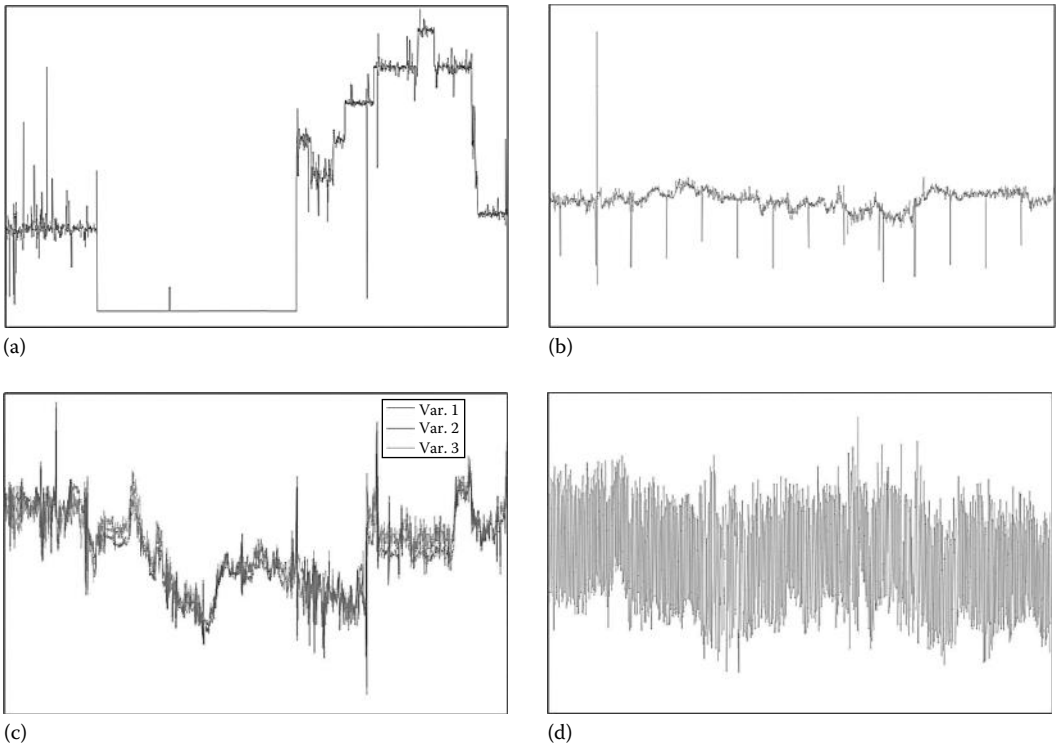


FIGURE 1.3 Common issues found in industrial data sets: (a) missing values; (b) outliers; (c) data colinearity; (d) measurement noise.

that some of the sensors occasionally will fail. A possible cause of missing data is related to the transmission of the data between the sensors and the database, errors in the database, problems in accessing the database, etc.

Since most of the techniques applied to data-driven soft sensing cannot deal with missing data, a strategy for their replacement usually has to be implemented. An approach, which is very primitive and not recommended but still commonly applied in practical scenarios, is to replace the missing values with the mean values of the affected variable. Another nonoptimal approach is to skip the data samples consisting of variable or variables with the missing values, that is, case deletion [25]. A more efficient approach to missing values handling takes into account the multivariate statistics of the data and thus makes the reconstruction of the missing values dependent on the other available variables of the affected samples (e.g., maximum-likelihood multivariate approach to missing values replacement [26]). These types of approaches are related to sensor fault detection and reconstruction. From another point of view, one can distinguish two different approaches for dealing with missing values [25]. These are (1) single imputation where the missing values are replaced in a single step (using, e.g., mean/median values) and (2) multiple imputation, which are iterative techniques where several imputation steps are performed.

In [27], an algorithm based on iteratively reweighed least squares is applied to deal with missing and noisy data. This algorithm is limited to the estimation of dynamic linear system parameters only. The authors show that the algorithm can deal with a situation where the probability of missing data is less than 50%, provided that a high number of samples are available.

Another technique for dealing with multiple imputation techniques for missing values handling was presented in [26,28].

2. *Data outliers*: Outliers are sensor values that deviate from the typical, or sometimes also meaningful, ranges of the measured values. One can distinguish between two types of outliers, namely, obvious outliers and nonobvious outliers [29]. Obvious outliers are those values that violate the physical or technological limitations. For example, the absolute pressure may not reach negative values or the flow sensor may not deliver values that exceed the technological limitations of the sensor. To be able to detect this type of outlier efficiently, the system has to be provided with the limiting values in the form of a priori information. In contrast to this, nonobvious outliers are even harder to identify because they do not violate any limitations but still do not reflect the correct variable states.

   Outlier detection as part of the data preprocessing remains very critical for soft sensor development because undetected outliers have a negative effect on the performance of the models. For example, the influence of a single outlier can be critical for the PCA [30–32]. Another problem of outlier detection is that even when applying automatic outlier handling preprocessing steps, usually the results have to be validated manually by the model developer. The goal of the manual inspection is to detect any possible outlier masking (i.e., false-negative detections, undetected outliers) and outlier swamping (i.e., false-positive detections, correct values labeled as outliers) [33].

   Typical approaches to outlier detection are based on the statistics of the historical data. The most simple approach is the 3_ outlier detection algorithm (see, e.g., [9,34]), which is based on univariate observations of the variable distributions. This method labels all data samples out of the range _(*x*) _ 3_(*x*), where _(*x*) is the mean value and _(*x*) the standard deviation of the variable *x*, as outliers. A more robust version of this approach is the Hampel identifier [35], which uses a more outlier resistant median and median absolute deviation (MAD) values [33,34] to calculate the limits.

   The influence of outliers on the identification of linear and nonlinear models is discussed in [33]. For the handled models, the Hampel identifier is found to be an effective approach for dealing with outliers. In [36], a moving window filter is combined with the Hampel identifier to obtain an outlier detection and removal system. In contrast to the univariate approaches, the multivariate methods use combinations of more features to detect the outliers. An example from this group based on the PCA is the Jolliffe parameter [37,38]. A two-stage outlier detection approach is discussed in [39]. The first stage is the application of the PCA, after this the Hotelling's T2 measure [18] can be used to detect outlier candidates that are located outside of the 99% confidence ellipse. These candidates are then further analyzed in the second step, where Scheffe's test [40] is applied to these points.
3. *Drifting data*: There are two types of drifting data, and dependent on the cause of the drifts, one can distinguish between system and sensor drifts. The causes of the system drift are the changes of the underlying system due to abrasion of equipment, aging of instruments, environmental conditions, etc. These are particularly prevalent in industries with a large number of mechanical elements that undergo steady abrasion during their operation. Another cause of system drifts can also be external influences like changing environmental conditions (e.g., weather influence), the purity of the input materials, and catalyst deactivation. These factors have not only an influence on the data but affect the process state as well. Therefore, the drifts should be recognized and reported, and appropriate actions should be taken to remove their source. This is different in the case of sensor drifts, which are caused by changes in the measuring devices and not by the process itself. The critical point is that this type of drifts, while still observed in the measured data, does not reflect any changes in the system. Therefore, in the case of sensor drifts, the action to be taken should be the recalibration of the measurement devices or the adaptation of the soft sensor.

   In terms of the effects of drifts on the data, one can observe changes in the means and variances of the single variables as well as changes of the correlation structure of the data [41].

   Distinguishing between the two different drift causes discussed is challenging and once again a lot of expert knowledge is needed in order to take appropriate action. Another challenging

aspect of dealing with drifting data is the fact that the changes may progress very slowly and may influence each other and thus leading to a nonlinearity.

The most common approach to deal with dynamics in the data is to apply the moving window technique. In this case, the model is updated on a periodic basis using only a defined number of the most recent samples. Some examples of the application of this technique in the context of soft sensor modeling were published in References [42–45].

4. *Data colinearity*: Another challenging issue for soft sensing is related to the structure of the data. Typically, the data measured in heavily instrumented industries are strongly colinear. These results from the partial redundancy in the sensor arrangement, for example, two neighboring temperature sensors, will deliver strongly correlated measurements. Such environments are often called data rich but information poor [46]. However, for soft sensing the requirements are different, in this case only informative variables are required. Anything else is unnecessarily increasing the input data dimensionality, which often has a negative effect on the model training and performance.

   There are two ways to deal with the colinearity problem. One way is by transforming the multiple input variables into a new reduced space with less colinearity as it is done in the case of the PCA and PLS. These two approaches are the most popular ones to deal with data colinearity in the process industry. Examples of applications where PCA is used are [9,43,47,48] and for the PLS [49,50,22]. Another way to handle colinearity is to select a subset of input variables that are less colinear. These approaches are summarized under the umbrella of variable (or feature) selection methods in the computational learning research. A general review of these methods is presented in [51]. Some feature selection methods in the context of soft sensing are also discussed in [38]. Among the discussed approaches are correlation- and partial correlation-based feature selection as well as Mallows' Cp statistics.
5. *Sampling rates and measurement delays*: Various sensors usually work at different sampling rates and thus one has to take care to synchronize them. The synchronization of the data is usually handled by the recording databases. The definition of such a threshold is another critical point, which influences the quality of the historical data. Soft sensors are often applied in multirate systems with several operating sampling rates. Such a scenario occurs in a system where some of the variables, usually critical for the process control, are evaluated in laboratories at a much lower sampling rate than the rest of the automatically measured data.

   Additional issues of the data are the process-related delays in the measurements. For instance, materials in the processes usually have a given run-time through the process (e.g., the dwell period within a reactor or distillation column), and thus, it is not reasonable to relate two different measurements taken at the same time at different locations within the process. Instead of this, the delays in the particular measurements should be compensated by synchronizing the variables. However, in order to perform the synchronization, an extensive knowledge about the system is required.
6. *Measurement noise*: Measurement noise is another common effect observed in industrial data. Most of the approaches to soft sensor development are trying to cope with measurement noise during the preprocessing stage of the data processing. This is achieved mainly by applying a smoothing (averaging) filter as a preprocessing step.

   The PCA is also a useful tool for dealing with measurement noise. As a least mean squares-based method, it can deal with measurement noise as long as it can be assumed as Gaussian noise, that is, randomly distributed with zero mean value [9]. In the same work the authors also highlight the application of robust statistics, that is, using the median instead of the mean operator and MAD instead of the standard deviation, for the normalization of noisy data.

   Zamprogna et al. have shown the robustness of the PLS method toward measurement noise in [22]. The authors have shown that there are only small changes of the prediction error of a PLS soft sensor with increasing noise levels. The explanation of this fact is that the noise influences mainly the higher-order latent variables that are normally skipped in practical application.

1.3 Data-Driven Techniques for Soft Sensor Development

This section describes the three most common methods for the development of soft sensors. The PCA can be either used as a data preprocessing method or in combination with a regression method as a full soft sensor, in such case the method is referred to as PCR. PLS is another popular method for soft sensing [2]. This method is in particular very useful for adaptive soft sensors. The third presented method is the artificial neural network (ANN) and in particular its popular form often used for nonlinear modeling called MLP.

1.3.1 Principal Component Regression

The PCA algorithm requires the matrix of the input data $X \in \mathbb{R}^{n \times m}$ (where n is the number of recorded time instances and m the number of input measurement streams). For the PCA, each measurement stream, that is, column of the matrix, has to be normalized to zero mean and unit variance:

$$\begin{aligned} &X' = \left(X - \mathbf{1}_n \mathbf{b}^T\right)\Sigma^{-1} \quad \text{with} \quad \mathbf{b} = \frac{1}{n} X^T \mathbf{1}_n \quad \text{and} \\ &\mathbf{1}_n = [1,1,\ldots,1]^T \in \mathbb{R}^{n \times 1}, \\ &\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_m) \end{aligned} \tag{1.1}$$

where σ_i is the standard deviation of the input variables. The scaling has to be done in order to assure equal influence of the input variables in the model. The normalized data X' is then transformed to 1D $l \leq m$ scores $T \in \mathbb{R}^{n \times l}$ using the following relation:

$$X' \in TP^T + E \tag{1.2}$$

where $P \in \mathbb{R}^{n \times m}$, $E \in \mathbb{R}^{n \times m}$ are the loadings and residuals matrices, respectively.

There are several ways to calculate the matrix P. In the case of the covariance approach, the correlation matrix C of the input data X has to be calculated:

$$C = \frac{1}{n-1}(X')^T X' \tag{1.3}$$

Next the eigenvalues and eigenvectors of this matrix are derived by solving

$$V^{-1} C V = eig(C) \tag{1.4}$$

where

$eig(C)$ is the diagonal eigenvalues matrix
V is the eigenvector matrix

The eigenvalues λ_i are then sorted in descending order such that $\lambda_1 > \lambda_2 > \lambda_3 > \ldots \lambda_m$. The columns of matrix P are then formed by the eigenvectors v_i corresponding to the l highest eigenvalues:

$$P = [v_1, v_2, \ldots, v_l] \quad \text{with} \quad v_i \in \mathbb{R}^{n \times l} \tag{1.5}$$

For the calculation of the eigenvalue decomposition, either the singular value decomposition (SVD) [37, p. 44], the nonlinear iterative partial least squares (NIPALS) algorithm [52] or other methods (see, e.g., [41]) can be used.

Having the matrix T, one can build a regression model using the LS algorithm:

$$\hat{y} = Ta \tag{1.6}$$

Applying the fact that the vectors in T are orthogonal, the previous equation can be simplified to

$$a = (T^T T)^{-1} T^T y = L^{-2} T^T y \quad \text{with} \quad L \in \mathbb{R}^{l \times l} \tag{1.7}$$

where L is a diagonal matrix with elements equal to $\sqrt{\lambda^i}$ (see Ref. [37, p. 168]).

1.3.2 Recursive Partial Least Squares

In this section, the most common soft sensing technique and its adaptive (recursive) version are outlined. The PLS method was originally proposed in [53]. The goal of the algorithm is to project the scaled and mean-centered input data X and output data $Y \in \mathbb{R}^{n \times p}$ (where p is the number of output variable streams that are supposed to be predicted) to separate latent variables:

$$X = TP^T + E \tag{1.8}$$

$$Y = UQ^T + F \tag{1.9}$$

where

$T \in \mathbb{R}^{n \times l}$ (with $l < m$ as the number of latent variables)
$U \in \mathbb{R}^{n \times l}$ is the score matrices
$P \in \mathbb{R}^{m \times l}$ and $Q \in \mathbb{R}^{p \times l}$ are the corresponding loading matrices
E and F are the input and output data residuals

The score matrices T and U consist of the so-called latent vectors:

$$T = [t_1, t_2, \ldots t_l] \quad \text{with} \quad u_i \in \mathbb{R}^{n \times l} \tag{1.10}$$

$$U = [u_1, u_2, \ldots u_l] \quad \text{with} \quad u_i \in \mathbb{R}^{n \times l} \tag{1.11}$$

The latent vectors, which are orthonormal to each other $\left(\text{i.e., } t_i^T t_j = 0 \ \forall_{i \neq j}\right)$, are a more compact description of the input data. The column vectors $p \in \mathbb{R}^{m \times l}$ and $q \in \mathbb{R}^{p \times l}$ of the loading matrices P and Q represent the contributions of the input and output variables to the latent vectors t and u, respectively. The previous two equations are also called the outer model of the PLS method [44].

Subsequently, the PLS method builds a regression model between the latent scores (the PLS inner model):

$$U = TB + R \tag{1.12}$$

where $B \in \mathbb{R}^{l \times l}$ is a diagonal matrix of regression weights that is calculated such as to minimize the regression residuals R. The estimates $\tilde{Y}$ of Y are consequently

$$\tilde{Y} = TBQ^T \tag{1.13}$$

There are several ways to calculate the required vectors *t*, *p*, *u*, *q*, and *b*. A particularly popular algorithm for the calculation of PLS is the NIPALS algorithm [52]. The NIPALS algorithm calculates one latent vector after the other in an iterative way. The explained covariance is removed from the data after the each iteration:

$$X_{i+1} = X_i - t_i p_i^T \tag{1.14}$$

$$Y_{i+1} = Y_i - u_i q_i^T \tag{1.15}$$

which is followed by the calculation of the next, that is, $(i + 1)th$ vectors for the PLS outer and inner models using the new data matrices X_{i+1} and Y_{i+1}. The number of calculated latent dimensions is usually established using cross-validation or some other parameter optimization techniques.

In the nonadaptive modeling scenario, the matrices *P, T, Q, U,* and *B* are calculated during the training phase based on the batch of historical data. However, as discussed in the introduction as well as in [44], this approach is often problematic because the process and its data are changing over the time. An advantage of the PLS algorithm is that it has the ability to incrementally integrate new data. The recursive partial least squares (RPLS) has been introduced in [54] and further clarified in [44]. It can be used to adapt the model in several ways: (1) on sample-by-sample basis, (2) by integrating a new batch of data, or (3) by merging two PLS models.

In the first case, which is exploited in this work, the model update is achieved by merging the old model represented by the matrices *P, B,* and *Q* with the latest data samples x_i, y_i:

$$X^{new} = \begin{bmatrix} \lambda P^T \\ x_i \end{bmatrix} \quad Y^{new} = \begin{bmatrix} \lambda B Q^T \\ y_i \end{bmatrix} \tag{1.16}$$

The forgetting factor λ defines the strength of the adaptation. The lower value this factor has, the smaller the influence of the previous model, which results in a faster adaptation to the new data. After the expansion, the PLS algorithm can be applied X^{new}, Y^{new} as usual. In order to be able to perform the previous updates, the number of latent variables has to be selected to be equal to the rank of X [44]. Practically, this condition can be assured by finding a number of latent variables $\boldsymbol{a}$, which fulfill

$$\|E_a\| \leq \varepsilon \vdots \quad \text{with} \quad \varepsilon \geq 0 \tag{1.17}$$

In general, the number of latent variables *a* required for the recursive operation can differ from the optimal number of latent variables required for the modeling.

1.3.3 Artificial Neural Networks

The original idea of ANN [12] was to build computational models motivated by the operation of biological neurons that are the basic information processing units in nervous systems. The task of both the biological and the artificial neurons is to collect signals at the inputs, to process this information, and then to output it. In case of soft sensors, the most commonly applied type of ANNs are the MLPs.

1. *MLP*: An MLP is a feed-forward ANN as shown in Figure 1.4. It consists of one input layer, one output layer, and at least one hidden layer. The role of the input layer is to collect the input data and provide it to the hidden layer for further processing. The number of units in the input layer is equivalent to the dimensionality of the input data. Each of the input units is connected to each

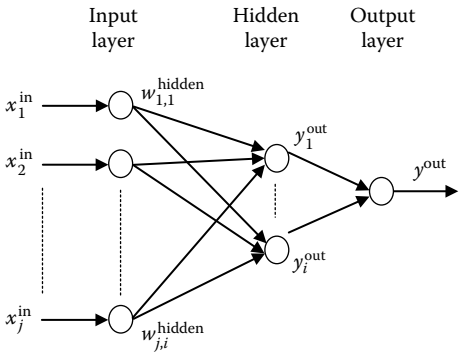


FIGURE 1.4 MLP structure.

hidden unit, and the connections between the units carry weights. The role of the hidden units is to collect the signals at their input, that is, the outputs of the preceding layer; multiply them by the connection weights; build a sum of them; and process them using the transfer function g^{hidden}:

$$x_i^{\text{hidden}} = g^{\text{hidden}}\left(\sum_j w_{j,i}^{\text{hidden}} x_j^{\text{in}}\right) \tag{1.18}$$

where

x^{out} is the output of the MLP
w_i^{out} is the weight between the ith hidden unit and the output unit
$g^{\text{out}}()$ is the transfer function of the output neuron

A remarkable theoretical property of the MLP is that they are universal function approximators, which means that provided enough training data and complex enough structure, they can be trained to approximate any function with any given accuracy [55].

Where x_j^{in} is the jth variable of the input sample, $w_{j,i}^{\text{hidden}}$ is the weight between the jth input unit and ith hidden unit and x_i^{hidden} the output of the ith hidden unit. The transfer function can be any differentiable nonlinear function. Often used are the sigmoid functions such as

$$g^{\text{hidden}}(x) = \frac{1}{1 + exp(-x)} \tag{1.19}$$

Once the signals are processed by the hidden layer, they are passed further to the next hidden layer, in the case when there is more than one hidden layer or to the output layer otherwise. The output layer can consist of one or more neurons. In the case of typical soft sensors, the output layer consists of only one unit:

$$x^{\text{out}} = g^{\text{out}}\left(\sum_i w_i^{\text{out}} x_{ij}^{\text{hidden}}\right) \tag{1.20}$$

Learning of the MLP (and ANN in general) is achieved by applying an algorithm for finding the optimal weights between the neurons. The most popular of these algorithms is the back-propagation algorithm originally proposed in [56].

1.4 Applications of Soft Sensors

The applications of soft sensors can be found across many fields of any industry. The most common application types are presented in the following sections of this work; please note that this is only short extract of the large number of applications of soft sensors. A comprehensive list of case studies can be found in [2,24,57].

1.4.1 Online Prediction

The most common application of soft sensors is the prediction of values that cannot be measured online using automated measurements. This may be for technological reasons (e.g., there is no equipment available for the required measurement), economical reasons (e.g., the necessary equipment is too expensive), etc. This often applies to critical values that are related to the final product quality. Soft sensors can in such scenarios provide useful information about the values of interest and in the case when the soft sensor prediction fulfills given standards, it can also be incorporated into the automated control loops of the process [58]. Data-driven soft sensors have been widely used in fermentation, polymerization, and refinery processes.

1.4.2 Process Monitoring and Process Fault Detection

Another application area of soft sensors is process monitoring. Process monitoring can be either an unsupervised learning or binary classification task. The systems can be either trained to describe/analyze the normal operating state or to recognize possible process faults [59–61]. Commonly, process monitoring techniques are based on multivariate statistical techniques like PCA, or more precisely on Hotelling's T2 [18] and *Q*-statistics [19]. These measures have, on the one hand, the advantage of considering all input features, that is, using multivariate statistics and, on the other hand, providing information about the contribution of the particular features to a possible violation of the monitoring statistics [20].

1.4.3 Hardware Sensor Backup

Soft sensors can also act as backup sensors for hardware sensors with a tendency to failure or with a requirement for periodic maintenance. In such a situation in order to prevent a disturbance of the whole system, soft sensors can be developed to replace the hardware sensors during their unavailability [21].

1.5 Conclusions

Undoubtedly, soft sensors will gain importance in the near future. For this reason, this work focused on the description of software-based sensors as a valuable alternative to hardware sensors. The greatest potential of soft sensors lies in situations where the application hardware sensor reaches its practical, physical, or economical limits. In such situations, soft sensors can often deliver better measurements. A prerequisite for the ability to deliver useful measurements is however the availability of high-quality and relevant data for the training of the soft sensor, which in practical applications often comes at the price of intensive manual data preprocessing only. This work described the process of soft sensor development and operation and discussed the main aspects that have to be considered during the development and application of a soft sensor. The most common techniques applied for soft sensors were also outlined in detail in this work.

Acknowledgment

Parts of the study presented in this chapter were carried out during the author's affiliation with Bournemouth University, United Kingdom.

References

1. P. Kadlec and B. Gabrys, Adaptive on-line prediction soft sensing without historical data, in *Proceedings of the IEEE International Joint Conference on Neural Networks*. Barcelona, Spain: IEEE, 2010.
2. P. Kadlec, B. Gabrys, and S. Strandt, Data-driven soft sensor in the process industry, *Computers and Chemical Engineering*, 33(4), 795–814, 2009.
3. R. O. Duda, P. E. Hart, and D. G. Stork, *Patten Classification*. New York: John Wiley & Sons, 2001.
4. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd edn., New York, Springer, 2009.
5. L. Fortuna, *Soft Sensors for Monitoring and Control of Industrial Processes*. London, U.K.: Springer Verlag, 2007.
6. W. Yan, H. Shao, and X. Wang, Soft sensing modeling based on support vector machine and Bayesian model selection, *Computers and Chemical Engineering*, 28(8), 1489–1498, 2004.
7. M. J. Arauzo-Bravo, J. M. Cano-Izquierdo, E. Gomez-Sanchez, M. J. Lopez-Nieto, Y. A. Dimitriadis, and J. Lopez-Coronado, Automatization of a penicillin production process with soft sensors and an adaptive controller based on neuro fuzzy systems, *Control Engineering Practice*, 12(9), 1073–1090, 2004.
8. D. Dong, T. J. McAvoy, and L. J. Chang, Emission monitoring using multivariate soft sensors, in *Proceedings of the American Control Conference*, Vol. 1, Seattle, WA, 1995.
9. B. Lin, B. Recke, J. Knudsen, and S. B. Jorgensen, A systematic approach for soft sensor development, *Computers and Chemical Engineering*, 31(5), 419–425, 2007.
10. H. Kaneko, M. Arakawa, and K. Funatsu, Development of a new soft sensor method using independent component analysis and partial least squares, *AIChE Journal*, 55(1), 87–98, 2009.
11. Y. Rotem, A. Wachs, and D. R. Lewin, Ethylene compressor monitoring using model-based PCA, *AIChE Journal*, 46(9), 1825–1836, 2000.
12. C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford University Press, 1995.
13. L. Fortuna, S. Graziani, and M. Xibilia, Comparison of soft-sensor design methods for industrial plants using small data sets, *IEEE Transactions on Instrumentation and Measurement*, 58(8), 2009.
14. V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.
15. J. S. R. Jang, C. T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing*. Upper Saddle River, NJ: Prentice Hall, 1997.
16. W. Zhong, D. Pi, and Y. Sun, SVM based soft sensor for antibiotic fermentation process, in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Vol. 1, Washington, DC, 2003.
17. J. J. Macias, P. Angelov, and P. X. Zhou, A method for predicting quality of the crude oil distillation, in *Proceedings of the International Symposium on Evolving Fuzzy Systems*, Ambleside, U.K., 2006, pp. 214–220.
18. H. Hotelling, The generalization of student's ratio, *The Annals of Mathematical Statistics*, 2(3), 360–378, 1931.
19. J. E. Jackson and G. S. Mudholkar, Control procedures for residuals associated with principal component analysis, *Technometrics*, 21(3), 341–349, 1979.
20. S. W. Choi, E. B. Martin, A. J. Morris, and I. B. Lee, Adaptive multivariate statistical process control for monitoring time-varying processes, *Industrial and Engineering Chemistry Research*, 45, 3108–3118, 2006.
21. G. D. Gonzalez, Soft sensors for processing plants, in *Proceedings of the Second International Conference on Intelligent Processing and Manufacturing of Materials (IPMM'99)*, Honolulu, HI, 1999.
22. E. Zamprogna, M. Barolo, and D. E. Seborg, Development of a soft sensor for a batch distillation column using linear and nonlinear PLS regression techniques, *Control Engineering Practice*, 12, 917–929, 2004.

23. N. Gallagher, B. Wise, S. Butler, D. White, and G. Barna, Development and benchmarking of multivariate statistical process control tools for a semiconductor etch process: Improving robustness through model updating, in *IFAC ADCHEM'97*. Banff, Alberta, Canada: Citeseer, 1997, pp. 78–83.
24. P. Kadlec, R. Grbić, and B. Gabrys, Review of adaptation mechanisms for data-driven soft sensors, *Computers and Chemical Engineering*, 35(1), 1–34, 2011.
25. J. Scheffer, Dealing with missing data, *Research Letters in the Information and Mathematical Sciences*, 3(1), 153–160, 2002.
26. B. Walczak and D. L. Massart, Dealing with missing data: Part ii, *Chemometrics and Intelligent Laboratory Systems*, 58(1), 29–42, 2001.
27. J. M. Chen and B. S. Chen, System parameter estimation with input/output noisy data and missing measurements, *IEEE Transactions on Signal Processing*, 48(6), 1548–1558, 2000.
28. B. Walczak and D. L. Massart, Dealing with missing data part i, *Chemometrics and Intelligent Laboratory Systems*, 58(1), 15–27, 2001.
29. S. J. Qin, Neural networks for intelligent sensors and control—Practical issues and some solutions, *Neural Systems for Control*, Academic Press, San Diego, CA, Chapter 8, 213–234, 1997.
30. B. Walczak and D. L. Massart, Robust principal components regression as a detection tool for outliers, *Chemometrics and Intelligent Laboratory Systems*, 27(1), 41–54, 1995.
31. I. Stanimirova, M. Daszykowski, and B. Walczak, Dealing with missing values and outliers in principal component analysis, *Talanta*, 72(1), 172–178, 2007.
32. S. Serneels and T. Verdonck, Principal component analysis for data containing outliers and missing elements, *Computational Statistics and Data Analysis*, 52(3), 1712–1727, 2008.
33. R. K. Pearson, Exploring process data, *Journal of Process Control*, 11(2), 179–194, 2001.
34. R. K. Pearson, Outliers in process modeling and identification, *IEEE Transactions on Control Systems Technology*, 10(1), 55–63, 2002.
35. L. Davies and U. Gather, The identification of multiple outliers, *Journal of the American Statistical Association*, 88(423), 782–792, 1993.
36. P. H. Menold, R. K. Pearson, and F. Allgower, Online outlier detection and removal, in *Proceedings of the Seventh Mediterranean on Control and Automation (MED99)*, Haifa, Israel, 1999, pp. 1110–1133.
37. I. T. Jolliffe, *Principal Component Analysis*. New York: Springer, 2002.
38. K. Warne, G. Prasad, S. Rezvani, and L. Maguire, Statistical and computational intelligence techniques for inferential model development: A comparative evaluation and a novel proposition for fusion, *Engineering Applications of Artificial Intelligence*, 17(8), 871–885, 2004.
39. G. D. Gonzalez, M. Orchard, J. L. Cerda, A. Casali, and G. Vallebuona, Local models for soft-sensors in a rougher flotation bank, *Minerals Engineering*, 16(5), 441–453, 2003.
40. E. Gomez, H. Unbehauen, P. Kortmann, and S. Peters, Fault detection and diagnosis with the help of fuzzy-logic and with application to a laboratory turbogenerator, in *Proceedings of the Thirteenth IFAC World Congress*, Vol. N, San Francisco, CA, 1996, pp. 175–180.
41. W. Li, H. H. Yue, S. Valle-Cervantes, and S. J. Qin, Recursive PCA for adaptive process monitoring, *Journal of Process Control*, 10(5), 471–486, 2000.
42. X. Wang, U. Kruger, and G. W. Irwin, Process monitoring approach using fast moving window PCA, *Industrial and Engineering Chemistry Research*, 44(15), 5691–5702, 2005.
43. L. Zhao and T. Chai, Adaptive moving window MPCA for online batch monitoring, in *Proceedings of the Fifth Asian Control Conference*, Vol. 2, Melbourne, Victoria, Australia, 2004.
44. S. J. Qin, Recursive PLS algorithms for adaptive data modeling, *Computers and Chemical Engineering*, 22(4–5), 503–514, 1998.
45. B. S. Dayal and J. F. MacGregor, Recursive exponentially weighted PLS and its applications to adaptive control and prediction, *Journal of Process Control*, 7(3), 169–179, 1997.
46. D. Dong and T. J. McAvoy, Nonlinear principal component analysis—Based on principal curves and neural networks, *Computers and Chemical Engineering*, 20(1), 65–78, 1996.

47. M. Amazouz and R. Pantea, Use of multivariate data analysis for lumber drying process monitoring and fault detection, in *Proceedings of the International Conference on Data Mining* (*ICDM*), S. F. Crone, S. Lessmann, and R. Stahlbock (eds.), Hong Kong, China, 2006, pp. 329–332.
48. S. Wang and J. Cui, Sensor-fault detection, diagnosis and estimation for centrifugal chiller systems using principal-component analysis method, *Applied Energy*, 82(3), 197–213, 2005.
49. O. Marjanovic, B. Lennox, D. Sandoz, K. Smith, and M. Crofts, Real-time monitoring of an industrial batch process, *Computers and Chemical Engineering*, 30(10–12), 1476–1481, 2006.
50. H. Zhang and B. Lennox, Integrated condition monitoring and control of fed-batch fermentation processes, *Journal of Process Control*, 14(1), 41–50, 2004.
51. I. Guyon and A. Elisseeff, An introduction to variable and feature selection, *Journal of Machine Learning Research*, 3(7–8), 1157–1182, 2003.
52. P. Geladi and B. Kowalski, Partial least-squares regression: A tutorial, *Analytica Chimica Acta*, 185, 1–17, 1986.
53. H. Wold, Nonlinear estimation by iterative least squares procedures, in *Research Papers in Statistics*. Festschrift for J. Neyman, F. David (eds.), Wiley, 1966, pp. 411–444.
54. H. E. Helland, Recursive algorithm for partial least squares regression, *Chemometrics and Intelligent Laboratory Systems*, 14(1–3), 129–137, 1992.
55. K. Funahashi, On the approximate realization of continuous mappings by neural networks, *Neural Networks*, 2(3), 183–192, 1989.
56. P. J. Werbos, Beyond regression: New tools for prediction and analysis in the behavioral sciences, PhD dissertation, 1974.
57. Y. Dote and S. J. Ovaska, Industrial applications of soft computing: A review, *Proceedings of the IEEE*, 89(9), 1243–1265, 2001.
58. K. Fujiwara, M. Kano, S. Hasebe, and A. Takinami, Soft-sensor development using correlation-based just-in-time modeling, *AIChE Journal*, 55(7), 1754–1765, 2009.
59. V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, A review of process fault detection and diagnosis part I: Quantitative model-based methods, *Computers and Chemical Engineering*, Elsevier, 27(3), 293–311, 2003.
60. V. Venkatasubramanian, R. Rengaswamy, and S. N. Kavuri, A review of process fault detection and diagnosis part II: Qualitative models and search strategies, *Computers and Chemical Engineering*, Elsevier, 27(3), 313–326, 2003.
61. V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin, A review of process fault detection and diagnosis part III: Process history based methods, *Computers and Chemical Engineering*, Elsevier, 27(3), 327–346, 2003.

2

Implantable Systems

Vincenzo Luciano
University of Brescia

Emilio Sardini
University of Brescia

Alessandro Dionisi
University of Brescia

Mauro Serpelloni
University of Brescia

Andrea Cadei
University of Brescia

2.1 Introduction

Nowadays, with the rapid development of bioengineering, implantable medical devices are being widely employed in treating human ailments. These devices, which are implantable in the human body, include pacemakers, smart prostheses, drug pumps, cochlear implants, implantable defibrillators, neurostimulators, bladder stimulators, nerve stimulators, and diaphragm stimulators. They perform a real-time control and/or monitoring of several physiological parameters.

Over 6 million heart patients worldwide have implanted pacemakers and about 150,000 pacemakers are being surgically implanted every year in the United States alone (Wang, 2003). These pacemakers, and defibrillators, help in treating several heart conditions including fibrillation, atrial and ventricular tachyarrhythmia, and bradycardia (Soykan, 2002).

These monitoring systems can communicate wirelessly with an external readout unit. By means of this architecture, the use of transcutaneous wires and the risk of infection can be avoided and necessary data can be collected easily. Wireless communication allows regular monitoring of several physiological parameters like the blood pressure, heart rate, and body temperature of patients during their daily activities at home (Tamura et al., 2011; Valdastri et al., 2004). These data can be collected and sent to a remote medical unit via Internet for the medical staff to access and assess and then assist the patient accordingly (Halperin et al., 2008).

As shown in Troyk et al. (2007), myoelectric sensors are another example of an implantable medical device. They are composed of several myoelectric electrodes implanted in the residual limb of patients with a prosthetic arm. These electrodes detect intramuscular myoelectric activity and provide signals that are elaborated to control the prosthesis.

Typically, implantable medical devices are powered by batteries, which is a severe limitation. In most cases, because of battery discharge, an implantable system will need battery replacement through a surgical operation. For example, although an implanted defibrillator lasts about 10 years, its battery

must be substituted after about 4 years 7 months (Wei and Liu, 2008). Hence, the patient must undergo a surgical operation that will not only cause physical and mental pain to the patient, but will also put an economic burden both on the patient and on the national health system. As stated earlier, the battery defines the lifespan of the implantable medical device. To avoid this problem, the implantable device must be powered using a telemetric technique or an energy-harvesting system that is implanted together with the device.

In the telemetric technique, an inductive powering system composed of two coils is used; the first coil is placed outside the body while the second one is implanted. The primary coil produces a magnetic field that is harvested by the secondary coil. In this way, energy is transferred through the human body. Using this technique, the circuits of the implanted medical devices are powered wirelessly, avoiding transcutaneous wires.

Examples of devices using telemetric techniques, transferring power inductively, are reported by Morais et al. (2009), Silay et al. (2011), and Riistama et al. (2007). Morais has described a smart hip prosthesis for the measurement of the joint forces and the temperature distribution in the prosthesis; the inductive link is performed through a coil placed in the stem of the insert. An inductive power system for cortical implant is reported in (Silay et al., 2011). The device mentioned is composed of a Class-E power amplifier, a matching network, and a rectifier, and two coils and is embedded in a biocompatible packaging that can be placed in a cavity of the skull. Another example of inductive power system is mentioned in Riistama et al. (2007), where an implant for the measurement of the electrocardiogram, with an operational range of approximately 16 mm, is described.

The telemetric technique requires the two coils, external and internal, to be close to each other, and that restricts the patient's movements. Furthermore, the operating frequency must be compatible with the tissue's absorption level; in particular, the power received by the implanted device, the tissue's absorption level, and the consequent warming of tissues are related to the radiation frequency. Hence, incorrect radiation frequency and energy absorption of the tissues can cause not only transmission problems, but also biocompatibility and safety issues. Table 2.1 shows maximum permissible exposure values of human tissues to magnetic, electric, and power density fields as suggested by the IEEE Standards for Safety Levels. As shown in the table, the magnetic-field strength and power density decreases as the radiation frequency increases. Ideal operating frequency for data and power transfer in telemetric systems is 125 kHz (Crescini et al., 2011).

Energy-harvesting technique, as mentioned earlier, is an alternative technique to power an implantable medical device; this technique, which is being researched and developed, will help avoid the use of a battery and, thereby, escape a battery replacement surgery, and can also allow the patient to move freely. Energy harvesting is a process by which energy is captured and stored from the ambient, in this case the human body. As reported by Starner (1996), the human body is a rich reservoir of energy, and the energy values reported are more than sufficient to supply for an implantable medical device. In Starner (1996),

TABLE 2.1 Maximum Permissible Values of Magnetic Field, Electric Field, and Power Density Field for Human Tissue

Frequency Range (MHz)	Magnetic Field Strength (H) (A/m)	Electric Field Strength (E) (V/m)	Power Density (S) E-Field, H-Field (mW/cm^2)
0.003–0.1	163	614	(100, 1,000,000)
0.1–3.0	$16.3/f$	614	$(100, 10{,}000/f^2)$
3–30	$16.3/f$	$1842/f$	$(900/f^2, 10{,}000/f^2)$
30–100	$16.3/f$	61.4	$(1, 10{,}000/f^2)$
100–300	0.163	61.4	1.0
300–3,000	—	—	$f/300$
3,000–15,000	—	—	10

Source: IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz, C95.1-1999.

the authors have calculated that energy of about 390 MJ is stored by an individual weighing 68 kg and with 15% body fat; this energy is converted into mechanical energy through muscles and is partially used up during daily activities. The paper reports the power utilized for some of the common activities; for example, the power spent for walking, upper limb motion, and breathing are approximately 67, 35, and 1 W, respectively.

The exploitable energy in the human body is classified into three different forms: thermal energy, mechanical energy, and chemical energy.

The thermal energy is harvested through a thermoelectric generator, which exploits the Seebeck effect. A thermoelectric generator is typically composed of several thermocouples, and it produces a voltage proportional to the thermal gradient across the thermoelectric generator itself. In Stark (2006), a compact thermoelectric generator composed of over 5000 thermocouples is described; this device has a volume less than 1 cm^3 and is able to produce 120 μW of power at 3 V with a thermal gradient of 5 K. Hence, an implantable medical device can be powered with one or more of these generators connected in series or in parallel. For the thermoelectric generator, obtaining a high thermal gradient in the human body is difficult; the highest thermal gradient occurs near the skin surface (Yang et al., 2007).

The chemical energy is harvested by means of microbiofuel cells, which are composed of a cathode, an anode, and an electrolyte. An example of an implantable glucose biofuel cell is described in von Stetten et al. (2006); the proposed device is composed of a hydrogel membrane that separates the electrodes realized with activated carbon. Following a series of in vitro tests, this fuel cell is able to generate a power density and a peak power of 2 μW/cm^2 and 20 μW, respectively, for a period of 7 days.

Mechanical energy is the most available and the most easily exploitable energy source in the human body; it is usually harvested with electrostatic, electromagnetic, and piezoelectric transducers. Some transducers, especially the electromagnetic and the electrostatic ones, can be modeled with a mass spring damper system working in resonance conditions with the input motion.

In a generic electromagnetic harvester, the human movement produces a displacement between a coil and a permanent magnet; this displacement generates a time-variable magnetic flux and, according to Faraday's law, an induced voltage on the coil. The permanent magnet and the coil can be placed in the resonant structure to maximize the magnet–coil relative motion, the induced voltage, and the stored energy by the harvester. An electromagnetic transducer for powering a pacemaker is presented in Goto et al. (1998); the heart muscle contractions generate a relative motion between a movable permanent magnet and a fixed coil. With a heart frequency between 0.5 and 2 Hz, the harvester is able to produce a maximum power of 200 μW, and it allows powering a pacemaker without battery. An electromagnetic harvester implanted in a hip prosthesis is reported in the work of Morais et al. (2011). The transducer is a resonant structure, composed of two external fixed coils and a Teflon tube in which a magnet swings during walking or whatever other activity, produces a hip movement. In Nasiri et al. (2011), a resonant electromagnetic transducer implantable in the diaphragm muscle is described. As the diaphragm muscle works continuously, the proposed transducer harvests energy even if the person is sleeping.

An electrostatic harvester is typically based on a variable precharged capacitor composed of a moving plate whose movement is produced by the activity of the human body. Examples of electrostatic transducers are reported in several studies (Miao et al., 2004; Tashiro et al., 2002). In Tashiro et al. (2002), the proposed transducer, composed of a honeycomb structure, exploits the heart muscle motion to power a cardiac pacemaker. In Miao et al. (2004), a MEMS electrostatic transducer is described; it produces a power of 24 μW at an input mechanical frequency of 10 Hz.

In a piezoelectric transducer, the movement of any body part can be exploited to deform a piezoelectric material in order to produce a voltage. Some piezoelectric transducers use a resonant structure to maximize the piezoelectric material deformation. In several studies (Almouahed et al., 2010, 2011a,b; Lahuec et al., 2011), a knee monitoring system embedded in a total knee prosthesis is presented. Four piezoceramics are placed in the tibial plate to measure the tibiofemoral forces and contemporaneously to harvest energy. During a single gait cycle, a total power of 7.2 mW is produced (1.8 mW for every piezoceramic). In Potkay et al. (2008), a blood pressure sensor is described.

The proposed device, composed of an arterial cuff in a thin piezoelectric film, converts the artery contraction/expansion into electric energy.

The design of energy harvesters for implantable medical devices is usually more complicated than that for industrial applications (Mitcheson, 2010). First of all, an energy harvester for an implantable medical device must have limited size. Furthermore, for mechanical transducers, the matching between the resonance frequency and the input mechanical frequency is difficult to obtain because the resonance frequency and the transducer size are, in general, inversely proportional. In particular, motion frequencies of the human body are generally less than 10 Hz (Romero et al., 2009); to lower the resonance frequency, using a linear spring-mass-damper system, a reduction in stiffness or an increase in mass is required, but, as previously stated, implantable devices must have limited dimensions; hence, it is difficult to satisfy both the requirements. In summary, from these considerations, on lowering the resonance frequency the transducer size increases, making the system unsuitable for applications in the human body. A material with nonlinear elastic characteristics can be used to solve this problem (Raman et al., 2010). Another important point is that an implantable medical device must be realized with biocompatible materials; otherwise, the different elements of the implantable medical device must be placed and fully sealed in a biocompatible material packaging. Two examples of architectures for batteryless implantable medical devices are presented in this chapter. In particular, a telemetric technique and an energy-harvesting system are described.

2.2 Architecture

Implantable systems can be defined as devices that execute their measurement functions in the human body autonomously. They are characterized by an autonomous power supply capable of measuring and transmitting data from inside the human body to a readout unit placed outside. They normally consist of a sensor module, conditioning electronics, a transmission module and a powering system. A block diagram showing the main elements of an implantable system is found in Figure 2.1. The common characteristics of each element are very low-power design, standalone configuration, minimal control and elaboration circuits resulting in less use of the microprocessor and power consumption, and minimal communication circuits, which require software with shorter and a more streamlined protocol, simple and quick to run with a low-power microprocessor.

An implantable system must be biocompatible with the tissues and cells of the environment in which it works. This requires using materials that are biocompatible to embed the sensor and electronic circuits inside a component of the prosthesis that is already biocompatible. Other requirements are the dimensions and frequency of the electromagnetic waves travelling through the human body. The antenna for the transmission must be small, for a telemetric system (Crescini et al., 2011) on the order of centimeters, and the transmission module must have frequencies that allow sending or receiving data through the human body.

Sensors and electronic blocks depend on the quantity to be measured. The supply block can be composed of a battery but, as mentioned in the introduction, a surgical operation will be required to substitute the battery. In the literature, other power sources are proposed, such as harvesting modules and inductive links. Each of these solutions determines a specific composition of the transmission and supply modules.

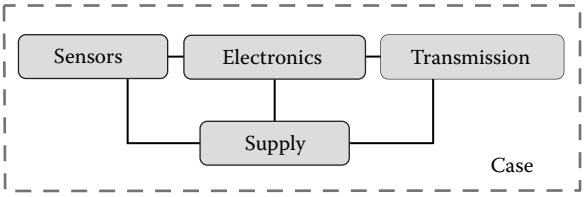


FIGURE 2.1 Block diagram of an implanted system.

The analysis of implantable systems has led to the definition of a classification according to the type of architecture; one class is "telemetric systems" and the other is "self-powered systems." "Telemetric systems" are defined as those that are powered inductively and interrogated wirelessly by a readout unit. "Self-powered systems" are those that have a power-harvesting module that scavenge energy for the functioning of the system from the environment. In the next section, the general architecture of telemetric and self-powered systems is described and discussed.

2.2.1 Telemetric Systems

The general architecture of a measurement system for telemetric systems is represented in Figure 2.2.

The telemetric system is composed of an implanted unit and a readout unit. The sensor is a block in the implanted unit inside the human body, while the readout unit is placed outside and the communication between the two units is done with telemetric techniques. A block diagram of the implanted unit and readout unit is found in Figure 2.2. It consists of different modules: a low-power sensor that measures the quantity under interest; a low-power microcontroller for the analog to digital conversion of the data, the storage in memory, and the telemetric operation; a transponder that transfers the data collected to the readout unit. The two elements are connected by wireless communication exploiting an electromagnetic field at typical frequencies of about 125 kHz (Crescini et al., 2011). The coil, connected to the transponder of the implanted module, is coupled to the external one, receiving the power

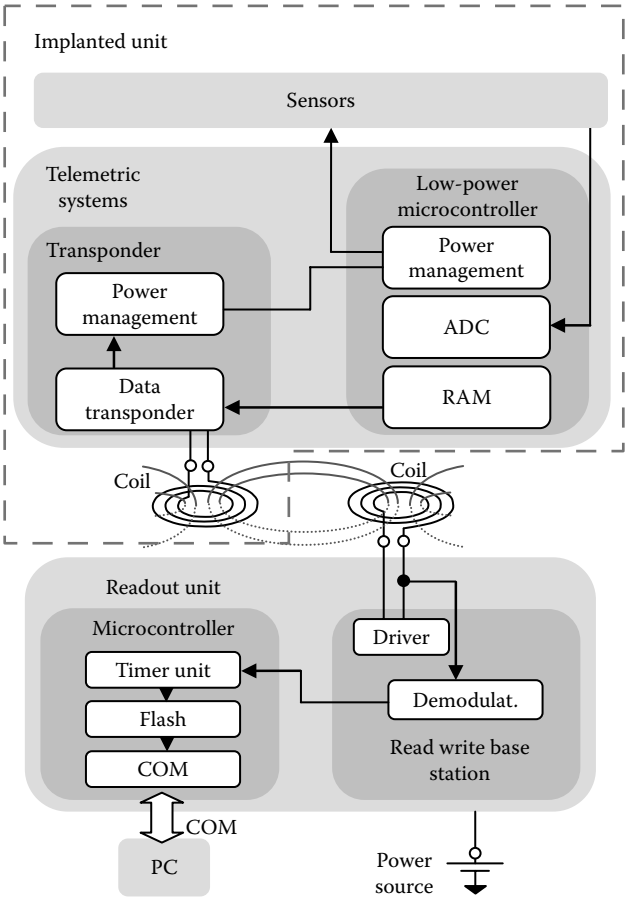


FIGURE 2.2 Block diagram of a telemetric system.

for the communication of the measurement data. Usually, the readout coil is bigger than the coil of the implanted unit. This is due, on the one hand, to the need to occupy a small space inside the human body and, on the other hand, to the difficulty of properly coupling the two windings. The readout coil therefore allows for having a greater area related to the coupling of the magnetic field, but at the expense of the efficiency of power transfer. To transmit data, the transponder of the implanted unit modulates the magnetic field using a damping stage. It modulates the coil voltage by varying the coil's load. A high level ("1") increases the current into the coil and damps the coil voltage. A low level ("0") decreases the current and increases the coil voltage. However, the current through the coil is never zero, so as to continuously provide the power supply. In particular, amplitude modulation is typical and so is the Manchester code. The transponder interface can also receive data: the readout unit modulates the emitted field with short gaps, and then a gap-detection circuit in the implanted unit reveals these gaps and decodes the signal. Furthermore, as specified earlier, the readout unit generates the power supply, which is handled via an electromagnetic field and the coil antenna of the transponder interface; then the voltage across the coil is rectified and managed by the power management module to generate a rectified voltage and current for the functioning of the electronic circuit. The power to all the internal modules of the implanted unit is supplied by the energy transmitted with the electromagnetic field generated by the read/write base station. The low-power microcontroller can include an ADC or a timer unit to measure the sensor signals. The microcontroller has a volatile memory to save the data before the transmission and timer units that permit to synchronize data transmission. To maintain low-power consumption, the bus frequency should be low, the ADC and transmitting unit have a low-power configuration, and all the unused peripherals should be switched off.

Since telemetric systems are wireless devices, transmitting not only the data but the power as well, the covered distance between the wireless device and the collecting data system must be short. The maximum transmitting distance depends on different factors; in Dalola et al. (2009) a maximum distance of about 8 cm is reported for open field transmission. For these reasons and for energy saving, point-to-point communication must be implemented. Point-to-point communication avoids the integration in the implanted system of circuits to manage the complexity of a network protocol and avoids complex communications such as those on multiple nodes that involve more complex software and therefore a longer time of execution of the same software, saving power supply and making the system compatible with the available low energy. Furthermore, the readout coil must be present and active during the measurement, conditioning, and transmitting phases; this means that the external coil, sometimes uncomfortable for the patient, must be placed close to or around the human body.

The external readout unit usually consists of a read/write base station able to supply power to the transponder driving the coil antenna and to demodulate the digital signal from the implanted unit. The readout unit is supplied by a line voltage, and no low-power characteristics are required, so the microcontroller, the bus frequency, and peripherals have no functioning limits. A timer unit is used to decode the demodulated signal, and the data collected are transferred to a personal computer (PC) using a serial communication interface (SCI).

2.2.2 Power-Harvesting Systems

A general architecture of a measurement system for passive autonomous sensors is shown in Figure 2.3.

Since the possibility of substituting batteries with the harvesting system is ecologically attractive and avoids surgical operations, our analysis shows that a self-powered system equipped with a harvesting system is a viable solution for implanted systems. These self-powered systems consist of one or more sensing elements and different modules: front-end electronics, an analog-to-digital converter, an elaboration unit to manage the internal tasks, a power management circuit, a wireless transceiver, and storage memories. In Figure 2.3, a block diagram of a typical self-powered system can be found. The power-harvesting module, usually separated from the circuit board, collects the energy present in

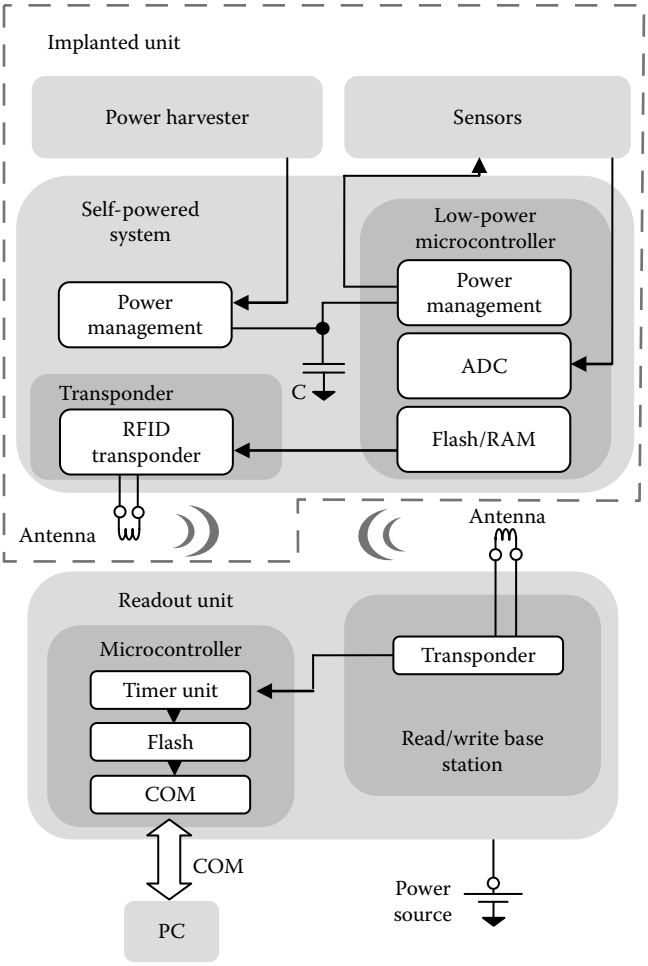


FIGURE 2.3 Block diagram of a self-powered system.

the environment of measurement in the form of mechanical energy, temperature difference, etc., and transforms it into electric energy.

The power-harvesting module must comply with very specific constraints, not only of space but also the compatibility of the materials and method of operation. For example, a system of power harvesting that uses the mechanical energy due to the movement of the human body cannot be excessively large; however, it must be able to operate with very low-frequency vibrations, of the order of a few hertz. The power harvester is connected to a power management block that is very important and essential. Since the voltage and current levels of the electronic circuits do not currently meet the possibility offered by the power harvesting system or sometimes even by batteries, management of the power supply is essential. Self-powered systems require a specific level of voltage and current obtainable by an appropriate power management block. This block commonly consists of dedicated circuits for the conditioning and/or storage of the energy harvested. First, a specific circuit can be used for matching the output electric impedance of the generator with the characteristics of the circuit load in order to have the maximum power transfer. Then, usually the power management circuit has a dedicated DC–DC converter or charge pump to provide a specific level of voltage and current at the circuit load.

The low-power microcontroller controls the sensor interface circuit, configures the front-end electronics and converts the data coming from the sensor interface circuit, and stores it in a nonvolatile

memory or directly sends the data at the transponder for communication. So, in this architecture two different strategies can be implemented. The first one, with a nonvolatile memory, saves the measurement data in the implanted unit and does not lose the data even when the device is not powered. This means that the external readout unit is not necessary during the measuring and saving phases, increasing the possible applications and the comfort for the patient, who does not need to "wear" the readout unit constantly. All the collected data can be downloaded in a second's time using different methods. The second strategy regards the possibility of measuring when enough power is scavenged and uses it to measure and transmit the data outside. In this configuration, the nonvolatile memory is not needed and the power for saving the data on the memory is not required. This means that the data, when the readout unit is not closed to the implanted unit, can be lost. Moreover, specific applications can be implemented to measure only when a specific event happens or only when requested. This leads to power savings and avoids the loss of data. In order to reduce the energy consumption for the data transmission, some smart compression algorithms on the measurement data can be implemented as well. In fact, the system deploys strategies to reduce the power consumption; the sensor module is designed to be triggered to transmission only when required, thereby consuming less power because unnecessary transmissions are avoided.

2.3 Force Measurement inside Knee Prosthesis

In this section, an example of an implantable sensor that monitors a total knee prosthesis (TKP) is described. It measures the forces applied on the knee prosthesis and exploits a telemetric technique for data communication and power supply (Crescini et al., 2011). Monitoring the prosthesis with an implantable sensor is very important in biomedical application. This provides several advantages such as the analysis of the wear conditions of the prosthesis caused by incorrect use or placement, the data collection to improve future design of the prosthesis, and a better control of the patient rehabilitation. Moreover, measurement devices for tibiofemoral contact stress give precise knowledge about articular movement behavior, and they can be used to refine surgical instrumentation, guide postoperative physical therapy, and detect human activities that can overload the implant. Therefore, monitoring the forces on the TKP it is possible to ensure the lifetime of the TKP is greater than that available with the current prosthesis.

The TKP consists of a tibial component and a femoral component, both made up of metal alloy. These components are attached to bone by using acrylic cement, and between the two components an ultra high molecular weight polyethylene (UHMWPE) insert has been embedded. The implantable sensor should be placed into the polyethylene insert avoiding biocompatibility problems.

The implantable sensor measures the forces applied on the knee insert by three magnetoresistive force transducers, which consist of magnetoresistors and permanent magnets as shown in Figure 2.4. The output resistance of each sensor depends on the distance between the permanent magnet and the magnetoresistor. Two magnetoresistors are placed in the areas where the two condyles of the femoral component transmit the forces between femur and tibia, and the third magnetoresistor is placed in the central part of the insert where the forces generated have no significant effect. The third magnetoresistor works as dummy for temperature compensation operation.

The force applied by the femoral component generates a deformation of the polyethylene insert, changing the distance between the magnet and the magnetoresistors and causing a resistance variation. The relationship between the force applied to the insert and the resistance variation of the magnetoresistors' output has been experimentally evaluated:

$$F \cong 2.15 \cdot (R - R_0) \tag{2.1}$$

Equation 2.1 shows the linear relationship between the force applied and the resistance measured, where F (N) is the force applied to the polyethylene insert, R (Ω) is the resistance value associated to the insert

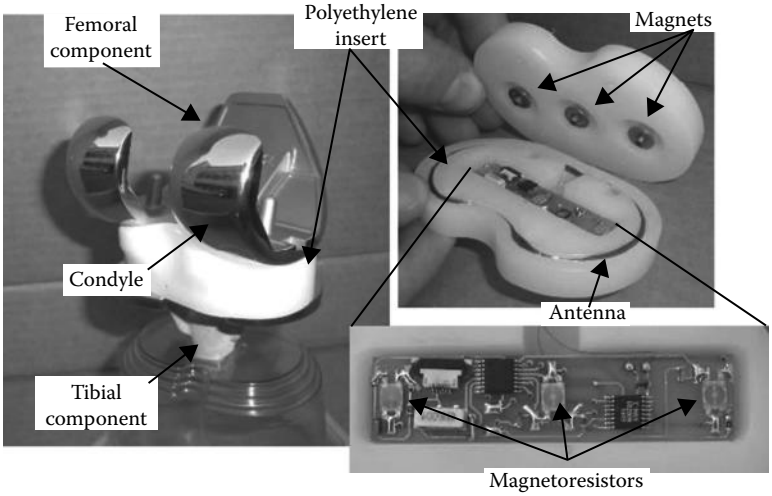


FIGURE 2.4 Example of total knee prosthesis and the implantable sensor into a polyethylene insert of the TKP.

deformation, and $R_0 = 1.25$ kΩ is the resistance value when the force applied is equal to 0 N. In this way, it has been possible to measure the forces applied to the insert by resistance measurements.

A significant issue with the implantable sensors is the relationship between the output of the sensors and the temperature. In fact, it is possible that an increase in temperature of about 3 °C, when the sensor is inside the human body, can be reached when a person walks for 45 min (Graichen et al., 1999). In this particular example, the magnetoresistors have a resistance drift equal to about 150 Ω/°C. For this reason, it is important to compensate the thermal drift so as to obtain an accurate measurement equation.

The control circuit of the implantable sensor is composed of a low-power microcontroller to acquire data with a 12 bit ADC, a 128 kB flash memory to store the force data, and a transponder working at frequency of 125 kHz. Figure 2.4 shows the antenna implanted into the insert, which communicates the data to the readout unit by RF and, at the same time, supplies the circuit coupling with another external antenna. A damping modulator is included in the implantable sensor to transmit the data in digital mode. Furthermore, a temperature sensor has been integrated in the implantable sensor so that the microcontroller is able to measure the temperature and eventually to compensate the resistance data during the measurement activity. Outside the knee, the readout unit consists of a transceiver to drive the coil antenna and to demodulate the digital signal received. The antenna is controlled by a readout unit localized around the knee as in Figure 2.5. Furthermore, the readout unit supplies the implantable system by telemetry. The readout unit is managed by a low-power microcontroller and powered by an external battery.

The readout unit contains a transponder to transmit and to receive data by modulating the magnetic field using a damping stage, in particular, with OOK (On-Off Keying) modulation and Manchester code. The RF communication between the readout unit transceiver and implantable sensor receiver is supported by a magnetic field generated applying to the antenna a sinusoidal voltage, whose magnitude is 80 Vpp at 125 kHz frequency. The capability to transfer data and energy through the human body is the main advantage of this solution.

The implantable sensor works based on three phases defined as stop, measure, and transmission modes. Figure 2.6 shows the three activities of the implantable sensor where the stop mode duration is 6 s, while the duration of the measure conversion and the transmit mode are about 6 and 7 ms, respectively. The communication protocol is composed of 6 strings of 16 bits each. The first two are synchronization strings, and, then, there are three strings that contain the resistance data of three magnetoresistors and one string that provides the temperature. Finally, the data collected are transferred to a personal computer (PC) by a serial interface so that the data can be analyzed by a qualified medical staff.

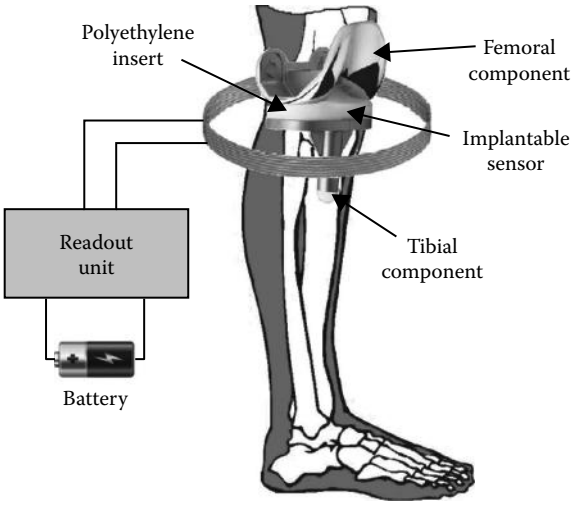


FIGURE 2.5 Prototype scheme of telemetric system to measure the forces in TKP prosthesis.

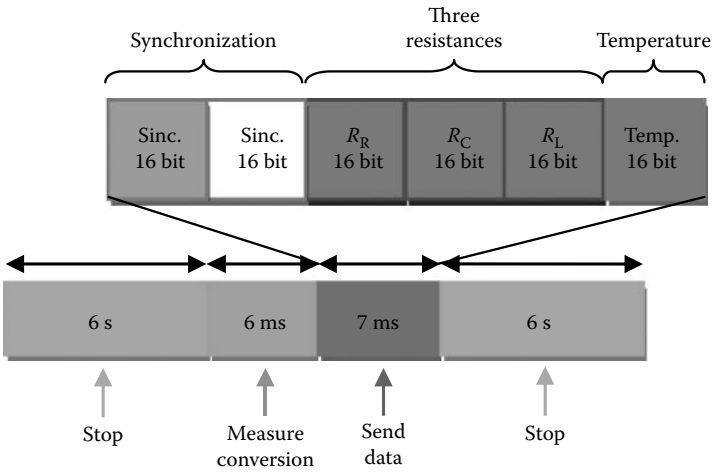


FIGURE 2.6 Functioning of implantable sensor.

2.3.1 Experimental Results

The force applied by the femoral component has been simulated using an Instron 8501 machine. This machine applies a linearly increasing force to the polyethylene insert from 0 N to 3 kN in 500 s. Figure 2.7 shows the relationship between the input orthogonal forces generated by the Instron on the TKP femoral component (solid line) and the force obtained by using Equation 2.1 considering the resistance value of the magnetoresistor output placed on the right side in the polyethylene insert (+ line). The samples were acquired every 6 s when the implantable sensor is in active mode.

The possible reasons for the difference between the two trends are mainly (1) nonlinear behavior of the magnetoresistor material, (2) no proper temperature compensation, (3) uncertainty of the experimental apparatus, and (4) the hysteresis effect due to the geometry and physical characteristics of the polyethylene insert. Figure 2.8 shows the wireless transmission signals monitored when the implantable

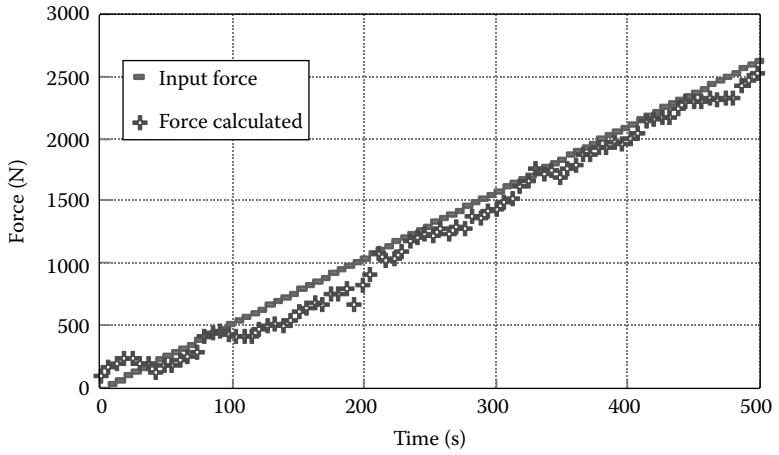


FIGURE 2.7 Comparison between the force values calculated by the resistance variation of the magnetoresistor on the right side of the insert of the implantable sensor and the input forces applied by the Instron.

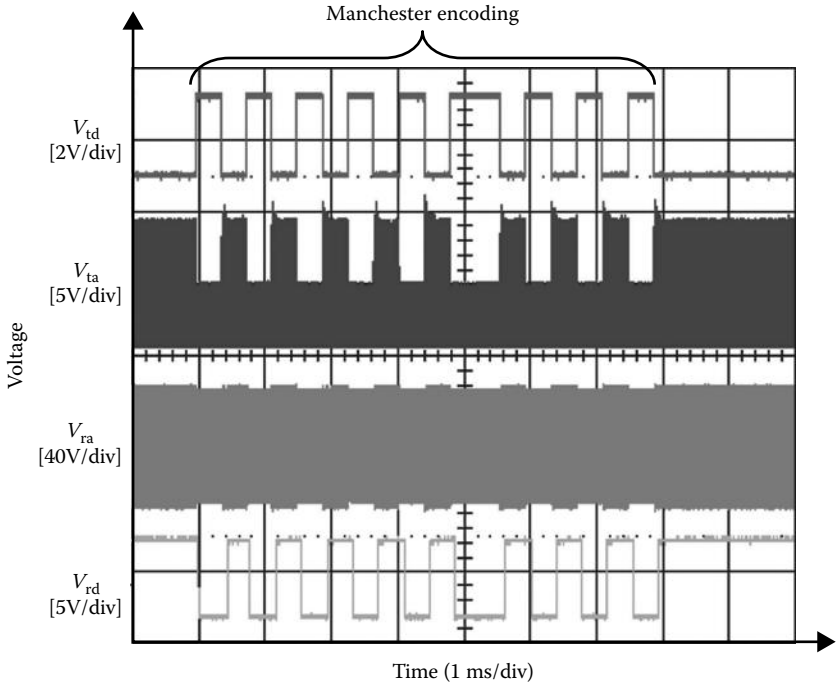


FIGURE 2.8 Wireless transmission signals of the converted data and typical implantable sensor activities.

sensor is in active mode: V_{td} is the transmitted data, V_{ta} is the transponder antenna voltage, V_{ra} is the signal of the differential voltage of the reader antenna, and V_{rd} is the received data.

Table 2.2 shows the power consumption during the different activities of the measurement system, when the external power from the readout unit is active. For example, when the implantable sensor measures and transmits the data to the readout unit, the microcontroller and the transponder require a power supply of about 1.7 mW, with a current consumption of approximately 850 μA and a voltage of 2 V.

TABLE 2.2 Power Consumption Measurements

Activity	Voltage (V)	Current (mA)	Power (mW)
Readout unit—transceiver communication	12	19.3	230.0
Implantable sensor—measurement and transmission	2	0.85	1.7
Implantable sensor—stop mode	2	0.16	0.3

2.4 Power Harvesting in Implantable Human Total Knee Prosthesis

The fundamental requirements for an implantable medical device are the capabilities of being self-powered and maintenance free. These ambitious goals excited the interest in a new and lively research field oriented toward the harvesting of energy from the human body with the aim of making the implantable (or wearable) system autonomous. Such a system, theoretically without human supervision, should provide information about the physiological parameters concerning its application. The rapid development and reduction in size, cost, and power consumption of the wireless communications devices allow for solving the important problem of the measurement of communication from the device to outside the human body. Considering the energy flow, an autonomous implantable device can be divided into two subsystems (López, 2010): the power-harvesting module and the load (power conditioning circuit, sensor, processor, and transceiver).

The first subsystem realizes the conversion of the energy from a particular domain (chemical, thermal, mechanical, etc.) to the electric domain. The second one carries out the mission of the implanted sensor, that is, the measuring and the transmission of the data.

This section shows the attempt to make autonomous a force sensor system inserted in a TKP (shown in Section 2.3), exploiting the mechanical energy produced by the human knee joint movement.

A TKP is composed of three components: the femoral component (condyles), the tibial plate, and the tibial insert constituting the contact surface of the tibia with respect to the femur. Figure 2.9 shows a 3D CAD model of the proposed solution: a copper coil is housed in a prominence of the tibial insert and six couples of block-shaped magnets are placed into each condyle. The axes of the magnets and the coil are parallel to the tibial insert.

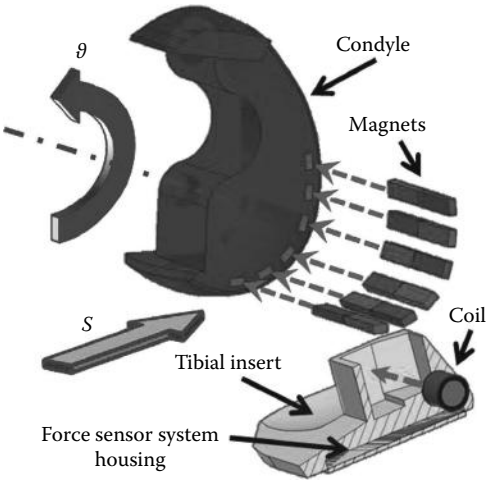


FIGURE 2.9 Cross section (in the sagittal plane) of a TKP with the energy-harvesting components. The angle ϑ and the displacement S are the degrees of freedom of the proximate relative motion of the femur (condyles) with respect to the tibial insert.

The energy conversion principle is based on the Faraday–Newmann–Lenz's law of induction: *a time-varying magnetic induction field B, linked to a conductive path c, leads to a potential difference to the extremities of c* (Woodson and Melcher, 1968). This way, when the femur moves with respect to the tibia, the magnetic induction field induces a time-varying flux and then a potential difference to the terminals of the coil.

In general, the relative motion of the condyles with respect to the tibia has six degrees of freedom and it is very difficult to reproduce. Because of this, the complex kinematics of the electromechanical system has been reduced by a reasonable simplification: only the relative rotation ϑ and translation S (Figure 2.9), in the sagittal plane, were considered in the design of the system.

A tailored motion control system allowed the reproduction of the gait conditions under the previous assumptions. In particular, the combined motion of translation and rotation of the TKP has been reproduced with the dedicated four-bar mechanism shown in Figure 2.10. An improved design of the four-bar mechanism allows for the translation S during the rotation ϑ with respect to the range of movement deduced by the literature (Masouros et al., 2010; Pinskerova et al., 2000).

During the gait cycle, in general, the angle ϑ is variable with the trend reported in Figure 2.11.

Due to the bigger amount of mechanical energy in the swing phase, with respect to the stance phase, in the following considerations and experimental tests, the analysis is limited only to the first one, that is, $\vartheta_{stance} = 0$, while ϑ_{swing} is supposed linearly variable in the time (i.e., $d\vartheta_{swing}/dt$ = constant) between $\vartheta_{swing} = 0°$ and $\vartheta_{swing} = 60°$ according to the trend reported in Figure 2.11.

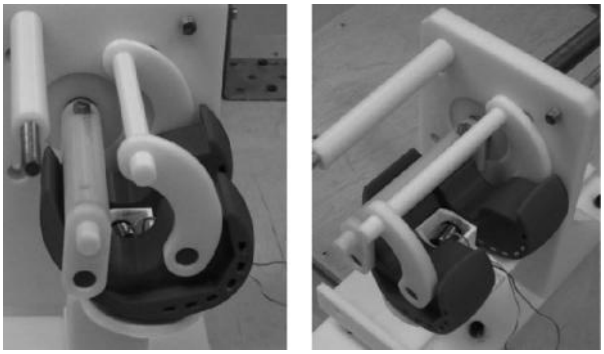


FIGURE 2.10 Prototype of the energy-harvester system implanted in the TKP.

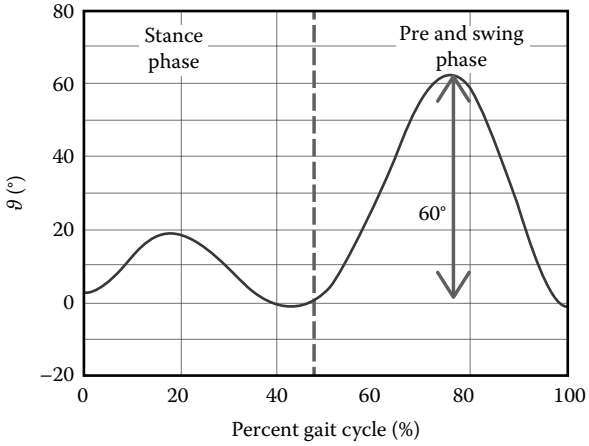


FIGURE 2.11 Knee joint angle ϑ as a function of the percent gait cycle.

The discontinuous nature of the human movement and its irregularity impose the design of a power and energy conditioning circuit for matching the power source and the energy requested by the measurement circuit. In fact, for example, depending on the technology by which the electronic circuits are realized, the voltage supply needs to respect a precise value or a proper range, while the voltage generated by the proposed energy-harvester system is very variable and discontinuous in the time domain, due to the strict dependence on the characteristics of the knee motion, as described earlier. Furthermore, the energy consumption of the system is related to the time requested to measurement and also to the techniques chosen for data processing and transmitting. Then it is necessary to establish a strategy for the functioning of the system, that is, the autonomous force sensor has a phase in which the power-harvesting module converts the mechanical energy from the knee movement, and, only when the required energy is available, a second phase in which the measurements are possible. For the proposed system, downstream of the power-harvesting module, a power conditioning circuit (*p.c.c.*) was realized with the aim to provide the energy supply to the load with the requested characteristics of voltage and duration.

Two experimental tests were performed. The first one was conducted considering the measurement of the open circuit voltage $V_{o.c.}(t)$ induced on the coil; in this test the energy-harvesting system is not connected to the *p.c.c.* and to the load. Figure 2.12 shows the oscillating nature of this voltage due to the different couples ($M1$, $M2$, …, $M6$) of magnets going near the coil (Luciano et al., 2012). In particular, the rising and falling ramps, between the broken boundary lines, delimit the complete passage, across the coil, of the generic couple Mn and the initial entrance of the following couple $M(n + 1)$.

The second test was performed, considering the complete autonomous system, that is, connecting the energy-harvesting system to the load using the realized *p.c.c.*. From Table 2.2, the mean power consumption of the force sensor system during a single cycle of data acquisition and transmission is 1.7 mW, while the time requested for the cycle is $T_{cycle} = 13$ ms, with a total energy consumption equal to $E_L = 22.1$ μJ. This load consumption was simulated using a resistive load $R_L = 2.2$ kΩ. The power conditioning circuit (Figure 2.13) is composed essentially by an impedance matching circuit, a charge pump (CP), which is turned on when the input voltage is 300 mV, and a startup capacitor. The startup capacitor is connected, by the CP, to the energy source, when the system harvests the energy, and to the load R_L, when it executes and transmits the force measures.

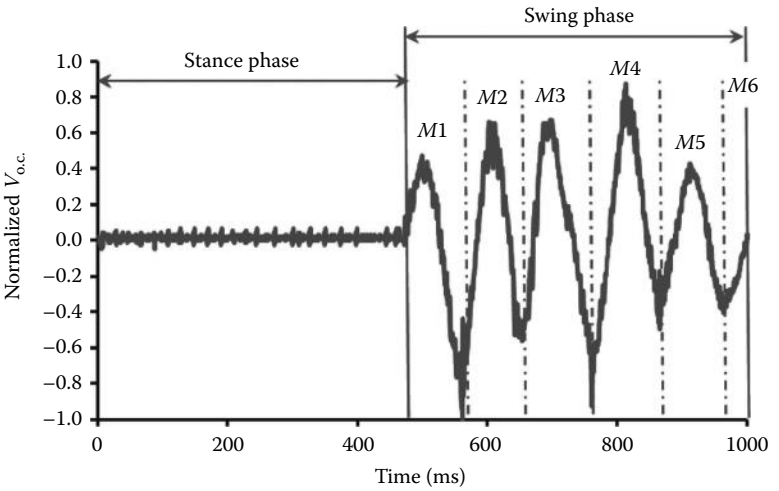


FIGURE 2.12 Normalized open circuit voltage during a gait cycle with a step frequency of 1.0 Hz.

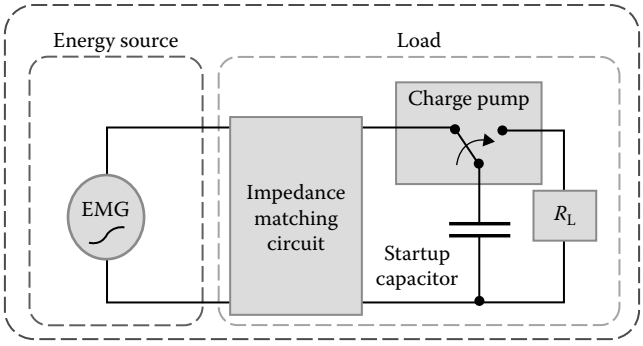


FIGURE 2.13 Block diagram of the autonomous force sensor system. EMG, electromechanical generator; R_L, resistive load equivalent to the force sensor system.

The CP has a discharge start output voltage $V_{CPout\_1} = 2.0$ V and a discharge stop voltage $V_{CPout\_2} = 1.4$ V. The charging time, T_c, requested for charging the startup capacitor is about $T_{c,zic} = 30.4$ s if the initial voltage of the startup capacitor is zero (zero initial condition), and $T_{c,ssc} = 7.6$ s when it is 1.4 V (steady state condition). The time requested of the patient, during the gait, for charging the startup capacitor, is more than acceptable. The capacitance $C_{startup}$ of the startup capacitor was deduced considering the minimum supply voltage of the processor ($V_{\mu C\_min} = 1.8$ V), the time necessary for the data acquisition ($T_{cycle} = 13$ ms), and the related energy consumption ($E_L = 22.1$ μJ). In particular, using the relation $E_L = 0.5 \cdot C_{startup}((V_{CPout\_1})^2 - (V_{\mu C\_min})^2)$, it is possible to deduce its minimum value ($C_{startup} = 58$ μF). Choosing a capacitor $C = 68$ μF, a discharge time $T_d = 16$ ms is necessary to decrease the CP output voltage from $V_{CPout\_1} = 2.0$ V to $V_{\mu C\_min} = 1.8$ V. This time is greater than the time $T_{cycle} = 13$ ms necessary for a single measurement cycle (Figure 2.14).

In conclusion, the energy-harvester system makes it possible to power supply a TKP implantable force sensor system making the system autonomous.

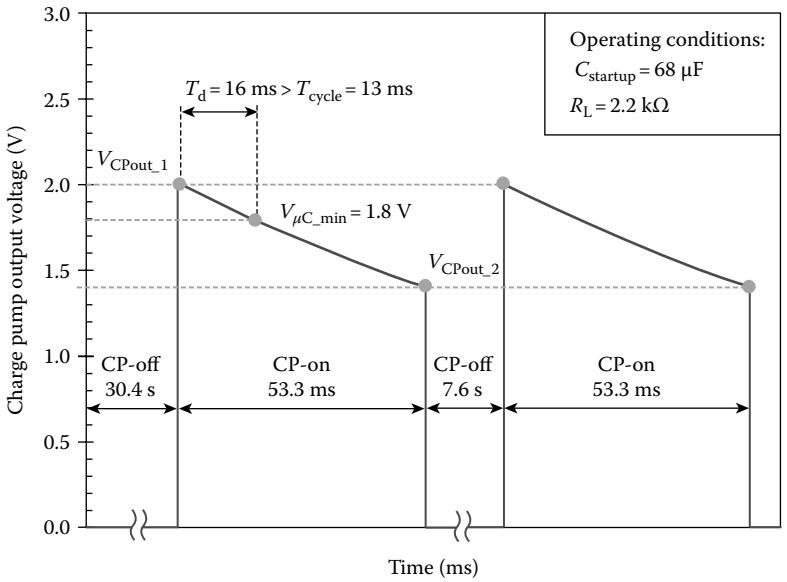


FIGURE 2.14 Operating conditions of the charge pump output voltage: CP-on, the CP connects the startup capacitor to R_L; CP-off, the CP connects the energy source to the startup capacitor.

2.5 Conclusions

Implantable medical devices are widely employed to monitor or to control different physiological parameters. Several implantable medical devices are powered by a battery, which constitutes a severe limitation because, in most cases, the battery defines the lifetime of the entire implantable medical device; in particular, because of the battery discharge, the implanted system must be surgically replaced. To obviate this problem, the implantable medical device should be powered through a telemetric technique or through an energy-harvesting system implanted together with the device. Several examples of implantable medical devices powered by telemetric or energy-harvesting technique are reported in the literature.

In this chapter, an example for each alternative to the battery is described; the first one is a force measurement system powered through the telemetric technique, while the second one is an energy-harvesting system that exploits the mechanical energy produced by the human knee joint movement.

The telemetric and the energy-harvesting systems can theoretically operate for an indefinite time, hence the implantable system must not be prematurely replaced. Furthermore, as the battery occupies significant space, batteryless implantable systems could be made smaller and more easily implanted. Otherwise, with a telemetric or an energy-harvesting system, energy is not always available and the measurement cannot be performed in a continuous way; the telemetric technique requires the presence of an external coil, while the energy-harvesting technique needs to store sufficient energy before powering the whole system.

The telemetric and the energy-harvesting techniques, especially the second one, represent two valid alternative solutions to power an implantable medical device. The energy-harvesting technique is undergoing research and development, and it could enable the patient to move in a free and autonomous way.

References

Almouahed, S., Gouriou, M., Hamitouche, C., Stindel, E., Roux, C., Self-powered instrumented knee implant for early detection of postoperative complications, *Engineering in Medicine and Biology Society*, Buenos Aires, 5121–5124, 2010.

Almouahed, S., Gouriou, M., Hamitouche, C., Stindel, E., Roux, C., Design and evaluation of instrumented smart knee implant, *Biomedical Engineering* 58:971–982, 2011a.

Almouahed, S., Gouriou, M., Hamitouche, C., Stindel, E., Roux, C., The use of piezoceramics as electrical energy harvesters within instrumented knee implant during walking, *Mechatronics* 16:799–807, 2011b.

Crescini, D., Sardini, E., Serpelloni, M., Design and test of an autonomous sensor for force measurements in human knee implants, *Sensors and Actuators A: Physical* 166(1):1–8, 2011.

Dalola, S., Ferrari, V., Guizzetti et al., Autonomous sensor system with power harvesting for telemetric temperature measurements of pipes, *Transactions on Instrumentation and Measurement*, 58(5):1471–1478, 2009.

Goto, H., Sugiura, T., Kazui, T., Feasibility of the automatic generating system (AGS) for quartz watches as a leadless pacemaker power source: A preliminary report, *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE,* Vol. 1, Brescia Goto, pp. 417–419, 1998.

Graichen, F., Bergmann, G., Rohlmann, A., Hip endoprosthesis for in-vivo measurement of joint force and temperature, *Journal of Biomechanics* 32:1113–1117, 1999.

Halperin, D., Kohno, T., Heydt-Benjamin, T.S., Fu, K., Maisel, W.H., Security and privacy for implantable medical devices, *Pervasive Computing, IEEE* 7(1):30–39, 2008.

IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz. C95.1-1999, Institute of Electrical and Electronics Engineers, New York.

Lahuec, C., Almouahed, S., Arzel, M., Gupta et al., A self-powered telemetry system to estimate the postoperative instability of a knee implant, *Biomedical Engineering* 58:822–825, 2011.

López, M.T.P., Methods and circuits for the efficient management of power and energy in autonomous sensors. Tesi doctoral. Universitat politècnica de Catalunya. Departament d'Enginyeria Electronica, 2010.

Luciano, V., Sardini, E., Serpelloni, M., Baronio, G., Analysis of an electromechanical generator implanted in a human total knee prosthesis, *Sensors Applications Symposium IEEE*, Hong Kong, pp. 1–5, 2012.

Masouros, S.D., Bull, A.M.J., Amis, A.A., Biomechanics of the knee joint, *Orthopaedics and Trauma* 24:84–91, 2010.

Miao, P., Holmes, A.S., Yeatman, E.M. et al., Micro-machined variable capacitors for power generation, *11th International Conference on Electrostatics:* 53–55, 2004.

Mitcheson, P.D., Energy harvesting for human wearable and implantable bio-sensors, *Engineering in Medicine and Biology Society*, Buenos Aires, 3432–3436, 2010.

Morais, R., Frias, C.M., Silva, N.M. et al., An activation circuit for battery-powered biomedical implantable systems, *Sensors and Actuators A: Physical* 156(1):229–236, 2009.

Morais, R., Silva, N.M., Santos, P.M. et al., Double permanent magnet vibration power generator for smart hip prosthesis, *Eurosensors XXIV*, Linz, Austria, September 5–8, 2010, 172(1):259–268, 2011.

Nasiri, A., Zabalawi, S.A., Jeutter, D.C., A linear permanent magnet generator for powering implanted electronic devices, *IEEE Transactions on Power Electronics,* 26(1):192–199, 2011.

Pinskerova, V., Iwaki, H., Freeman, M.A., The shapes and relative movements of the femur and tibia at the knee, *Orthopade* 29(Suppl 1):S3–S5, 2000.

Potkay, J., Brooks, K., An arterial cuff energy scavenger for implanted microsystems. *Bioinformatics and Biomedical Engineering*, Shanghai, 1580–1583, 2008.

Raman, R., Brennan, M.J., Mace, B.R., Kovacic, I., Potential benefits of a non-linear stiffness in an energy harvesting device, *Nonlinear Dynamics* 59:545–558, 2010.

Riistama, J., Väisänen, J.H.S., Harjunpää, H., et al., Wireless and inductively powered implant for measuring electrocardiogram, *Medical and Biological Engineering and Computing* 45(12):1163–1174, 2007.

Romero, E., Warrington, R.O., Neuman, M.R., Body motion for powering biomedical devices, *Engineering in Medicine and Biology Society*, Minneapolis, 2752–2755, 2009.

Silay, K.M., Dehollain, C., Declercq, M., 2011. Inductive power link for a wireless cortical implant with two-body packaging. *Sensors Journal, IEEE* 11(11):2825–2833.

Soykan, O., Power sources for implantable medical devices, *Business Briefing: Medical Device Manufacturing & Technology*, World Markets Research Centre, London, 76–79, 2002.

Stark, I., Invited talk: Thermal energy harvesting with thermo life, *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006.* International Workshop, Cambridge, MA, 19–22, 2006.

Starner, T., Human-powered wearable computing *IBM Systems Journal* 35(3–4):618–629, 1996.

Tamura, T., Mizukura, I., Sekine, M., Kimura, Y., Monitoring and evaluation of blood pressure changes with a home healthcare system, *Information Technology in Biomedicine, IEEE Transactions on* 15(4):602–607, 2011.

Tashiro, R., Kabei, N., Katayama, K., Tsuboi, E., Tsuchiya, K., Development of an electrostatic generator for a cardiac pacemaker that harnesses the ventricular wall motion, *Journal of Artificial Organs,* 5(4):0239–0245, 2002.

Troyk, P.R., DeMichele, G.A., Kerns, D.A., Weir, R.F., IMES: An implantable myoelectric sensor, *Engineering in Medicine and Biology Society, 2007, EMBS 2007, 29th Annual International Conference of the IEEE*, pp. 1730–1733, 2007.

Valdastri, P., Menciassi, A., Arena, A., Caccamo, C., Dario, P., An implantable telemetry platform system for in vivo monitoring of physiological parameters, *IEEE Transactions on Information Technology in Biomedicine,* 8(3):271–278, 2004.

Von Stetten, F., Kerzenmacher, S., Lorenz, A. et al., A one-compartment, direct glucose fuel cell for powering long-term medical implants, *MEMS 2006—19th IEEE International Conference on Micro Electro Mechanical Systems,* January 22–26 2006, Istanbul, Turkey.

Wang, F., Hua, W., Zhang, S., Hu, D., Chen, X., Clinical survey of pacemakers 2000–2001, *Zhonghua Xin Lv Shi Chang Xue Za Zhi 2003* 7:189–191, 2003.

Wei, X., Liu, J., Power sources and electrical recharging strategies for implantable medical devices, *Frontiers of Energy and Power Engineering in China* 2(1):1–13, 2008.

Woodson, H.H. and Melcher, J.R. *Electromechanical Dynamics*, 3 vols., 1968. (Massachusetts Institute of Technology: MIT Open Course Ware). http://ocw.mit.edu. License: Creative Commons Attribution-NonCommercial-Share Alike, 1968.

Yang, Y., Wei, X., Liu, J., Suitability of other thermoelectric power generator for implantable medical electronic devices, *Journal of Physics D: Applied Physics* 40:5790–5800, 2007.

3

Bio-Inspired and Life-Inspired Sensors

Cesar Ortega-Sanchez
Curtin University

Halit Eren
Curtin University

3.1 Introduction

Living beings have always inspired the imagination of inventors and scientists throughout history. According to Greek mythology, Daedalus invented a pair of wings made out of waxed feathers so that he and his son Icarus could escape from the labyrinth of Crete. They succeeded, but Icarus flew so close to the sun that the wax melted and his wings dismantled, with tragic consequences. In renaissance Italy, Leonardo da Vinci invented flying machines and submarines inspired by birds and fish more than 500 years ago, but his inventions never came into real practice. These attempts failed not because they were fundamentally flawed, but mainly because technology at the time was far too immature for any realistic implementation of such projects.

What are the distinctive features of biological systems that we would like to incorporate in our man-made systems? Here is a partial list:

- *Resilience*. The ability to continue functioning even when the system has suffered damage (fault-tolerance).
- *Healing*. The ability to recover from damage and return the system to a state similar, if not identical, to the state it was before damage occurred.
- *Low power*. Living organisms are extremely efficient in transforming energy (food) into growth and behaviors.
- *Autonomy*. Plants and animals do not need external control to survive. Their behaviors are either innate, or learned from interactions with their environment.
- *Evolution, adaptation, and learning*. Biological organisms have these mechanisms that allow them to adapt, up to certain extent, to changes in their environment.
- *Intelligence*. The ability to perceive the environment and interpret it to produce behaviors in response to possible outcomes.

The remarkable increase in computational power and, more recently, advances in the architecture and size of field-programmable logic devices (FPLDs) have made it possible to build devices inspired by mechanisms found in nature. Artificial neural networks (ANN) and evolutionary algorithms (EA) are many examples of bio-inspired techniques that have been successfully developed and incorporated in commercial systems. In this chapter, we will discuss fundamentals of bio-inspired sensors and give some examples of their applications.

With today's technology it is possible to implement a complex system on a single chip with capabilities listed previously. This possibility is enhanced due to integration of the wireless communication systems and complex networking capabilities on that chip. These new types of systems are feasible and affordable for applications that require miniaturization, high performance, low-power dissipation, and distant communications. There are many advanced devices available to transmit and process measurement information from mobile objects, from inaccessible places, or remote devices and machines. This prompts stimulus for development of much rigorous and more complex niche and application-specific systems. Most of these application-specific systems contain integrated embedded microprocessor circuits that have included intelligence, augmented reality, and wearable or implanted components. It is worth mentioning that they are and must be extensively supported by a range of appropriate software.

3.2 Bio-Inspired Systems

For the purpose of studying life looking for inspiration, living organisms can be described at four fundamental structural levels, as in Figure 3.1. The features at one level of the hierarchy emerge from the interaction of elements at lower levels. Understanding life in any depth requires knowledge of all four levels shown in Figure 3.1. A fifth level can be suggested in which organisms from the same or different species interact with each other in the form of populations and ecosystems, respectively. In current research, hardware systems are used to study the organ, system, and organism levels. Cellular and population levels are studied mainly with the software systems.

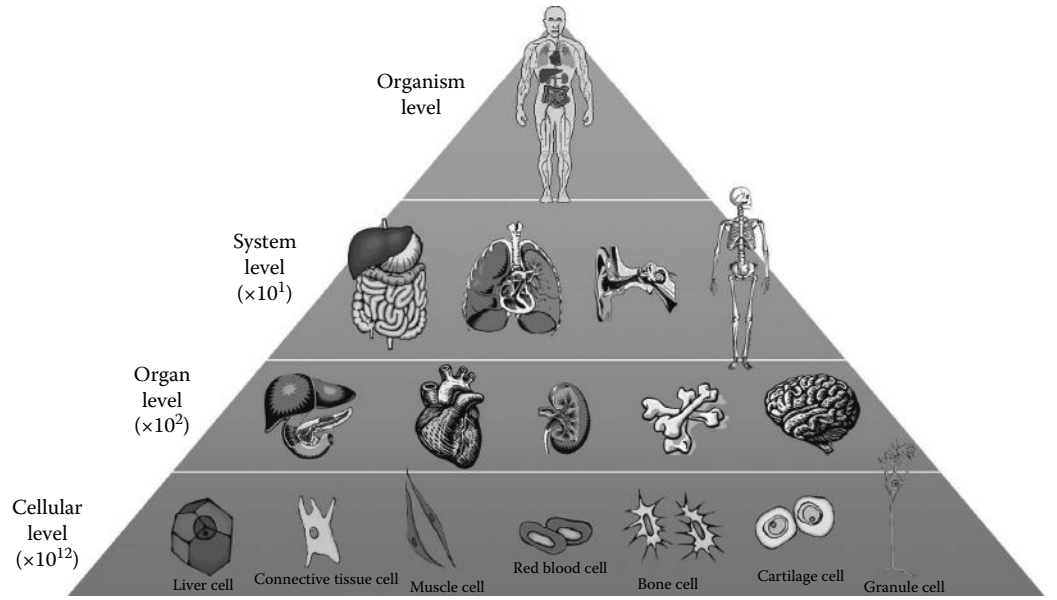


FIGURE 3.1 Hierarchical structures of multicellular organisms. (Authors' artwork incorporating Microsoft™ clip arts.)

Even though elements from every level in Figure 3.1 are very different in size and behavior, there are certain features common to all:

- Interaction of simple components that generates complex emergent behaviors
- Asynchronous interaction: As most interactions are event-driven, there is no explicit synchronization
- High level of redundancy

Current research in bio-inspired systems focuses in one, or more, of these features. In order to exploit these features, a commonly studied but very fundamental area is the bio-inspired semiconductor sensors and will be discussed in this chapter in detail.

3.3 Life-Inspired Systems

The paradigm of the life-inspired systems originates from the observations with the aim that the microelectronic-based systems should have characteristics resembling the characteristics of individual life organisms or organized populations.

Engineers and scientists aim for autonomous, self-contained, robust, self-organizing, self-adapting, self-regulating dynamic evolutionary systems that resemble real organisms. The products should be composed of autonomous systems made from diverse subsystems, but each of which has its own functionality to be integrated with the other subsystems and the central organizer just like the brain or nervous system in living organisms. The aim is to come up with systems that do not limit themselves to the basic functions of collecting, transmitting, storing, processing, and presenting information in relation to some external systems, but in addition they are able to solve complex problems, make and implement difficult decisions, learn, discover new ideas, and evolve.

To efficiently develop the new systems closely resembling intelligent life organisms, new system architecture concepts and design methods and tools are developed. Challenges in biologically and life-inspired systems created a new field termed as *bio-inspired engineering*. This new discipline applies biological principles to develop new engineering solutions for industry, medicine, environment, and many other fields that have not been previously investigated. The emergence of this discipline is the culmination of the unification of the life sciences with engineering and physical sciences. Hence, bio-inspired engineering involves deep exploration into the living cells, tissues, and organisms that can build, control, regenerate, recycle, and adapt to their environment.

A well-studied area in bio-inspired and life-inspired systems is artificial intelligence (AI). Many AI techniques have been developed and applied using computer-based technologies. Various computational models and knowledge-based systems have been applied in industrial, economic, and many other areas for the purpose of automated reasoning, learning, and problem-solving. There are thousands of publications on AI. As an example, an advanced study on AI is web intelligence (WI). Since the web is the largest network, the application of AI on this large scale is demanding and challenging. Studies such as WI brain informatics (BI) are leading to profound advances in the analysis and understanding of data, information, knowledge, and wisdom, as well as their interrelationship, organization, and creation process. Research and development in WI is making it possible to develop human-level WI [1–3].

3.4 Semiconductor Sensors

Since the early 1970s, silicon technology has been used to detect chemical changes such as hydrogen ions, pH levels, chemical substances, and odor detection. Since then many more sophisticated sensors have been developed based on complementary metal oxide semiconductor (CMOS), ion-sensitive field effect transistor (ISFET), metal oxide semiconductor field effect transistor (MOSFET), and other similar technologies. Many advanced types of semiconductor sensors are available, some of which are known as chemical field effect transistors (CHEMFETs) and enzyme field effect transistors (ENFETs)

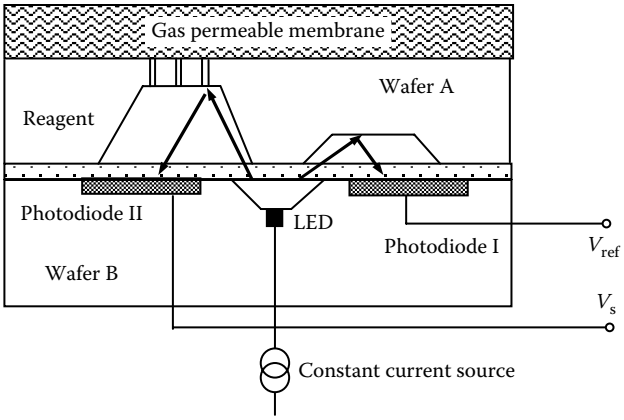


FIGURE 3.2 An optical semiconductor chemical sensor.

for sensing various analytes, elements such as potassium, sodium, and complex chemical structures such as glucose in blood and urea.

Chemical sensors have been leading to semiconductor-based devices to lead to bio-inspired sensors. Figure 3.2 illustrates an example of semiconductor-based optical chemical sensor. This sensor consists of two chambers, one acting as a reference containing a reference concentration chemical and the other as the sensor. These chambers are illuminated by a common light-emitting diode (LED). The surfaces of chambers are metalized to improve internal reflectivity, and the bottom of chambers is covered by glass. One of the chambers has slots covered with a gas-permeable membrane. The slots allow the penetration of gas to be measured (e.g., CO_2) into the chamber. Wafers A and B form optical waveguides. The chamber filled with reagent is used to monitor the optical absorbency for comparison with that obtained from the reference chamber.

Another example semiconductor-based sensor is the ISFET sensor is illustrated in Figure 3.3. In this example, the sensor is chemical based, and the signals generated by the sensor are amplified and processed in the same chip.

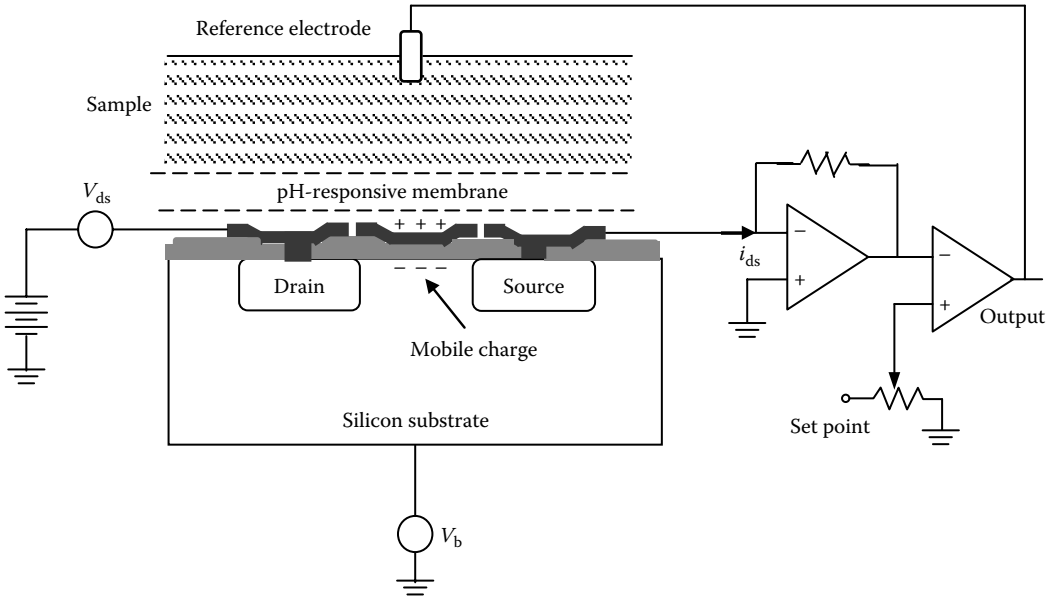


FIGURE 3.3 A pH sensor based on ISFET.

In addition to simple single-chip chemical sensors, there are many complete systems that have extensive and complex information processing features all built on a single chip or in a single package. The implementation of complete systems on a chip (SoCs) or in a system in package (SiPs) enables development of new generation sensors in a diverse range of application [4–6]. Nowadays, the use of semiconductor technology in biomedical systems is an accepted norm and very common. One of the most advanced technologies is DNA sequencing where the DNA and variations in the DNA are used to trigger sensors in microchips and go further by taking the necessary actions to be followed.

3.5 Biomedical and Biological Sensors

Semiconductor-based sensors for biomedical and biological applications have been researched and developed extensively. They find many applications, including the following:

- Biopotential and electrophysiology determination
- Blood pressure measurements
- Blood flow measurements
- Body temperature measurements
- Body weight and composition measurements

Traditionally, biological detectors require human intervention in a laboratory environment. However, in recent years, automatic devices and robots have involved biological applications, such as detection of microorganisms and their concentration levels. For example, for the detection of microorganisms in air, three different common methods are as follows:

1. *Biochemical*, which detect a DNA sequence and protein that are unique to a bio-agent through its interaction with test modules
2. *Chemical*, for example, mass spectrometry, which works by breaking down a sample into its components such as amino acids and then comparing their weights with those of known bio-agents and other molecules
3. *Biological tissue-based systems*, in which a bio-agent or biotoxin affects live mammalian cells, causing them to undergo some measurable response

Some classes of biosensors rely on comparing the DNA taken from a microorganism with the DNA of a known agent [7–9]. In other classes, there are microfluidic devices, which contain capillary channels, valves, and chambers in a single chip. Once the microorganism is in the chip, their cells are cracked open ultrasonically. Then polymerase chain reaction (PCR) is applied by means of small thin film heaters.

Some semiconductor sensors apply either optical and fluorescent methods or magnetic methods, as illustrated in Figure 3.4. These devices comprise an array of wirelike magnetic field microsensors. These sensors are coated with single-stranded DNA probes specific for a gene from a bio-agent. Once a strand of bio-agent DNA in a sample binds with a probe, the resulting double strand binds a single magnetic microbead. When there are magnetic beads in the sensor, the resistance of sensor decreases in proportion to the number of microbeads.

Among many other methods, in one application multiple array biosensors are multiplexed and electrochemical detection is realized by using an electronically active substrate, which has been constructed with a standard CMOS technology.

Another commonly used method for biosensing is by the use of live tissues. Many toxins trigger measurable or differentiable reactions in living cells. Mammalian cells, such as heart cells, are cultured in a lab and then seeded into a cartridge containing a microelectrode array. When a biotoxin is introduced, the cells create voltages detectable in millivolts at the electrodes.

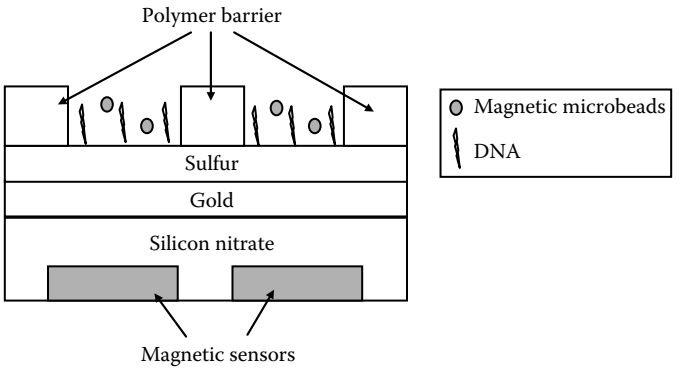


FIGURE 3.4 A magnetic biosensor.

3.5.1 Advanced Biosensors

There are many types of advanced biosensors, some of which automatically collect samples and then make multiple copies of the DNA using PCR. In this method, the samples are heated and cooled periodically to produce copies of the DNA. As the DNA is copied, the resulting strands are mixed with fluorescent DNA probes. The probe bound with the bio-agent glows under ultraviolet light. By using different probe markers several PCR reactions can run at the same time, thus enabling detection of more than one agent at a time.

In general, semiconductor sensors are extensively used in healthcare for a number of purposes. With the help of wireless systems, the CMOS-based systems are commonly implemented for physiological condition monitoring of athletes or patients (ECG, respiration, temperature, muscular conditions, etc.) They also are used for point-of-care diagnostic systems and genetic sequencing. Coupled with their potential for scalability, miniaturization, and low costs, they have applications in bio-inspired prosthetics where CMOS-based circuits are designed and implemented to replicate biology to configure implantable devices for personalized therapies such as for diabetes [10,11].

In human health, bio-inspired sensors are common in areas such as cochlear implants to restore hearing, retinas to cure blindness, glucose monitoring and control for diabetes, and vestibular feedback to restore balance.

Apart from microelectronics there are other bio-inspired technologies currently investigated and developed. For example, some polymers and polymer composites are capable of self-healing. In some applications, polymers are used for crack-mending initiated by an external thermal-, photo-, mechanical-, or chemical-induced stimulus. Successful crack healing has been demonstrated through both molecular diffusion and thermally reversible solid-state reactions. In another approach, damage in the form of a crack triggers the release of healing agents stored in the material, as also occurs for fracture events in biological systems.

3.6 Biomimetic Sensors

Bar-Cohen [12,13] explains that organisms use numerous receptors (senses) to control every aspect of their life and functions. Due to millions of years of evolution, biological sensors became very effective for providing external and internal information to warn about dangerous conditions, sense locations, control growth, and functionality of organs. These near-perfect capabilities of biological sensors prompted scientists and engineers to make efforts to understand and mimic them as closely as possible.

Scientists and engineers are studying biological sensors to copy, adapt, and inspire new capabilities and apply them in diverse ranges. These studies require interdisciplinary collaboration of researchers from the field of biology, materials science, mathematics, engineering, medicine, chemistry, and physics as well as involvement of social scientists [14–16].

The investigated biological sensors include the capabilities of self-healing, reconfigurability, replication, balancing chemical contents, and sensing of pH, temperature, pressure, and other life essentials. Armed with the observability of these sensors, studies concentrate on the understanding of the purposes, the rules, the concepts, the mechanisms, and the principles of biological sensors for developing new methods of manufacturing, processes, mechanisms, materials, and algorithms.

Biological receptors evolved for external sensing provide inputs to the central nervous system about the environment such that the muscles are commanded for action-based analysis of the received information. Some of the senses provide the equivalent of alarm conditions in the form of pain, others provide a *feel* of temperature, pressure, seeing, hearing, smelling, and tasting. However, these types of sensing have already been used extensively in human-made systems. For example, advanced robots can mimic biological models by sensing pressure and temperature. By using optical or acoustic sensors, they can determine their positions or recognize other objects, avoid collisions, interact with humans, and so on [17].

The common design requirements of engineering structures are quite similar to those of biological systems including the need to have a structure that is lightweight, consumes minimum power, and is durable over the life of the product. Mimicking biological sensors and using them help development of new technologies offer numerous benefits to human lives. Everyday improved methods and tools are emerging, finding many applications [18–20]. We can summarize the current applications of biomimetic sensors in the following areas:

- *Touch and tactile sensing*: robotics, material deposition, machining, angle sensing, pressure profiles, special resolutions, bruise sensing, and damage sensing military and commercial aircrafts
- *Sensing heat and IR*: domestic and industrial applications and forest fire detection
- *Smell and artificial nose*: odor detection, food industry, pharmaceutical, environmental monitoring, quality control, gases, chemical compounds, and space explorations
- *Taste and artificial tongue*: environmental monitoring, explosives, weapon detection, and healthcare
- *Vision and optics*: navigation systems, cameras, pattern recognition, defense systems, crime reduction, surveillance systems, endoscopes, and visual prosthetics
- *Proximity sensing*: robotics, shape identification, shape details, manufacturing, and transport industry
- *Hearing and sonar imaging*: cochlear implants, speech recognition in noisy environments, identification of localized sound sources, direction finding, submarines, oceanography, and various navigation systems
- *Flow detection*: flow and force detection and direction finding

3.7 Signal Processing

Bio-inspired sensors commonly come as microchips, and they largely generate analog signals. As in the case of all analog sensors, the signal processing requires analog-to-digital and digital-to-analog conversions, illustrated in Figure 3.5.

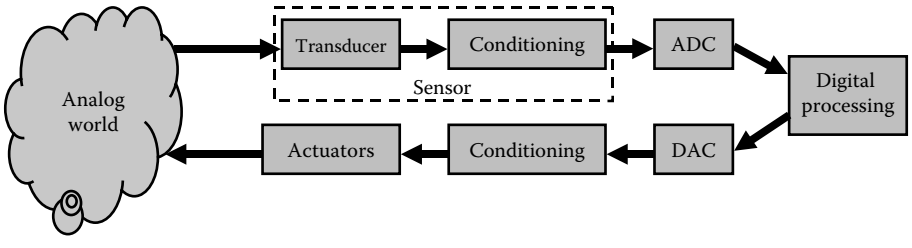


FIGURE 3.5 A bio-inspired data acquisition and control array.

The system in Figure 3.5 comprises hardware and software elements. Each one of them can potentially be biologically inspired. The number of interesting mechanisms and behaviors that could be adapted from nature into man-made systems is vast; therefore, it is possible to apply bio-inspiration to any of the blocks.

The basic principles of sensors can be temperature variations in physical (dimensions, resistivity, capacitance, etc.) or chemical (phase, color, pH, etc.). Hence, the signal processing requires appropriate selection of techniques suitable to the nature of the signals as they can be inductive, capacitive, and resistive, or they can be current source or voltage source [21,22]. In most cases the outputs of the sensors are not linear; therefore, extensive signal conditioning is required, such as the following:

- Linearization
- Impedance-matching circuitry
- Amplification
- Filtering
- Integration
- Differentiation
- Modulations

A typical example of signal processing requirements in biosensors is implantable medical devices, which require extreme low-power operation to avoid periodic replacement of batteries for long-term maintenance-free operation. For real-time monitoring of health condition, implantable medical devices also need to transmit the vital information outside of the human body for further diagnostics and processing. In an application, a data generator block takes the signal from sensors and transforms the signal into frequency-modulated digital pulses by employing relaxation oscillators. The next block such as the impulse generator converts the digital pulses into impulse signals for impulse-radio-based wireless telemetry.

AI is a commonly used technique for signal processing information from biosensors. For example, in an application a biosensor was simulated employing a feed-forward neural network with three layers and trained using a back-propagation (BP) algorithm [23]. Spectra generated from an optical phenolic biosensor at selected wavelengths were used as input data for an ANN. The network architecture of 5 inputs neurons, 21 hidden neurons, and 1 output neuron was found suitable for this application. The results had been compared with the information obtained from the operational biosensors.

A new trend in bio-inspired sensors is the use of field-programmable gate arrays (FPGAs) as they inherently offer greater flexibility in system configuration, application of AI, and flexibility in software development. In a study attempt has been made to replace microcircuits by FPGAs in portable biosensor devices [24]. Thin film technology electrochemical-based sensor chips such as pH, oxygen partial pressure, impedance, electrical potentials, and temperature can be used for the simultaneous in vitro measurement of metabolic, morphologic, and electrophysiological parameters. All these can be combined and the signals processed by using mixed signal electronic and local (point-of-care) analyses.

3.8 Bio-Inspired Sensors in Industry

Many of the interesting behaviors that can be found in nature are the result of nervous systems. Nervous systems include the senses and muscles to interact with the environment. Among species, vertebrates have evolved the most sophisticated nervous systems. Hence, they are a good source of inspiration for the investigation of bio-inspired systems.

The nervous system of mammals is made out of neurons with different functions according to where they are located in the sensory–motor cycle. In its simplest form, the sensory–motor cycle comprises three functional levels: the peripheral nervous system (PNS) that perceives and interacts with the environment, the spinal cord that transmits electrochemical signals, and the brain that receives signals generated by

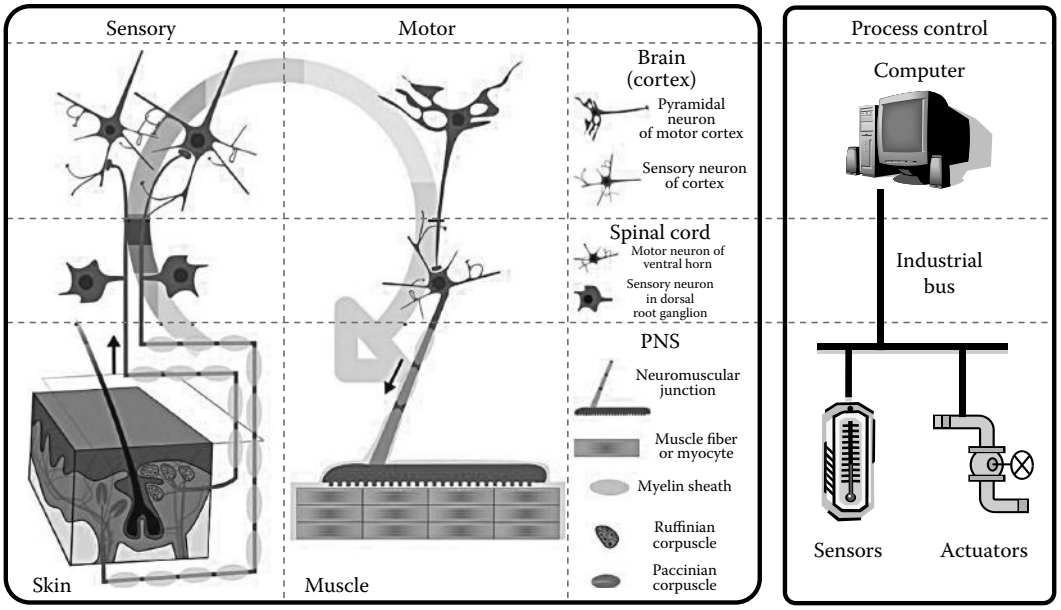


FIGURE 3.6 Sensory–motor and process control systems. (Courtesy of Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Nervous_system_organization_en.svg)

sensory cells, processes them, and produces signals to control motor cells in the PNS. These three levels can be associated with the main components of a conventional control system, as shown in Figure 3.6.

Figure 3.6 suggests that in a first approximation, conventional control systems already have the hardware needed to implement bio-inspired systems. Bio-inspiration may apply to the way elements are interconnected and how software deals with sensor information to produce control signals. Good examples of this approach can be found in distributed control systems where intelligent sensors and actuators take low-level fast decisions at node-level while transmitting information to a central computer for high-level coordination. Properties like redundancy, event-driven processing, and fault-tolerance can be incorporated using additional hardware or software.

To achieve fault-tolerance, researchers are studying artificial immune systems (AIS) [25]. By using concepts like "the self" and "antigens," systems are given the ability to recognize unexpected states (*antigens*) that do not belong in an expected state set (*the self*) and apply corrective measures to return the system to a "*healthy*" state.

3.8.1 Bio-Inspired Sensors in Control Systems

Process control is a typical area where bio-inspired systems find applications. In modern control practices, it is ever more frequent to find what are called "intelligent instruments." These devices usually incorporate transducers, conditioners, analog-to-digital converters (ADCs), and a small local computer that can perform simple signal preprocessing and control. Information is transmitted from intelligent instruments to central computers using industrial buses and communication protocols. Figure 3.7 shows a typical control system that incorporates intelligent instrumentation.

Actuators are typical examples of bio-inspired systems [26,27]. They are essential elements within mechatronic systems due to their important role in motion control systems, and hence there is a need to develop new trend of actuators that can be inspired from biological actuation systems in nature and can be associated with different levels of control. Making intelligent devices by biologically inspired modern

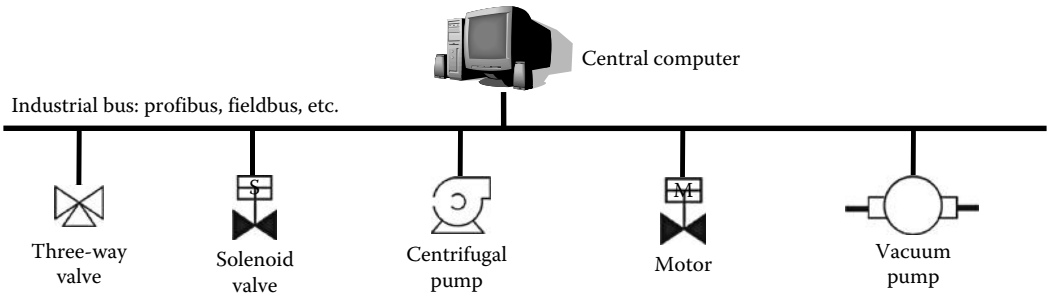


FIGURE 3.7 Industrial control system using intelligent instruments and network buses.

actuators and artificial muscles can create a new reality with greater potential in comparison to conventional techniques. Electroactive polymers (EAPs) are a good example of materials that find usage as bio-inspired motion systems in industrial applications. EAP actuators are electrically responsive materials that have common characteristics with natural muscles and hence are suitable for "artificial muscles" in biomimetic motion applications. EAP materials can be activated either by electrically induced transport of ions or by electrostatic forces. EAPs are finding numerous industrial applications.

In one study, algorithms for bio-inspired tracking systems have been developed as the control algorithm for sensor-mediated chemical plume tracking in a turbulent flow environment. This study focused on development of a signal processing strategy capable of replicating behavioral responses of actively tracking strategies of blue crabs to chemical stimuli. The geometric arrangement of the sensor array is inspired by the location of blue crab sensor populations. Upstream motion is induced by a binary response to suprathreshold spikes of concentration, and cross-stream steering is controlled by contrast between bilaterally separated sensors [28].

References

1. Zhong, N. et al. Web intelligence meets brain informatics, *First WICI International Workshop, WImBI 2006*, December 15–16, 2006, Beijing, China; Springer-Verlag, Berlin, Germany, 2006.
2. Gupta, S. and Thakur, N., Using capture-recapture method for web intelligence, *2010 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC 2010)*, Coimbatore, India, 5pp., 2010.
3. Dzitac, I. and Moisil, I., Advanced AI techniques for web mining, *Proceedings of the 10th WSEAS International Conference on Mathematical Methods, Computational Techniques and Intelligent Systems (MAMECTIS '08)*, Stevens Point, WI, pp. 343–346, 2008.
4. Khan, A., Recent developments in high-performance system-on-chip IC design, *2004 International Conference on Integrated Circuit Design and Technology*, Austin, TX, pp. 151–158, 2004.
5. Kato, Y. and Mukai, T., A real-time intelligent gas sensor system using a nonlinear dynamic response, *Sensors and Actuators B (Chemical)*, 120(2), 514–520, January 10, 2007.
6. Hassanzadeh, A. and Linquist, R. G., Single chip solution to capacitive liquid crystal chemical and biological sensor, *IEEE Sensors Applications Symposium (SAS)*, San Antonio, TX, pp. 1–6, 2011.
7. Levine, P. M., Gong, P., Levicky, R., and Shepard, K. L., Real-time, multiplexed electrochemical DNA detection using an active complementary metal-oxide-semiconductor biosensor array with integrated sensor electronics, *Biosensors and Bioelectronics*, 24(7), 1995–2001, Elsevier Science Ltd., U.K., 2009.
8. Purushothaman, S., Toumazou, C., and Georgiou, J., Towards fast solid state DNA sequencing, *2002 IEEE International Symposium on Circuits and Systems*, Scottsdale, AZ, Vol. 4, pp. IV-169–IV-172, 2002.
9. Rothberg, J., Hinz, W., Rearick, T., Schultz, J., Mileski, W., Davey, M., Leamon, J., Johnson, M., Milgrew, K., and Edwards, M., An integrated semiconductor device enabling non-optical genome sequencing, *Nature*, 475(7356), 348–352, 2011.

10. Georgiou, J. and Toumazou, C., A 126-μw cochlear chip for a totally implantable system, *IEEE Journal of Solid-State Circuits*, 40(2), 430–443, 2005.
11. Steil, G. M., Panteleon, A. E., and Rebrin, K., Closed-loop insulin delivery—The path to physiological glucose control, *Advanced Drug Delivery Reviews*, 56(2), 125–144, 2004.
12. Bar-Cohen, J., Biological senses as inspiring model for biomimetric sensors, *IEEE Sensor Journal*, 11(12), 3194–3202, 2011.
13. Bar-Cohen, J., ed., *Biomimetics: Nature Based Innovation*. Boca Raton, FL: CRC Press, 2011.
14. Johnson, E. A. C., Ronser, R. H. C., and Jeronimidis, G., Technologies recent advances in biomimetic sensing, *Philosophical Transactions of the Royal Society A*, 367, 1559–1569, 2009.
15. Nemat-Nasser, S., Nemat-Nasser, S., Plaisted, T., Starr, A., and Vakil-Amirkhizi, A., Multifunctional materials, Chapter 12, in *Biomimetics—Biologically Inspired Technologies*, Y. Bar-Cohen, ed., Boca Raton, FL: CRC Press, pp. 309–340, 2005.
16. Bleckmann, H., Schmitz, H., and von der Emde, G., Nature as a model for technical sensors, *Journal of Comparative Physiology A, Neuroethology, Sensory, Neural and Behavioral Physiology*, 190(12), 971–981, 2004.
17. Wettels, N., Santos, V. J., Johansson, R. S., and Loeb, G. E., Biomimetic tactile sensor array, *Advanced Robotics (RSJ), Special Issue on Tactile Feedback for Humanoids and Humans*, 22(7), 829–849, June 2008.
18. Carlson, J., Ghaey, S., Moran, S., Tran, C. A., and Kaplan, D. L., Biological materials in engineering mechanisms, Chapter 14, in *Biomimetics—Biologically Inspired Technologies*. Boca Raton, FL: CRC Press, pp. 365–380, 2005.
19. Johnson, E. A. C., Ronser, R. H. C., and Jeronimidis, G., Technologies recent advances in biomimetic sensing, *Philosophical Transactions of the Royal Society A*, 367, 1559–1569, 2009.
20. Hubbard, A., Cohen, H., Karl, C., Freedman, D., Mountain, D., Ziph-Schatzberg, L., Nourzad Karl, M. et al., Biologically inspired circuitry that mimics mammalian hearing, in *Proceedings of the Bio-Inspired/Biomimetic Sensor Technologies and Applications Conference*, Proceedings of SPIE, Orlando, FL, N. F. Fell and V. S. Swaminathan, Eds., 2009.
21. Bayley, H. and Martin, C. R., Resistive-pulse sensing-from microbes to molecules, *Chemical Reviews,* 100, 2575–2594, 2000.
22. Kwon, H., Knee, P., Spanias, A., Goodnick, S., Thornton, T., and Phillips, S., Transform-domain features for ion-channel sensors, in *Signal Processing, Pattern Recognition, and Applications*, Innsbruck, Austria, 2008.
23. Abdullah, J., Ahmad, M., Lee Yook H., Karuppiah, N., and Sidek, H., Evaluation of an optical phenolic biosensor signal employing artificial neural networks, *Sensors and Actuators B (Chemical)*, 134(2), 959–965, 2008.
24. Schmidhuber, M., Bähr, J., Ilchmann, F., Wiest, J., and Wolf, B., Space saving mixed signal FPGAs for improving processing power and memory capacity as a replacement for μCs in portable biosensor devices, *World Congress on Medical Physics and Biomedical Engineering, IFMBE Proceedings*, Munich, Germany, Vol. 25 IFMBE, No. 8, pp. 37–40, 2009.
25. Samir, M. E., Reda, A., and Sanguthevar, R., Artificial immune systems: Models, applications, and challenges, *Proceedings of the 27th Annual ACM Symposium on Applied Computing (SAC' 12)*, ACM, New York, ©2012, pp. 256–258, 2012.
26. Carpi, F., Kornbluh, R., Sommer-Larsen, P., and Alici, G., Electroactive polymer actuators as artificial muscles: Are they ready for bio-inspired applications? *Bioinspiration & Biomimetics*, 6(4), 040201, 2011.
27. Habib, M. K., Watanabe, K., and Nagata, F., Bioinspiration and emerging actuator technologies, *Artificial Life and Robotics*, 17(2), 191–196, 2012.
28. Webster, D. R., Volyanskyy, K. Y., and Weissburk, M. J., Bioinspired algorithm for autonomous sensor-driven guidance in turbulent chemical plumes, *Bioinspiration & Biomimetics*, 7(3), 036023, 2012.

4

Image Sensors

Mehdi Habibi
University of Isfahan

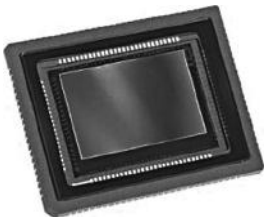
4.1 Introduction

Image sensors play an important role both in everyday life and in highly specialized tasks. Image sensors are used in webcams, cell phones, and digital cameras. As shown in Figure 4.1, their application can be confined to tasks such as low-powered miniaturized sensors, high-resolution imagers, high-speed acquisition systems, and machine vision processing. Each application requires certain specifications and characteristics of the image sensor. For example, in one instance, the power consumption can be the dominant factor, while in another application the sensor speed is critical. In portable devices such as mobile phones, low power consumption is preferred [1]. High-speed machine vision task used in industrial plants for food grain separation based on the grain size is an example where high-speed imagers with possible in-pixel processing are appropriate [2].

Conventionally, image sensors were designed using charged-coupled device (CCD) technology. CCDs have the ability to accumulate and transfer charge using potential wells produced by the device gate, but it is not possible to implement active devices that provide signal processing power. Currently, CMOS technology has an advantage over the CCD process in image sensor implementation due to the fact that it is possible to integrate active circuits in the pixel structure. For example, the integration of a simple source follower amplifier in the pixel structure can help reduce noise due to cross talk and neighbor cell signals. This structure is known as the active pixel sensor (APS) in contrast to passive pixel sensors (PPSs), which do not incorporate any in-pixel amplifiers. CMOS image sensor technology benefits from several other advantages such as higher sensor speed due to in-pixel processing of data, higher signal data range, and lower power consumption. CCD technology has the advantage of better charge transfer characteristics, which is important in correlation procedures. Thus, in certain correlation image sensors that are used in 3D range estimation, combinational CCD/CMOS processes are also used. The three different processes of CCD, CMOS, and CCD/CMOS are shown in Figure 4.2. Due to the diverse variations that exist in CMOS image sensor design, this chapter mainly focuses on CMOS image sensors. In the following sections, three groups of image sensors are discussed. Section 4.2 explains the main trends in CMOS imager design. Section 4.3 focuses on digital image sensors and different approaches in digital imagers. Section 4.4 presents image sensors with basic machine vision processing power. Finally, in the last section, some of the remaining image sensors with specialized functions are introduced and concluding remarks are presented.

Sony minature mobile camera

(a)



Foveon X3

(b)

NAC hotshot

(c)

Multigrain sorter from orange

(d)

FIGURE 4.1 Different applications of image sensors. They can be used for general-purpose tasks such as (a) mobile phone cameras, (b) camcorders, (c) high-speed acquisitions, or in (d) dedicated functions such as separation and sorting of food grains.

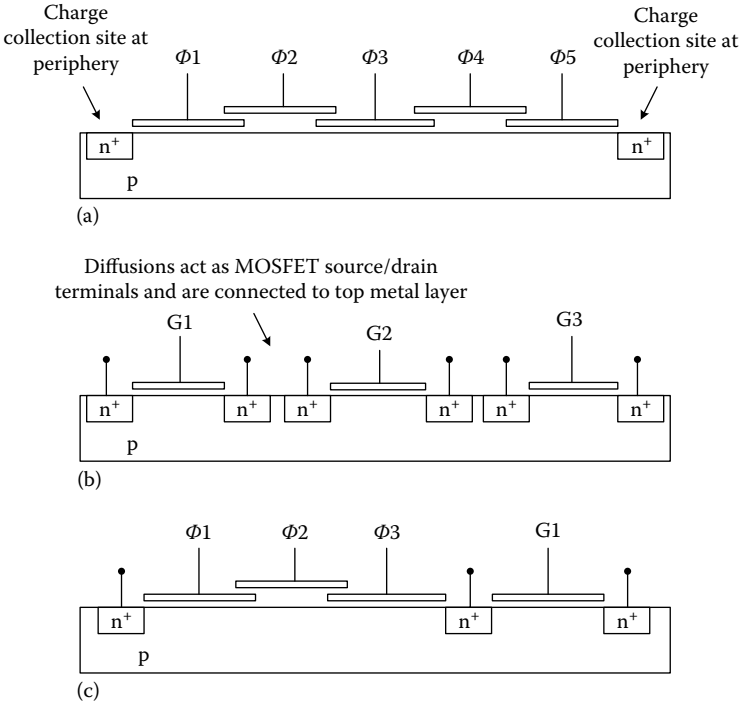


FIGURE 4.2 Basic structure of the (a) CCD, (b) CMOS, and (c) combinational CCD/CMOS technology.

4.2 Basic Pixel Trends

The basic pixel of an image sensor consists of a photoreceptor and photocircuitry. Common photoreceptors include photodiodes and photogates. Photodiodes have a simpler structure, are easier to operate, and have higher quantum efficiency. In contrast, photogates have lower quantum efficiency due to the absorbance effect of the gate layer but have better photo charge transfer characteristics. Currently, photodiodes are commonly used in image sensors built using CMOS technology. The current–voltage characteristic of a photodiode is shown in Figure 4.3. As the figure shows, there are three possible operation regions for the photodiode. The first is forward bias that is not used in image sensor designs since the photocurrent is completely insignificant compared to the main diode forward current and subsequently forward random noise. The second is the solar cell region where the photodiode is able to deliver power. This region is also not commonly used in image sensors since the capacitance effect of the photodiode is high, which results in a low-speed response of the photodiode to incident illumination. The region best suited for the photodiode to operate as a photoreceptor is the reverse bias region where the amount of current is proportional to the incident illumination.

The other part of a basic pixel is the photocircuitry, which includes devices attached to the photoreceptor and is used to convert the photocharge to an electric voltage or current signal. In its simplest form, the photocircuitry can be a simple reset switch that precharges the photodiode to a predetermined voltage. After the reset phase, the photocharge is integrated on a capacitor (namely, the capacitance of the photoreceptor itself). Subsequently, after an appropriate integration period, the amount of voltage on the capacitor is proportional to the photocharge. The photocharge in turn is proportional to the photocurrent, which is dependent on the amount of received illumination. Several issues have to be considered for the simple reset switch. For example, the implementation of the reset switch similar to the photoreceptor can significantly reduce the size of the reset switch (since no extra well is required) and hence increase the fill factor (ratio of pixel photoreceptor area to the overall pixel area) [3]. However, this configuration places the reset switch in the active region, and thus the precharge voltage will experience a threshold voltage difference from the supply potential. This voltage difference limits the output signal swing and can be important especially in submicron technologies where the supply potential decrease is faster than the threshold voltage decrease. Furthermore, as the precharge voltage increases, the current delivered by the reset switch decreases, which means that if the reset time is not sufficiently large, the amount of final precharge potential will be dependent on the previous potential of the photoreceptor. This effect results in an effect known as lag which can be suppressed using correlated double sampling (CDS) explained later. The use of a complementary reset device can solve both the precharge potential and voltage lag but suffers the drawback of smaller fill factor [4]. Another drawback of the complementary reset switch compared with the ordinary reset switch is as follows: In deep submicron technology, the short channel devices show considerably high leakage currents even at zero gate–source potentials. The solution is to use a negative gate–source voltage (for an nMOS device) to reduce the leakage current in the off state. The use of negative potentials is possible for an nMOS reset switch, but for pMOS switches, charge pump circuits are required to increase the gate voltage above the power supply

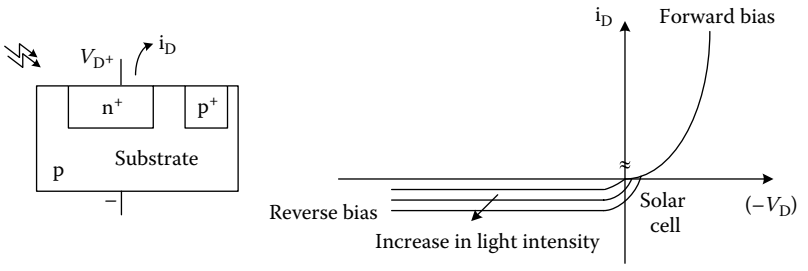


FIGURE 4.3 Photodiode current–voltage characteristics.

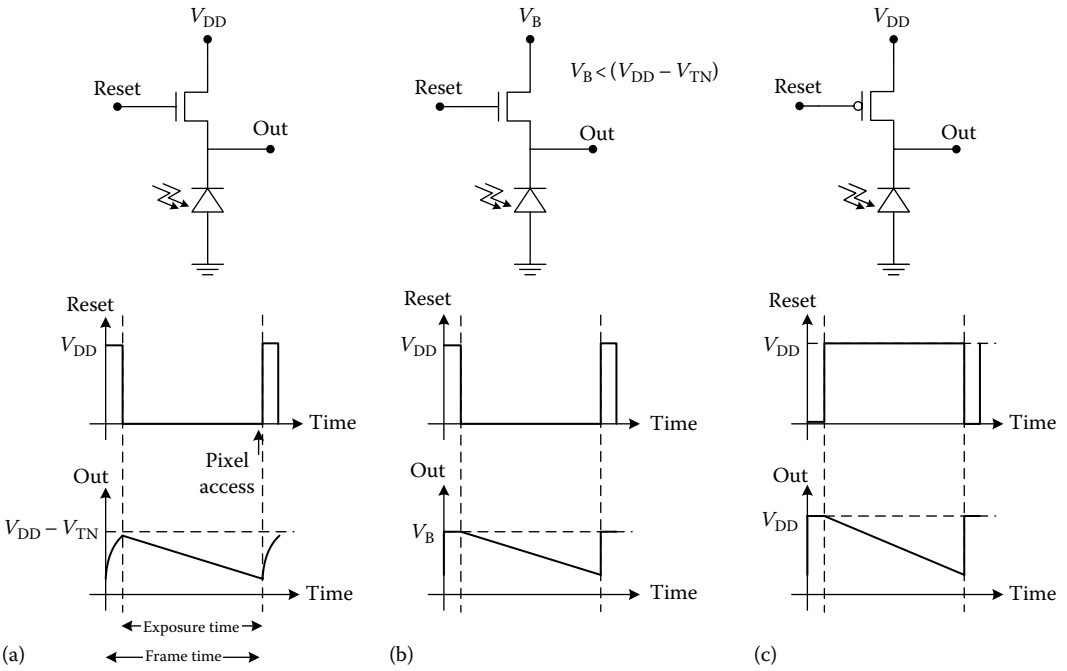


FIGURE 4.4 Reset switch configuration: (a) nMOS reset switch with supply voltage precharge potential, (b) nMOS reset switch with V_B precharge potential, and (c) pMOS reset switch.

precharge potential, or a precharge potential lower than the supply voltage should be provided. Note that if a lower potential is provided for the nMOS reset switch, the problem of image lag can be solved since the reset switch will be placed on the triode region. Figure 4.4 shows the three configurations, their control waveforms and output signals.

Diode-connected loads can also be used as photocircuits, which logarithmically convert the photocurrent into a voltage signal. This structure covers a wider dynamic range of illuminations (due to the logarithmic effect) but lower sensitivity is obtained, since a wider illumination range is mapped to a limited voltage swing range. Furthermore, since the photocurrent is relatively limited, the photodiode load MOSFET will operate in the subthreshold region, and thus mismatches in different load devices result in substantially different signal conversion characteristics from one pixel to another, which produces fixed pattern noise (FPN). There have been circuits presented that compensate this FPN using in-pixel methods [5,6]. The speed of diode-connected photocircuits can be increased by gain boosting methods.

Currently, the most common method remains to be the integration setup due to lower random noise. The reason is that the capacitor will act as an element that calculates the average current and thus random noise effects will decrease with higher integration periods.

Other common components in a basic pixel include a select switch MOSFET, which places the signal data on the common column bus. In the PPS array of Figure 4.5, no amplifier is used in the pixel circuit. A single reset switch can be used for the entire column of a PPS structure [7].

With the incorporation of a common drain MOSFET that acts as a unity gain voltage buffer in the pixel structure, an APS is achieved. This configuration is shown in Figure 4.6a. If the pixel voltage is to be placed on the common column bus directly (similar to the PPS case), charge charging happens between the column data bus capacitance and the photoreceptor capacitance, and thus the final column signal will be dependent on the previous pixel data on the column bus. Also, the parasitic capacitances between the column bus and the photodiode will change the actual photodiode potential during the column bus activity and create substantial cross talk. With the voltage buffer, the photoreceptor crosstalk

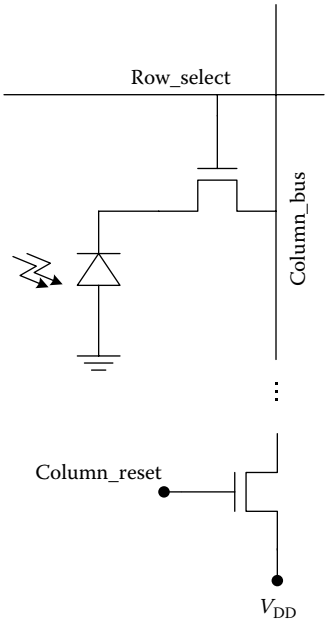


FIGURE 4.5 PPS sensor structure with a common reset switch per column.

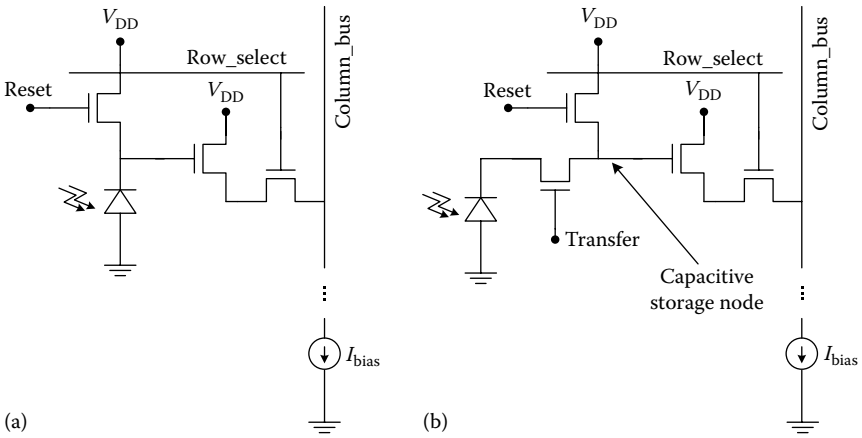


FIGURE 4.6 APS sensor: (a) Without storage site and (b) with an extra analog storage site and transfer gate.

and column bus charge sharing is eliminated. Still, mismatches between the source followers result in FPN effects, which can be compensated by CDS methods.

Analog pixel arrays usually have a common reset signal. After the pixel array readout, all cells are precharged simultaneously. After the desired amount of exposure time, the pixel data needs to be accessed. To avoid further charge accumulation on the remaining pixels during the readout cycle, a mechanical shutter can be used. If the pixel access time is short, the use of this shutter will not be necessary. Another option to avoid the mechanical shutter is the use of electronic shutter. As shown in Figure 4.6b, in the electronic shutter method, an extra transfer gate and storage site are required to keep the analog photodiode signal while other pixels are accessed [8]. The use of photogate in Figure 4.7 can enhance the transfer operation since by changing the photogate potential, the charge displacement is controlled more effectively. The timing waveform of Figure 4.8 shows one possible approach to control the electronic shutter pixel. In the figure, the difference between exposure time and frame

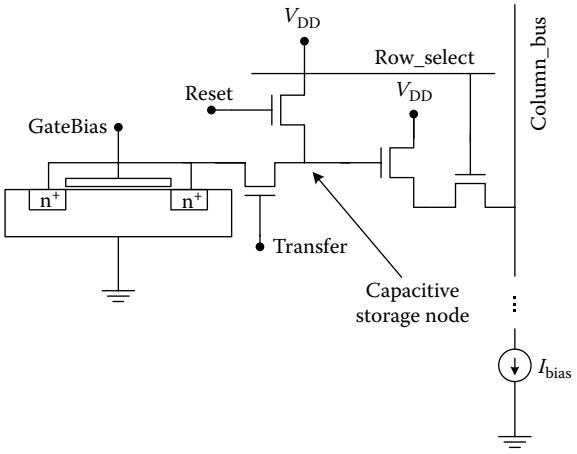


FIGURE 4.7 Electronic shutter implementation with photogate.

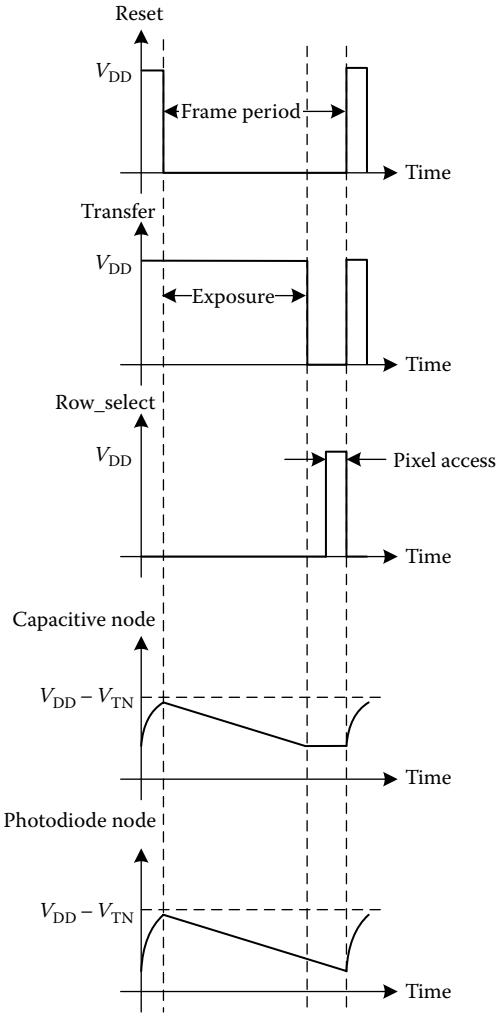


FIGURE 4.8 Electronic shutter control waveform.

rate should be noted. The exposure period is the effective amount of time that the photodiode integrates photocurrent, but the frame rate is the inverse of average time where a frame is transferred to the output and a new frame capturing sequence begins. The maximum frame rate will be limited by the required exposure time.

In order to increase the fill factor, the number of required transistors per pixel in an APS device can be reduced by using one amplifier for a group of photodiodes. This configuration is shown in Figure 4.9. The different colors of a pixel are an appropriate choice for sharing a common amplifier. Note that as the number of pixels per common amplifier increases, the charge sharing and cross talk also increase. Amplifier sharing of 4 or 5 photodiodes results in 1.75 and 1.5 transistors per pixel in APS devices, respectively [9,10].

Another important issue in the basic pixel structure is the color consideration. A photodiode's photocurrent varies with wavelength (or color), but a single photodiode itself cannot distinguish one color from another since the photocurrent also varies with intensity. Color filters are an effective solution to assign certain colors to different pixels. These filters are repeated throughout the sensor in an appropriate color pattern. For example, the Bayer RGB is the one that is usual among image sensors. Different polymeric dye masks are available for this purpose. Another alternative for color implementation is the use of metal grids that act as gratings [11,12]. These gratings can separate different wavelengths of light and focus each color on appropriate photodiodes. The advantage with this technique is that it can be implemented with the standard CMOS process using the metal layers. Another option for implementing color pixels is to use multiple photodiodes with different quantum efficiencies within a single pixel [13,14]. For example, a pixel might contain three different photodiodes where each has peak quantum efficiency at red, green, and blue wavelengths. Considering the signal produced on the three photodiodes, the pixel circuitry can extract the three essential intensity, color, and saturation information of that pixel. Fortunately, the implementation of different photodiodes can be easily implemented by implanting multiple p–n junctions at different depths, which is also compatible with the standard CMOS technology [15]. An advantage with this configuration is that unlike the color mask [16] and grating-based method, each

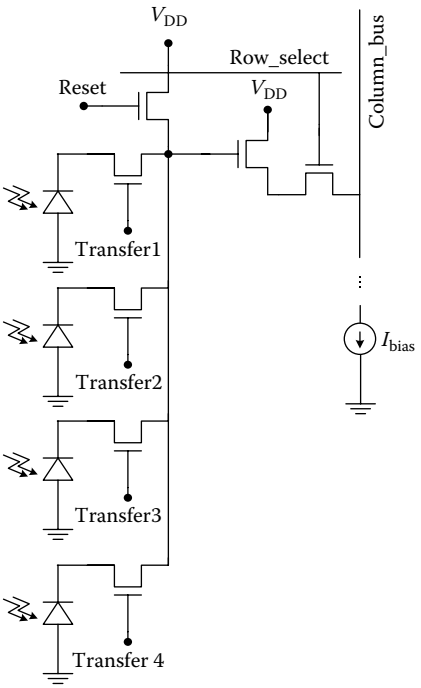


FIGURE 4.9 Pixel structure with amplifier sharing.

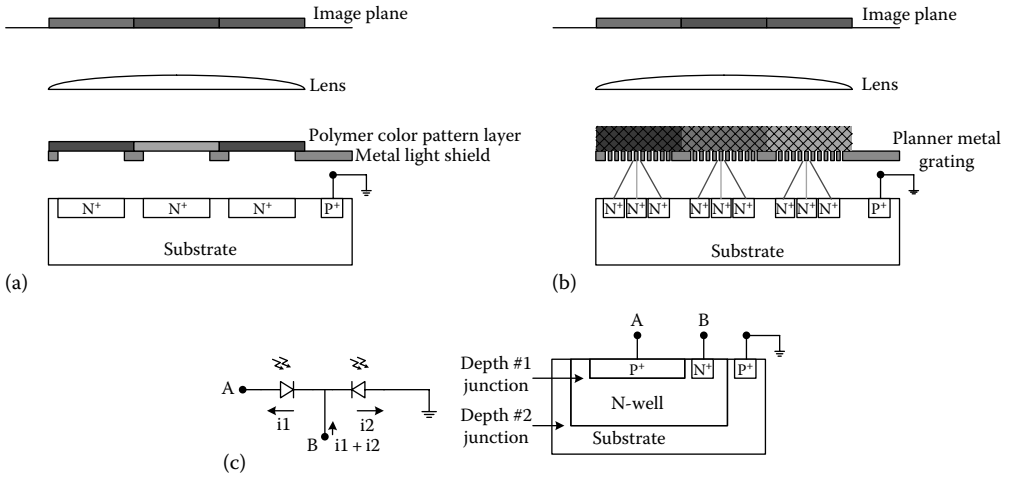


FIGURE 4.10 Color filtering techniques: (a) Polymeric dye color masks, (b) planner metal grid gratings, and (c) multidepth junctions.

pixel can produce the data signals of all the required three colors. Also, no extra metal or dye layers are required, which reduces the chip cost. Figure 4.10 shows some of the different possible configurations for color sensor implementation.

An extra layer that is occasionally used in basic pixel structure is the microlens layer [17]. These lenses act to focus the received illumination directly on the photodiode. This approach is effective in increasing the sensitivity of pixels with low fill factor. With pixels that include significant in-pixel circuitry and the fill factor is low, the microlens can help direct the incident pixel illumination onto the actual photodiode. Planner microlenses that are implemented with metal layers are also used in some sensors [18]. Figure 4.11 shows some of the different structures used in microlens implementation.

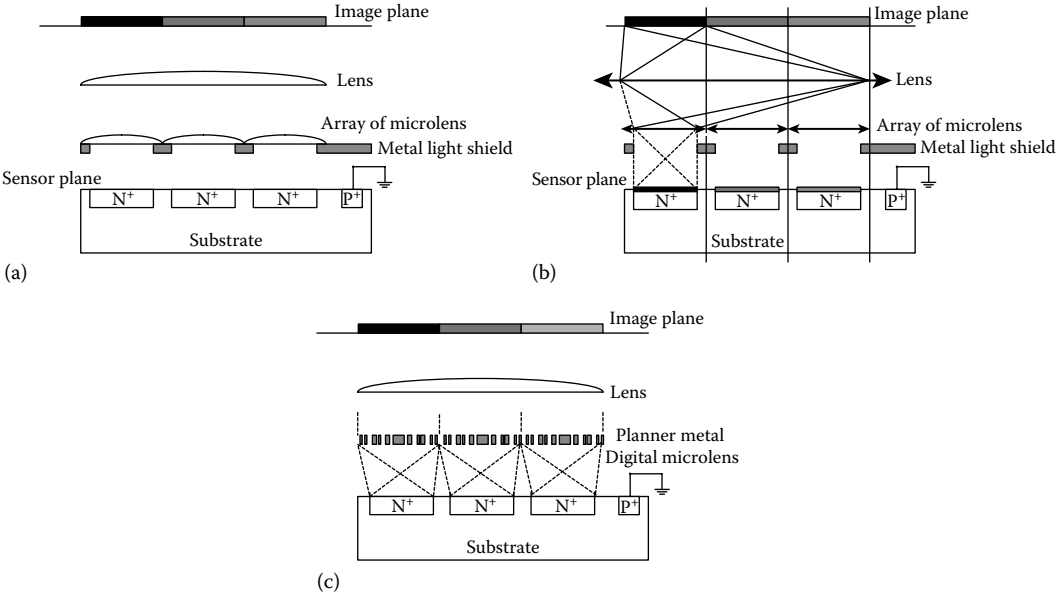


FIGURE 4.11 Microlens implementation techniques: (a) Conventional microlens fabrication using resins, (b) operational principle of the microlens for directing light to the actual photodiode, and (c) planner metal structure for microlens implementation.

FPN and random noise are other important aspects in a basic pixel that need to be addressed. If an image is captured in complete dark conditions, it will contain two components. One is the FPN, which is referred to as the part that is constant in successive image captures, and the other is random noise, which randomly changes in successive captures. These two noise sources have different natures. Since the FPN is constant, it can be easily removed in further processing steps. There have been sensors designed that can eliminate FPN. Since the random noise has nondeterministic characteristics, it mainly determines the sensor noise floor and specifies the minimum sensor detectable signal.

The random noise originates from the photodiode shot and thermal noise and also amplifier flicker noise, shot noise, and thermal noise. The reset switch also produces random noise but can be removed with appropriate methods. The FPN originates from device mismatches and difference in MOSFET parameters. Since the device mismatches are formed during fabrication and are constant afterward, the noise effect caused by this factor is considered fixed. For example, the most important factor that produces FPN in APS pixels is the threshold voltage variations of the source follower MOSFET. The nMOS reset switch will also produce FPN due to the threshold voltage drop variations on the photodiode precharge potential. The pMOS reset switch however does not contribute to FPN since it operates in the triode region and directly precharges the photodiode node to the supply voltage.

The CDS method is an effective solution for eliminating FPN and also random reset noise by capturing two samples before and after integration. The subtraction of the two samples eliminates the noise due to the reset switch random and fixed noise and also the amplifier fixed noise. An issue with this approach is that if the samples are to be taken at the beginning and end of integration, an analog memory array is required to store all samples captured at the beginning of the integration and wait for the integration to complete and start the second capture period and subtraction. The solution for this problem is to use the structure of Figure 4.6b, reset the photodiode, wait for the integration period to complete (give the photodiode the required exposure time), and perform the two captures subsequently before and after the transfer phase [19,20]. With this method, the correlation double sampling can be performed as each pixel is accessed and no extra storage site is required. The corresponding signal waveform required for CDS is shown in Figure 4.12.

Section 4.3 will focus on image sensors that are able to convert the pixel analog signal to a digital data.

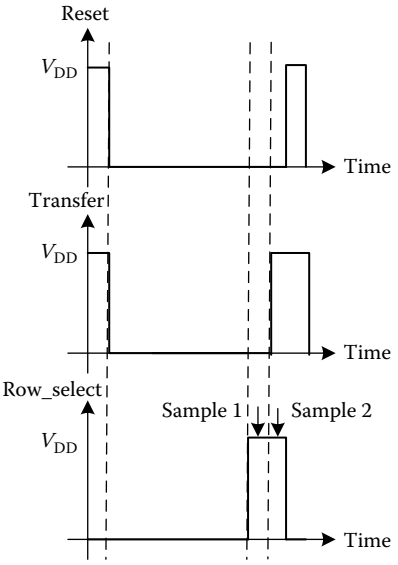


FIGURE 4.12 CDS control signals.

4.3 Digital Image Sensors

As pointed out earlier, a significant advantage of CMOS image sensors is that they can incorporate processing circuitry in the sensor structure. The processing circuitry can help amplify the pixel signal, increase the signal dynamic range, reduce FPN, etc. The processing circuitry can also be used to convert the analog pixel data into digital form.

The analog-to-digital conversion of data in a digital image sensor can be performed on-chip, in-column, or in-pixel. With the on-chip configuration, an embedded analog-to-digital converter (ADC) is used to convert the scanned pixel array analog data into digital code. Since most signal processing systems work with digital data, system costs will be reduced by eliminating the external ADC. Furthermore, basic and common digital signal processing stages such as color and white balance correction, focus enhancement, and 50/60 Hz flicker correction can also be imbedded in the sensor chip [1].

With the column parallel ADC structures, an ADC is placed at each column, and analog data are converted to digital in a column parallel sequence. The speed of the sensor can be significantly increased since the main bottleneck in an ordinary image sensor is the time required for each conversion. Usually with the column parallel ADC, no mechanical or electronic shutter mechanism is required, since all pixels can be accessed and converted to digital in a short time after the desired duration of exposure period. Furthermore, image sensors with column parallel ADC have been presented, which can accept extended image dynamic range by changing the shutter (integration) period for different pixels [21]. This is accomplished by using multiple exposure captures on the array. On each column, the sensor uses a single-slope ADC for one long-exposure capture and a variable-slope ADC for multiple short-exposure captures.

The column parallel approach can increase the frame rate, but for high-resolution and high-speed imagers, reduction of the data conversion time remains a challenge. In this case, the in-pixel conversion of pixel data is appropriate. This pixel configuration is also known as the digital pixel sensor (DPS). Versatile and different DPSs have been proposed, which operate on different principles. They mainly differ in complexity, dynamic range, and error. In a DPS, the photodiode potential is usually compared against a reference voltage to create a digital match signal. As shown in Figure 4.13, they can be divided into major oscillatory and nonoscillatory groups. In the nonoscillatory approach, the pixels match signal changes once in each frame. After the reset phase (with the starting of each frame), the procedure is repeated. The match signal pulse width can be used to evaluate the analog signal intensity. In the oscillator category, the match signal is used in a positive feedback loop to reset the photodiode, and thus a periodic output signal during one frame capture is produced. The pixel design is such that the pixel photocurrent changes the periodic digital output pulse characteristics. Figure 4.14 shows the different output signals that can be produced based on the two methods.

Figure 4.15 shows some of the different methods that can be used to create a match signal in the nonoscillatory approach. In Figure 4.15a, the reference signal is initially constant until the integration

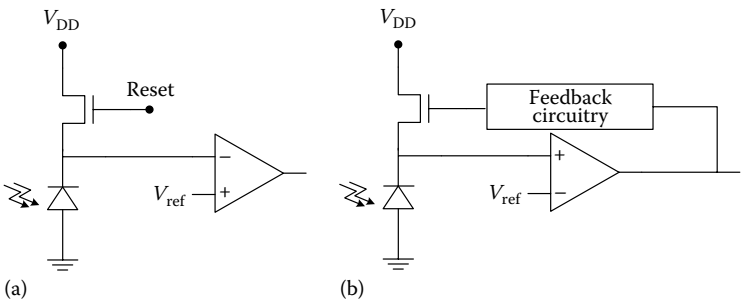


FIGURE 4.13 DPS pixel concept: (a) In-pixel circuitry for the oscillatory approach and (b) in-pixel circuitry for the nonoscillatory approach.

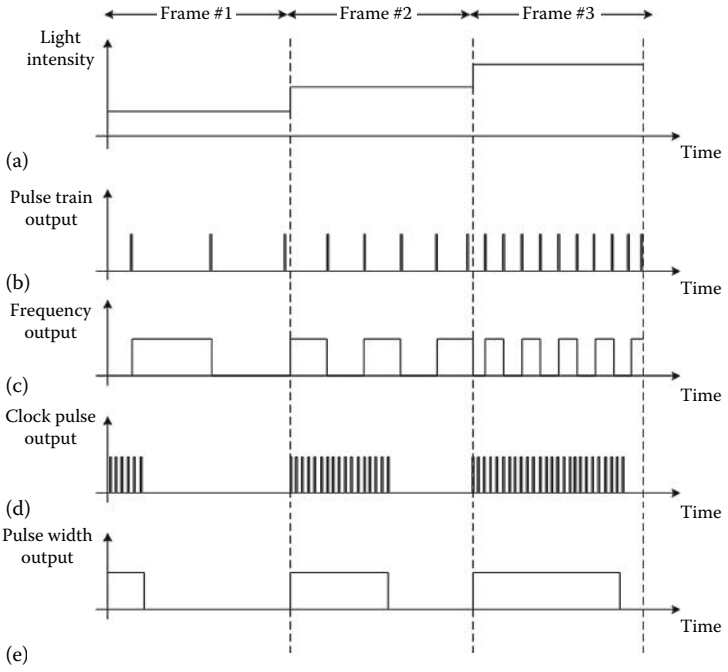


FIGURE 4.14 Different output waveforms that can be produced by the oscillatory and nonoscillatory methods. Waveform in (a) shows an arbitrary incident illumination in different frames. In each frame, the incident illumination is dependent on (b) the number of output spikes, (c) the frequency of the output pulse, (d) the number of clock pulses transferred to the output, and (e) the pulse width.

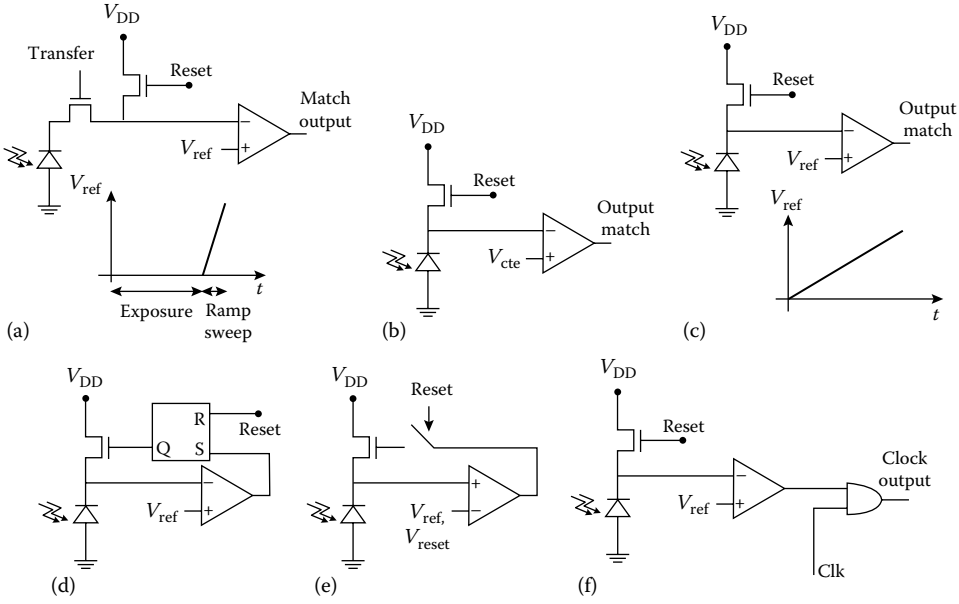


FIGURE 4.15 Different methods of the nonoscillatory approach: (a) Simple ramp sweep, (b) photodiode voltage sweep, (c) combined ramp sweep and photodiode voltage sweep, (d) reset latch method, (e) comparator offset voltage compensation method, (f) clock gating technique.

period is completed on all pixels. Then the reference signal is swept across the photodiode voltage output range using a ramp waveform. This method produces an output match signal where the pulse width is linearly proportional to the light intensity [22]. The drawback is that all pixels should have the same integration period, and hence the supported light dynamic range can be limited. In Figure 4.15b, the reference signal is kept constant, and the match signal pulse width is the time required for the descending photodiode voltage to reach the reference value. In fact, the photodiode ramp signal is used to produce the match output. Since the photodiode ramp signal steepness is dependent on the photocurrent, the output signal width will vary with the incident light intensity. In this method, the light dynamic range is relatively extended, but the pulse width is not linearly dependent on the photocurrent. The drawback of this method is that the conversion resolution is low for small photocurrents (since at low light intensities and illuminations, the ramp signal variation is small). The pixel structure of Figure 4.15c combines the two methods and uses both the photodiode ramp signal and a ramp signal on the reference voltage to increase the output signal resolution at low illumination levels. DPS structures with multislope ramp signals have also been presented to achieve both high dynamic range and desired resolution. In Figure 4.15d, a feedback is used to reset the photodiode as soon as the match signal occurs [23,24]. Due to the latching effect of the RS flip-flop feedback, this method does not produce any oscillations but will enhance the sensor's performance by preventing photodiode saturation and blooming (due to the photodiode entering forward bias and producing a negative potential on the photodiode output node) and also allowing more time for the reset switch to bring the photodiode node potential to its final value and preventing image lags. A major issue that needs to be considered in the design of DPS structures is the offset voltage of the comparator. The comparator can show substantial offset voltage specifically at submicron channel lengths. Since the comparator should be fitted inside each pixel, the use of small MOSFETs will give more space to the photodiode and hence increase fill factor. With small MOSFETs, however, random parametric variations and offset voltages are greater. The method used in Figure 4.15e to overcome this issue is to reset the photodiode to its precharge potential through a negative feedback loop which contains the comparator [25]. With this method, the comparator's offset voltage will be included in the reset potential, and the amount of time required for the photodiode to reach the reference voltage plus the offset value will be independent of the offset itself. As explained briefly, each method gives different dynamic range, complexity, fill factor, and power consumption. The comparison signal provided by this method can be used in a number of ways. One is to use this signal to store the corresponding time stamp in an in-pixel memory. The time stamp is provided by a global digital counter, and when the comparison signal occurs, the counter value is stored in the corresponding memory, producing a digital code per pixel. Other sensors use the comparison signal to generate an address event for the pixel readout [26]. The time stamp of the generated signal is a representation of the pixel value.

If the match signal is used to continuously reset the photodiode node, the oscillatory approach is achieved. In Figure 4.16a, a number of inverters are used in the feedback loop to produce some delay [27]. The delay gates provide enough time for the photodiode node to reach its final recharge value. In Figure 4.16b, a clocked delay is used for the same purpose. In the oscillatory approach, it is difficult to use a variable reference signal for the pixels since the resetting of the pixels are not synchronous and thus the ramp signals should be produced individually for each pixel; however, in [28], an in-pixel ramp generator has been presented to solve this issue and obtain an in-pixel pulse train converter with enhanced operation at low illumination levels. The pixel structure is shown in Figure 4.16c.

The sigma–delta approach uses constant charge packets to keep the photodiode potential oscillating around a constant voltage value [29]. The charge packets are transferred with a synchronous clock signal. The result is a constant pulse width signal where its repetition rate varies with the photocurrent. Since the photodiode potential continuously oscillates around a constant dc value, the structure is independent of the comparator offset properties. The pixel structure is illustrated in Figure 4.17.

Usually, in DPSs, a digital code is required for each pixel. As explained earlier, in the nonoscillatory approach where the match signal only changes once in each frame, in-pixel digital memories are used to store the time where the match signal changes state. The digital time stamp is provided by a binary

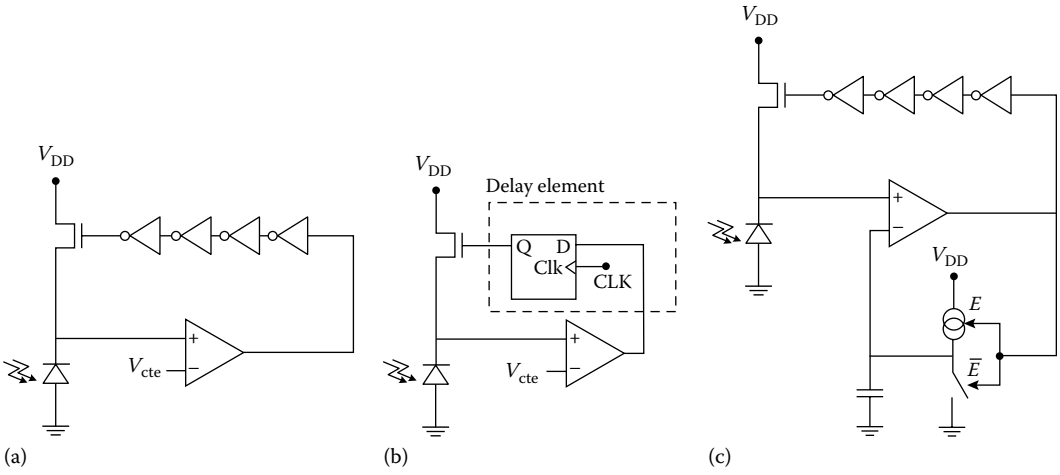


FIGURE 4.16 Different configurations of the oscillatory approach. As can be seen, all the structures contain a positive feedback loop that causes the oscillations: (a) Feedback loop with delay gates, (b) feedback loop with clocked delay, and (c) in-pixel ramp generator.

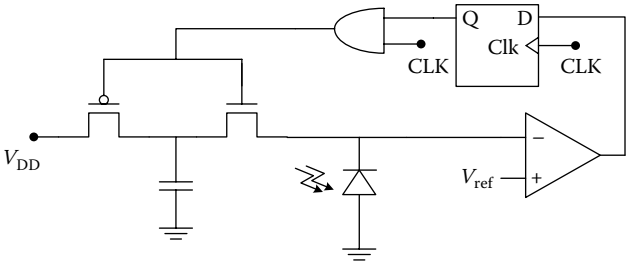


FIGURE 4.17 Sigma–delta approach and concept.

counter routed to all the pixels. Both dynamic and static memories can be used for this purpose. In event-driven sensors, the match signal is used to create an address event and the external sensor processor uses the time stamp where the event has been generated to store the digital value in an external memory. In large pixel counts, this method is less intuitive since the arbitration of several events at the same time is complicated. As will be explained later, the event-driven method is more effective in identifying pixels with value change compared with the previous frame since it can significantly reduce the data to be transferred to the output. In the oscillatory approach where the pixel match signal frequency is changed with the light intensity, in-pixel counters are used to produce the required digital code. Both binary counters [30] and linear feedback shift registers (LFSRs) [31] are used as in-pixel counters. LFSRs do not produce binary codes, but they need less silicon space.

The in-pixel conversion of analog data also benefits from the fact that each pixel can have an independent exposure period. This means that some pixels can be adjusted to accept low illumination levels while others are set for higher values; hence, a higher dynamic illumination range can be supported by the sensor. This is in contrast to an analog array pixel where the exposure time of all pixels should be adjusted simultaneously (i.e., high shutter speeds should be used for high illuminated environments and low shutter speeds at low illuminated conditions).

The main bottleneck for the DPS speed is the photocurrent random noise and the minimum time required for the photodiode to accumulate enough charge that is considerably higher than the charge produced by random and shot noise effects. PIN photodiodes produce less noise, but the current commercial high-speed sensor speeds are around a few 10 kfps. As the shutter speed decreases, more

illumination will be required to accumulate the required charge. Therefore, high-speed image sensors usually require a special illuminated environment. Ultrahigh sensitivity photoreceptors also exist that are able to acquire images at normal illuminations and low shutter exposure time, which will be discussed later.

4.4 Vision Sensors with Processing Power

With the possibility of designing pixels which include active devices, in situ image processing has also become an option.

If static power dissipation is avoided, analog processing circuits can significantly reduce power since analog processing circuits do not require an array of bits and are typically smaller with lower capacitance effect and thus lower dynamic power dissipation.

Image processing functions and machine vision tasks tend to be time consuming, process intensive, and diverse. CMOS image sensors have the power and ability to perform all or part of the required processing using active MOSFET devices. The processing can be on-chip, column parallel, or in-pixel. These types of image sensors with processing capabilities are known as vision chips. Vision chips are able to considerably increase the image processing speed and reduce power consumption. Since image processing and especially machine vision tasks are diverse, usually only fundamental and basic tasks are implemented in the sensor structure. Some of the common tasks implemented on vision sensors are image correction and enhancement functions, edge detection, optical flow extraction, change detection, kernel filtering, morphologic operations, and image compression. Figure 4.18 shows some of the fundamental image processing operations.

A number of general-purpose programmable vision sensors have also been presented, which can be configured to perform most of the basic and fundamental image processing tasks. In the following, some of the structures used in programmable vision sensors will be investigated.

The tasks usually supported by typical programmable vision sensors include kernel convolution for use in different types of common filters (Gaussian, Laplacian, Laplacian of Gaussian, and edge detectors), motion detection, and morphologic operations. Most of these tasks can be categorized and implemented in parallel using the cellular neural network (CNN) concept. These tasks are primarily used in all machine vision applications as the initial processing steps required to prepare the data for further classification, recognition, and decision-making steps. The initial steps have the properties of being performed in parallel and are thus ideal for in-pixel vision chip implementation to increase the processing speed. They can be easily implemented using a single instruction, multiple data architecture model where a single instruction is issued by the sensor for all of the pixels. The pixels then perform the same operation on their own and possibly neighbor data. This type of operation is usually referred to

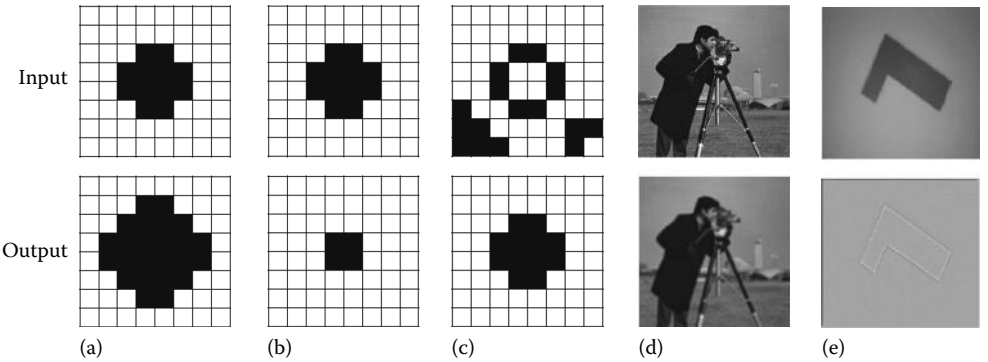


FIGURE 4.18 Fundamental image processing functions usually used in the initial steps of machine vision tasks: (a) Erosion, (b) dilation, (c) flood fill of center circle, (d) Gaussian filter, and (e) Laplacian filter.

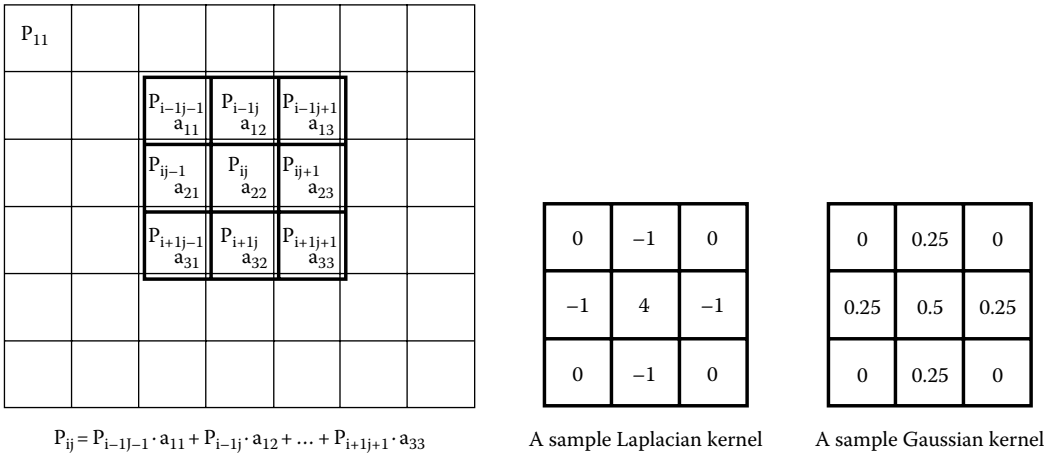


FIGURE 4.19 Kernel convolution process.

as pixel-based local functions. The final processing steps and specially object-based tasks might require sequential programming and are thus left to be performed out of the vision chip sensor. Although there have been dedicated sensors presented that even perform the final object processing step inside the chip with the aim of eliminating mass data output from the sensor, increasing speed, and reducing power consumption [32], such sensors usually have application-specific usages.

The kernel convolution procedure involves the weight sum of each pixel and its neighbors at a specific distance (usually nearest neighbors). The task is performed by multiplication and summation of the 2D matrix known as the kernel on each of the image pixels and its neighbors as shown in Figure 4.19. Edge detection involves extraction of pixel sites where high special intensity change exists in the image. This procedure can be accomplished by calculating the sum of absolute differences of each pixel with its neighbors or using an appropriate kernel convolution sequences like Laplacian of Gaussian. Motion detection is the extraction of pixels with temporal intensity change.

CNN operations refer to tasks that involve the calculation of a result for each image pixel value by considering a function (usually sum, multiplication, and hard limiter) of the pixel value itself, the pixel value of the neighbors, the state of the pixel itself in the previous frame(s), and the state of the neighbor pixel in the previous state. This definition, in general, forms the CNN operations that can accomplish all of the aforementioned procedures. Furthermore, it is possible to perform morphologic operations using multiple iteration executions of a CNN operation on one image capture. Morphologic operators are also fundamental building blocks of machine vision algorithms in the initial step. Examples of such operation include erosion, dilation, region fill, and others. For example, in the fill operation, a desired region is highlighted until the region edges are reached. In this way, a specific object can be separated from the rest of the image. CNN operations need more hardware since the result from the previous states should be stored if the result of previous iterations is required.

Different structures can be used in programmable vision chips. They can be categorized as follows: vision sensors with 1 bit image processing power where the image is initially binarized and all subsequent tasks are performed with single-bit in-pixel digital logic circuits [34]. With the second category, the image data are converted to digital code, and subsequent processing is performed using multibit in-pixel processing circuitry [34–38]. In the third category, digital processing is performed on bio-inspired pulse trains [39,40]. In these types of sensors, a parallel data bus that stores the digital intensity value of the pixel does not exist; instead, the pixel intensity value is represented by a single-bit digital data line using pulse trains. The processing of data is also performed in the pulse train domain by eliminating or adding extra pulses. Figure 4.20 shows a conceptual illustration of the pulse train processing element.

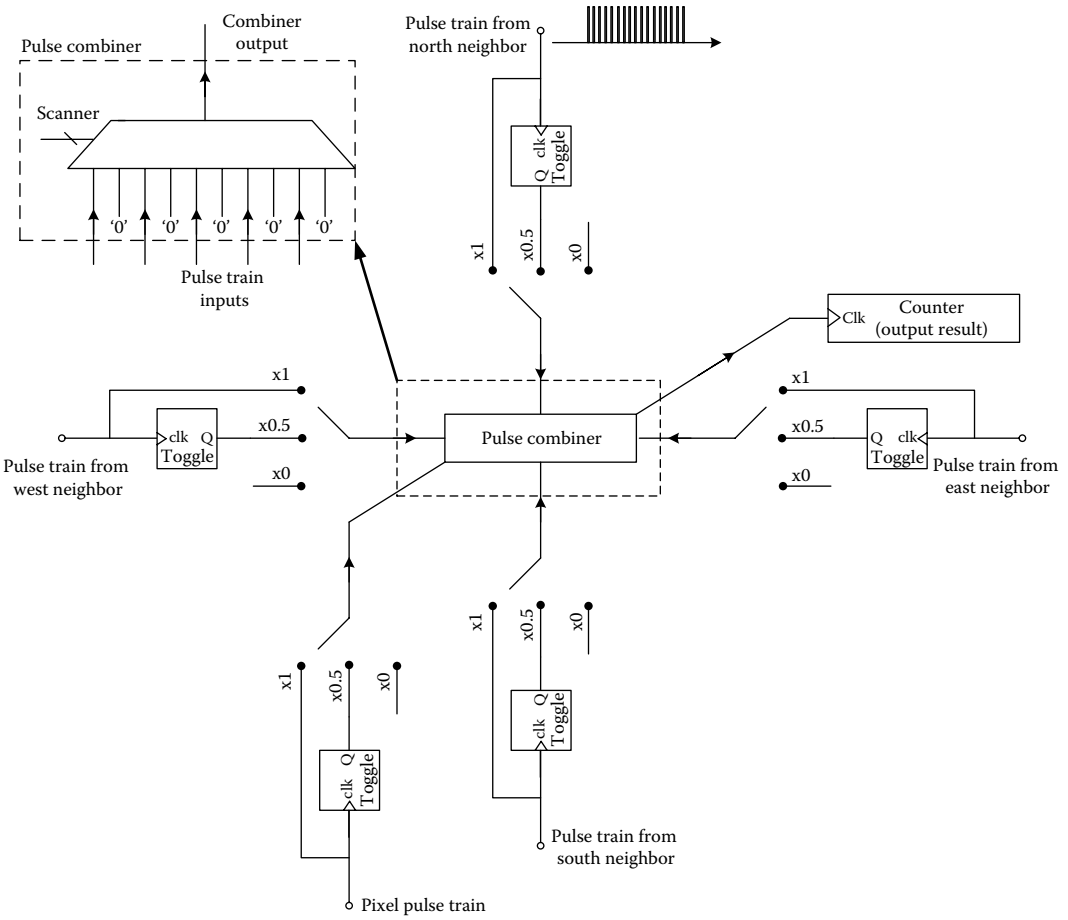


FIGURE 4.20 In-pixel pulse train processing element concept.

In the simple structure, three different coefficients can be applied to each of the pulse train signals, namely, 0, 0.5, and 1. The final pulse signals are added using a combiner that scans the different inputs.

The in-pixel pulse train itself is produced by the circuits presented in the DPS section. This type of processing is more appropriate for use in kernel convolution and is not suitable for CNN tasks since the storage of the previous state is not feasible on pulse trains. The fourth category performs the required processing on analog photodiode data directly [41–43]. Different mixed signal circuits can be used in this case and several options are incorporated: Figure 4.21 shows the voltage processing technique where switch capacitor circuits perform summation [42]. For multiplication, the capacitor weight is adjusted either by using digital code or by repeating the addition step.

Figure 4.22 shows the current processing approach [43]. In this method, addition is performed by current mirrors and current summation in a common node. Multiplication is achieved using multiple addition steps or by adjusting the current mirror ratio by a digital code weight. In this case, special current memories are required to store the current value of previous states. The SCAMP3 programmable image processor is based on this type of architecture.

Figure 4.23 shows the SCAMP-3 commercial product and its associated software for programming the configurable vision chip.

In some designs, differential pairs are used as multipliers where the photocurrent is used as the tail current source and the coefficients are applied as the differential inputs using analog voltage levels stored on floating gate devices. The basic concept is illustrated in Figure 4.24.

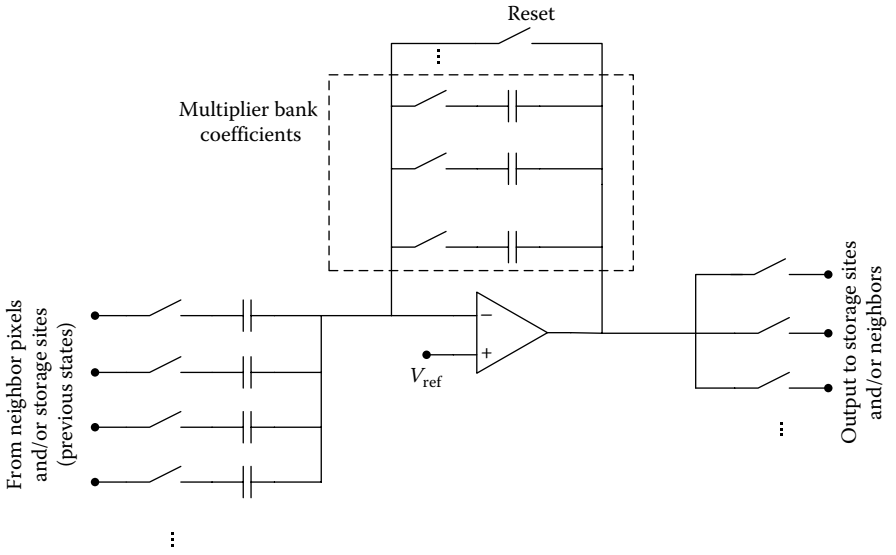


FIGURE 4.21 Voltage processing approach for programmable vision chip.

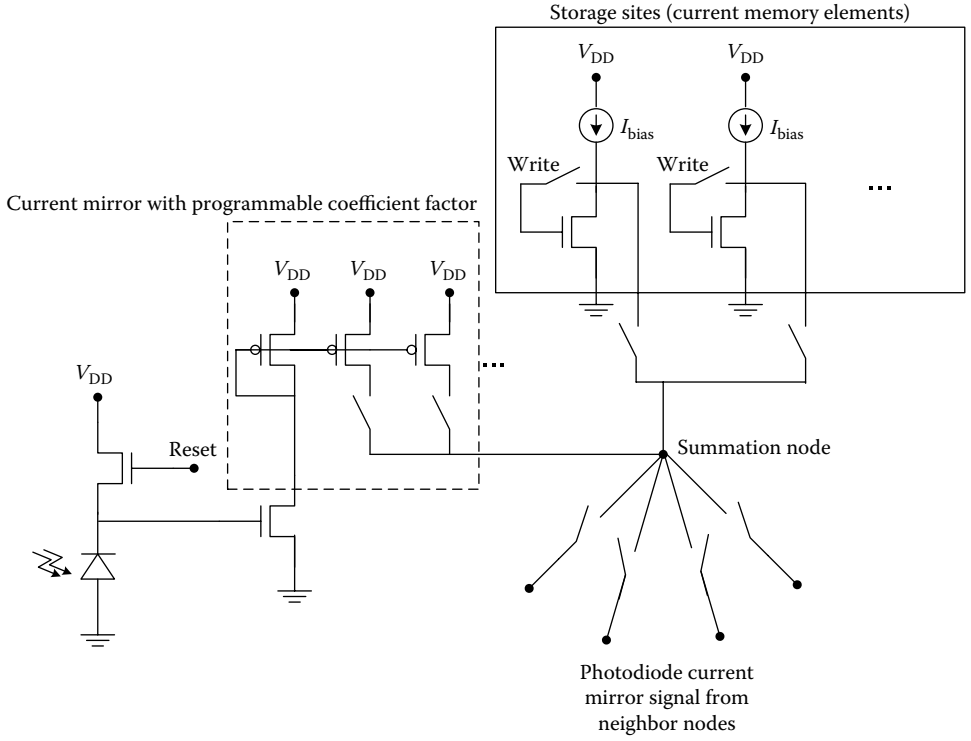


FIGURE 4.22 Current processing approach for programmable vision chip.

For the on-chip and column parallel cases, digital processing circuits are more appropriate since the larger size of the processing circuitry (especially at higher bit counts) does not sacrifice the pixel fill factor.

In-pixel digital processing circuitry is advantageous compared with on-chip digital processors for two main reasons: first, the processing speed can be increased significantly and second, the dynamic

FIGURE 4.23 SCAMP-3 programmable vision chip.

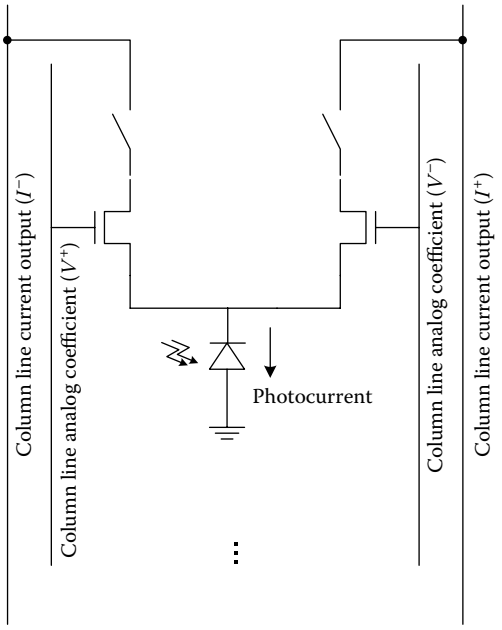


FIGURE 4.24 Current multiplier concept.

power dissipation can be considerably reduced. The reduction of dynamic power is because with multiple in-pixel digital processors, the clock pulse frequency can be reduced. To this point and with a constant processing speed, no advantage will be gained since although the clock frequency is reduced, the number of processors and capacitance effect has increased ($P_{diss} = f \cdot C_L \cdot V_{DD}^2$). But the point is that slower digital processing circuits can be operated with lower supply voltage. Subsequently, with a lower clock frequency, the supply voltage can be reduced, resulting in overall lower-power dissipation.

The problem of low photodiode fill factor due to digital circuits is getting less important with submicron devices. The photodiode cannot be decreased in size since random noise effects will increase, however, the digital circuit area can be decreased in sub micron devices which results in higher fill factors.

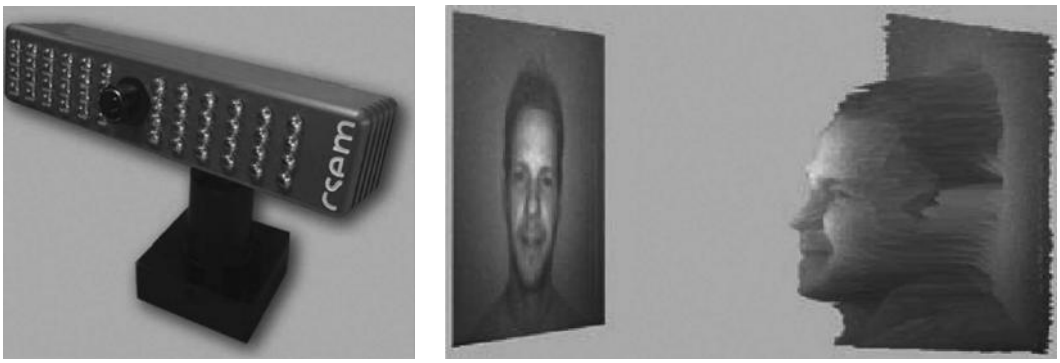


FIGURE 4.25 Swissranger 3D range image sensor.

4.5 Final Remarks

Currently, CMOS image sensors are undergoing extensive progress, and each year, many new enhancements and advances are achieved in sensitivity, resolution, frame rate, dynamic range, processing power, etc. In this chapter, some of the important image sensors that are currently being incorporated in practical applications were presented; still, there are many other image sensors continuously developed for specific tasks. An important image sensor that needs to be addressed here is the correlation image sensor. These sensors are able to route the photocharge to different storage sites. If different weights are considered for the voltages stored at the storage sites, the procedure can be equivalent to multiplication and integration (namely, correlation) of the photocurrent by a reference signal. This procedure can be significantly effective for tasks such as measurement of the phase delay of a transmitted illumination source and the reflected light and thus the detection of object distance and finally extraction of range image. For range image extraction, high-frequency correlation is required, and thus CMOS/CCD process is more appropriate [44]. Figure 4.25 shows a 3D range imager from Swissranger.

The correlation procedure is also effective in identifying and separating modulated beacons and electronic light tags from the rest of the scene [45,46]. Compression image sensors are another field where image sensors are experiencing development [47]. Since the compression is time consuming and also a good candidate for parallel processing, it is suitable for in-pixel vision chip implementation with low-power and high-speed features. Image sensors with power delivery capabilities [48,49] are also of certain interest. They are specifically important in applications with limited energy resources such as insect robots, miniaturized unmanned flying vehicles (UFVs), and prosthesis sensors. Ultrahigh sensitive sensors, another important field in imager design, are used in high-speed imagers, night vision, and 3D image captures [50,51]. Currently, research still continues on enhancement of different image sensors and further developments should be expected in both the sensors' specifications and designs.

References

1. K. Yoon, C. Kim, B. Lee, and D. Lee, Single chip CMOS image sensor for mobile applications, *IEEE Journal of Solid State Circuits*, 37(12), 1839–1845, 2002.
2. T. Pearson, Hardware-based image processing for high-speed inspection of grains, *Computers and Electronics in Agriculture*, 69(1), 12–18, 2009.
3. H. Y. Cheng and Y. C. King, An ultra-low dark current CMOS image sensor cell using n+ ring reset, *Electron Device Letters*, 23(9), 538–540, 2002.
4. C. Xu, W. H. Ki, and M. Chan, A low-voltage CMOS complementary active pixel sensor (CAPS) fabricated using a 0.25 μm CMOS technology, *Electron Device Letters*, 23(7), 398–400, 2002.

5. S. Kavadias, B. Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts, and J. Bogaerts, A logarithmic response CMOS image sensor with on-chip calibration, *IEEE Journal of Solid-State Circuits*, 35(8), 1146–1152, 2000.
6. D. Joseph and S. Collins, Modeling, calibration and correction of nonlinear illumination dependent fixed pattern noise in logarithmic CMOS image sensors, *IEEE Transactions on Instrumentation and Measurement*, 51(5), 996–1001, 2002.
7. I. L. Fujimori, W. Ching-Chun, and C. G. Sodini, A 256 × 256 CMOS differential passive pixel imager with FPN reduction techniques, *IEEE Journal of Solid-State Circuits*, 35(12), 2031–2037, 2000.
8. S. Matsuo, T. J. Bales, M. Shoda, S. Osawa, K. Kawamura, A. Andersson, M. Haque et al., 8.9-Megapixel video image sensor with 14-b column-parallel SA-ADC, *IEEE Transactions on Electron Devices*, 56(11), 2380–2389, 2009.
9. M. Mori, M. Katsuno, S. Kasuga, T. Murata, and T. Yamaguchi, 1/4-Inch 2-Mpixel MOS image sensor with 1.75 Transistors/Pixel, *IEEE Journal of Solid State Circuits*, 39(12), 2426–2430, 2004.
10. M. Kasano, Y. Inaba, M. Mori, S. Kasuga, T. Murata, and T. Yamaguchi, A 2.0-μm pixel pitch MOS image sensor with 1.5 transistor/pixel and an amorphous Si color filter, *IEEE Transactions on Electron Devices*, 53(4), 611–617, 2006.
11. P. B. Catrysse and B. A. Wandell, Integrated color pixels in 0.18-mm complementary metal oxide semiconductor technology, *Journal of the Optical Society of America A*, 20(12), 2293–2306, 2003.
12. P. Catrysse, B. Wandell, and A. El Gamal, An integrated color pixel in 0.18 μm CMOS technology, In *Electron Devices Meeting,* 559–562, 2001.
13. G. F. Dalla Betta, N. Zorzi, P. Belluttil, M. Boscardin, and G. Soncini, Triple- junction colour sensor fully compatible with CMOS Technology: Results of a test chip, *Proceedings of the IEEE 2002 International Conference on Microelectronic Test Structures*, 15, 217, 2002.
14. K. M. Findlater, D. Renshaw, J. E. D. Hurwitz, R. K. Henderson, M. D. Purcell, S. G. Smith, and T. E. R. Bailey, A CMOS image sensor with a double-junction active pixel, *IEEE Transactions on Electron Devices*, 50(1), 32–42, 2003.
15. R. B. Merrill, Color separation in an active pixel cell imaging array using a triple-well structure, US Patent 5,965,875, 1999.
16. T. H. Hsu, Y. K. Fang, D. N. Yaung, S. G. Wuu, H. C. Chien, C. S. Wang, J. S. Lin, C. H. Tseng, S. F. Chen, C. S. Lin, and C. Y. Lin, Color mixing improvement of CMOS image sensor with air-gap-guard ring in deep submicrometer CMOS technology, *IEEE Electron Device Letters*, 26(5), 301–303, May 2005.
17. T. H. Hsu, Y. K. Fang, D. N. Yaung, S. G. Wuu, H. C. Chien, C. H. Tseng, L. L. Yao, W. D. Wang, C. S. Wang, and S. F. Chen, A high-efficiency CMOS image sensor with air gap in situ microlens (AGML) fabricated by 0.18 μm CMOS technology, *IEEE Electron Device Letters*, 26(9), 634–636, September 2005.
18. K. Onozawa, K. Toshikiyo, T. Yogo, M. Ishii, K. Yamanaka, T. Matsuno, and D. Ueda, A MOS image sensor with a digital-microlens, *IEEE Transactions on Electron Devices*, 55(4), 2008.
19. K. Matou and Y. Ni, Precise FPN compensation circuit for CMOS APS, *Electronics Letters*, 38(19), 1078–1079, 2002.
20. K. Yonemoto and H. Sumi, A CMOS image sensor with a simple fixed pattern noise reduction technology and a hole accumulation diode, *IEEE Journal of Solid State Circuits*, 35(12), 2038–2043, 2000.
21. M. Sasaki, M. Mase, S. Kawahito, and Y. Tadokoro, A wide dynamic range CMOS image sensor based on multiple resolution column parallel ADC, *IEEE Journal of Solid State Circuits*, 40(12), 2787–2795, 2005.
22. S. Kleinfelder, S. Lim, X. Liu, and A. El Gamal, A 10000 Frames/s CMOS digital pixel sensor, *IEEE Journal of Solid State Circuits*, 36(12), 2049–2059, 2001.
23. A. Kitchen, A. Bermak, and A. Bouzerdoum, A digital pixel sensor array with programmable dynamic range, *IEEE Transactions on Electron Devices*, 52(12), 2591–2601, 2005.

24. A. Bermak and Y. F. Yung, A DPS array with programmable resolution and reconfigurable conversion time, *IEEE Transactions on VLSI Systems*, 14(1), 15–22, 2006.
25. C.-H. Lai, Y.-C. King, and S.-Y. Huang, A 1.2-V 0.25 μm clock output pixel architecture with wide dynamic range and self-offset cancellation, *IEEE Sensors Journal*, 6(2), 398–405, 2006.
26. E. Culurciello, R. Etienne-Cummings, and K. A. Boahen, A biomorphic digital image sensor, *IEEE Journal of Solid State Circuits*, 38(2), 2003.
27. L. G. McIlrath, A low power low noise ultrawide dynamic range CMOS imager with pixel parallel A/D conversion, *IEEE Journal of Solid State Circuits*, 36(5), 846–853, 2001.
28. Y. Chen, F. Yuan, and G. Khan, A wide dynamic range CMOS PFM digital pixel sensor with in-pixel variable voltage reference, *Analog Integrated Circuits and Signal Processing*, 61, 287–299, 2009.
29. B. Fowler, A. El Gamal, and D. X. D. Yang, A CMOS area image sensor with pixel level A/D conversion, In ISSCC Digest of Technical Papers, 1995.
30. S. Chen, A. Bermak, and F. Boussaid, A compact reconfigurable counter memory for spiking pixels, *IEEE Electron Device Letters*, 27(4), 255–257, 2006.
31. X. Wang, W. Wong, and R. Hornsey, A high dynamic range CMOS image sensor with inpixel light-to-frequency conversion, *IEEE Transactions on Electron Devices*, 53(12), 2570–2573, 2006.
32. M. Habibi and M. Sayedi, Geometric centre tracking of tagged objects using a low power demodulation smart vision sensor, *IET Circuits Devices and Systems*, 4(1), 67–77, 2010.
33. W. Miao, Q. Lin, W. Zhang, and N.-J. Wu, A programmable SIMD vision chip for real-time vision applications, *IEEE Journal of Solid State Circuits*, 43(6), 1470–1473, 2008.
34. A. Lopich and P. Dudek, A SIMD cellular processor array vision chip with asynchronous processing capabilities, *IEEE Transactions on Circuits and Systems—I: Regular Papers*, 58(10), 2420–2431, 2011.
35. P. S. Mandolesi, P. Julian, and A. G. Andreou, A scalable and programmable simplicial CNN digital pixel processor, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 51(5), 988–996, 2004.
36. S. Mertoguno and N. G. Bourbakis, A digital retina like low level vision processor, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 33(5), 782–788, 2003.
37. L. Camunas-Mesa, A. Acosta-Jimenez, C. Zamarreno-Ramos, T. Serrano-Gotarredona, and B. Linares-Barranco, A 32 × 32 pixel convolution processor chip for address event vision sensors with 155 ns event latency and 20 Meps throughput, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 58(4), 777–790, 2011.
38. W. Zhang, Q. Fu, and N.-J. Wu, A programmable vision chip based on multiple levels of parallel processors, *IEEE Journal of Solid State Circuits*, 46(9), 2132–2147, 2011.
39. K. Kagawa, K. Yasuoka, D. C. Ng, T. Furumiya, T. Tokuda, J. Ohta, and M. Nunoshita, Pulse-domain digital image processing for vision chips employing low-voltage operation in deep- submicrometer technologies, *IEEE Journal of Selected Topics in Quantum Electronics*, 10(4), 816–828, 2004.
40. R. C. Galan, F. Jimenez-Garrido, R. Dominguez-Castro, S. Espejo, T. Roska, C. Rekeczky, I. Petras, and A. Rodriguez-Vazquez, A bio-inspired two-layer mixed signal flexible programmable chip for early vision, *IEEE Transactions on Neural Networks*, 14(5), 1313–1336, 2003.
41. A. Rodriguez-Vazquez, G. Linan-Cembrano, L. Carranza, E. Roca-Moreno, R. Carmona-Galan, F. Jimenez-Garrido, R. Dominguez-Castro, and S. E. Meana, ACE16K: The third generation of mixed-signal SIMD-CNN ACE chips toward VSoCs, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 51(5), 851–863, 2004.
42. N. Massari, M. Gottardi, L. Gonzo, D. Stoppa, and A. Simoni, A CMOS image sensor with programmable pixel-level analog processing, *IEEE Transactions on Neural Networks*, 16(6), 1673–1684, 2005.
43. P. Dudek, SCAMP-3: A vision chip with SIMD current-mode analogue processor array, A. Zarandy (ed.), in: *Focal-Plane Sensor-Processor Chips*, Springer Science, New York, pp. 17–43, 2011.
44. S. Kawahito, I. A. Halin, T. Ushinaga, T. Sawada, M. Homma, and Y. Maeda, A CMOS time-of-flight range image sensor with gates-on-field-oxide structure, *IEEE Sensors Journal*, 7(12), 1578–1586, 2007.
45. M. Habibi, Analysis and optimization of the asynchronous modulated light detection pixel, *Analog Integrated Circuits*, 67(3), 339–352, 2011.

46. M. Habibi, Analysis, enhancement, and sensitivity improvement of the correlation image sensor, *IEEE Transactions on Instrumentation and Measurement*, 61(3), 708–718, 2012.
47. S. Chen, A. Bermak, and W. Yan, A CMOS image sensor with on-chip image compression based on predictive boundary adaptation and memoryless QTD algorithm, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 19(4), 538–547, 2011.
48. C. Shi, M. K. Law, and A. Bermak, A novel asynchronous pixel for an energy harvesting CMOS image sensor, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 19(1), 118–129, 2011.
49. F. Tang and A. Bermak, An 84 pW/frame per pixel current-mode CMOS image sensor with energy harvesting capability, *IEEE Sensors Journal*, 12(4), 720–726, 2012.
50. R. Shimizu, M. Arimoto, H. Nakashima, K. Misawa, T. Ohno, Y. Nose, K. Watanabe, and T. Ohyama, A charge-multiplication CMOS image sensor suitable for low-light-level imaging, *IEEE Journal of Solid State Circuits*, 44(12), 3603–3608, 2009.
51. C. Niclass, C. Favi, T. Kluter, M. Gersbach, and E. Charbon, A 128 × 128 single-photon image sensor with column-level 10-bit time-to-digital converter array, *IEEE Journal of Solid State Circuits*, 43(12), 2977–2989, 2008.

5

Principles and Technology of SQUIDs

Robert L. Fagaly
Quasar Federal Systems

Superconducting quantum interference devices (SQUIDs) have been a key factor in the development and commercialization of ultrasensitive electric and magnetic measurement systems. In many cases, SQUID instrumentation offers the ability to make measurements where no other methodology is possible. In addition to measuring magnetic fields, SQUID sensors can be configured to measure a wide variety of electromagnetic properties.

Because of their superconducting nature, SQUID sensors operate at cryogenic temperatures. Although SQUID electronics have the capability to operate well above 1 MHz, most applications tend to be at lower frequencies. Most commercial SQUID systems utilize the SQUID sensor as a current-to-voltage amplifier, often connected to a separate detection coil for detection of magnetic fields. The ability of a SQUID sensor to measure changes in magnetic fields and currents is based on four effects: superconductivity, the Meissner effect, flux quantization, and the Josephson effect.

5.1 Superconductivity

At temperatures approaching absolute zero, certain materials undergo a transition to what is known as the superconducting state. In 1911, Kamerlingh-Onnes [1] discovered that the resistance of mercury when cooled below 4.2 K dropped to an immeasurably small resistance. This transition from normal resistance to resistanceless behavior takes place over a narrow temperature range—about 0.001 K for pure, strain-free metals and a degree or more for alloys and ceramics. Below this temperature, known as the transition temperature (T_c), the material is characterized by a complete lack of electrical resistance.

Subsequent investigations indicated that a large number of materials undergo a similar superconducting transition. In 1986, Bednorz and Müller [2] discovered a new class of ceramic oxides that became

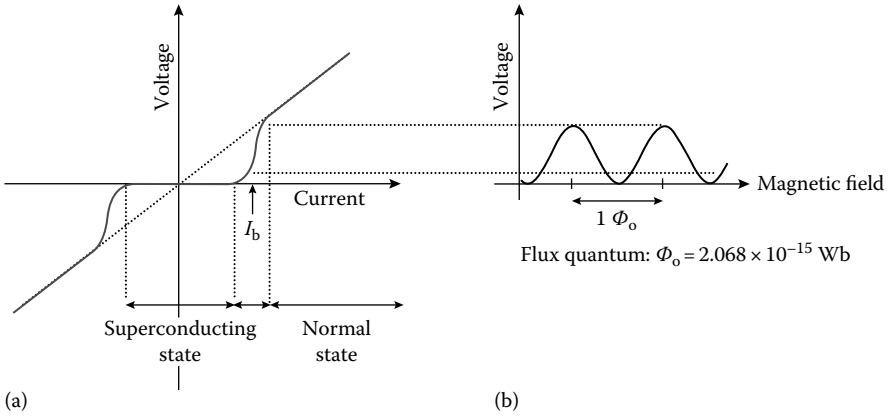


FIGURE 5.1 (a) Bias point (I_b) for the Josephson junction and (b) voltage vs. external flux at constant bias current.

superconducting near 30 K—significantly warmer than any previously known superconductor. Since then, materials have been developed with the superconducting transition temperatures above 130 K, well above the boiling point of liquid nitrogen. $YBa_2Cu_3O_{7-\delta}$, often referred to as YBCO, with $T_c > 90$ K is the most commonly used superconducting ceramic oxide. To distinguish between the types of materials used in making SQUID sensors, we denote the (primarily) metallic superconductors that typically operate at liquid helium temperatures (4.2 K) as low-temperature superconductors (LTSs) and the (primarily ceramic oxide) materials that can operate at liquid nitrogen temperatures (77 K) as high-temperature superconductors (HTSs), although some (e.g., MgB_2) have T_c's around 40 K.

Another property of superconductors—referred to as the Meissner [3] effect—is that they do not allow any magnetic field to penetrate into them (beyond a very thin surface layer known as the skin depth). To prevent magnetic fields from penetrating, a current is induced (via Lenz's law) that opposes any external magnetic field or field changes. A third property is that if a superconducting ring is constructed, one cannot change the level of magnetic field in the ring in a continuous manner. Instead, the magnetic field is quantized in multiples of flux quanta (Φ_o) whose magnitude is 2.068×10^{-15} Wb ($T \cdot m^2$). That is, for a ring with an area of 1 cm^2, the interior magnetic field can only exist in discrete steps of 2.068×10^{-11} T. In the situation where two superconducting (resistanceless) regions are separated by a normal (resistive) region, the Josephson effect [4] permits current (paired electrons) to tunnel through the barrier without any resistance (Figure 5.1a). Above a critical current, characteristic of the junction, the junction behaves like a normal resistor, following Ohm's law.

5.2 Superconducting Quantum Interference Devices (SQUID)

These effects have been used to fabricate a superconducting amplifier. This device known as a SQUID acts as a magnetic flux-to-voltage amplifier. The addition of an input coil allows current to be transformed into a magnetic flux that is inductively coupled into the SQUID loop, transforming the SQUID into a current-to-voltage amplifier.

The SQUID loop is connected to circuitry for detecting changes in the flux penetrating the loop. SQUIDs are operated as either radio-frequency (rf) or direct current (dc) SQUID. The dc SQUID differs from the rf SQUID in the manner of biasing the Josephson junction and the number of junctions. The prefix rf or dc refers to whether the Josephson junction(s) is biased with an rf alternating current (ac) or dc. In the former case, flux changes are detected by a resonant tank circuit that is inductively coupled to the SQUID loop (the rf SQUID). In the latter, the SQUID loop contains two Josephson junctions (Figure 5.2) and a dc is applied (the dc SQUID). Flux changes are detected by monitoring the voltage across the junctions.

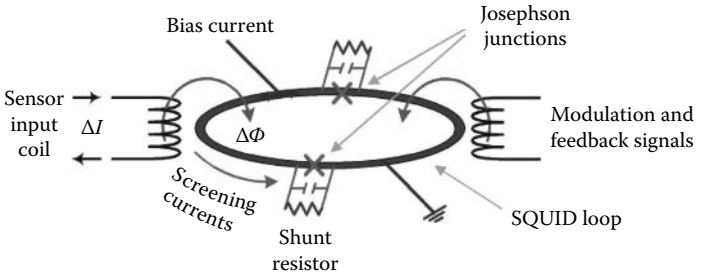


FIGURE 5.2 Dual junction (dc) SQUID loop. The capacitor represents the self-capacitance of the junction.

Since rf SQUID noise is proportional to the square root of the bias frequency (initially 19 MHz), energy sensitivities (E_N) were on the order of 10^{-28} J/√Hz. The major limiting factor in the noise of an rf SQUID is the bias frequency (f_B) used to excite the tank circuit with rf SQUID noise proportional to $1/\sqrt{f_B}$. As f_B increases, the complexity of the electronics also tends to increase.

The minimum noise energy for a dc SQUID is given by [5]

$$E_N \approx 4k_B T\sqrt{\pi L_{loop} C} \tag{5.1}$$

where

k_B is Boltzmann's constant
T is the temperature of the SQUID
L_{loop} is the inductance of the SQUID loop
C is the self-capacitance of the junction

Substituting appropriate numbers indicates that the minimum E_N for a dc SQUID is on the order of $h/2$ and devices with these sensitivities have been constructed [6]. Most commercial dc SQUIDs have $E_N \approx 5 \times 10^{-31}$ J/√Hz. Configured as an ammeter, this translates to pA/√Hz current sensitivities.

From a historical viewpoint, although the dc SQUID was the first type of SQUID magnetometer made, the early development was with rf SQUIDs. The rf SQUID had the advantage of needing only a single point-contact Josephson junction that was essentially handmade. With modern thin film fabrication techniques that allow the fabrication of SQUID devices with two Josephson junctions having identical weak link characteristics, dc SQUIDs have come to dominate the commercial landscape. Additional information on rf SQUIDs can be found in Ref. [7].

External currents are transformed into a magnetic flux and inductively coupled flux into the dc SQUID loop via an input coil. A dc equal to twice the critical current (I_c) of the junction is used to bias the SQUID at an operating point where the junctions are midway between resistanceless and resistive behavior (Figure 5.1a).

Shunt resistors are used to prevent hysteretic behaviors in the I–V curve [5]. Inductively coupling magnetic flux into the SQUID loop creates screening currents ($I_{loop} = \Phi_{loop}/L_{loop}$) that will effectively increase or decrease I_c, depending on the direction of the induced flux. A change in the magnetic flux applied through the SQUID loop induces a wave function phase change that enhances the current through one Josephson junction and reduces the current through the other. This asymmetry, which is periodic in Φ_o, is used to provide a feedback current that nulls the flux penetrating the SQUID loop. The voltage generated by the bias current (across the shunt resistors) is used to monitor the effect of any externally coupled signals. Negative feedback (inductively coupled) allows detection of changes in flux penetrating the SQUID loop to better than $10^{-6}\ \Phi_o$.

The system output voltage is the voltage drop across the feedback resistor in a negative feedback loop controlled by the SQUID electronics (Figure 5.3). The feedback signal is generated in response

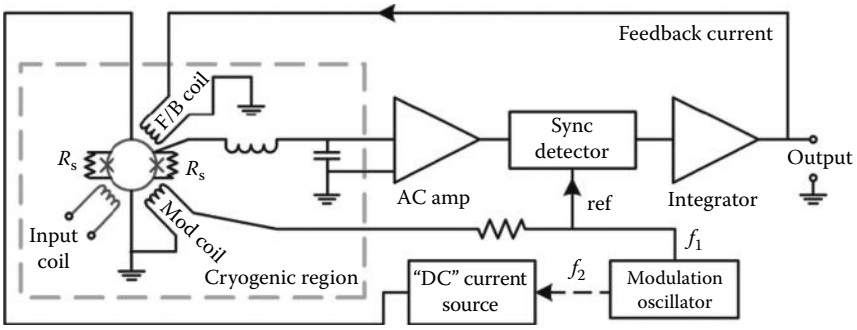


FIGURE 5.3 Schematic of dc SQUID sensor and electronics.

to changes in the output signal of the SQUID sensor. The output of the SQUID sensor is periodic in the field coupled into the SQUID loop (Figure 5.1b). Negative feedback (similar to a phase-locked loop technique) is used to maintain the system operating point at a particular (and arbitrary) flux quantum. When operated in this mode, the system is in a flux-locked loop. The transfer function (V–Φ) for rf SQUIDs differs from the dc SQUID in that—rather than sinusoidal—it has a triangular shape (similar to Figure 5.5b). References [8–10] give a more detailed explanation of the principles of SQUID sensors and the underlying concepts behind the Josephson effect along with qualitative descriptions of the operating principles of SQUID sensors and the properties of materials used to fabricate SQUID sensors.

One important factor of SQUID design is such that the feedback electronics is able to follow changes in the shielding currents. If the shielding current changes so fast that the flux in SQUID loops changes by more than one-half a flux quantum, it is possible that the feedback electronics will lag behind the rapidly changing flux. When the electronics finally "catch up," they can lock on an operating point (Figure 5.1b) different from the original. In this case, the SQUID has "lost lock" because the SQUID has exceeded the maximum slew rate of the electronics. Typical slew rates for SQUIDs are in the range of 10^5–10^6 Φ_o/s. The ability to have multiple slew rates is a convenience that allows for greater rejection of noise (due to a narrower bandwidth of the feedback electronics) for signals that are slowly varying.

Most LTS dc SQUID sensors are configured with an integrated input coil (Figure 5.3) whose input impedance is purely superconducting and can be characterized as an inductance, $L_{input} \sim 1$ μH. This allows the SQUID sensor to act as an ammeter with pA/√Hz sensitivities. Connecting a sensing coil to the SQUID sensor allows it to act as a magnetometer (cf. Section 14.2). Because of the inability to make superconducting HTS joints, HTS SQUID sensors are configured as pure magnetometers using just the Josephson loop as the magnetic sensing area or an inductively coupled input coil that acts as a magnetometer.

5.2.1 Sensitivity

SQUID noise is often presented as the spectral density of the equivalent flux noise $S_\Phi(f)$ as a function of frequency or noise energy per unit bandwidth $E_N(f) = S_\Phi(f)/2L_{input}$ where Φ_N is the flux noise of the Josephson device. To allow devices with differing input inductances to be directly compared, the sensitivity of SQUID devices is best discussed in terms of the energy sensitivity:

$$E_N = \frac{1}{2} L_{input} I_N^2 = \frac{\Phi_N^2}{2L_{input}} \tag{5.2}$$

where I_N is the current noise of the sensor. Because HTS SQUIDs are either pure magnetometers or planar gradiometers, their sensitivities are noted in either field (T/√Hz) or field gradient (T/m · √Hz) values.

There also exists a low-frequency contribution (which typically has a 1/*f* energy spectrum) to the frequency-independent (white) system noise. Note that this low-frequency noise is roughly the same for all types of SQUID sensors (rf vs. dc, thin film vs. point contact, etc.) and is thought to be due to a fundamental noise contribution that is inherent in SQUID sensors.

A substantial contribution to this noise in some dc SQUIDs can arise from the presence of the dc bias. By chopping the dc bias (f_2 in Figure 5.3) in combination with the conventional flux modulation techniques, it is possible to reduce this added 1/*f* noise. This double-modulation approach [11] separates the original signal waveform from the noise associated with the dc bias and is responsible for decreased 1/*f* noise at very low frequencies. While not normally used with LTS SQUIDs (which tend to have 1/*f* knees below 1 Hz), ac biasing is needed for HTS SQUIDs where the 1/*f* knee is well above 100 Hz.

In addition, the relative amount of 1/*f* noise can be dependent on the ambient magnetic field when the SQUID sensor is cooled. When cooled in the Earth's magnetic field (~50 μT), the point at which the 1/*f* noise equals the white (frequency-independent) noise is typically 1–3 Hz. Cooling the SQUID sensor in low ambient magnetic fields (less than 1 μT) may improve the 1/*f* performance by as much as an order of magnitude.

Normally, SQUID electronics utilize an internal feedback technique (Figure 5.3) to cancel changes in magnetic flux. An alternative approach to feedback may be made by coupling the feedback signal to the input coil (Figure 5.4) rather than to the Josephson loop. This can avoid large feedback currents (e.g., due to 50 or 60 Hz) that could exceed I_c for the Josephson junction(s). A second method uses an externally generated signal in addition to the SQUID feedback electronics. One example is a circuit to compensate for an external ac field being applied near the detection coil(s) in a susceptibility experiment.

Several factors are responsible for the excellent performance of SQUID systems including the following:

1. The SQUID operates at a very low temperature minimizing Nyquist noise.
2. In most applications, the detection circuit is superconducting and thus contributes no additional noise.
3. The periodic nature of SQUID devices, due to flux quantization, is responsible for their superb linearity.
4. Because a superconducting circuit has zero resistance and, consequently, the current induced in the sense coil does not decay in time, the frequency response of superconducting circuits can extend down to dc.
5. SQUID sensors offer flat frequency and phase response. This can allow for seamless integration of data from multiple sensors.

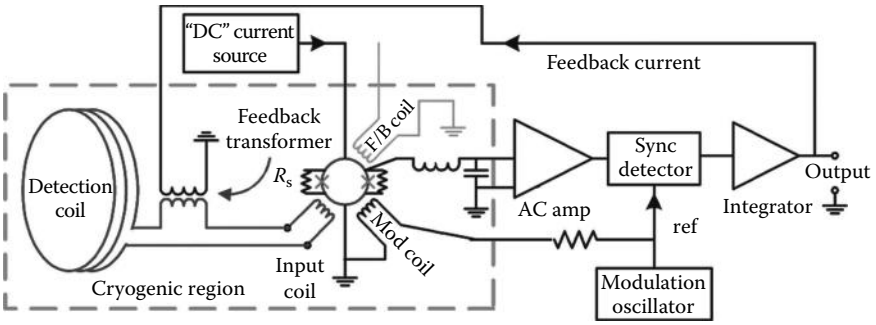


FIGURE 5.4 External feedback circuit. Note that the internal feedback circuitry shown in (Figure 5.3) is not used.

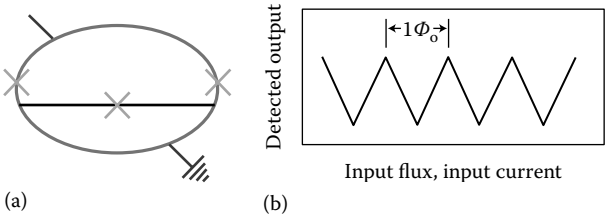


FIGURE 5.5 (a) Schematic of bi-SQUID (shunt resistors not shown) and (b) transfer function (V–Φ).

5.2.2 Bi-SQUID

The standard dc SQUID (Figure 5.2) can be modified by adding a nonlinear inductance, in the form of a third Josephson (Figure 5.5a) junction [12]. This third junction, when combined with the main inductance in the loop, acts as a single-junction SQUID yielding a device (known as a bi-SQUID) that has a triangular V–Φ transfer function (Figure 5.5b) with significantly improved linearity.

5.2.3 Superconducting Quantum Interference Filters

An array of different superconducting loop sizes, in-series or in-parallel configuration or in a 2-D combination of both, acting as grating structures, can be used to form a superconducting quantum interference filter (SQIF) [13]. This type of interferometer is based on the phase-dependent superposition of currents flowing through a nonperiodic multiloop network (series, parallel, or series–parallel as in Figure 5.6a) of Josephson junctions, where the loop areas (A_{ij}) are designed to be nonidentical. The effect of such an arrangement is that the contributions of the loops to the output signal mutually cancel each other for any finite value of the ambient magnetic field. For zero magnetic fields a mutual enhancement occurs by means of the coherent superposition yielding a unique dip at zero field. In contrast to conventional SQUIDs, the characteristic flux dependence of the voltage output (V–Φ) of a SQIF is nonperiodic (Figure 5.6b).

The unique dip in the V–Φ curve allows for the absolute field magnetometry [14] and high-precision rf applications like amplifiers and mixers [15]. The sensitivity [16] and dynamic range of an SQIF scale proportionally to the square root of the number of loops ($N \times M$) in the array. With even a moderate number of loops, this can allow absolute field measurements to be made at sensitivities approaching a few fT/√Hz. Another advantage of SQIFs is that the Josephson junction parameters need not be well matched, making fabrication of SQIF devices from HTS materials significantly easier. SQIFs also have the ability to operate at frequencies above 20 GHz.

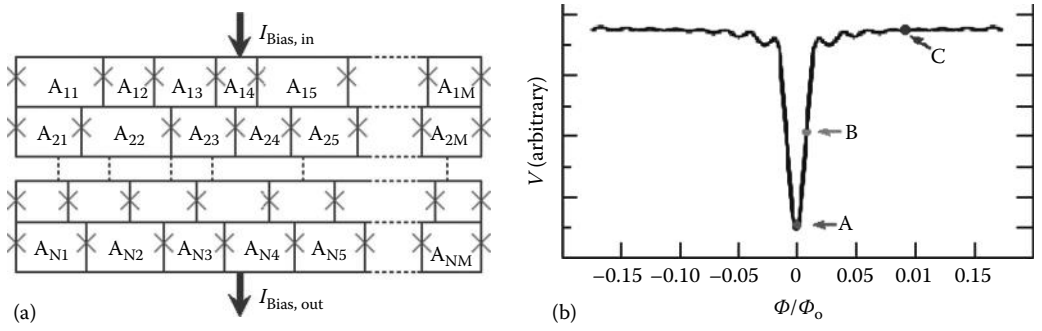


FIGURE 5.6 (a) Schematic of a 2-D (N series by M parallel elements) SQIF array (A_{ij} is the area of the ijth Josephson loop) and (b) SQIF transfer function [(A) mixing mode (V–Φ very nonlinear), (B) detection mode (V–Φ is "linear"), (C) not useful ($dV/d\Phi$ very small)].

5.3 Cryogenic Requirements

The superconducting nature of SQUIDs requires them to operate well below their superconducting transition temperature (9.3 K for niobium and 93 K for $YBa_2Cu_3O_{7-\delta}$). Ideally, the cryogenic environment should provide stable cooling, have no time-varying magnetic signature, be reasonably compact and reliable, and, if mechanical in nature, introduce neither mechanical vibration nor a magnetic signature into the detection system. The thermal environment for the SQUID sensor and detection coil has typically been liquid helium or liquid nitrogen contained in a vacuum-insulated vessel known as a dewar.

As an alternative to the use of liquid cryogens, closed cycle refrigeration [17] can be desirable for several reasons. These include reduction of operating costs, use in remote locations, operations in nonvertical orientations, avoiding interruptions in cryogen deliveries, safety, and the convenience of not having to transfer every few days. Parameters governing suitability include physical size, absence of periodic replacement of cryogenic fluid, and most importantly vibration and magnetic signature. There are two main obstacles to using closed cycle refrigeration with SQUIDs. The first is the mechanical movement that (ultimately) causes the detection coils to move in the Earth's magnetic field. The second is the magnetic signature due to the moving parts of the cryocooler's cold head and compressor. Moving the compressor far away will reduce its relative signature.

If the experiment involves measurements interior to the dewar (*cf.* Figure 5.10), then a metallic dewar is preferable. Metallic dewars offer significant shielding from environmental noise at frequencies above 10–100 Hz. If the system is to measure magnetic fields exterior to the dewar, the dewar must be magnetically transparent and metallic construction is not appropriate. Dewars for external field measurements (e.g., geophysical or biomagnetic) are normally constructed of nonmetallic, low-susceptibility materials to minimize their magnetic interactions with the SQUID sensors and detection coils. Materials used are typically glass-fiber–epoxy composites such as G-10.

5.3.1 Limitations on SQUID Technology

SQUIDs offer the ability to measure at sensitivities unachievable by other magnetic sensing methodologies. However, their sensitivity requires proper attention to cryogenics and environmental noise. This applies not just to laboratory applications but also to every potential use of SQUID sensors. When utilizing SQUID-based measurement systems and data reduction algorithms, it is important to bear in mind several fundamental limitations:

5.3.1.1 Differential Measurements

SQUIDs are sensitive to relative (field or current) changes only. This is a consequence of the fact that the output voltage of a SQUID is a periodic function (Figure 5.1b or 5.5b) of the flux penetrating the SQUID loop. The SQUID is "flux-locked" on an arbitrary point on the V–Φ curve, and the SQUID output is sensitive to flux changes relative to this lock point. If absolute field measurements are required, then a SQIF device should be used.

5.3.1.2 Slew Rate Limitations

If the signal changes faster than the feedback electronics can follow (i.e., the slew rate is exceeded) and the total signal change exceeds 1/2 Φ_o, it is possible for the operating point to shift by one or more flux quanta (Figure 5.1b). If high bandwidths are needed, it is possible to operate the electronics in a limited range mode where the raw output is amplified without use of a feedback signal. Although the SQUID has an intrinsic bandwidth of several GHz, when operated with standard flux-locked loop electronics using ac flux modulation, the maximum usable bandwidth of most commercially available electronics is typically 50–100 kHz.

In reality, the electronics may need to track multiple frequencies, placing further limitations on achievable slew rates. Actual slew rate values tend to be much less than the theoretical (or advertised) values (which are usually based on a single sine wave) due to the multispectral nature of most signal and noise sources.

5.3.1.3 1/*f* Noise

Another limitation is the presence of 1/*f* noise. The use of ac biasing [11] in HTS SQUIDs limits their maximum bandwidth to less than half the bias reversal frequency. If the bias reversal frequency is too high, noise can be induced due to voltage spikes in the transformer-coupled preamplifier input circuit. Because of this, the maximum bandwidth of commercially available HTS SQUIDs is usually limited to ~50 kHz. If MHz bandwidths are required, the ac bias is not used; however, there will be excess (i.e., 1/*f*) noise below 1 kHz.

5.3.1.4 HTS Limitations

A significant difference between LTS and HTS materials is that LTS materials (e.g., NbTi) are ductile and—in wire form—can be made into complex 3-D structures such as axial gradiometers with extremely small bend radii. Additionally, using NbTi (or Nb_3Sn) allows detection coils to be in high-field regions (>9 T), while the actual LTS SQUID sensor can be placed in a low-field environment. Because of the inability to make a truly superconducting flexible 3-D structure, axial HTS gradiometers are not possible, although thin-film planar gradiometers are. Even if it was possible to make a separate HTS coil, the inability to make HTS superconducting joints (or joints with contact resistances at the sub-pΩ level) due to the shorter coherence length of HTS materials prevents true dc response in discrete element HTS circuits [18] (i.e., the in-series resistance acts as a low-frequency high-pass filter). It also prevents the ability to have detection coils remote from the SQUID sensor. The only way to couple flux between two HTS circuits (and maintain their superconducting nature) is to inductively couple them. Reference [19] gives an excellent overview of electrical contact issues in superconductors.

5.3.1.5 Vector Nature of SQUID Magnetometers

SQUID magnetometers are vector magnetometers. For a pure magnetometer operating in the Earth's magnetic field (~50 μT), a 180° rotation can sweep out a total field change of up to 100 μT. If the magnetometer has a sensitivity of 10 fT/√Hz, tracking the total field change requires a dynamic range of 100 μT/10 fT = 200 dB, well beyond the capabilities of current electronics. In addition, the rotational speed must not cause the current flowing through the SQUID sensor to exceed its slew rate limitations. Pure gradiometers would be insensitive to uniform fields and not suffer this dynamic range limitation. In reality, gradiometers are not perfect and have some magnetometer component that can place motion restrictions even on well-balanced gradiometers. If it is necessary to measure absolute fields, then a SQIF device should be used.

5.4 Applications

A simple SQUID system (Figure 5.7) consists of room temperature electronics and a low-temperature probe containing the SQUID sensor, configured with an integrated input coil to measure changes in current. Connecting the input coil to a detection coil allows measurement of magnetic fields. The low-temperature environment is necessary due to the requirement that the SQUID sensor be below its critical temperature.

Whether an rf or dc SQUID, we can consider a SQUID system as a black box that acts like a current-to-voltage amplifier with extremely high gain. In addition, it offers extremely low noise,

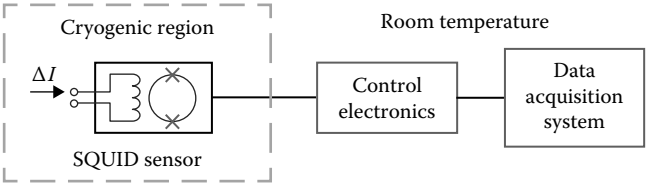


FIGURE 5.7 Schematic diagram of typical SQUID input circuit configured as an ammeter.

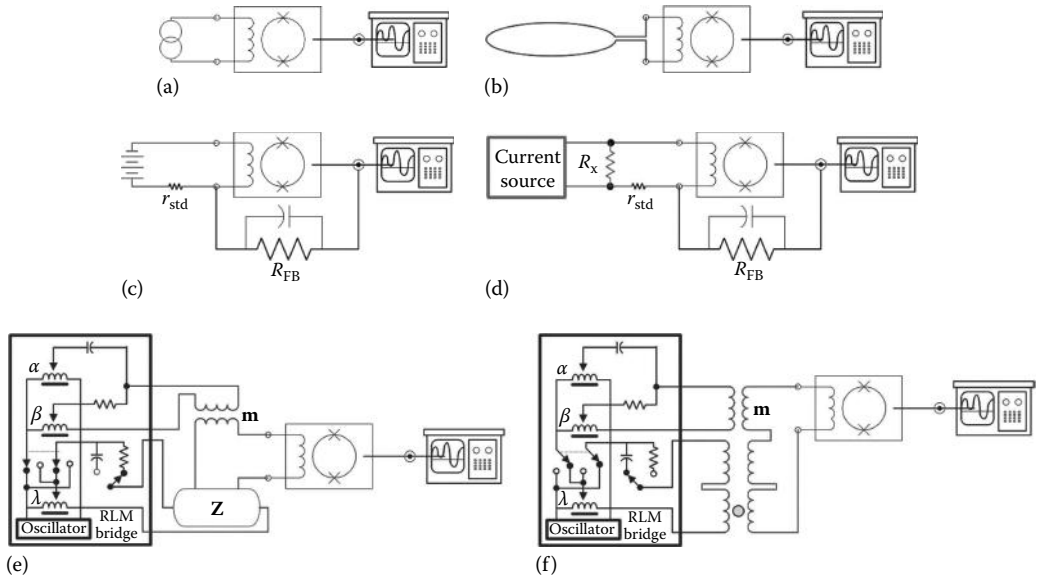


FIGURE 5.8 (a) AC and dc, (b) magnetic field, (c) dc voltage, (d) dc resistance, (e) ac resistance/inductance bridge, and (f) ac mutual inductance (susceptibility bridge).

TABLE 5.1 Typical Sensitivities of SQUID Instruments

Measurement	Sensitivity
Current (Figure 5.8a)	10^{-12} A/√Hz
Magnetic fields (Figure 5.8b)	10^{-15} T/√Hz
DC voltage (Figure 5.8c)	10^{-14} V
DC resistance (Figure 5.8d)	10^{-12} Ω
Mutual/self-inductance (Figure 5.8e)	10^{-12} H
Magnetic moment (Figure 5.8f)	10^{-10} emu

high dynamic range, excellent linearity, and a wide bandwidth that can extend down to dc. A typical SQUID system can provide a transfer function in excess of 10^7 V/A with an input sensitivity better than 1 pA/Hz. Linearities can exceed 1 part in 10^7 and bandwidths typically extend from dc to hundreds of kHz. Specialized SQUID systems have been made to operate with bandwidths extending to 10 MHz.

SQUIDs can also be configured to measure a wide variety of electromagnetic properties (Figure 5.8). As mentioned earlier, there is no method for making true superconducting connections to SQUIDs with HTS wire. As a result, commercially available HTS devices are currently in the form of magnetic sensing rather than current sensing devices.

Table 5.1 shows typical capabilities of SQUID-based instruments. Additional information on laboratory applications of SQUID systems can be found in Refs. [7,8,20,21].

5.4.1 Current Measurements

One common use of a SQUID is as an ammeter (Figure 5.8a). The input can be connected to an experiment at liquid helium temperatures or to room temperature. If the signal is to be inductively coupled to a detection coil that is connected to the SQUID input, then the circuit must be superconducting if dc response is desired.

5.4.1.1 Cryogenic Current Comparator

A current comparator can be used to compare currents and/or to determine current ratios with highest accuracy. Two currents (I_1 and I_2) are fed through separate wires, which are led through a superconducting tube. The Meissner effect causes a screening current that is induced on the inner surface of the tube, flowing opposite to and being exactly as large as the sum of the currents inside the tube. This shielding current exactly cancels the magnetic field inside the tube produced by the currents in the wires. The screening current flows back across the outer surface of the tube, giving rise to a magnetic field in the room outside of the tube, which is detected by a SQUID magnetometer used as a null meter. When the external field is at a null, $I_1 = I_2$. Commercially available [22], sensitivity is at the pA level with an uncertainty of a few parts per million (ppm) [23]. Laboratory systems have achieved sensitivities as low as 0.08 fA/√Hz.

If the measurement is of a current that passes through the detection coil, a toroidal geometry for the detection coil has the advantage of extremely good coupling to the source while rejecting contributions due to external sources. Such configuration has been designed to measure antiproton beam currents [24] and monitor cyclotron beam currents [25]. Another configuration [26] has a room temperature conductor passing through a reentrant dewar with a current sensitivity of better than 100 μA at a current of 250 A.

So far, the majority of applications use superconducting circuits. There are, however, a number of applications where resistive circuits are used. One example [27] passes the current to be (inductively) measured through a room temperature toroid, which was directly coupled into an rf SQUID input coil. Sensitivities of 40 nA/√Hz with a usable frequency response of 3 Hz to above 1 kHz have been achieved. The use of a dc SQUID could improve sensitivity by an order of magnitude.

5.4.2 Voltage Measurements

Voltage (and resistance) measurements are examples where nonsuperconducting elements are used in the measurement circuitry. Figure 5.8c shows the schematic of a SQUID picovolt measuring system. When a voltage (V_{input}) is applied across the input terminals, a current is generated in the SQUID input coil. In this situation, the feedback current (I_{FB}) that would normally be applied to the SQUID loop via the feedback coil is fed back via R_{FB} through r_{std} until the voltage drop across r_{std} is equal to V_{input} and there is no net current through the SQUID. The output voltage (V_{out}) measures the voltage drop across R_{FB} and r_{std} with $V_{input} = V_{out} \times r_{std}/(R_{FB} + r_{std})$. The voltage gain of the system is determined by the ratio of R_{FB}/r_{std}. Typical values for R_{FB} and r_{std} are 3 and 30 μΩ, respectively, giving a voltage gain of 10^8. The standard resistor r_{std} is (like the SQUID) at a temperature of 4.2 K. The voltage source however may be at a completely different temperature. For source resistances less than 10^{-4} Ω and temperatures less than 4.2 K, voltages of 10^{-13} V can be resolved.

Typical applications of picovoltmeters include noise thermometry, measurements of thermopower, and infrared bolometer detectors.

5.4.3 Resistance Measurements

By using a precision, ultralow-noise current source (Figure 5.8d), a SQUID picovoltmeter can be turned into a pico-ohmmeter. Replacing V_{input} with a resistor (R_X) and a highly stable constant current source (I_{CCS}) [28] creates a voltage drop across the resistor. The resistance is determined by $R_X = V_{output} \times r_{std}/[(R_{FB} + r_{std})\, I_{CCS}]$. Resolutions of 10^{-11} Ω can be achieved for resistances as high as 10^{-2} Ω.

5.4.4 AC Impedance Measurements

The SQUID can also be used as the null detector in an ac bridge circuit (Figure 5.8e) to measure both resistive and reactive components of complex impedance. The unknown impedance **Z** is excited by a current generated by an oscillator voltage, which is attenuated by a precision ratio transformer (λ).

The difference between the voltage developed across the unknown impedance **Z** and that developed in the secondary of a nulling mutual inductor **m** is applied to the input of the SQUID circuit. The primary current in **m** is proportional to the oscillator voltage and defined by the setting of the ratio transformer (α). An additional reactive current is supplied by a second ratio transformer (β) that causes the primary current to be passed through a capacitor rather than a resistor, thus generating a 90° phase shift in the voltage applied to **m**. The amplified off-balance signal, which appears at the output of the SQUID control electronics, can be displayed by means of a lock-in amplifier tuned to the oscillator frequency.

The sensitivity is limited inherently by Johnson noise in the resistive components of the unknown (including the potential connections) and by the device noise of the SQUID sensor. Assuming a SQUID noise of 8 pA/√Hz (equivalent to a noise temperature of 5 μK), the system is capable of measuring resistances between 0.5 and 10^{-10} Ω and inductances (both self and mutual) between 10^{-12} and 10^{-3} H. Its resolution depends on the value of **Z** but can be as high as 10 ppm. Using a current comparator as the bridge excitation **V**, a four-terminal resistance bridge with a resolution of 0.1 ppm can be constructed [29].

5.4.5 Magnetic Susceptibility

The SQUID ac measurement system can be modified to measure magnetic susceptibility by the addition of an excitation coil around the sample.

Figure 5.9 shows a typical experimental setup for measurement (Figure 5.8f) of ac susceptibility. Such a configuration can be used for thermometry at mK and even μK temperatures [21] by measuring the low (mT)-field magnetic susceptibility of paramagnetic salts such as $Ce_2Mg_3(NO_3)_{12} \cdot 24\ H_2O$—usually referred to as CMN [7].

5.4.6 Variable Temperature SQUID Susceptometers

Instead of using a secondary ac excitation coil (Figures 5.8f and 5.9b), a dc field can be used to magnetize samples. Typically the field is fixed and the sample moved into the detection coil's region of sensitivity (Figure 5.9c). The change in detected magnetization is directly proportional to the magnetic moment of the sample.

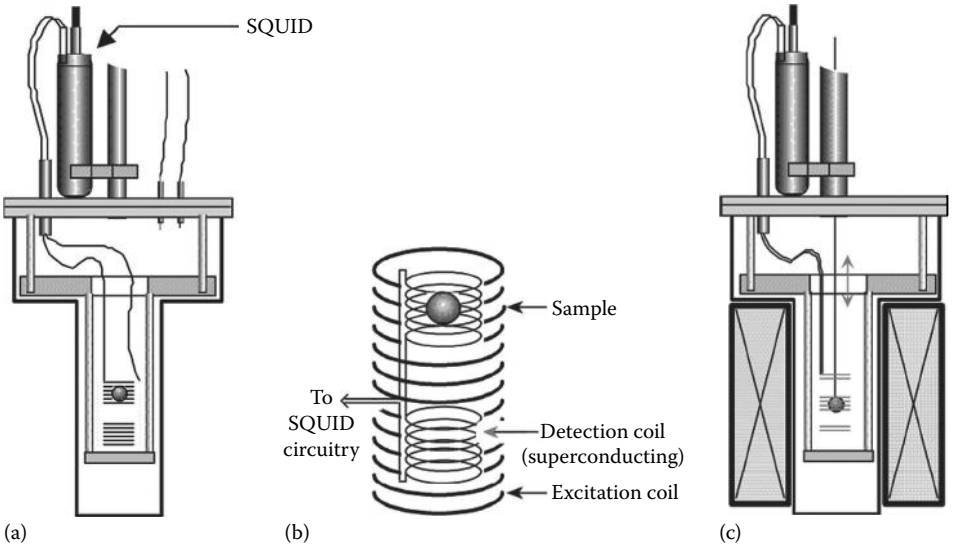


FIGURE 5.9 Magnetic susceptibility measurement apparatus (liquid helium dewar not shown): (a) AC susceptibility, (b) signal and excitation coil details, and (c) second-derivative oscillating magnetometer for dc measurements with external dc field coils.

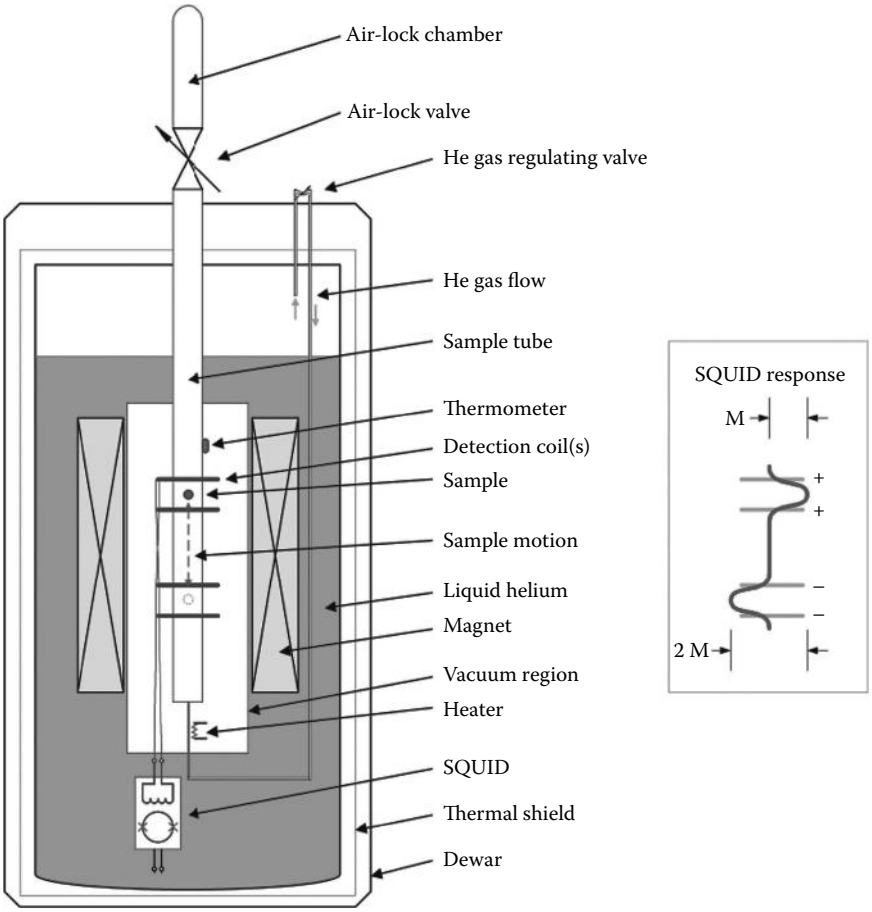


FIGURE 5.10 Variable temperature susceptometer (various electrical leads omitted for clarity). The trace on the right shows the response of the detection coil(s) as a function of sample position height.

The use of a variable temperature insert can allow measurements to be made over a wide range (1.8–400 K). Typically, the sample temperature is controlled by helium gas flowing slowly past the sample (Figure 5.10). The temperature of the gas is regulated using a heater located below the sample-measuring region and a thermometer located above the sample region. This arrangement ensures that the entire region has reached thermal equilibrium prior to data acquisition. The helium gas is obtained from normal evaporation in the dewar and its flow rate is controlled by a precision regulating valve. The use of an oven probe can raise the sample temperature to 800+ K (at the expense of smaller sample measurement volume).

Commonly referred to as SQUID magnetometers, these systems are properly called SQUID susceptometers. They have a homogeneous superconducting magnet (as large as 9 T) to create a uniform field over the entire sample-measuring region and the superconducting pickup loops. The magnet induces a moment allowing a measurement of magnetic susceptibility. The superconducting detection loop array is rigidly mounted in the center of the magnet. This array is configured as a gradient coil to reject external noise sources. The detection coil geometry determines what mathematical algorithm is used to calculate the net magnetization. Oppositely paired Helmholtz coils [30] (Figure 5.10), transverse detection coils, and first- and second-derivative gradiometers have all been successfully used. The system shown in Figure 5.10 is capable of 10^{-8} emu sensitivities.

The most popular detection coil configuration is that of a second-order coil configuration (Figure 5.9c) [31]. In this situation, the detected output is compared with that of an ideal response to determine the magnetization. This technique has the advantage of eliminating noise. Additionally, the intentional choice not to use a NbTi shield allowed swept field scans to be performed. Coupling two axial channels of differing gradient orders (e.g., first and second) can significantly improve noise rejection. The detection coil does not necessarily need to be in an axial configuration. Planar gradiometers have been used to detect the onset of superconductivity in diamond pressure cells [32]. Placement of secondary excitation coils can allow ac susceptibility measurements approaching 10^{-8} emu to be made in the presence of a significant dc bias field. Rock magnetometers for paleomagnetism use a configuration similar to variable temperature susceptometers but without the superconducting magnet.

5.4.7 Gravimetry

Superconducting accelerometers have been in gravity wave detectors [33] to detect dimensional changes smaller than 10^{-19} m on test masses exceeding 10^4 kg. SQUIDs were also used as ultrahigh-resolution (nanoradian) angular position detectors in NASA's Gravity Probe B program [34], which tested several predictions of Einstein's general theory of relativity. Airborne superconducting gravity gradiometers with sensitivities at the 10^{-12} m/s^2/m level have been used for oil and mineral deposit detection. Although SQUIDs are not normally used in superconducting gravimeters [35], they have the potential to increase sensitivities beyond the current 10^{-12} m/s^2 sensitivity (0.01% linearity) of commercial systems.

5.4.8 Other SQUID Measurement Techniques

In addition to the detection of magnetic fields, SQUID detection coils can also be used for detection of nuclear magnetic resonance (NMR). NMR signals [36] can be measured by placing a sample (e.g., protons or ^{19}F) in the center of SQUID detection coils and either sweeping the external field or applying an rf excitation to the sample. The same concept can be used to measure electron paramagnetic resonance (EPR) signals. Although limited to 100 ppm field uniformities, SQUID susceptometers (e.g., SHE VTS [Figure 5.10], Quantum Design MPMS [Figure 5.9c], or Conductus χMAG) are excellent platforms for basic demonstrations of NMR measurements. SQUIDs can also be used as detectors for low-field MRI.

In addition, there are many exotic uses for SQUID sensors. Because SQUID magnetometers are vector devices, they can detect rotational movement of a magnetometer coil in the Earth's magnetic field (~50 μT) as small as 10^{-3} arc-s. SQUIDs have been used for more esoteric applications including temperature measurements [37] with resolution near 10^{-12} K. SQUIDs have been used in searches for dark matter such as weakly interacting massive particles (WIMPs) [38] and axions [39] along with attempts at detecting magnetic monopoles [40] and free quarks [41]. Other exotic applications are proposed experiments to use SQUID systems for the detection of neutrinos [42].

While this chapter is focused on low-frequency applications, there is a need for higher frequency SQUID instruments. In particular, applications such as radio astronomy and NMR are examples where SQUIDs and SQIFs offer superior sensitivity at frequencies that can exceed 20 GHz [43].

References

1. H. Kamerlingh-Onnes, *Akad. van Wetenschappen,* Amsterdam, the Netherlands, **14**: **113**, 818 (1911).
2. J.G. Bednorz and K.A. Müller, Possible high T_c superconductivity in the Ba-La-Cu-O system, *Z. Phys.*, **B64**, 189–193 (1986).
3. W. Meissner and R. Oschsenfeld, Ein Neuer Effekt bei Eintritt der Supraleitfahigkeit, *Naturwissenschaften*, **21**, 787–788 (1933).
4. B.D. Josephson, Possible new effect in superconductive tunneling, *Phys. Lett.*, **1**, 251–253 (1962).

5. C.D. Tesche and J. Clarke, DC SQUID: Noise and optimization, *J. Low Temp. Phys*., **29**, 301–331 (1982).
6. R.T. Wakai and D.J. Van Harlingen, Signal and white noise properties of edge junction dc SQUIDs, *Appl. Phys. Lett*., **52**, 1182–1184 (1988).
7. R.P. Giffard, R.A. Webb and J.C. Wheatley, Principles and methods of low-frequency electric and magnetic measurements using a rf-biased point-contact superconducting device, *J. Low Temp. Phys*., **6**, 533–610 (1972).
8. R.L. Fagaly, Superconducting quantum interference device instruments and applications, *Rev. Sci. Instrum*., **77**, 101101 (2006).
9. J. Clarke and A.I. Braginski, *The SQUID Handbook: Volume I: Fundamentals and Technology of SQUIDs and SQUID Systems*, New York: Wiley, (2004).
10. T. Van Duzer and C.W. Turner, *Principles of Superconductive Devices and Circuits*, New York: Elsevier, (1981).
11. M.B. Simmonds and R.P. Giffard, Apparatus for reducing low frequency noise in DC biased SQUIDs, United States Patent No. 4,389,612, (1983).
12. I.I. Soloviev, V.K. Kornev, N.V. Klenov, and O.A. Mukhanov, Superconducting josephson structures with high linearity of transformation of magnetic signal into voltage, *Phys. Solid State*, **52**, 2252–2258 (2010).
13. J. Oppenländer, Ch. Häussler, T. Träuble, and N. Schopohl, Highly sensitive magnetometers for absolute magnetic field measurements based on quantum interference filters, *Physica C*, **368**, 119–124 (2002).
14. P. Caputo, J. Tomes, J. Oppenländer, Ch. Häussler, A. Friesch, T. Träuble, and N. Schopohl, Superconducting quantum interference filters as absolute magnetic field sensors, *IEEE Trans. Appl. Supercond.,* **15**, 1044–1047 (2005).
15. P. Caputo, J. Tomes, J. Oppenländer, Ch. Häussler, A. Friesch, T. Träuble, and N. Schopohl, Quadratic mixing of radio frequency signals using superconducting quantum interference filter, *Appl. Phys. Lett*., **89**, 062507 (2006).
16. N. Schopohl, Superconducting quantum interference filters operated in commercial miniature cryocoolers, *IEEE Trans. Appl. Supercond*., **15**, 936–939 (2005).
17. G. Walker, *Miniature Refrigerators for Cryogenic Sensors and Cold Electronics*, Oxford, U.K.: Clarendon Press, (1989).
18. R.B. Stephens and R.L. Fagaly, High temperature superconductors for SQUID detection coils, *Cryogenics,* **31**, 988–992 (1991).
19. J. Talvacchio, Electrical contact to superconductors, *IEEE Trans. Comp. Hybrids, Manuf. Technol*., **12**, 21–31 (1989).
20. J. Clarke and A.I. Braginski, *The SQUID Handbook: Volume II: Applications of SQUIDs and SQUID Systems*, New York: Wiley, (2004).
21. O. Lounasmaa, *Experimental Principles and Methods below 1K*, London, U.K.: Academic Press, (1974).
22. Cryogenics Ltd., London, England, http://www.cryogenic.co.uk/products/cryogenic-current-comparator-ccc, last accessed August 5, 2013.
23. B. Steck, A. Gonzalez-Cano, N. Feltin et al., Characterization and metrological investigation of an R-pump with driving frequencies up to 100 MHz, *Metrologia*, **45**, 482–491 (2008).
24. M. Kuchnir, J.D. McCarthy, and P.A. Rapidis, SQUID based beam current monitor, *IEEE Trans. Magn*., **MAG-21**, 997–999 (1985).
25. T. Watanabe, S. Watanabe, T. Ikeda, M. Kase, Y. Sasaki, T. Kawaguchi, and T. Katayama, Prototype of highly sensitive cryogenic current comparator with HTS SQUID and HTS magnetic shield, *Supercond. Sci. Technol*., **17**, S450–S455 (2004).
26. B.R. Bernard, High-resolution SQUID monitor system for large direct currents, *J. Phys. E: Sci. Instrum*., **17**, 1031 (1984).
27. J.P. Wikswo, Improved instrumentation for measuring the magnetic fields of cellular action currents, *Rev. Sci. Instrum*., **53**, 1846–1850 (1982).

28. B. Levy and A.J. Greenfield, Constant-current supply of 3 ppm stability and resettability; application for a SQUID, *Rev. Sci. Instrum.*, **50**, 655–658 (1979).
29. D.L. Edmunds, W.P. Pratt, and J.A. Rowlands, 0.1 ppm four-terminal resistance bridge for use with a dilution refrigerator, *Rev. Sci. Instrum.*, **51**, 1516 (1980).
30. J.S. Philo and W.M. Fairbank, High-sensitivity magnetic susceptometer employing superconducting technology, *Rev. Sci. Instrum.,* **48**, 1529–1536 (1977).
31. Products—Magnetic Property Measurement Systems—*Model MPMS*, San Diego, CA: Quantum Design, http://www.qdusa.com/index.html, last accessed August 5, 2013.
32. K. Amaya, K. Shimizu, M.I. Eremets, T.C. Kobayashi, and S. Endo, Observation of pressure-induced superconductivity in the megabar region, *J. Phys.: Condens. Matter.*, **10**, 11179–11190 (1998).
33. N. Solomonson, W.O. Hamilton, W. Johnson, and B. Xu, Construction and performance of a low noise inductive transducer for the Louisiana State University gravitational wave detector, *Rev. Sci. Instrum.*, **65**, 174–181 (1977).
34. J.M. Lockhart, B. Muhlfelder, G.M. Gutt, M. Luo, R.C. Clappier, T.R. McGinnis, and G.R. Smith, Optimization of a SQUID system for space, *IEEE Trans. Appl. Supercon.*, **7**, 2354–2357 (1997).
35. J.M. Goodkind, The superconducting gravimeter, *Rev. Sci. Instrum.*, **70**, 4131–4152 (1999).
36. R.A. Webb, New technique of improved low-temperature SQUID NMR measurements, *Rev. Sci. Instrum.*, **48**, 1585–1594 (1977).
37. B.J. Klemme, M.J. Adriaans, P.K. Day, D.A. Sergatskov, T.L. Aselage, and R.V. Duncan, PdMn and PdFe: New materials for temperature measurement near 2 K, *J. Low Temp. Phys.*, **116**, 133–146 (1999).
38. D.S. Akerib et al. (CDMS Collaboration), First results from the cryogenic dark matter search in the Soudan underground laboratory, *Phys. Rev. Lett.*, **93**, 211301 (2004).
39. R. Bradley, J. Clarke, D. Kinion, L.J. Rosenberg, K. van Bibber, S. Matsuki, M. Mück, and P. Sikivie, Microwave cavity searches for dark-matter axions, *Rev. Mod. Phys.*, **75**, 777–817 (2003).
40. B. Cabrera, First results from a superconductive device for moving magnetic monopoles, *Phys. Rev. Lett.*, **48**, 1378–1381 (1982).
41. L.W. Jones, A review of quark search experiments, *Rev. Mod. Phys.,* **49**, 717–752 (1977).
42. A. Drükier and L. Stodolsky, Principles and applications of a neutral-current detector for neutrino physics and astronomy, *Physical Review D*, **30**, 2295–2309 (1984).
43. M. Mück and J. Clarke, Flux-bias stabilization scheme for a radio astronomy amplified based on a superconducting quantum interference device, *Rev. Sci. Instrum.*, **72**, 3691–3693 (2001).

Further Information

Commercial vendors of SQUID sensors and measurement systems

1. Quantum Design, San Diego, CA, http://www.qdusa.com/, accessed on August 2, 2012.
2. Star Cryoelectronics, Santa Fe, NM, http://www.starcryo.com/, accessed on August 9, 2012.
3. Supracon, Jena, Germany, http://www.supracon.com/
4. Applied Physics Systems, Mountain View, CA, http://www.appliedphysics.com/, accessed on August 11, 2012.
5. Tristan Technologies, San Diego, CA, http://www.tristantech.com/, accessed on August 8, 2012.
6. Hypres, Elmsford, New York, http://www.hypres.com/, accessed on August 10, 2012.
7. Cryogenics Ltd., London, England, http://www.cryogenic.co.uk/, accessed on August 9, 2012.
8. GWR Instruments, San Diego, CA, http://www.gwrinstruments.com/, accessed on August 7, 2012.

6

Next Generation of Smart Sensors

Michael E. Stanley
Freescale Semiconductor

Stéphane Gervais-Ducouret
Freescale Semiconductor

6.1 Introduction

IEEE STD 1451.0 (IEEE, 2005) defines standard interfaces in support of plug and play operation of smart sensors and actuators on a network. Key components of the standard include

- A format for transducer electronic data sheet (TEDS)
- Transducer interface modules (TIMs)
- Network capable applications processor (NDAP)

The standard is optimized for network-based sensors and trade-off complexity and overall system power for interoperability. It is not optimized for sensor subsystems deeply embedded in portable consumer electronics. A new generation of smart sensors is evolving to meet the needs of this market. These include the following features (Gervais-Ducouret, 2011):

- Competitive in terms of cost, area, and power.
- Virtual sensors can be created by fusing data from multiple physical transducers.
- Ability to detect events.
- Ability to determine current context (both spatial and activity).
- Sensor subsystems must be easy to integrate by product designers.

Nine- and ten-axis sensor subsystems will be used as the basis for discussions to follow. They include the following components:

- 3 axes (*XYZ*) of acceleration (*A*) data
- 3 axes (*XYZ*) of magnetic field (*M*) data
- 3 axes (*XYZ*) of angular velocity (*G*) data
- 1 axis of air pressure (*P*) data

Next-generation devices will incorporate combinations of these and other sensors, as well as the ability to abstract the measured data to make it more useful to the user.

6.2 Need for Sensor Fusion

Each of the sensors in the sensor subsystem outlined in the previous section comes with its own set of strengths and weaknesses. As an example, let us consider the case of an accelerometer and a magnetometer at rest. Both provide the magnitude and direction, in the sensor's frame of reference, for physical quantities being measured. In this case, these are gravity and earth magnetic field, respectively. Both sensor types are unable to detect rotation about the vector in question when one of *X*, *Y*, or *Z* sensor axes is aligned with that vector. Since the earth's magnetic and gravity vectors are guaranteed to be non-colinear,* we can combine the two sets of measurements to provide a complete description of the device orientation in space. The ability to extract knowledge from the sum of multiple sensor outputs that could not be determined from any individual sensor is a key feature of *sensor fusion*.

Table 6.1 expands on this theme to outline key strengths and weaknesses of the various sensors contained in our representative sensor subsystem. The sensors shown represent only a small (but important) fraction of device types that might be included in specific products.

Figure 6.1 summarizes basic market requirements for sensor fusion at the time of writing. They include the following outputs:

1. "Corrected" sensor outputs in which sensor errors have been minimized
2. Current device orientation, which can be expressed in several ways
   a. Rotation matrix to take the device from a defined reference orientation to the current device orientation
   b. A quaternion representation of the same rotation
   c. Euler angle representation of the same rotation (Φ, θ, and ψ)
3. Tilt-compensated compass heading

TABLE 6.1 Trade-Offs between Sensor Types

Sensor	Strengths	Weaknesses
Accelerometer	Inexpensive Extremely low power Very linear Very low noise	Measures the difference of gravity and acceleration We need them separate
Magnetometer	The only sensor that can orient itself with regard to "north" Insensitive to linear acceleration	Subject to magnetic interference Not constant with location within office environments
Gyroscope	Relatively independent of linear acceleration Can be used to "gyro-compensate" the magnetometer	Power hog Long start-up time Zero rate offset drifts over time
Pressure sensor	The only stand-alone sensor that can give an indication of altitude	A "relative" measurement Subject to many interferences/ environmental factors

* Except at the north and south magnetic poles. Since the human population of both is nominally zero, we can discount this case.

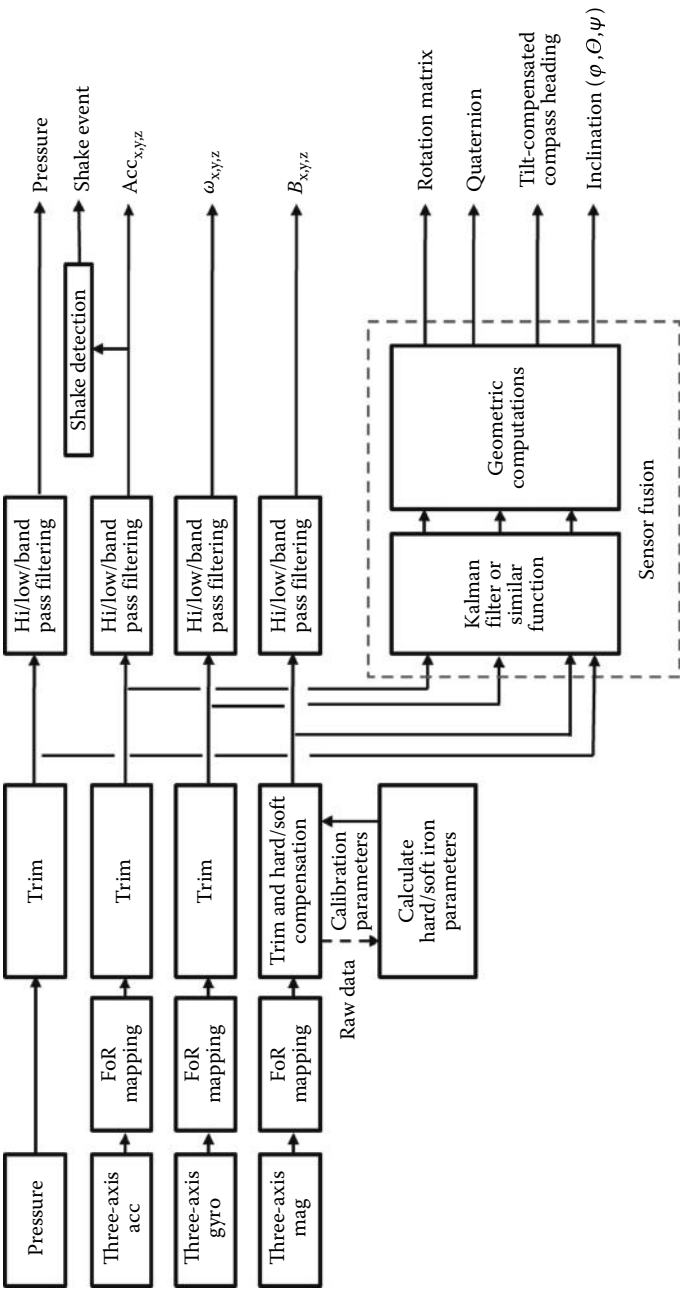


FIGURE 6.1 Base-level sensor fusion for consumer devices. FoR = Frame of reference mapping. (Stanley M. and Kelliher T., *Xtrinsic Sensor Fusion Part 1: Terms, Trends, Challenges and Advantages*, Freescale Technology Forum, San Antonio, TX, June 2012.)

The outputs of Figure 6.1 represent a basic set of foundation functions upon which higher levels of software abstraction are layered. These would include

- Pedometry
- Gesture recognition
- Head tracking
- Motion capture
- Fall detection
- VR tracking
- Augmented reality
- Navigation

6.3 Use Cases

The sensor subsystem outlined in the previous sections can support a variety of features, applications, and use cases. Table 6.2 presents a variety of features and applications versus the minimum set of sensors required for implementation. In almost every case, inclusion of additional sensor types over the minimum shown will improve accuracy.

It is common for combinations of the previously mentioned features to be executing concurrently on the same device. Some applications may continue to run even when the end user is not actively interacting with the device. Pedometry is a good example of this: you probably want to count steps all day, even when the smartphone on which it is running is not in "active use."

6.4 Power Considerations

The time between battery charges is a primary figure of merit for any portable consumer device; hence, every milliwatt is precious. A properly designed sensor subsystem can enable major improvements in battery life.

Figure 6.2 illustrates the case where the main applications processor (AP) in a system is responsible for sensor fusion. Most commonly, the AP bus interface is comprised of one or more serial interfaces (I^2C or SPI). Every sensor sample must be individually read across the bus interface by the AP, which is subject to a high interrupt rate.

APs are typically manufactured using state-of-the art complementary metal-oxide-semiconductor (CMOS) processes. These are designed to minimize run mode currents for the AP; however, this is

TABLE 6.2 Minimum Sensor Sets for Various Applications

Feature/Application	Accelerometer	Magnetometer	Gyroscope	Pressure Sensor
Portrait/landscape, tap detection, fall detection	X			
Pedometry, vibration analysis	X			
Electronic compass, pointing/remote control, augmented/virtual reality, gesture recognition	X	X		
Virtual gyroscope	X	X		
Gyro-compensated electronic compass	X	X	X	
Activity monitors	X	X		
	X		X	
Motion capture	X	X		
3D mapping and localization	X	X		X
Image stabilization, gesture recognition	X			

Note: X indicates that the indicated sensor type is a minimum requirement for the feature/application.

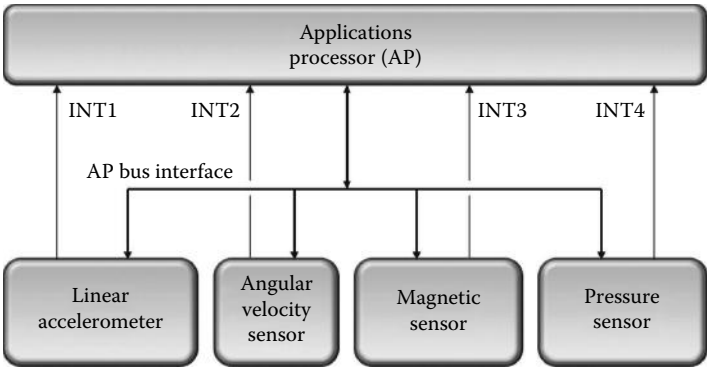


FIGURE 6.2 Sensor fusion via AP.

normally at the expense of increased start-up latency coupled with multiple modes of operation (each with their own power trade-offs). Deep pipelines for the AP result in substantial power wasted during context switching. If the AP is otherwise idle, additional power is burned during wake-up/sleep transitions.

Another problem with the system of Figure 6.2 is that the operating system running on the AP doesn't normally support real-time operation. The AP may have, at best, a rough idea of exactly when each sensor sample was taken. This uncertainty in the time domain leads directly to errors in sensor fusion algorithms performed on the AP.

In contrast to the previous text, with the system of Figure 6.3, the sensor fusion is performed by a local compute engine residing on an intelligent sensor hub. The compute engine may be a microcontroller, dedicated fusion logic implemented in gates, or a combination of the two. Use of a dedicated microcontroller offers maximum flexibility for enhancing the fusion feature mix over time. Once specific algorithms have proved to be matured, they can be moved into gates. Logic-based solutions are almost guaranteed to have lower power than other architectures, but should only be used when algorithm maturity has been proven. We expect a mixture of both methods to be used for some years to come.

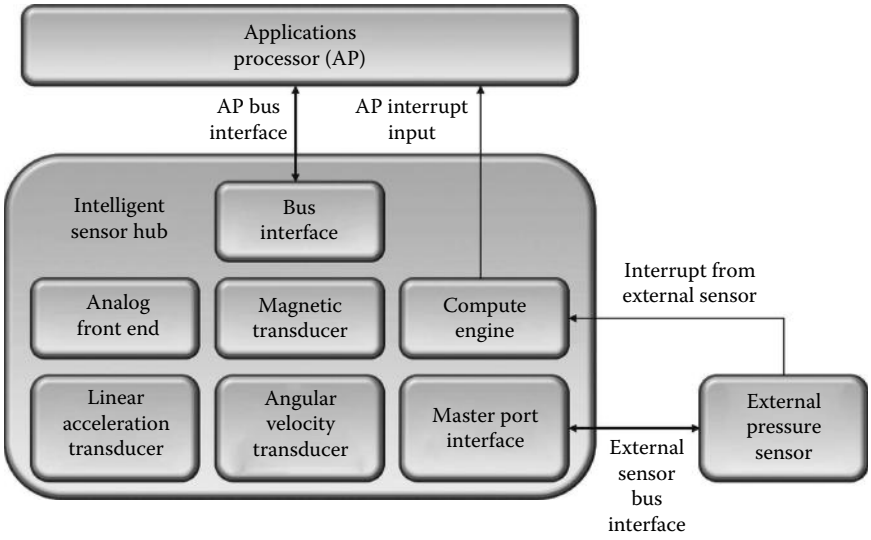


FIGURE 6.3 Sensor fusion via an intelligent sensor hub.

A number of advantages are obtained by moving sensor fusion out of the main AP:

- Sample times are tightly coordinated in time.
- MCU-based sensor hubs can utilize a real-time OS (RTOS) to manage external sensor inputs and algorithm timings.
- AP bus traffic can be orders of magnitude less than the system of Figure 6.2 because only processed and aggregated data need to be passed up to the AP.
- Because the local compute engine does not have the context/power mode switching overhead of the AP, milliwatt per computation is minimized.
- The AP can be completely asleep while the sensor hub continues to process sensor-related functions. Consider a smartphone equipped with a sensor hub: the application software for a pedometer can be executed completely on the sensor hub. Therefore, the AP can remain asleep, even while steps are counted.
- The sensor hub can use heuristic methods to manage power for the overall subsystem without any oversight from the AP. For the case shown in Figure 6.3, the magnetometer and gyroscope could be completely powered down until movement is detected via accelerometer inputs.

6.5 Integration and Partitioning Issues

A variety of physical mechanisms can be used to measure the previously mentioned components. Microelectromechanical systems (MEMS) are commonly used to measure acceleration, angular velocity, and pressure for consumer-grade devices. Magnetic sensors can be fabricated using a number of different technologies, including tunneling magnetoresistance (TMR) and anisotropic magnetoresistance (AMR).

First-generation MEMS devices included only one or two measurement axes in a single packaged device. As time proceeds, the trend is to consolidate more axes into a single package. Three-axis sensors are now common, and 2012 has seen the introduction of 6 (*X*/*Y*/*Z* of both accelerometer and gyroscope) and 9 axes (*X*/*Y*/*Z* of accelerometer, magnetometer, and gyroscope) devices. There are a number of items that need to be taken into consideration Stanley (2012) when selecting components of this type:

- Equations of motion incorporating acceleration are simpler when the sensor is located in the center of mass of the device in question.
- Magnetic sensors are subject to interference from ferrous and magnetic materials within the unit. Typically, a magnetic sensor performs best when situated near the center of one edge of the device because the magnetic field will suffer less distortion in that area.
- The first two bullets suggest different sensor placement on the PCB. If a single device incorporating both magnetic and acceleration transducers is used, the fusion algorithms need to include the ability to specify sensor offset from center of mass.

From a physical perspective, the trend is clearly toward more and more integration into smaller and smaller packages. In 2010, discrete packages for accelerometer, gyrometer, and magnetometer were the norm. $3 \times 3 \times 1$ mm packaged sensors were the state of the art. At the time of writing, 2×2 mm is standard for 3-axis accelerometers, and gyroscopes are at 4×4 mm and trending to 3×3 mm. Combinational devices are trending to 15–20 mm^2, including the sensor hub. Issues being managed by sensor providers are discussed in the following paragraphs.

6.5.1 Consolidation of MEMS Devices onto a Single Die

Stand-alone consumer accelerometers do not usually have their chambers evacuated. Gas molecules within the MEMS chamber provide a damping effect that is advantageous in shaping the transfer function of the accelerometer. Conversely gyroscopes work better when motion damping is not present.

The Coriolis effect leveraged by gyroscopes requires a proof mass that oscillates at high speeds. The same effect that is helpful for accelerometers increases power required for the gyro drive circuitry. To support higher levels of integration, MEMS structures will most likely be designed to operate at the same pressure. The accelerometer circuitry will need to be designed for force feedback (to ensure the accelerometer proof mass stays centered) if vacuum is used for both gyro- and accelerometer. Conversely, additional drive capabilities may be required if the gyroscope pressure is raised. Finding the right "sweet spot" will be even more complicated when we consider inclusion of a pressure sensor on the same die.

6.5.2 Magnetic Interference

By their very nature, magnetometers are sensitive to electronic currents. Placing them in close proximity to a microcontroller or logic core requires care. The Biot–Savart law can be used to estimate the effect of wire/trace currents on magnetic sensors. For long wires and traces, it can be simplified to

$$|B| = \frac{\mu_0 I}{2\pi r}$$

where

$|B|$ is the magnetic field strength in T (teslas)
μ_0 is the $4\pi \times 10^{-7}$ T m/A
I is the current in amps
r is the distance from wire/trace to sensor in meters

This can be rearranged to yield

$$I \leq 5\ r\,|B|$$

where

$|B|$ is the magnetic field strength in μT that you can afford to ignore
I is the current in milliamps
r is the distance from wire/trace to sensor in mm

As an example, if B = 0.1 μT and r = 0.1 mm, then I must be roughly less than 50 μA in order to not affect the accuracy of the result. While these numbers may sound small, they must be considered when designing within the context of an MCU-equipped smart sensor in a 3 × 3 × 1 mm package.

Today, the magnetic sensor component of a composite device is commonly implemented as a separate die that is stacked on top of the controller Application Specific Integrated Circuit (ASIC). Depending upon the technology used, it should be possible to place the magnetic sensor on the same die as the compute engine. Ultimately, a 9-axis fused solution will probably only require two stacked dies.

Figure 6.4 illustrates this process of continuing integration. Notice that at any point in time, several alternative solutions may exist. These may be driven by both logistical (supply) and technical considerations. Typically, no one solution fits all problem spaces.

6.5.3 Soft Partitioning

A benefit of the sensor hub partitioning is that fusion algorithms can be packaged with the hub itself, simplifying hardware and software system design and integration. System integrators are increasingly looking for a single source capable of providing sensors, compute engine, and associated software. Microsoft Windows 8 is architected with this division in mind. Microsoft supports specific bus

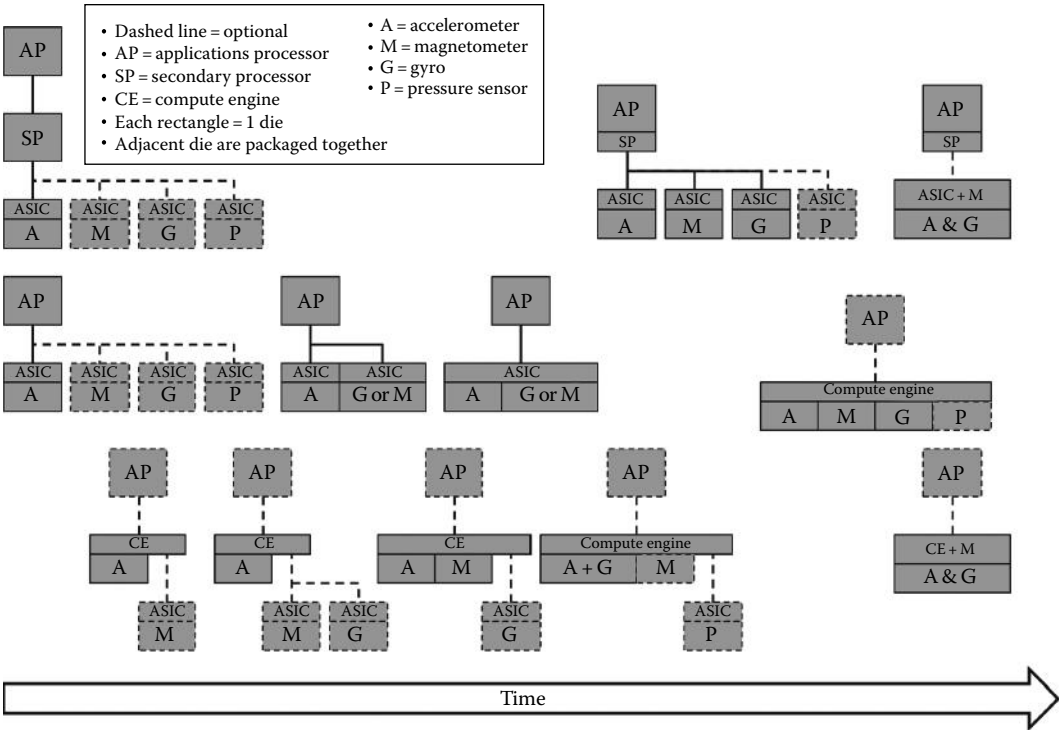


FIGURE 6.4 Increasing levels of integration over time.

protocols for communications with sensor subsystems. These include HID* over Universal Serial Bus (USB) and HID over I²C. Microsoft requires vendors to meet all Windows 8 Hardware Certification Requirements (Microsoft, 2012c) before the vendor's solution can be marketed as Microsoft-approved. Google has created the Android Compatibility Test Suite (CTS) for similar purposes (Android, 2012).

A high-level view of the Android software stack utilizing an intelligent hub is shown in Figure 6.5. By agreeing to a common communications protocol over I²C, SPI, or USB, sensor providers and system integrators can cleanly divide software development tasks. An alternate approach that is even more hardware/software agnostic could be based upon the Multicore Communications API (MCAPI, 2012). This standard message passing protocol could be used to completely abstract the smart sensor interface at the software level. Regardless of which approach wins out in the end, industry-wide definitions will enable the ability to swap out sensor subsystems with the substitution of a single component—with no impact to system software.

Physically, the systems of Figures 6.3 and 6.5 are the same. They support a clean division of efforts between sensor providers and system integrators and improved performance (both power and accuracy) over previous generations of devices. However, power can be improved even more by moving the compute engine onto the same die as the AP. This is shown in Figure 6.6.

The controller IC for the Figure 6.3 intelligent controller will be built upon CMOS processes that include support for flash memory. This allows fusion software to be stored directly in the sensor hub, simplifying distribution and integration of that software. The disadvantage is that flash-based technologies tend to "lag" the more advanced processes used for APs in terms of instructions/mW. In Figure 6.6, the secondary CPU can be designed for real-time operation with fast transitions to/from sleep mode. Improvements in power come from using the same advanced CMOS technology for the secondary CPU

* HID, human interface device. See Microsoft (2012b).

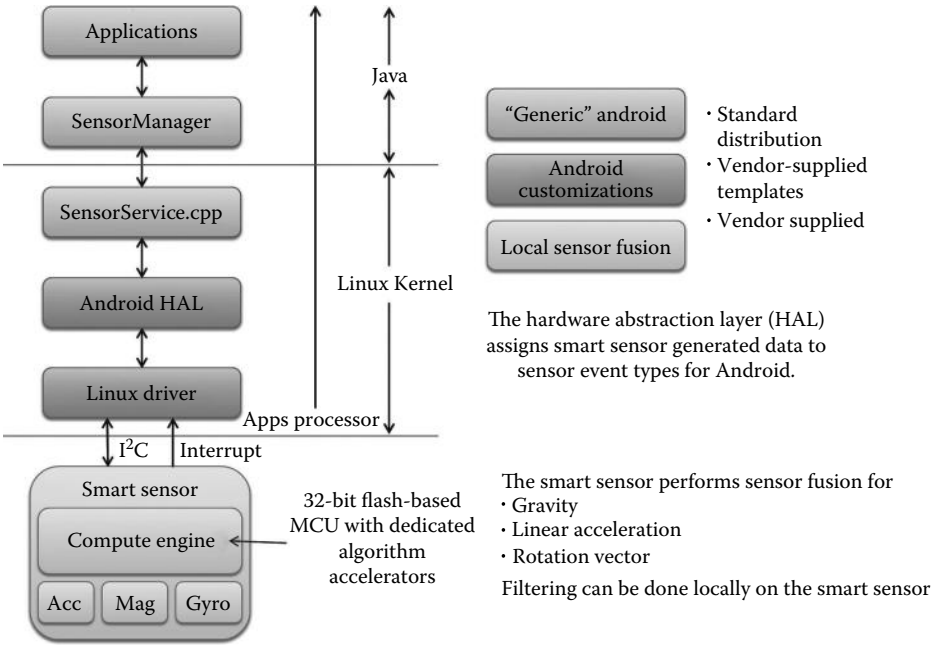


FIGURE 6.5 Software stack based upon an intelligent sensor hub.

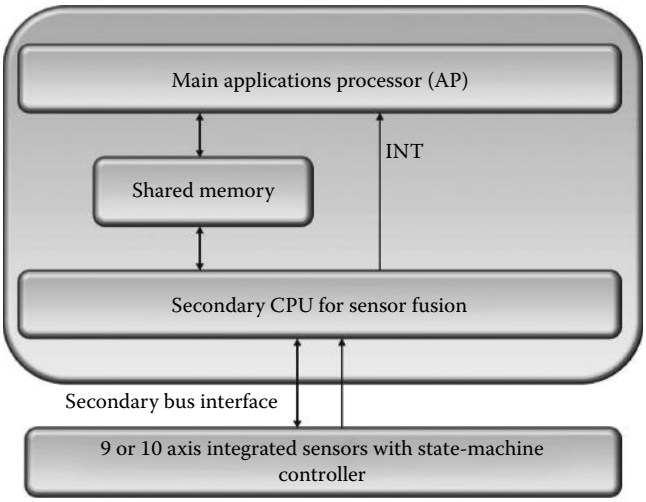


FIGURE 6.6 Sensor fusion as a secondary engine within the AP.

as for the main AP. The architecture of Figure 6.6 allows the system designer to choose best in class options for each individual sensor, providing a more "a la carte" option to the partitioning shown in Figure 6.3, which is more likely to be provided by a single vendor. This comes at the expense of additional traffic on the secondary bus and a tighter coupling between the sensor subsystem and the rest of the system from a software perspective. The latter may be reduced by implementing a software protocol such as MCAPI, which hides details of the physical transport mechanism from the application. Ultimately, the market will determine which of the two sensor hub approaches wins out.

Trade-offs between the systems of Figures 6.2, 6.3, and 6.6 are outlined in Table 6.3.

TABLE 6.3 Trade-Offs Associated with Choice of Processor for Sensor Fusion

System Option	Pros	Cons
Figure 6.2: Fusion on the AP	Least expensive option Supported by 2012 technology	Higher power Highest software burden on-system designer
Figure 6.3: Fusion within the smart sensor hub	Highest level of decoupling between top-level OS and sensor subsystem	Significantly improved power over 100% discrete option
Figure 6.6: Fusion on a secondary processor on the AP	Best power	Physical (and possibly software) partitioning between the main AP and secondary processor will tend to be different for each AP design

6.6 Trend and Market Segments

Improvements in sensor performance and intelligence are driven by new applications that are themselves enabled by the availability of affordable, small, and low-power sensors. An overview of some of the main trends of automotive, industrial, healthcare, and consumer markets is provided as follows.

6.6.1 Automotive

The automotive market segment has used sensors for many years for monitoring and controlling engine performance. The explosion of sensor usage in this market has been driven by the need for increased safety, starting with airbags (which utilize accelerometers for triggering purposes) and tire pressure monitoring systems (TPMSs). The current trend is to further improve automotive security by adding active safety features to the existing passive safety feature set. Electronic stability control (ESC) can prevent some crashes due to loss of control by actively managing the braking of each wheel. ESC uses a combination of a gyroscope (1–3 axes) and accelerometer (2–3 axes) in a common package with an ASIC to efficiently monitor yaw, pitch, and rollover. The trend is clearly driving further integration of sensors and sensor fusion for proactive safety. Adoption of 6-axis sensors (3-axis gyroscope and 3-axis accelerometer) and the fusion of these sensors should increase in this market.

6.6.2 Industrial

The industrial market also benefits from smart sensors. Usage of sensors is as diverse as the industrial applications themselves. For instance, physical tamper detection, utilizing extremely low-power 3-axis accelerometers running for years on battery power, can now be seen in smart meters and other devices. Accelerometers are also used for vibration monitoring to detect machine defects and trigger maintenance prior to a real failure. Associating sensors with low-power wireless transceivers may enable the most pervasive technology for industrial: building and home automation. The main benefits are energy saving (heater, light, HVAC, etc.), cost saving (manpower, cables, maintenance, etc.), and security (alarm, detection of fire and gas, malfunction of devices, etc.). Smart sensors are crucial for such applications since power consumption is critical: 10 years is the minimum battery life for building automation sensors. Dedicated processing at the sensor level is one of the solutions used to minimize data flow and communication and therefore power consumption. Association of smart sensors together with very low-power wireless transceivers, such as ZigBee Smart Energy (Zigbee, 2012) and even energy harvesting (using differential temperature or vibration), will dramatically increase the spread of wireless sensor networks for better management of our energy.

6.6.3 Healthcare

Health monitoring applications benefit from smart sensor's low-power and quality of data. For instance, seamless monitoring of blood pressure, physical activity, and heart rate can be achieved using smart

accelerometers and pressure sensors. Data can be recorded and displayed on any device acting as a health hub (e.g., smartphone). Accurate and low-power monitoring is possible thanks to smart sensor features such as First-In-First-Out (FIFO) buffers, which allow the transmission of measured data in burst mode. This minimizes the number of context changes required by the main processor, again, minimizing overall system power consumption.

Smart sensors and resulting fused data will be increasingly used to enable seamless and more detailed health monitoring. Data can be used for sport activity monitoring and accident prevention. It can also be collected and analyzed on a continuous basis, opening the possibility of adapting medication regimens to each individual patient's daily routine.

6.6.4 Consumer

The consumer market benefits from smart sensors more than any other, since these devices enable high-volume applications such as gaming, intuitive user interfaces, augmented reality, localization, and context awareness. The 10-axis sensor system discussed in the introduction to this chapter forms the basis for many consumer applications.

Figure 6.7 breaks applications into three levels of complexity. Bottom tier applications are supported by a 3-axis accelerometer and 3-axis magnetometer. This tier includes basic gesture recognition such as tilt or tap detection.* It also requires functions for soft and hard iron compensation, which must be performed prior to electronic compass calculations. The first level of smart sensor and sensor fusion provides improved sensor performance in terms of drift, calibration, and ease of use. It forms the foundation for applications in the upper levels.

The middle level in Figure 6.7 includes applications such as a high-performance pedometer (with an accelerometer to detect steps, a gyroscope to remove parasitic movement linked to the handling of the

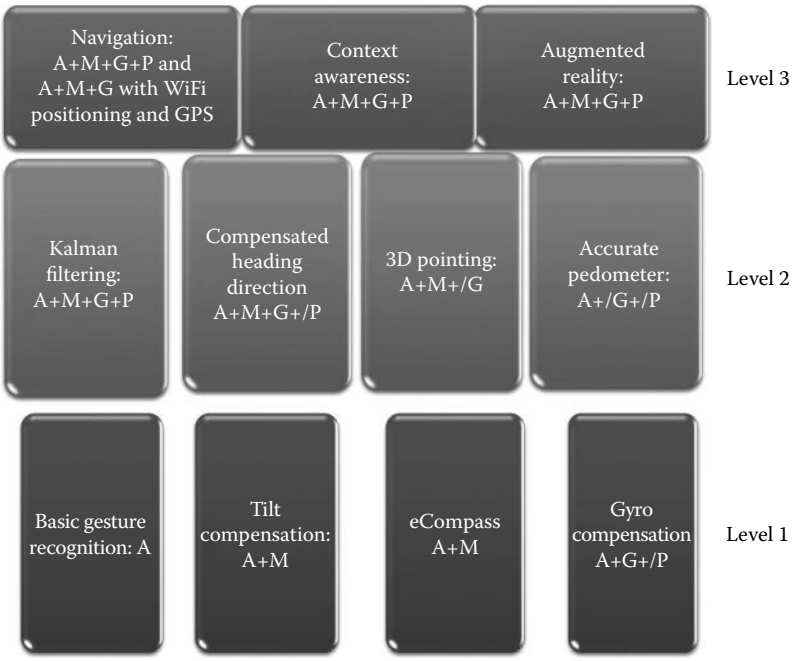


FIGURE 6.7 Sensor fusion complexity levels/indicates optical sensors.

* Basic functions like tilt and tap are often implemented in dedicated CMOS logic today. As other algorithms increase in maturity, it will be natural that many of these functions also move into "hard logic."

device, and a pressure sensor to detect altitude), pointing device (used in remote controls to improve intuitive interface for smart TV), and Kalman filtering to get reliable and more accurate sensor data as described in earlier sections.

Level 3 applications leverage sensor and general *data fusion*. Navigation applications can use the outcome of sensor fusion to lower the power consumption of GPS by extrapolating the trajectory between two navigational fixes. However, the most dramatic improvements are for indoor localization. These systems utilize inertial/magnetic sensor fusion coupled with diverse techniques such as Wi-Fi positioning, cell IDs/triangulation, and soft maps to enable indoor navigation via smartphones. Going forward, they will also combine pressure readings with map/GPS data to estimate altitude.

Smart sensors play a key role in decreasing power consumption in these systems, since the power consumed by GPS/Wi-Fi is of two orders of magnitude higher than other sensors. They also help improve accuracy of position estimates. Today's consumer sensors are not accurate enough to support purely inertial navigation equations. But by coupling pedometry techniques with orientation data provided by the same sensors, reasonable (linear) bounds can be placed on position estimates.

Another new trend is inclusion of context awareness. This involves gathering information from different sensors and data sources to characterize the user's environment and activity within that environment. Examples might include entering/exiting an elevator, recognizing that user is at rest, eating, and walking. Indeed, understanding the context is as important as knowing the location of a person for location-based services (LBSs), contextual services, and menus. Knowing if a person is stressed, rushing about town, or just browsing in a shopping mall can be achieved by doing pattern analysis of sensor outputs—primarily accelerometers but also gyroscopes.

Modern operating systems include the ability to register software functions as *event listeners*. Think of these as callback functions that are invoked when something interesting has happened in the physical world. To enable innovative applications, the sensor fusion layer should provide ready to use data for pattern recognition, which then provides trigger events for application-level software to log context changes and/or perform event-triggered actions. The goal is to continually raise the level of data abstraction to make it easy for software developers to develop applications that interact in the world in natural and innovative ways. Hence, data fusion of smart and low-power sensors combined with wireless connectivity is enabling new applications for multiple markets. Indeed, since the role of smart sensors in context awareness, positioning, healthcare, and safety is essential, this technology is becoming pervasive for most of related IoT (Internet of Things) applications.

6.7 Summary and Conclusions

This chapter has examined trends in smart sensors using a basic 10-axis sensor subsystem as a basis for discussion. Clear trends include

- Continued focus on reducing cost and power
- Continued improvement in raw sensor performance, as well as performance improvements resulting from intelligent sensor fusion
- Consolidation of more sensors into fewer dies within fewer packages, utilizing less board space
- Localized computation
- Standardizing interfaces to sensor subsystems
- Raising levels of data abstraction, including inclusion of higher-level functions like gesture recognition and context awareness

References

Android. 2012. Android compatibility definition document (CDD) and android compatibility test suite (CTS), http://source.android.com/compatibility/. Accessed on August 15, 2012.

Gervais-Ducouret S. 2011. *Next Smart Sensors Generation*, Sensors Applications Symposium (SAS), IEEE, February 22–24, 2011, pp. 193–196.

IEEE. 2005. IEEE 1451.0, IEEE Standard for a smart transducer interface for sensors and actuators—Common functions, Communication protocols, and transducer electronic data sheet (TEDS) formats, Instrumentation and Measurement Society, TC-9, Institute of Electrical and Electronic Engineers, New York, pp. 10016–15997, 2005.

MCAPI. 2012. Multicore communications API working group (MCAPI®), http://www.multicore-association.org/workgroup/mcapi.php. Accessed on August 15, 2012.

Microsoft, HID Sensor Usages—Annotations for Windows HID Sensor Class Driver, Microsoft Corporation, http://msdn.microsoft.com/en-us/library/windows/hardware/hh975383.aspx. Accessed on August 14, 2013.

Microsoft, Integrating Motion and Orientation Sensors with PC Hardware Running Windows 8, Microsoft Corporation, http://feishare.com/attachments/upload/integrating-motion-and-orientation-sensors.pdf. Accessed August 14, 2013.

Microsoft. 2012c. Windows 8 Hardware Certification Requirements, http://msdn.microsoft.com/en-us/library/windows/hardware/hh748200.aspx?ppud=4. Accessed on August 15, 2012.

Stanley M. 2012. *Physical Design Considerations for Products Incorporating Inertial & Magnetic Sensors*, Presentation at Design West, San Jose, CA, March 2012.

Stanley M. and Kelliher T. 2012. *Xtrinsic Sensor Fusion Part 1: Terms, Trends, Challenges and Advantages*, June 2012, Freescale Technology Forum, San Antonio, TX.

Zigbee. 2012. Zigbee smart energy overview, http://www.zigbee.org/Standards/ZigBeeSmartEnergy/Overview.aspx. Accessed on August 15, 2012.

Partial List of Sensor Manufacturers

AKM, manufacturer of magnetometers and 6-axis combo eCompass.

Analog Devices, manufacturer of accelerometers, pressure sensors, gyroscopes, microphones, 6-axis combo sensors, sensor fusion, and inertial modules.

Bosch Sensortec, manufacturer of accelerometers, pressure sensors, magnetometers, microphones, sensor fusion, and 6-axis combo sensors.

Freescale Semiconductor, manufacturer of accelerometers, gyroscopes, pressure sensors, magnetometers, touch sensors, sensor fusion, 6-axis sensors, and intelligent sensor platform.

Infineon, manufacturer of magnetometers and pressure sensors.

Honeywell, manufacturer of magnetometers.

Invensense, manufacturer of gyroscopes and combo sensors of a gyroscope with an accelerometer and/or magnetometer and sensor fusion.

Kionix, manufacturer of smart sensors: accelerometers, gyroscopes, 6-axis combo sensors, and sensor fusion.

Knowles Electronic, manufacturer of MEMS microphones.

Memsic, manufacturer of accelerometers, magnetometers, and flow sensors.

Murata, manufacturer of accelerometers, pressure sensors, gyroscopes, 6-axis combo sensors, and sensor fusion.

Panasonic, manufacturer of accelerometers, gyroscopes, and RF MEMS.

STMicroelectronics, manufacturer of MEMS and smart sensors: accelerometers, pressure sensors, magnetometers, gyroscopes, microphones, 6-axis combo sensors, and inertial modules.

7

Needle-Type Environmental Sensors

Woo Hyoung Lee
University of Central Florida

Jin-Hwan Lee
Intel Corporation

Woo-Hyuck Choi
Globalfoundries Incorporation

Ian Papautsky
University of Cincinnati

Paul L. Bishop
University of Rhode Island

7.1 Introduction

The development of environmental microsensor techniques is a revolutionary advancement in the measurement of both absolute levels and changes in chemical species in engineered and natural aquatic systems. This advancement has utilized microscopic monitoring to contribute a greater understanding of biological mechanisms in many environmental processes for decades. Currently, microelectromechanical systems (MEMS) microfabrication technology, which is intimately connected to a microelectronic circuit for amplifying, processing, and transmitting microelectrode (ME) signals, is being successfully applied to the development of multianalyte sensor systems for in situ monitoring.

The needle-type electrochemical microsensor is one of the most prominent, novel methods for studying biofilms. Biofilms are colonies of microbial cells in an organic polymeric matrix (usually polysaccharides) attached to a surface in water systems. They have complex structures consisting of highly stratified microbial communities. The ability of bacteria to attach to surfaces and to form biofilms often is an important competitive advantage for them over bacteria growing in suspension. Biofilms allow for mixed microbial communities, concentration of nutrients, and protection from antibiotics. The formation of biofilms has been associated with a broad range of industrial problems at an annual cost of billions of dollars. For example, biofilms are ubiquitous in water distribution systems,

and the control of their growth has been a great challenge, with many water utilities in the United States reporting biofilm survival in water distribution systems despite the continuing presence of disinfectants. Biofilm may also harbor various types of microorganisms including opportunistic pathogens and thus can threaten public health. On the other hand, biofilm plays a major role in many water reclamation and reuse technologies as well as bioremediation of contaminated soil. Therefore, whether we want to promote good biofilms or eliminate bad biofilms, it is important to understand how they work and what works to control them [1]. They are critical for effective design and operation of many environmental systems.

The tiny tip (3–15 μm diameters) of the microsensors makes them very attractive experimental tools for direct measurement by penetrating biological samples (e.g., biofilm sediments, pore water, activated sludge flocs). Most commercial sensors have a large tip (e.g., 3 cm diameter), so they are difficult to use for microscopic analysis. Using a microsensor, a concentration profile is measured at a microscale, and parameters are estimated from the measured concentration profiles to determine rates of activity and transport of biological and chemical compounds. The important kinetic parameters at a given location in a floc or biofilm include constituent flux (J), net specific consumption (or production) rates, diffusion coefficient (D), maximum reaction rate (K_s), and disinfectant penetration rate. Therefore, the microsensors are very useful for the analyses of dynamics and kinetics in complex environmental systems such as biofilm processes and bioremediation.

In this chapter, we describe needle-type environmental microsensors that can measure at high spatial resolution and that are yet sensitive enough to detect concentration changes of a target compound (or constituent) that are relevant in environmental settings. Subsequently, the applications of environmental microsensors are demonstrated in engineered and natural stratified microbial communities (e.g., nitrifying biofilm, marine sediments, activated sludge flocs, microbial mats).

7.2 Microelectrode Sensors

In many cases, the laboratory analysis of environmental samples is either undesirable or not possible. In situ analysis is often necessary to preserve sample integrity, while in other cases on-site analysis is needed to ensure that temporal or chemical changes to samples do not occur. Microsensors for in situ environmental applications must achieve four major requirements in order to accurately assess complex environmental samples:

- Good sensitivity in the concentration range of interest
- High spatial resolution at the microscale level
- Robustness sufficient to penetrate biological samples (e.g., sediment, biofilm, activated sludge flocs)
- High selectivity against possible interferences (e.g., minimization of pH, oxygen, or other ionic interferences must often be achieved in the presence of relatively high concentrations of interfering cations [or anions])

Key parameters of an optimal microsensor include response time, stability over time, lower detection limits, and selectivity (independence of response). Most recent improvements are based on these parameters. Generally, there are two types of electrochemical microsensors (Figure 7.1): (1) amperometric sensors and (2) potentiometric sensors. The selection depends on cost factors, required data reliability, the materials to be analyzed, and available technologies. Amperometric sensors measure current under a controlled potential and convert it to the target concentration. Potentiometric sensors measure the voltage output (i.e., Nernstian electrode response), which is proportional to the concentration of a target ion. Table 7.1 summarizes current developments in environmental microsensors according to their type and measured analyte. Figure 7.2 shows schematic descriptions of various types of environmental microsensors. Most of them are not commercially available and thus have been fabricated in the laboratory. Table 7.2 summarizes the representative materials and equipment necessary to fabricate needle-type environmental ME sensors.

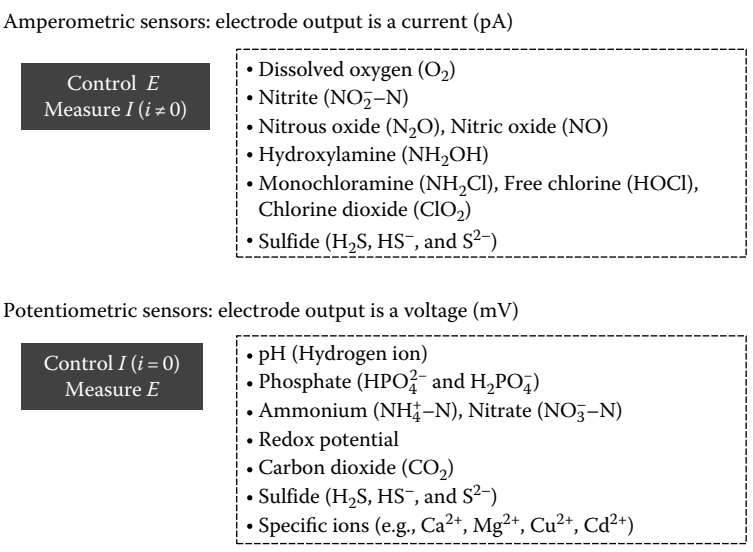


FIGURE 7.1 Principles of needle-type electrochemical environmental microsensor operation. *E*, potential (mV); *I*, current (pA).

A number of microsensors (e.g., oxygen, sulfide gas, hydrogen, nitrous oxide [N_2O], nitric oxide, redox, and nitrite) are commercially available from Unisense A/S (www.unisense.com) or other manufacturers. Typical environmental microsensors have diameters of about 3–15 μm (Figure 7.2), while a specific combined Clark-type microsensor, which needs an outer chamber (or casing), has a diameter of 50–80 μm. Smaller sensors have the advantage of causing less disturbance when introduced, but there is some decrease in sensitivity [2].

7.3 Amperometric Microsensors

The most well-known amperometric microsensor is the oxygen sensor (Figure 7.2a) because oxygen has been used as a critical indicator for microbial activity in many environmental processes [3,6,9,24–26]. The method used for ME sensor fabrication depends on the type of microsensor being constructed. Oxygen (or dissolved oxygen [DO]) MEs are made by filling a micropipette with a low-melting-point alloy; the tip is then electroplated with a thin layer of gold (Au) [27]. It measures DO amperometrically based on oxygen reduction in the cathode reaction [6]. The current is proportional to the bulk oxygen concentration (Equation 7.1):

$$I = \frac{nFASDP_{O_2}}{Z} = kP_{O_2} \tag{7.1}$$

where

I is the current
n is the number of electrons involved in oxygen reduction
F is the Faraday number (96,487 °C/mol)
A is the cathode area (cm^2)
P_{O_2} is the oxygen partial pressure
Z is the oxygen diffusion layer thickness
S and *D* are solubility coefficient and oxygen diffusivity, respectively

TABLE 7.1 Overview of Various Needle-Type Environmental Microsensors

Measurement	Type	Analytical Methods/ Mechanisms	Slope[a], Response Time	Conc. Range, (LOD[b])	Key Materials/Compound/ Procedures	References
Oxygen (O_2)	Clark[c], Solid[d]	• Amperometry • Oxygen reduction in the cathode reaction at Au	13 pA/(mg O_2/L), <5 s	0–8.3 mg O_2/L (23°C)	• Low-melting-point (47°C) bismuth alloy • Au electroplating • Polarization	[3–5]
pH	ISE[e]	• Potentiometry • Nernstian electrode response from neutral carrier	45–60 mV/pH	pH 6–10	• Salinization of the tip • Electrolyte: pH 7 buffer solution • Liquid membrane: Fluka 95293 • Cellulose acetate and protein layer at the tip	[3]
			55–59 mV/pH	pH 1.3–9.8	• Salinization of the tip • Electrolyte: NaCl (100 mmol/L), KH_2PO_4 (250 mmol/L), and Na_2HPO_4 (250 mmol/L) • Liquid membrane: Fluka 95297	[6]
	Solid	• Potentiometry • Nernstian electrode response	70–80 mV/pH, <1 min	pH 1–12	• Iridium wire • IrOx formed by potential cycling in 0.5 M sulfuric acid • H_2S interference	[7]
Ammonium (NH_4^+–N)	ISE	• Potentiometry • Nernstian electrode response (NH_4^+–N) from neutral carrier	54.5 mV/dec	10^{-4}–1^{-1} M ($10^{-5.8}$ M)	• Salinization of the tip • Electrolyte: 0.01 M NH_4Cl • Liquid membrane: Fluka 09879 • Cellulose acetate and protein layer at the tip	[8]
Nitrate (NO_3^-–N)	ISE	• Potentiometry • Nernstian electrode response (NO_3^-–N) from neutral carrier	55.6 mV/dec	10^{-4}–10^{-1} M (20 μM)	• Salinization of the tip • Electrolyte: 0.05 M KNO_3 + 0.05 M KCl • Liquid membrane: Fluka 72549 • Cellulose acetate and protein layer at the tip	[9]

Nitrite (NO_2^-–N)	Solid	• Amperometry • Direct electrochemical oxidation of nitrite (NO_2^-N) on solid electrode surface	242 pA/(mg N/L)	0–20 mg N/L (0.02 mg N/L)	• Carbon fiber • Polarization	[10–12]
Nitric oxide (NO)	Clark	• Amperometry • Oxidation of nitric oxide at electrocatalytic surfaces	2 pA/μM, ~1 s	0–4 μM (~30 nM)	• Carbon fiber • Deposition of Ni-TMPP[f] • Internal electrolyte (100 mM sodium phosphate, 100 mM KCl at pH 7.2) • Polarization • H_2S interference	[13]
N_2O	Clark	• Amperometry • N_2O reduction	0.2 pA/μM, 40–50 s	0–1.2 mM (1 μM)	• Polarization (–0.75 V) • Use of an ion-tight silicon (Si) membrane in the oxygen guard • Use of alkaline ascorbate solution to eliminate oxygen interference	[2]
Phosphate (P)	Solid	• Potentiometry • Oxidation corrosion potential (mV) measurement • Reaction between CoO and phosphate ion ($H_2PO_4^-$ and HPO_4^-) between pH 6.7 and 8.4	33–84 mV/dec, <1 min	10^{-6}–10^{-1} M (2.2×10^{-6} M)	• Co wire (50–100 μm diameter) or Co electroplating • Conditioning in DI water to generate CoO layer at the tip • Required simultaneous DO measurement	[14–16]
Free chlorine (Cl_2)	Solid	• Amperometry • Reaction between PtOH and HOCl (and/or ClO^-)	1 pA/μM <1 s	0–7.5 mM	• Pt wire • Polarization • Conditioning in DI water to generate platinum hydroxide (PtOH) at the tip	[17]

(*continued*)

TABLE 7.1 (continued) Overview of Various Needle-Type Environmental Microsensors

Measurement	Type	Analytical Methods/ Mechanisms	Slope[a], Response Time	Conc. Range, (LOD[b])	Key Materials/Compound/ Procedures	References
NH_2Cl	Solid	• Amperometry • Reaction between PtOH and NH_2Cl	18 pA/(mg Cl_2/L), <5 s	0–10 mg Cl_2/L (0.3 mg Cl_2/L)	• Pt wire • Polarization • Controlled condition where predominant chlorine species are NH_2Cl (i.e., Cl_2:N = 4:1, pH 8.0)	[18]
Redox potential	Solid	• Potentiometry • Nernstian electrode response	1 mV/mV, <1 min	—	• Pt wire • Etching in 2 M KCN solution • Within 10 mV of the nominal redox potentials (ASTM D 1498-93)	[3]
Sulfide (S^{-2})	Solid	• Potentiometric measurement of sulfide (S^{-2}) at alkaline pH and total dissolved sulfide ($S_T = H_2S + HS^- + S^{-2}$) at neutral pH • Nernstian response	31 mV/dec (pH >13.5) 51 mV/dec (pH 7.2)	5×10^{-7}–5×10^{-3} M (pH >13.5) 5.8×10^{-6}–5.8×10^{-3} M (pH 7.2)	• Silver wire • Etching in 2M KCN solution • Chemical treatment of Hg^{2+} ($HgCl_2$) to form the sulfide sensing layer • Sulfidization of the silver surface with Na_2S • Required simultaneous pH measurement	[3]
	Clark	• Amperometry • Oxidization of H_2S by $K_3[Fe(CN)_6]$, resulting in the formation of elemental sulfur and ferrocyanide	1.2–0.6 pA/μM H_2S, <100 ms	2–300 μM H_2S (2 μM)	• Polarization (+85 to +150 mV) • Internal electrolyte: ferricyanide ($K_3[Fe(CN)_6]$) as redox mediator • Si membrane	[19]
Carbon dioxide (CO_2)	ISE, Clark	• Potentiometry • Nernstian response from changes in pH (buffer action)	100 mV/dec, ~10 s (or 2–5 min)	0–2 mM (3 μM)	• LIX membrane (pH measurement) • Internal electrolyte: 2 mM $NaHCO_3$ + 0.5 M NaCl • Si membrane	[6,20]

Copper (Cu_2^+)	Solid	• Potentiometry • Reaction of selenium with the copper surface as adsorbed hydrogen reduces $HSeO_3^-$ ions to selenium	29 mV/dec, 10 s	10^{-6}–10^{-1} M (10^{-7} M)	• Copper wire (75 µm diameter) • Copper selenide (CuSe) formed by cathodic deposition (2 mA/cm² from a 0.1 M Na_2SeO_3, pH 4–6) • pH interference	[21]
	ISE	• Potentiometry • Nernstian electrode response (Cu^{2+}) from neutral carrier	33.1 mV/dec	10^{-9}–10^{-4} M (2×10^{-9} M)	• Membrane containing ionophore, NaTFPB[g], *o*-NPOE[h], and PVC • THF[i] • Inner filling solution: Et_3NNO_3[j] and $Cu(NO_3)_2$ • Conditioning for 1 day in 10^{-5} M $Cu(NO_3)_2$	[22]
Cadmium (Cd_2^+)	ISE	• Potentiometry • Nernstian electrode response from neutral carrier	25.2–27.7 mV/dec, 15–30 s	>0.001 mM (10^{-10} M)	• Salinization of the tip • Electrolyte: 10 mM $Cd(NO_3)_2$ + 100 µM KCl • Liquid membrane (200 µm): 10% Fluka 20909 + 10% Fluka 60588 + 80% Fluka 73732 • Cellulose acetate and protein layer at the tip	[23]

[a] Slope of linear regression indicating sensitivity.

[b] Limit of detection.

[c] The Clark-type microsensor is a combined microsensor. In this microsensor, the sensing cathode (or anode) is placed behind an electrically insulating gas-permeable Si membrane, which is extremely permeable to oxygen. The cathode (or anode) is bathed in an electrolyte solution into which a Ag/AgCl reference electrode is immersed [24].

[d] The solid-type microsensor is a single barrel microsensor (i.e., single-anode or cathode-type ME) with the solid metal tip sealed with a glass micropipette and with external reference electrode placed in the sample.

[e] ISE: Ion-selective electrode that contains LIX membrane.

[f] Ni-TMPP: nickel (II) tetrakis (3-methoxy-4-hydroxyphenyl) porphyrin.

[g] 2-Nitrophenyl octyl ether (*o*-NPOE).

[h] Sodium tetrakis [3,5-bis(trifluoromethyl)phenyl] borate (NaTFPB).

[i] Tetrahydrofuran (THF).

[j] Tetraethylammonium nitrate (Et_4NNO_3).

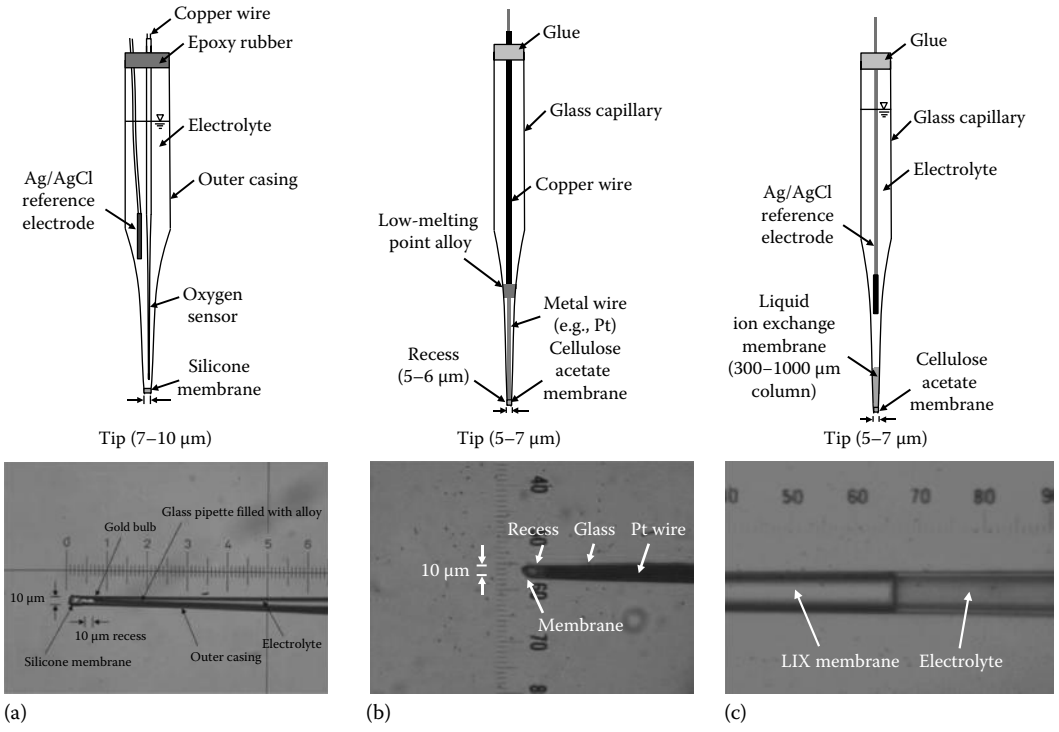


FIGURE 7.2 Schematic descriptions and images of needle-type environmental microsensors: (a) Clark-type combined microsensor (i.e., amperometric oxygen microsensor) and a tip of a combined oxygen microsensor, (b) solid-state microsensor (i.e., amperometric NH_2Cl microsensor) and a tip of a NH_2Cl microsensor, and (c) LIX microsensor with an image of a liquid membrane that is separated from an electrolyte.

Both platinum (Pt) and Au exhibit sensitivity to oxygen at negative potential (i.e., −0.75 V vs. Ag/AgCl), but Au has a higher sensitivity than Pt. As Au wire has different heat expansion coefficients and it is too soft to insert into the glass micropipette, Au electroplating methods have been used [3,5]. For the use of a bare Au wire (50 μm), a modified procedure is needed: the etched Au wire connected to a copper wire (for electric connection) is inserted into the glass micropipette and then pulled out using a micropipette puller (P-1000, Sutter Instrument) with an appropriate setting (e.g., heat: 460, pull: 120, vel: 100, delay: 1, pressure: 500, and ramp: 445). The response time of an amperometric microsensor is very short (within few seconds), and the shelf-life time of a solid-state microsensor (e.g., oxygen microsensor, free chlorine microsensor) is typically 1–2 years, depending on its contamination status.

In amperometric sensing, each microsensor must operate at a specific applied potential (or polarization potential) (e.g., −0.75 V for oxygen sensor, +0.15 V for monochloramine [NH_2Cl] sensor, and +1.2 V for nitrite sensor) depending on the electrochemical reaction between the sensing material and the analyte to be measured. Cyclic voltammetry (CV) has been used to evaluate the applied potential of the microsensors (Figure 7.3). Generally, oxygen reduction in a noble metal occurs at a negative potential: therefore, most microsensors, except oxygen sensors, have positive potentials for their operation unless they have a special treatment to remove the oxygen interference (e.g., ascorbate chamber in front of the microsensor [2]).

TABLE 7.2 Materials and Devices Necessary for Needle-Type Environmental Microsensors Fabrication[a]

Item	Use	Description	Sources
Metal wires	• Sensing materials for solid-state working electrodes	• 99.99% pure 50–100 μm diameter • Noble metals (Pt or Au) • Other metals (Co, iridium, or silver)	• California Fine Wire (Pt: 100-896, Ag: 100-183, Au: 100-542) • Alfa Aesar (Au wire: 0.1 mm, 10968) • Sigma-Aldrich (Pt wire: 26716-3, Co wire: 52344-5, discontinued)
Carbon fiber	• Nitrite microsensor • Nitric oxide microsensor	• PAN-based carbon fiber (7 μm) • Carbon finer (30 μm)	• Cytec Industries, Inc. (T300 12K) • WPI[b]
Nichrome wire	• Heating filaments for meting glass pipettes	• Nichrome wire, 22 gauge	• Fisher Scientific (S49085) • VWR (66258-066)
Silver wire	• Reference electrode	• 99.99% pure, 0.25 mm diameter	• Sigma-Aldrich (32703-4)
Graphite rods	• Counter electrodes for etching process	• High density, nonporous (e.g., glassy graphite)	• Sigma-Aldrich (49654-5)
Glass pipette	• Outer casing of combined sensors	• Disposable Pasteur pipettes	• Fisher Scientific (13-678-20C)
Glass micropipette	• Glass capillary tubing	• Low electric conductivity glass[c] • Borosilicate glass • Corning 8161 (potash rubium lead): amperometric microsensor • Corning 7056: LIX microsensor	• Warner Instruments (G86150T-4, G75150T-4) • Schott (8533) • Macalaster Bicknell Co. • Sutter Instrument (B12069-15) • WPI (PG10150-4)
Au solution	• Au electroplating	• $HAuCl_4 \cdot 3H_2O$ • Au plating solution	• Sigma-Aldrich (G4022-5G) • Sifco Metachemical (Dalic 3020)
KCN	• Electrochemical etching	• Potassium cyanide ACS grade	• Sigma-Aldrich (31252-250G)
Cellulose acetate	• Membrane used on amperometric microsensor	• Cellulose acetate solution 5%–10% (w/v) in acetone	• Sigma-Aldrich (180955)
Bismuth alloy	• Conductive filling material for electroplating Au for DO microsensors	• 117°F low-melting-point alloy: 44.7% bismuth, 22.6% lead, 19.1% indium, 8.3% tin, and 5.3% cadmium	• Belmont Metals, Inc. (2451A)
Si	• Membrane for DO or N_2O microsensor	• Silastic medical adhesive Si type A	• Dow Corning (Silastic or 732)
Injection needle	• Backfilling of internal reference solution	• Fine injection needle	• WPI (MF28G-5)
Liquid membrane	• LIX[d] membranes for ion-selective microsensors	• Hydrogen ionophore • Ammonium ionophore • Nitrate ionophore	• Fluka, Sigma-Aldrich (95293, 95297, 09879, or 72549)
Salinization solution	• Used to make glass pipette surface hydrophobic	• Salinization solution • *N*,*N*-dimethyltrimethylsilylamine	• Fluka, Sigma-Aldrich (85120, 41716-5ml)
Puller	• Pulling glass micropipette	• Flaming/brown micropipette puller	• Sutter Instrument (P-1000)

(*continued*)

TABLE 7.2 (continued) Materials and Devices Necessary for Needle-Type Environmental Microsensors Fabrication[a]

Item	Use	Description	Sources
Illuminator	• Light source	• NOVAFLEX fiber optic illuminator	• WPI (NOVA)
AC power supply	• Etching process • Heating process	• Variable transformer	• Staco (3PN1010B)
DC power supply	• Electroplating • Construction of Ag/AgCl	• Variable single output dc power supply (0–18 V dc at 0–3 A)	• Omnitron Electronics (HY1803D)
Micromanipulators	• Micromovement during construction	• Manual micromanipulator	• WPI (M3301)
Micropipette beveller	• Grinding ME tips	• Grinder, microscope, micromanipulator, and light source	• Sutter Instrument (BV-10)
Faraday cage	• Screen electromagnetic fields	• Bench top Faraday cage	• Technical Manufacturing (81-334-06)
Vibration-free platform	• Prevent vibrations	• Vibration-free platform	• WPI (VFP) • BenchMate (Kinetic Systems)
Stereomicroscope	• Monitor heat element	• Precision stereo zoom binocular microscope	• WPI (PZMIV-BS)
Uptight light microscope	• Monitoring and inspection of ME tips	• Profession grade microscope: 4×, 10×, 40×	• WPI (W30S)

[a] The information in the table is for reference purposes only. We do not promote any of the products specified in the table.
[b] World Precision Instruments (WPI).
[c] There is a need for a special type of glass with high insulation properties (i.e., Schott 8533 green glass) in the fabrication of oxygen microsensors in order to achieve a low background current.
[d] LIX, liquid ion exchange.

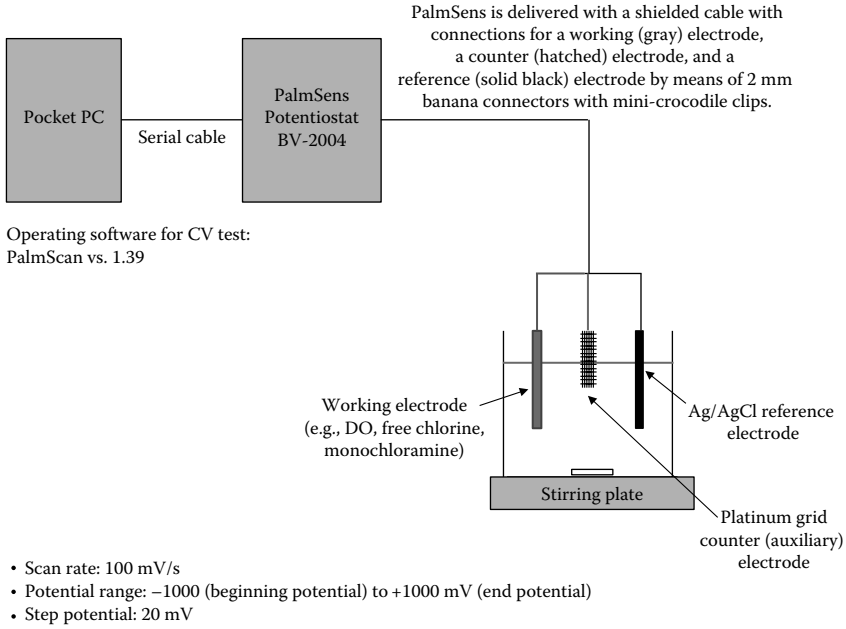


FIGURE 7.3 Schematic drawings of a CV test of amperometric microsensors. An example of the operation range and the connection is described.

7.4 Potentiometric Microsensors

Most ion-selective microsensors are potentiometric sensors where the voltage (mV generated vs. Ag/AgCl) is theoretically dependent on the logarithm ionic activity, according to the Nernst equation (Equation 7.2):

$$\begin{aligned} E_{\text{cell}} &= E_{\text{external ref.}} - E_{\text{internal ref.}} + E_{\text{membrane}} + E_{\text{junction}} \\ &= \Delta E_{\text{ref}} + \frac{RT}{ZF} \ln \frac{A_{i(\text{external})}}{A_{i(\text{internal})}} + E_j \\ &= k + \frac{RT}{ZF} \ln A_{i(\text{sample})} = k' + \frac{0.059}{Z} \log C_{i(\text{sample})} \end{aligned} \tag{7.2}$$

where

E_{cell} is the potential difference between the ion-selective microsensor and the reference electrode
E_{membrane} is the membrane potential
$E_{\text{external ref.}}$ and $E_{\text{internal ref.}}$ are potentials of the external and internal reference electrodes, respectively
E_{junction} is the liquid junction potential
A_i is the ionic activity [A(activity) = γ(activity coefficient) × C(concentration)]
z is the number of electrons involved in the reaction
F is the Faraday constant
R is the gas constant
T is the absolute temperature

The response time of a potentiometric microsensor depends on the concentration of the target ions. Therefore, it varies from a few seconds (at high concentration) to several minutes (low concentration). For example, the response time of a copper ion-selective microsensor varies from less than 60 s ($\geq 10^{-6}$ M Cu^{2+}) to 10 min ($<10^{-6}$ M Cu^{2+}) [22].

Liquid-ion-exchange (LIX) membrane ion-selective microsensors (Figure 7.2c) have been fabricated and used to measure specific ions such as pH, ammonium, and nitrate in biological samples. The key to the ion-selective microsensor is the LIX membrane, which contains a specific neutral carrier for the target ion. This neutral carrier makes it possible to measure the potential of the target ion's activity in the presence of other ions. Therefore, it can detect the target ion without any ion interference. Figure 7.2c shows the LIX membrane of the microsensor that is separated from the electrolyte. Liquid membrane length is generally 300 (50–1000) μm, and its lifetime depends on the length of the membrane, varying from 1 day to several weeks. A typical lifetime of an ion-selective microsensor for pH, ammonium, and nitrate is 2–3 days with a 500–1000 μm length of liquid membrane. Most ionophores for ion-selective MEs can be purchased from Fluka, Sigma-Aldrich (Switzerland) under the trademark Selectophore™ (http://www.sigmaaldrich.com/analytical-chromatography/analytical-reagents/sensoric-applications.html). Recently, ion-selective microsensors with a built-in reference electrode have been introduced to remove outer environmental interferences (e.g., low conductivity) during microprofile measurements [28]. Detailed fabrication procedures are described elsewhere [5,24]. One of the important steps is the silanization step. If LIX is not contained at the tip of the microsensor, this indicates that the silanization step was not done correctly. One possible reason is the contamination of the furnace (i.e., if the furnace is a general use one, organics may be deposited and the silanization will fail. Avoid any dirt, grease, or organics on the case before silanization).

7.5 Data Acquisition and Processing

The microsensor is positioned above the sample (e.g., biofilm) perpendicular to its surface when microprofiles are measured using a microsensor. Then it is moved to the bottom of the biofilm and is used to measure concentration profiles at a microscale (Figure 7.4). From the measured concentration profiles, we can

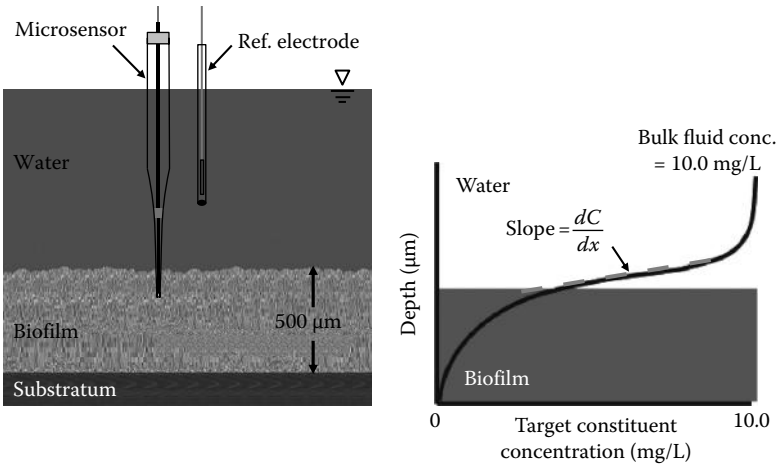


FIGURE 7.4 Schematic drawings of microprofile measurements using a needle-type microsensor and a measured concentration microprofile.

understand transport processes in the biofilm by calculating the rates and net fluxes of constituents (e.g., oxygen, ammonia, and nitrate) and/or diffusion coefficients using Fick's first law of diffusion (Equation 7.3):

$$J = -D_s \frac{dC}{dX} \tag{7.3}$$

where

J is the net flux (nmol/cm^{-2} s)
D_s is the molecular diffusion coefficient of compound C in water
dC/dX is the concentration gradient

The general experimental setup for microprofile measurements [27] includes a flow chamber (custom-made), stereo microscope (Nikon, SMZ-2T; PZMTIV-BS-CCTV, WPI, Sarasota, FL) with CCD camera (Model JE-3662 HR, Javelin Elec., Torrance, CA; 501381, WPI, Sarasota, FL), video display (501385, WPI, Sarasota, FL), data acquisition system (Diamond General Corp, Product 1231 for amperometric sensors; mV meter [Model 215, Denver Instruments] for potentiometric sensors; microsensor multimeter [Unisense A/S, Denmark] with SensorTrace Pro 2.0 software [Unisense A/S, Denmark] for multisensors), 3-D micromanipulators (Model MO-203, Narishige, Japan; MMS, Unisense A/S, Denmark), Faraday cage (Technical Manufacturing Co., Peabody, MA), vibration isolation table, and peristaltic pump systems to flow the medium.

7.6 MEMS Sensor Arrays

As discussed earlier, MEs can be fabricated in a number of ways [29,30], most commonly by pulling a glass micropipette, inserting a noble metal wire such as Pt or Au, and then filling with a low-melting-point alloy [3,4]. Although these ME fabrication methods are well established, there are a number of inherent disadvantages including low success rate, poor reproducibility, fragility, and difficulty in making a multianalyte sensor device [31–33]. Further, these ME sensors are susceptible to electric interference and have to be operated in specialized laboratories under highly controlled conditions. Therefore, there is a critical need for robust, sensitive, and easy-to-fabricate sensors for in situ measurements.

ME miniaturization and integration using MEMS technologies can offer many advantages for overcoming current difficulties for fabrication and integration of sensor components [16,27,33–43].

The key advantages of MEMS microfabrication technologies are (1) the reduction of costs due to batch fabrication, (2) an increase in integration, (3) reduction in power consumption due to smaller size, (4) reduction in complexity, and (5) increased reproducibility, thus producing a large number of uniform sensors. However, the most important advantage of using MEMS fabrication is the increase in sensor reliability due to redundancy and better process control.

Many researchers that apply MEMS technologies to solve sensor problems have focused on microfluidic lab-on-a-chip (LOC) systems [44–46]. Such systems typically contain microfluidic channels for sample collection, preparation, or transport with planar sensing areas for specific target analytes. Nevertheless, the key drawback of such systems is that samples still must be extracted from the site of interest, which often is not acceptable for in situ measurements. MEMS technologies have also been used to develop penetrating 3D ME sensors for neuroscience applications [47–50]. However, all of these needle-type MEMS electrodes have been designed for neural recording and can only measure potential (e.g., neuronal activity).

The recently developed needle-type environmental MEMS sensors, as summarized in Table 7.3, were tapered using the HF etch process, which is a typical MEMS fabrication process. The basic sensing principles are the same as described in Figure 7.1 for these miniaturized MEMS sensors. Figure 7.5 shows schematic diagrams and SEM images of various needle-type environmental MEMS sensors. There are currently no environmental MEMS sensors that are commercially available. The representative materials and fabrication equipment for needle-type environmental MEMS sensors are summarized in Table 7.4.

7.7 Amperometric MEMS Sensor Arrays

The DO MEMS sensor has a recess created at the tip of the ME, as shown in Figure 7.5b [27,36,37,40,43]. The fabrication process has five major steps:

- Dicing to form the array structure
- Etching to sharpen the probes
- Metal deposition for the conductive layer
- Packaging for electric connection
- Sensor tip formation for recessed DO MEMS sensor applications

The following sections are a brief description of the MEMS sensor fabrication process.

7.7.1 Dicing to Form the Array Structure

Glass cover slips are diced to form the array structure of the glass probes. The probe length is 2 cm with 900 μm center-to-center spacing between each glass probe. The cut wafer is then annealed at 550°C for 10 min to relieve stress from the dicing process.

7.7.2 Etching to Sharpen Probes

The probes are first etched in the HF-based etchant solution to smooth the diced surface and reduce the probe dimensions from 175 μm in thickness × 190 μm in width to 85 μm × 90 μm. Next, the glass probes are gradually pulled out to taper them down to 20 μm × 20 μm at the tip. In the final etching step, the glass probes are sharpened to ~200 nm tips using chemical meniscus etching, which forces surface tension at the glass–etchant interface. The process is schematically illustrated in Figure 7.6a.

7.7.3 Metal Deposition for Conductive Layer

The sharpened probes are deposited with a conductive layer, Ti/Au (20 nm/200 nm thick), on all sides by thermal evaporation. Figure 7.6b shows that twelve ME arrays containing 48 MEs were all metalized at the same time.

TABLE 7.3 Overview of Needle-Type Environmental MEMS Sensors

Measurement	Type	Analytical Methods/Mechanisms	Slope[a], Response Time	Conc. Range (LOD[b])	Key Materials/Compound/Procedures	References
Redox potential	Solid	• Potentiometry • Electrode potentials between a metal electrode and a standard reference electrode given by the Nernst equation • Measures the oxidized and reduced forms of all chemical species in the solution (nonselective)	1 mV/mV, <1 s	—	• Core material is a glass or a Si • Au thin film as a working electrode • Meniscus etching (chemical sharpening process) to sharpen penetrating needle-type ME • American Society for Testing and Materials (ASTM) recommends four calibration redox standard and reference solutions within 10 mV of the nominal redox potentials (ASTM D 1498-93)	[27,35,42]
Oxygen (O_2)	Recessed[c]	• Amperometry • Diffusing DO is reduced at the cathode to create a measurable electric current. • Linear relationship between the oxygen concentration and the electric current	~147 pA/mg/L DO, <20 s	0–8.3 mg O^2/L (23°C)	• Parylene C layer (1.5 μm thick) to insulate individual MEs • The recess process: Parylene C deposition, beveling tip end, and etch glass/Au • Au thin film as a working electrode • Apply −750 mV polarization voltage	[27,36,37,40,41]
Phosphate (P)	Solid	• Potentiometry • Co metal electrodes detecting the inorganic[d] and organic[e] phosphate • Oxidation corrosion potential (mV) measurement on CoO surface • Reaction between CoO and phosphate ion depending on the solution pH	96 mV mV/dec, <1–30 s	10^{-6}–10^{-1} M (2.2×10^{-6} M)	• Co electroplating using Co plates as anodes • Oxidize by immersing in DI water along with a Ag/AgCl reference electrode to generate CoO layer at the tip • Oxygen affect for activity of phosphate ions	[16,38,40–42]
pH	Solid	• Potentiometry • IrOx electrode without an ion-selective membrane • Electrode potentials given by the Nernst equation	~62.0 mV/pH, <5 s	pH 2–12	• Electroplate IrOx using IrOx electrolyte solution • Plasma oxygen etch to remove Parylene C insulate layer	[27,41,42]

[a] Slope of linear regression indicating sensitivity.

[b] Limit of detection.

[c] Three types of DO ME are used for DO measurements: the Clark-type, the open-type, and recessed DO ME. The Clark-type DO ME consists of a working electrode with a gas-permeable membrane and an internal reference electrode in the electrolyte. The open-type and recessed DO MEs use an external reference electrode; thus, the miniaturization of these sensor types can be achieved easily with simpler fabrication and smaller working electrode size.

[d] Inorganic phosphate including $H_2PO_4^-$, HPO_4^{2-}, and PO_4^{3-}.

[e] Organic phosphate including disodium adenosine 5′-triphosphate (ATP), disodium adenosine 5′-diphosphate (ADP), and disodium adenosine 5′-monophosphate (AMP).

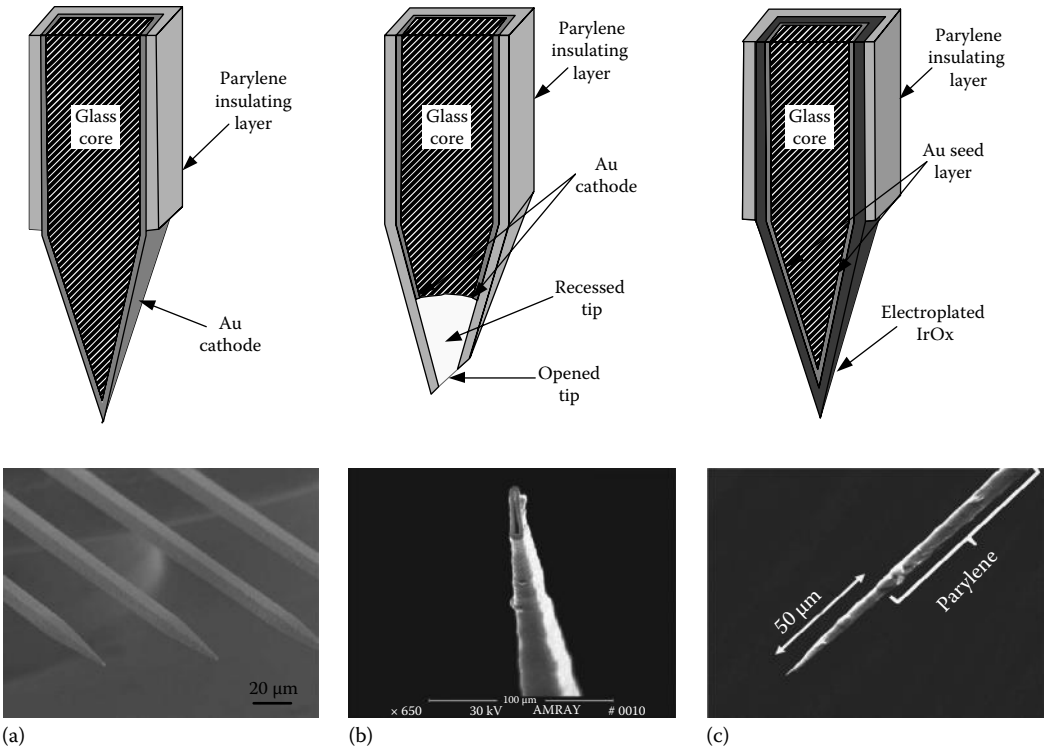


FIGURE 7.5 Schematic diagrams and SEM images of various MEMS sensors: (a) ORP MEMS sensor; (b) DO MEMS sensor: amperometric recessed DO sensor. (From Lee, J.H. et al., *J. Micromech. Microeng.*, 19, 025022, 2009.); and (c) IrOx pH MEMS sensor: solid-state pH sensor. (From Lee, W.H. et al., *Meas. Sci. Technol.*, 22, 042001, 2011.)

7.7.4 Packaging for Electric Connection

To establish electric connections with individual sensors, ME arrays are packaged on a printed circuit board (PCB) with conductive silver epoxy. Parylene C is coated over the entire substrate to insulate individual MEs.

7.7.5 Sensor Tip Formation

To create the recess, the HF-based etchant is used to etch the glass core, and the Ti is exposed by beveling. The beveling and etching steps permit precise control of the recess opening size and depth to relocate the Au sensing area inside the formed recess.

The finished MEA sensor consists of four 2 cm long sharpen MEs packaged on a PCB carrier as is shown in Figure 7.6d. The tip sizes and the overall dimensions of the MEMS sensor array are compared with those of commercially available millielectrodes (COMs) and ME sensors in Figure 7.6c. Then, the DO MEMS sensors are polarized with an applied −0.75 V and calibrated with a commercial Ag/AgCl reference electrode. The completed MEMS sensor has a high sensitivity of ~147 pA/mg/L DO, with less than 20 s required reaching 90% response (Table 7.3).

7.8 Potentiometric MEMS Sensor Arrays

Table 7.3 shows three potentiometric MEMS sensors including redox potential (oxidation–reduction potential [ORP]), phosphate, and pH sensors. These MEMS sensors are good examples for potentiometric sensors with/without an ion-selective membrane. The ORP sensor is the most basic electrode

TABLE 7.4 Materials and Devices Necessary for Needle-Type Environmental MEMS Sensor Array Fabrication

Item	Use	Description	Sources
Borosilicate glass wafer	• Core structure material for working electrode	• 175 μm thick, 45 mm × 50 mm	• Erie Scientific, Pittsburgh, PA
Dicing saw	• Cut individual glass probes	• 45 μm diamond grit resinoid blades (MicroSwiss)	• K&S (740)
Programmable box furnace	• Anneal glass wafers to relieve stress from the dicing process • Use cooling steps to reduce thermal shock	• Programmable target temperature and hold time • Reliable uniform temperature inside furnace	• Thermo Scientific, Norwood, MA (Lindberg/Blue M)
Motorized linear translation stage	• Control glass probes to pull out gradually from glass etchant • Taper glass probes	• Computer-controlled (via LabView v.7) translation stage • Minimum step size is 1.57 μm	• Newport Corp., Irvine, CA
Organic layer	• Modify the contact angle at the glass–etchant interface • Prevent vaporizing toxic HF-based etchant solution	• Vegetable oil • Paraffin oil	• Fisher Scientific
Metal targets	• Evaporation targets	• 99.99% pure • Titanium (Ti), chromium (Cr), Au, Pt	• Sigma-Aldrich
E-beam evaporator	• Metal deposition	• Evaporation Ti, Cr, Au, Pt	• Airco Temescal (FC-1800)
PCB	• Establish electric connections with individual sensors • Carriers for easier handling	• Copper-clad laminate glass epoxy • Laminate (790 μm thick) • Copper (35 μm thick layer) and dry film negative photoresist (33 μm thick layer) • Photolithographically pattern and etch in ferric chloride to define designed traces on the PCB surface	• D&L Products, Inc.
UV-cured epoxy	• Fix MEs on PCB carriers	• Epoxy bond	• Loctite, Rocky Hill, CT (3301)
Conductive silver epoxy	• Conductive filling material for electroplating Au for DO microsensors	• Conductive silver-based low curing temperature (120 °C)	• Emerson & Cuming, Billerica, MA (Ablebond 8700E)
Parylene C	• Insulate individual MEs	• A well-known biocompatible polymer material • Inert and optically transparent material with the pinhole-free surface	• Specialty Coating Systems, Indianapolis, IN

TABLE 7.4 (continued) Materials and Devices Necessary for Needle-Type Environmental MEMS Sensor Array Fabrication

Item	Use	Description	Sources
Parylene coater	• Deposit Parylene C easily and uniformly on any substrate	• Simple vapor deposition	• Specialty Coating Systems (PDS 2010 Parylene Labcoter)
Micropipette beveller	• Bevel ME tips	• Beveler, microscope, micromanipulator, and light source	• Sutter Instrument (BV-10)
Potentiostat	• Electroplate metal layers • Supply the polarization voltage and current	• Galvanostatic mode • Control current and voltage	• Gamry, Inc. (Reference 600) • Diamond General Development Corp. (No 1231)—Chemical Microsensors II potentiostat
pH/mV meter	• Measure pH and potential	• Data acquisition system • Integrate with BalanceTalk software (Labtronics, Inc.)	• Denver Instrument (model 225)
IrOx	• Use solid-state working electrode for pH sensor	• Electroplate IrOx film	• Sigma-Aldrich
IrOx electrolyte solution	• Pulling glass micropipette	• Mixture of iridium tetrachloride ($IrCl_3$), hydrogen peroxide (H_2O_2), oxalic acid ($C_2H_2O_4$), and potassium carbonate (K_2CO_3) in DI water	• Sutter Instrument (P-1000)
Cobalt (Co) electrolyte solution	• Grinding ME tips	• Mixtures of 33 g $CoSO_4$ and 3 g H_3BO_3 in 100 mL of DI water	• Sigma-Aldrich
Piranha etchant	• Clean glass surface	• H_2SO_4 and H_2O_2 in a 7:3 (v/v) ratio	• Fisher Scientific
HF-based etchant solution	• Glass etchant • Smooth the diced surface • Reduce the probe dimensions	• HF, HNO_3, and H_2O in a 10:7:33 (v/v/v) ratio	• Fisher Scientific
Au etchant	• Etch Au	• Mixtures of I_2, KI, and H_2O (1:4:40 (m/m/v)) (1:4:40 (m/m/v))	• Fisher Scientific
Ag/AgCl reference electrode	• Construction of Ag/AgCl	• Commercial Ag/AgCl reference ME	• Microelectrodes Inc. (MI-401)
Thick photoresist buffer	• Buffer for beveling and plasma ashing	• 5–8 μm	• Clariant Corporation (AZ 4562)
Plasma oxygen ashing	• Etching tip of electrodes	• 50 μm	• March CS-1701 RIE

and is used to measure the oxidized and reduced forms of all the chemical species (nonselective) in the solution. The working electrode just needs to have its exposed conductive area in contact with a solution without the need for an ion-selective membrane, as shown in Figure 7.5a [35,42]. ORP electrodes directly measure the potential of a solution between a metal electrode and a standard reference electrode, given by the Nernst equation (Equation 7.2).

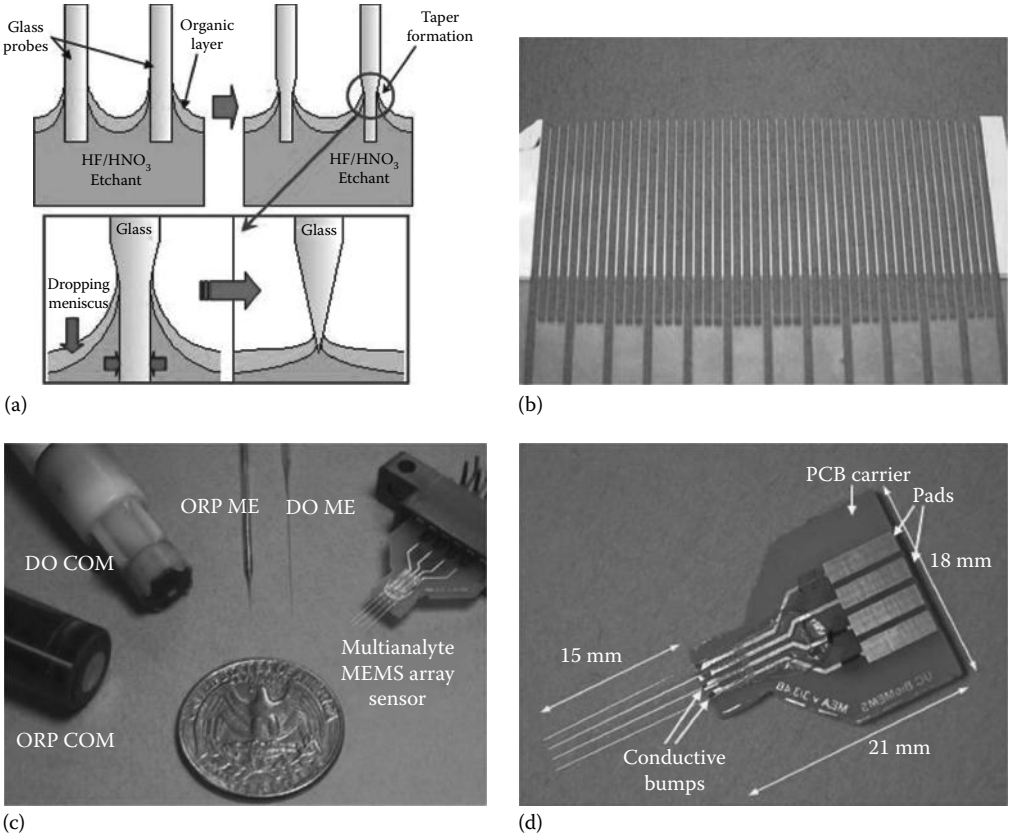


FIGURE 7.6 (a) Schematic diagram of fabrication of sharp MEMS sensor array tips using meniscus etching (meniscus height falls dynamically to form a taper with geometry dependent on the initial meniscus height) [27], (b) a metalized cover slip glass with sharpened probe tips, (c) the fabricated multianalyte MEMS sensor arrays packaged on a PCB carrier compared with COMs and ME sensors, and (d) integrated multianalyte MEMS sensor arrays with ORP, DO, and phosphate sensor. (From Lee, W.H. et al., *Meas. Sci. Technol.*, 22, 042001, 2011.)

For phosphate MEA sensors, cobalt (Co) is electroplated to form a Co film as an ion-selective membrane filling the empty recessed area, as depicted in Figure 7.5b [38]. Two Co plates are used as anodes in the prepared Co electroplating electrolyte solution. Co is electrodeposited to a ~0.2 μm thickness on the Au tips. The MEMS phosphate sensor measures oxidation corrosion potential (mV) on the cobalt oxide (CoO) surface, in conjunction with a reference electrode.

The MEMS pH sensor uses an electroplated iridium oxide (IrOx) film as a working electrode in place of an ion-selective membrane [42]. To achieve a more durable sensor for penetrating into a sample, the IrOx film is used in a solid-state electrode without the attachment of an ion-selective membrane. The fabrication process for the needle-type pH sensor is composed of six major steps: dicing, etching, metallization, IrOx electroplating, packaging, and sensor tip formation. The first three steps are the same as the batch fabrication processes previously described in Section 7.3. The final three steps to form a solid-state electrode for the pH MEMS sensor are discussed in detail later.

7.8.1 IrOx Electroplating

After the Ti/Au conductive layer deposition, IrOx is electroplated on the entire Au seed layer. Dark blue IrOx films are obtained with a thickness of 100–150 nm. Fresh IrOx-coated sensors are stabilized in pH 7 buffer for 2 days to reduce potential drift during measurements.

7.8.2 Packaging for Electric Connection

ME arrays are packaged on a PCB, followed by IrOx electrodeposition. Parylene C is coated over the entire substrate to insulate MEs.

7.8.3 Sensor Tip Formation

Two steps are used to expose a ~50 μm long IrOx sensing tip. The first step is to cover the entire sensing tip using photoresist, and then ~50 μm long sensing tips are stripped by acetone. Secondly, the plasma oxygen process is used to etch Parylene C on the exposed sensor tip end. The schematic diagram of the MEMS pH sensor is illustrated in Figure 7.5c.

7.9 Multianalyte MEMS Sensor Arrays

The individual needle-type MEA sensor for measuring ORP, DO, and phosphate, respectively, can be integrated into a single multianalyte sensor array as shown in Figure 7.6d, exhibiting higher sensitivity, faster response time, and higher stability with a smaller tip size than the conventional sensors. The major advantages of these new MEMS sensors include the ability to penetrate samples to perform measurements, the small tip size for in situ measurements, array structure for higher robustness, and the possibility of multianalyte detection. The sensors demonstrate the ability to monitor local concentration changes in small structures with a high spatial resolution and offer the versatility of the ME technique as well as the capability for repetitive measurements.

The integration process was aimed at easier and simpler fabrications for mass production with a minimum number of steps. Following the metallization step, the integration process begins with the application of Parylene C insulation to the ME tips. All four MEs in an array were beveled at the same time to expose glass and Ti/Au layers. The recess etching process was used to form the recessed structure. Since Au exposed to the solution enables ORP measurements, the recessed structure can measure not only ORP but also DO. For phosphate sensors, Co was electrodeposited on the exposed Au tips. Ultimately, the multianalyte MEMS sensor array will enable in situ measurements in a wide variety of small sample applications in environmental engineering and life sciences.

7.10 Data Acquisition and Processing

MEMS sensors can use the same general experimental setup for microprofile measurement described in the Section 7.2. However, the ability to perform an in situ measurement outside of the controlled laboratory environment of a well-grounded Faraday cage is based on the integration of the MEs with a complementary metal oxide semiconductor (CMOS)-based integrated circuit (IC) for noise cancellation.

The CMOS chip has been developed to perform both signal acquisition and processing [35] and then integrated with the MEMS sensor. The circuit for potentiometric measurements with MEMS sensors is based on the CMOS ASIC design. The circuit block is divided into three stages: a low pass filter, an isolation amplifier, and an instrumentation amplifier. The built-in noise cancellation design eliminates both environmental and instrumental noise, permitting measurement of voltages in the 10^{-3}–10^{-6} V range outside the Faraday cage.

7.11 Application for Environmental Processes

Figure 7.7 shows the areas where microsensors are or could be applied in engineered and natural water systems. Biofilm control issues in drinking water have been one of the high-priority research areas. A NH_2Cl microsensor has been used to evaluate biofilm control strategies and nitrification in

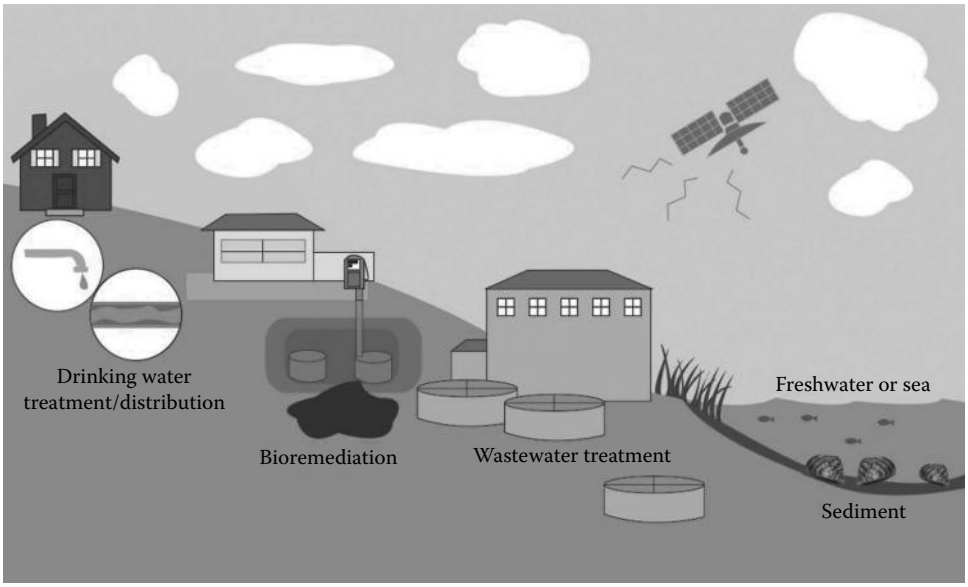


FIGURE 7.7 Various needle-type environmental microsensor applications for engineered and natural water systems. (Graphic created by Emily Nauman.)

chloraminated drinking water distribution systems, by combining with environmental microbiological approaches (e.g., confocal laser scanning microscopy [CLSM], cryo-cross sectioning, fluorescent in situ hybridization [FISH]) and other microsensors (e.g., oxygen, free chlorine, ammonium, pH, and nitrate) [25,26]. Microsensor applications can extend to the characterization and evaluation of corrosion and biocorrosion (e.g., localized corrosion and electrochemical activity distribution on pipe surfaces in real time with high spatial resolution) during disinfection in water infrastructure (e.g., ductile/cast iron pipe, copper, lead, and concrete pipe materials) and also provide significant impact on other industrial water infrastructures (e.g., ship hulls, ballast water disinfection, and cooling water systems), which experience corrosion and biocorrosion. One example of future applications is an envisioned field-deployable version of real-time monitoring of chloramine species in drinking water distribution systems (including premise plumbing).

In the bioremediation processes, it is important to investigate the dynamics of soil biofilm structure and function, related kinetics, and reaction mechanisms for better understanding of bioremediation processes such as degradation of hydrophobic petroleum compounds [51]. The appropriate modeling kinetic parameters can be determined using microsensors for better biological processes.

In wastewater treatment plants, environmental microsensors have been used for decades to understand the mechanics of relative biological processes and/or lead to the improvement of modeling, design, and operation of existing biological water/wastewater treatment processes including membrane processes for nutrient removal [3,8,13]. Microsensor applications for membrane biofilm processes [52] and microbial fuel cells (MFCs) [28,53] have also increased to improve engineering processes that produce renewable energy while reducing the energy requirements for a treatment plant. In situ monitoring using microsensors also prevents possible contamination of natural systems and/or treatment systems, and thus we can readily and continuously monitor water quality and prevent an impending event such as toxicity breakthrough or operational malfunction.

Microsensors have also been used in the fields of ecology and environmental science, such as to study the microenvironments in lake (or marine) sediments [54], with zebra mussel issues (e.g., nitrification in shell biofilms) [55], with invertebrates (e.g., denitrification in the gut) [56], with rhizosphere soil [57], and with roots of submerged plants [58] where important constituents' transport

(e.g., oxygen, phosphate) or nitrogen cycling occurs. The sensor techniques can be extended to in-field integrated water quality monitoring systems (e.g., monitoring of global climate change and eutrophication control) over a large area, providing real-time water quality data by combining with geographic information system (GIS) technology.

7.12 Outlook of Needle-Type Environmental Sensors

Needle-type environmental microsensors have been used to monitor environmental conditions and contaminant profiles in situ in various environmental samples (e.g., biofilm, pore water, and sediments). Their use for protection of human health and the environment has increased with the advanced fabrication techniques and improvements for stable data acquisition. In this chapter, we focused on needle-type microsensors for environmental monitoring, but they can also be applied to industrial and biomedical areas. Consequently, the development and application of microsensors have a great potential for collaboration with a wide range of disciplines. Potential research areas for new environmental microsensors would be hydroxylamine (NH_2OH), polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs), and heavy metal (e.g., lead, mercury, and arsenic) monitoring sensors. For example, NH_2OH is a key intermediate during the process of the nitrogen cycles (e.g., ammonia oxidizing bacteria) and the production of N_2O, a major greenhouse gas. Therefore, a sensor for quantitative determination of NH_2OH will be very important, both in studies of biological processes and for industrial purposes. Antifouling methods (e.g., functionalized and controlled surface design and biomimetics) are also challenge areas to prevent biofouling during in situ monitoring and will provide more reliable sensor surface technology. Furthermore, there is a critical need for highly sensitive and selective low-power chemical sensors for a broad range of applications from environmental monitoring, food safety, health diagnostics, as well as global security and global warming. Although many sensitive transducer platforms have been developed, including microsensors, the long-standing fundamental problem of sufficient sensitivity (sub-ppm) and selectivity has not yet been solved. The key component for accomplishing ultrahigh sensitive sensors is an innovative improvement of sensor technology. Nanoelectromechanical systems (NEMS) technologies can permit the integration of different mechanical elements, sensors, and electronics in a small device and therefore can be suitable for the development of multianalyte sensors. The current roadblock to the implementation of such ultrahigh sensitive multianalyte sensors is the lack of nanotransducer platforms, which can be integrated with various reliable membrane materials.

Acknowledgments

The authors gratefully acknowledge financial support of this work by grants from the National Science Foundation (BES-0228603, BES-0529217), the National Institute of Environmental Health Sciences (P42ES04908-16/Project 5), the US Environmental Protection Agency (EP-C-05-056), and the University of Cincinnati Institute for Nanoscale Science and Technology.

References

1. E. M. Hermann Eberl, C. P. Daniel Noguera, M. v. L. Bruce Rittmann, and A. O. Wanner, *Mathematical Modeling of Biofilms*. London, U.K.: IWA Publishing, 2006.
2. K. Andersen, T. Kjær, and N. P. Revsbech, An oxygen insensitive microsensor for nitrous oxide, *Sensors and Actuators B: Chemical*, 81, 42–48, 2001.
3. T. Yu and P. L. Bishop, Stratification and oxidation-reduction potential change in an aerobic and sulfate-reducing biofilm studied using microelectrodes, *Water Environment Research*, 73, 368–373, 2001.
4. T. C. Zhang and P. L. Bishop, Experimental-determination of the dissolved-oxygen boundary-layer and mass-transfer resistance near the fluid-biofilm interface, *Water Science and Technology*, 30, 47–58, 1994.

5. N. P. Revsbech, An oxygen microsensor with a guard cathode, *Limnology and Oceanography*, 474–478, 1989.
6. Z. Lewandowski and H. Beyenal, *Fundamentals of Biofilm Research*. New York: CRC Press, 2007.
7. P. VanHoudt and Z. Lewandowski, Iridium oxide pH microelectrode, *Biotechnology and Bioengineering*, 40, 601–608, 1992.
8. P. L. Bishop and Y. Tong, A microelectrode study of redox potential change in biofilms, *Water Science and Technology*, 39, 179–185, 1999.
9. J. Li and P. L. Bishop, *In situ* identification of azo dye inhibition effects on nitrifying biofilms using microelectrodes, *Water Science and Technology*, 46, 207–214, 2002.
10. W. H. Lee, D. G. Wahman, P. L. Bishop, and J. G. Pressman, Carbon-fibre nitrite microsensor for in situ biofilm monitoring, in *Proceedings of the World Water Congress & Exhibition*, *International Water Association (IWA)*, Busan, South Korea, September 16–21, 2012.
11. S. Silva, C. Alves, S. Muchado, L. Mazo, and L. Avaca, Electrochemical determination of nitrites in natural waters with ultramicroelectrodes, *Electroanalysis*, 8, 1055–1059, 1996.
12. B. Piela and P. K. Wrona, Oxidation of nitrites on solid electrodes, *Journal of the Electrochemical Society*, 149, E55–E63, 2002.
13. F. Schreiber, L. Polerecky, and D. de Beer, Nitric oxide microsensor for high spatial resolution measurements in biofilms and sediments, *Analytical Chemistry*, 80, 1152–1158, 2008.
14. W. H. Lee and P. L. Bishop, In situ microscale analyses of activated sludge flocs in the enhanced biological phosphate removal process by the use of microelectrodes and fluorescent in situ hybridization, *Journal of Environmental Engineering*, 136, 561–567, 2010.
15. W. H. Lee, Y. Seo, and P. L. Bishop, Characteristics of a cobalt-based phosphate microelectrode for in situ monitoring of phosphate and its biological application, *Sensors and Actuators B: Chemical*, 137, 121–128, 2009.
16. J. H. Lee, W. H. Lee, P. L. Bishop, and I. Papautsky, A cobalt-coated needle-type microelectrode array sensor for in situ monitoring of phosphate, *Journal of Micromechanics and Microengineering*, 19, p. 025022, 2009.
17. D. De Beer, R. Srinivasan, and P. S. Stewart, Direct measurement of chlorine penetration into biofilms during disinfection, *Applied and Environmental Microbiology*, 60, 4339–4344, 1994.
18. W. H. Lee, J. G. Pressman, D. G. Wahman, and P. L. Bishop, Characterization and application of a chlorine microelectrode for measuring monochloramine within a biofilm, *Sensors and Actuators B: Chemical*, 145, 734–742, 2010.
19. P. Jeroschewski, C. Steuckart, and M. Kühl, An amperometric microsensor for the determination of H_2S in aquatic environments, *Analytical Chemistry*, 68, 4351–4357, 1996.
20. D. De Beer, A. Glud, E. Epping, and M. Kuhl, A fast-responding CO2 microelectrode for profiling sediments, microbial mats, and biofilms, *Limnology and Oceanography*, 42, 1590–1600, 1997.
21. G. Papeschi, S. Pinzauti, P. Gratteri, and M. Larini, Copper (II) ion-selective sensor with electrolytically plated chalcogenide coating, *Sensors and Actuators B: Chemical*, 7, 544–548, 1992.
22. Z. Szigeti, I. Bitter, K. Toth, C. Latkoczy, D. J. Fliegel, D. Gunther, and E. Pretsch, A novel polymeric membrane electrode for the potentiometric analysis of Cu_2^+ in drinking water, *Analytica Chimica Acta*, 532, 129–136, 2005.
23. P. Jin, The impact of cadmium on a multi-species biofilm degrading naphthalene and the role of hydrogen peroxide in cadmium-biofilm interaction, PhD dissertation, Environmental Engineering, University of Cincinnati, Cincinnati, OH, 2007.
24. N. P. Revsbech and B. B. Jørgensen, Microelectrodes: Their use in microbial ecology, in K. C. Marshall, Ed., *Advanced in Microbial Ecology*, Vol. 9, New York: Plenum Press, 1986, pp. 293–352.
25. W. H. Lee, D. G. Wahman, P. L. Bishop, and J. G. Pressman, Free chlorine and monochloramine application to nitrifying biofilm: Comparison of biofilm penetration, activity, and viability, *Environmental Science and Technology*, 45, 1412–1419, 2011.

26. J. G. Pressman, W. H. Lee, P. L. Bishop, and D. G. Wahman, Effect of free ammonia concentration on monochloramine penetration within a nitrifying biofilm and its effect on activity, viability, and recovery, *Water Research*, 46, 882–894, 2012.
27. W. H. Lee, J. H. Lee, W. H. Choi, A. A. Hosni, I. Papautsky, and P. L. Bishop, Needle-type environmental microsensors: Design, construction and uses of microelectrodes and multi-analyte MEMS sensor arrays, *Measurement Science and Technology*, 22, p. 042001, 2011.
28. J. T. Babauta, H. D. Nguyen, and H. Beyenal, Redox and pH microenvironments within Shewanella oneidensis MR-1 biofilms reveal electron transfer mechanism, *Environmental Science and Technology*, 45, 6654–6660, 2011.
29. R. M. Wightman, Probing cellular chemistry in biological systems with microelectrodes, *Science*, 311, 1570–1574, 2006.
30. C. G. Zoski, Ultramicroelectrodes: Design, fabrication, and characterization, *Electroanalysis*, 14, 1041–1051, 2002.
31. R. A. Linsenmeier and C. M. Yancey, Improved fabrication of double-barreled recessed cathode O2 microelectrodes, *Journal of Applied Physiology*, 63, 2554–2557, 1987.
32. P. R. Warburton, R. S. Sawtelle, A. Watson, and A. Q. Wang, Failure prediction for a galvanic oxygen sensor, *Sensors and Actuators B: Chemical*, 72, 197–203, 2001.
33. H. Pang and T. C. Zhang, Fabrication of redox potential microelectrodes for studies in vegetated soils or biofilm systems, *Environmental Science and Technology*, 32, 3646–3652, 1998.
34. J. H. Lee, A. Jang, P. R. Bhadri, R. R. Myers, W. Timmons, F. R. Beyette, P. L. Bishop, and I. Papautsky, Fabrication of microelectrode arrays for in situ sensing of oxidation reduction potentials, *Sensors and Actuators B: Chemical*, 115, 220–226, 2006.
35. A. Jang, J. H. Lee, P. R. Bhadri, S. A. Kumar, W. Timmons, F. R. Beyette, I. Papautsky, and P. L. Bishop, Miniaturized redox potential probe for in situ environmental monitoring, *Environmental Science and Technology*, 39, 6191–6197, 2005.
36. J. H. Lee, T. S. Lim, Y. Seo, P. L. Bishop, and I. Papautsky, Needle-type dissolved oxygen microelectrode array sensors for in situ measurements, *Sensors and Actuators B: Chemical*, 128, 179–185, 2007.
37. J. H. Lee, Y. Seo, T. S. Lim, P. L. Bishop, and I. Papautsky, MEMS needle-type sensor array for in situ measurements of dissolved oxygen and redox potential, *Environmental Science and Technology*, 41, 7857–7863, 2007.
38. W. H. Lee, J. H. Lee, P. L. Bishop, and I. Papautsky, Biological application of MEMS microelectrode array sensors for direct measurement of phosphate in the enhanced biological phosphorous removal process, *Water Environment Research* 81, p. 748, 2009.
39. P. Srinivasan, F. R. Beyette, and I. Papautsky, Micromachined arrays of cantilevered glass probes, *Applied Optics*, 43, 776–782, Feb 2004.
40. J. H. Lee, Y. Seo, W. H. Lee, P. Bishop, and I. Papautsky, Needle-Type Multi-Analyte MEMS Sensor Arrays for in situ Measurements in Biofilms, in *Emerging Environmental Technologies*, Vol. II. Springer, New York, 2010, 115–145.
41. W. H. Choi and I. Papautsky, Fabrication of a needle-type pH sensor by selective electrodeposition, *Journal of Micro/Nanolithography, MEMS, and MOEMS*, 10, 020501, 2011.
42. W. H. Choi, J. R. Shann, and I. Papautsky, Multi-analyte needle-type sensor for measurement of pH, phosphate, and redox potential in soil, in *Proceedings of the 2010 IEEE Sensors*, Kona, HI, 2010, pp. 931–935.
43. W. H. Choi, W. H. Lee, P. L. Bishop, and I. Papautsky, Integrated dissolved oxygen-compensated needle-type sensor for phosphate analysis of biological and environmental samples, in *Proceedings of the International Conference on Chemical and Biochemical Analysis Systems (μTAS)*, Jeju, Korea, November 1–5, 2009.
44. C. H. Ahn, J. W. Choi, G. Beaucage, J. H. Nevin, J. B. Lee, A. Puntambekar, and J. Y. Lee, Disposable Smart lab on a chip for point-of-care clinical diagnostics, in *Proceedings of the IEEE*, 2004, pp. 154–173.

45. E. Verpoorte, Microfluidic chips for clinical and forensic analysis, *Electrophoresis*, 23, 677–712, 2002.
46. L. J. Kricka, Microchips, microarrays, biochips and nanochips: Personal laboratories for the 21st century, *Clinica Chimica Acta*, 307, 219–223, 2001.
47. T. A. Fofonoff, S. M. Martel, N. G. Hatsopoulos, J. P. Donoghue, and I. W. Hunter, Microelectrode array fabrication by electrical discharge machining and chemical etching, *IEEE Transactions on Biomedical Engineering*, 51, 890–895, 2004.
48. P. S. Motta and J. W. Judy, Multielectrode microprobes for deep-brain stimulation fabricated with a customizable 3-D electroplating process, *IEEE Transactions on Biomedical Engineering*, 52, 923–933, 2005.
49. S. Takeuchi, T. Suzuki, K. Mabuchi, and H. Fujita, 3D flexible multichannel neural probe array, *Journal of Micromechanics and Microengineering*, 14, 104–107, 2004.
50. K. E. Jones, P. K. Campbell, and R. A. Normann, A glass/silicone composite intracortical electrode array, *Annals of Biomedical Engineering*, 20, 423–437, 1992.
51. S. Grötzschel, J. Köster, R. M. M. Abed, and D. de Beer, Degradation of petroleum model compounds immobilized on clay by a hypersaline microbial mat, *Biodegradation*, 13, 273–283, 2002.
52. L. Downing and R. Nerenberg, Effect of bulk liquid BOD concentration on activity and microbial community structure of a nitrifying, membrane-aerated biofilm, *Applied Microbiology and Biotechnology*, 81, 153–162, 2008.
53. C. Butler and R. Nerenberg, Performance and microbial ecology of air-cathode microbial fuel cells with layered electrode assemblies, *Applied Microbiology and Biotechnology*, 86, 1399–1408, 2010.
54. S. P. C. Tankéré, D. G. Bourne, F. L. L. Muller, and V. Torsvik, Microenvironments and microbial community structure in sediments, *Environmental Microbiology*, 4, 97–105, 2002.
55. I. M. H. Nanna B. Svenningsena, M. Sigby-Clausena, L. H. Larsena, L. P. Nielsena, P. Stiefb, and A. Schramma, Shell biofilm nitrification and gut denitrification contribute to emission of nitrous oxide by the invasive freshwater mussel dreissena polymorpha (Zebra Mussel), *Applied and Environmental Microbiology*, 78, 4505–4509, 2012.
56. A. Schramm, Microsensors for the study of microenvironments and processes in the intestine of invertebrates, in H. Konig and Varma, A., Eds., *Soil Biology*, Vol. 6, Berlin, Germany: Springer-Verlag, 2006, pp. 463–473.
57. T. C. Zhang and H. Pang, Applications of microelectrode techniques to measure pH and oxidation–reduction potential in rhizosphere soil, *Environmental Science and Technology*, 33, 1293–1299, 1999.
58. P. B. Christensen, N. P. Revsbech, and K. Sand-Jensen, Microsensor analysis of oxygen in the rhizosphere of the aquatic macrophyte littorella uniflora (L.) ascherson, *Plant Physiology*, 105, 847–852, 1994.

8

Single-Chip Chemical and Biological Sensors

Alireza Hassanzadeh
Shahid Beheshti University, Tehran, Iran

Robert G. Lindquist
The University of Alabama in Huntsville

8.1 Introduction

Detection of very small amount of chemical and biological agents had been an active area of research for many years. From carbon monoxide detection for environmental applications to biological and biomedical monitoring, missile fuel leakage detection and security applications require high-sensitivity sensors that can measure low parts per billion (ppb) or part per million (ppm) of chemical or biological agents. Recent development of microelectronic technology has provided a good opportunity for fabrication of sensor and the interface circuit on a single chip. Single-chip fabrication of the sensor and the interface circuit provides higher-resolution, better accuracy, and lower-noise signal conditioning and amplification. In addition, better control over the sensor operating condition such as temperature and linearity can be provided. When the sensor is close to the interface circuit, higher sensitivity and lower noise can be achieved by omitting long wire capacitance and electromagnetic interface. The array of sensors can be built on a single chip for minimizing false alarms and selectivity of the sensor.

Biosensors that include transducers based on integrated circuit are often referred as biochips. A biochip normally consists of array of biosensors that can be monitored individually and can be used for the analysis of multiple analyte. The interaction between the analyte and the bioreceptor is designed to produce a measurable effect. A transducer converts the measurable effect into a signal (usually electrical) that can be monitored and recorded. The most common biosensors include antibody/antigen interaction, nucleic acid interactions, enzymatic interactions, cellular interaction (microorganism, proteins), and interaction using biomimetic materials (synthetic bioreceptor). For transducers, conventional methods are optical measurement (luminescence, absorption, surface plasmon), electrochemical, and mass-sensitive measurement (surface acoustic waves, microbalance) [1]. In the next section, some of the reported methods used for on-chip transducers will be reviewed. The principle of operation of single-chip chemical and biological sensors will be introduced, and liquid crystal (LC) chemical and biological sensors as a special case will be discussed in detail.

8.2 Transducers

Different transducers have been reported for gas sensors; mass-sensitive, capacitive, calorimetric, and electrochemical nanowire (NW) FETs have been reported for on-chip implementation [1–7]. Many of the sensors are based on a change in a polymer material due to exposure to a chemical or biological agent. The physical property of the polymer changes due to the analyte, and these changes can be detected using a microtransducer fabricated on the chip. The transducer shown in Figure 8.1 responds to the mass change of a cantilever beam due to the presence of a chemical agent. Resonating micromachined cantilever is a promising mass-sensitive chemical sensor especially when fabricated with the interface circuit on a single chip. The absorption of the agents in the chemically sensitive polymer causes a shift in the resonance frequency of the cantilever transducer. The change of the frequency of the cantilever or plate can be detected in a bridge circuit and simply measured using a frequency counter.

The transducer in Figure 8.2 is sensitive to change of heat when absorption happens in the polymer due to the presence of agents. The thermoelectric calorimeter is based on Seebeck effect. This transducer detects the heat when a change of enthalpy happens on absorption (condensation heat) or desorption (vaporization heat) due to the presence of chemical agents. An array of 256 polysilicon/aluminum thermocouples connected in series has been used in [2] to record temperature change in milli-Kelvin range. The calorimetric transducer only detects transitions and change of concentration of analyte in contrast to microcantilever and capacitive type.

In the Figure 8.3 transducer, the change in the dielectric of a material is detected using a capacitive transducer. The capacitive sensor relies on interdigitated electrodes, which corresponds to a basic two-plate capacitor that monitors dielectric change of a polymer sensitive to a chemical agent. The capacitor electrodes can be made of metal layers available on CMOS IC processes. For increasing transducer sensitivity, the number of electrodes per volume of the dielectric can be increased. Electroplating can also

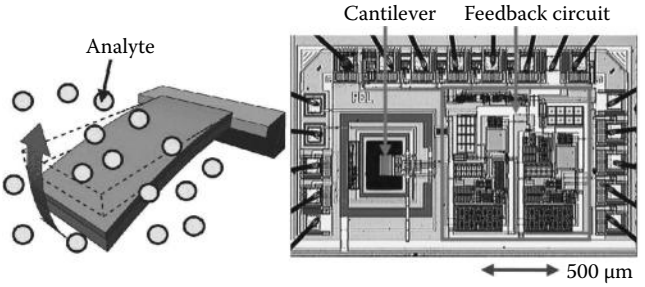


FIGURE 8.1 Cantilever mass-sensitive chemical transducer and the microchip. (From Hierlemann, A. and Baltes, H., *R. Soc. Chem*., 128, 15, 2003.)

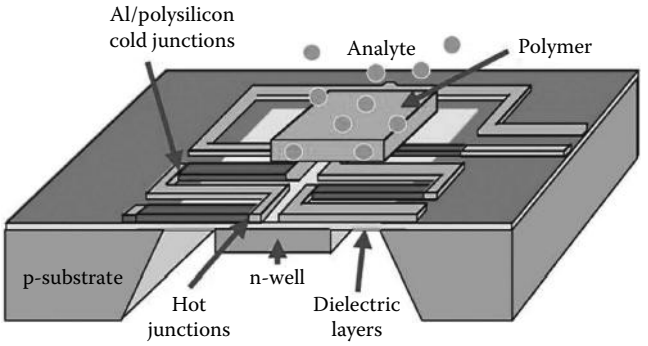


FIGURE 8.2 Calorimetric transducer. (From Hierlemann, A. and Baltes, H., *R. Soc. Chem*., 128, 15, 2003.)

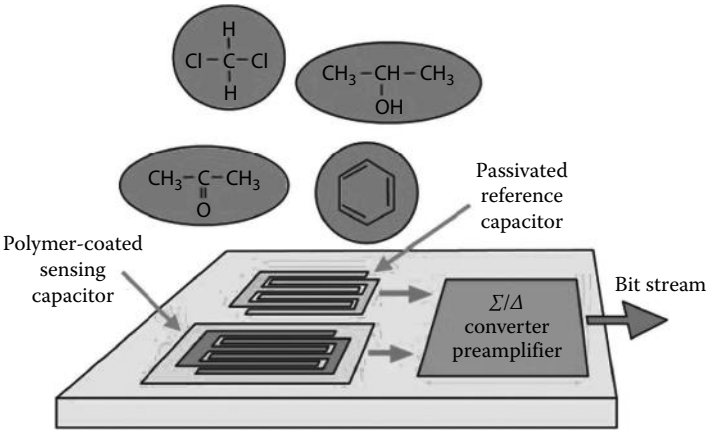


FIGURE 8.3 Capacitive chemical transducer. (From Hierlemann, A. and Baltes, H., *R. Soc. Chem.*, 128, 15, 2003.)

be used to elevate electrodes and improve capacitive sensor sensitivity. More details about interdigitated capacitor (IDC) transducer are provided in the LC section.

NWs have also been reported as a compatible technology for the detection of chemical species at low concentration of protein and viruses. Detectors based on NWs are field effect transistors (FETs) that show conductivity change when certain chemicals are attached to the NWs. In an FET, the current in the transistor is dependent on the electric field in the channel. Semiconductor NWs can function as FET devices, and silicon NWs are one of the best characterized that can be built as small as 2–3 nm for detecting applications. The NW acts as the gate of the transistor and the chemicals attaching to it having positive (or negative) net charge affect conduction in the channel. Macromolecules such as proteins have a net positive or negative charge that when attached to the NW will change the surface potential resulting in a change in the FET conductivity. Microfluidic channels can be used to carry the solution to the specific location on the chip area. Figure 8.4 shows fabricated NWs used for chemical and biological detection in this case single virus detection by the Harvard University researchers [4].

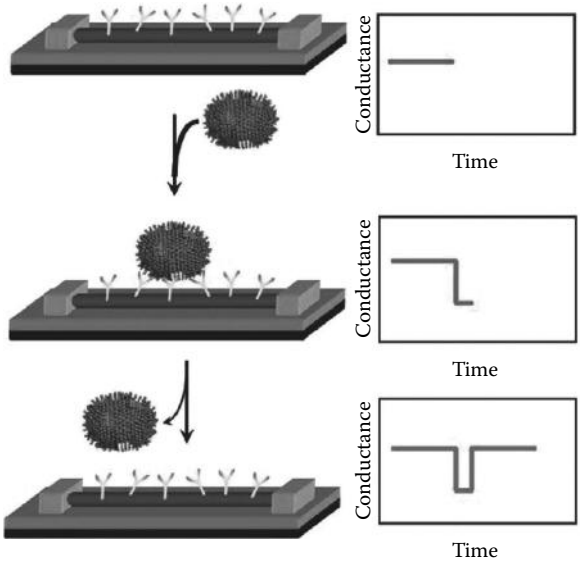


FIGURE 8.4 Single virus detection using NWs. (From Patolsky, F. et al., *Anal. Chem.*, 78, 4260, 2006.)

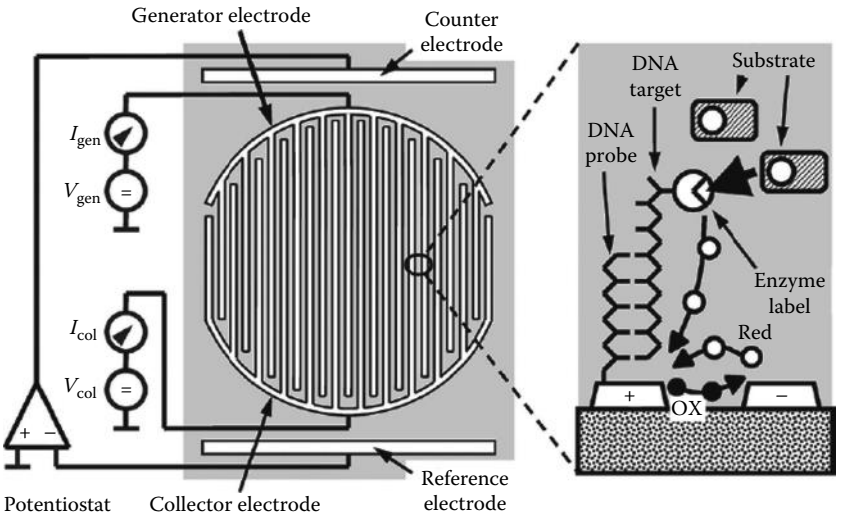


FIGURE 8.5 Interdigitated electrodes used in DNA detection in a sensor array. (From Schienle, M. et al., *IEEE J. Solid State Circuits*, 39(12), 2438, 2004.)

Similar techniques have been used such as extended gate FET that works on the same principle for the detection of chemicals [5]. Similarly, carbon nanotubes (CNTs) have also been fabricated on a chip for chemical and biological detection as reported in [6].

A DNA sensor microchip has been reported in [7] that is based on interdigitated gold electrodes with immobilized single-stranded DNA receptor molecules on them. When certain DNA attaches to the already existing DNAs, it will cause a change of the current in the sensor electrode detected and measured by an on-chip A/D converter. The method is called redox-cycling detection method for DNA detection. Figure 8.5 shows the principle of operation of the sensor as used in the 128 sensor array on a single chip.

In the next section, LC chemical and biological sensors will be discussed in detail.

8.3 Liquid Crystal Chemical and Biological Sensors

LC systems have proved to be an excellent candidate for reliable and sensitive detection for chemical and biological agents. They have been used to detect mustard gas, pesticides, and various viruses (West Nile, influenza, etc.) and have been used to amplify protein-binding events at an interface. A biological agent causes a change in the orientation of LC molecules; small changes at the nanoscale are amplified to micron level changes in the LC film. Therefore, the detection of very low level of chemical and biological agents is possible [8–15].

Recently, a novel technique developed by Dr. Lindquist and his colleagues in NMDC (Nano and Micro Device Center) at UAHuntsville uses a capacitive detection technique instead of optical inspection of LC alignment change [16,17]. This method brings a lot of worthy advantages over the conventional optical detection method. First, it provides greater insight into the fundamental distortion occurring in LC film by tracking the deformation, rather than just simply sensing the distortion by the optical method. Second, the removal of the optical transparency requirement of the substrate allows for a broader class of substrate selection. For instance, one of the substrates could be silicon to allow for the integration of electronics, and the other could be a porous ceramic to help facilitate the movement of chemical or biological agents to the sensing interface. The integration of the capacitive sensing method with signal processing on a chip leads to higher sensitivity and minimization of false positives.

8.3.1 Liquid Crystal Material

LC is a phase of material between crystalline solid and isotropic liquids. An Austrian botanist named Friedrich Reinitzer in 1888 discovered LC material. He found that an organic substance has two melting points; at the first melting point, it changes from solid to a cloudy liquid crystalline, and at the second point, it changes from LC to a clear liquid. Molecules of these kinds of materials are rod-like (calamitic) or disc-like (discotic) and typically are in the range of few nanometers (nm). The ratio between the length and diameter in rodlike or diameter and thickness for disklike materials is usually about 5 or greater [18,19]. Figure 8.6 shows two typical LC molecules, rodlike and disklike.

In addition to positional order, these molecules also may possess orientation order in the crystal. Figure 8.7 shows different states of an LC material, from solid to isotropic liquid. Different LCs may pass through different states before they reach isotropic liquid.

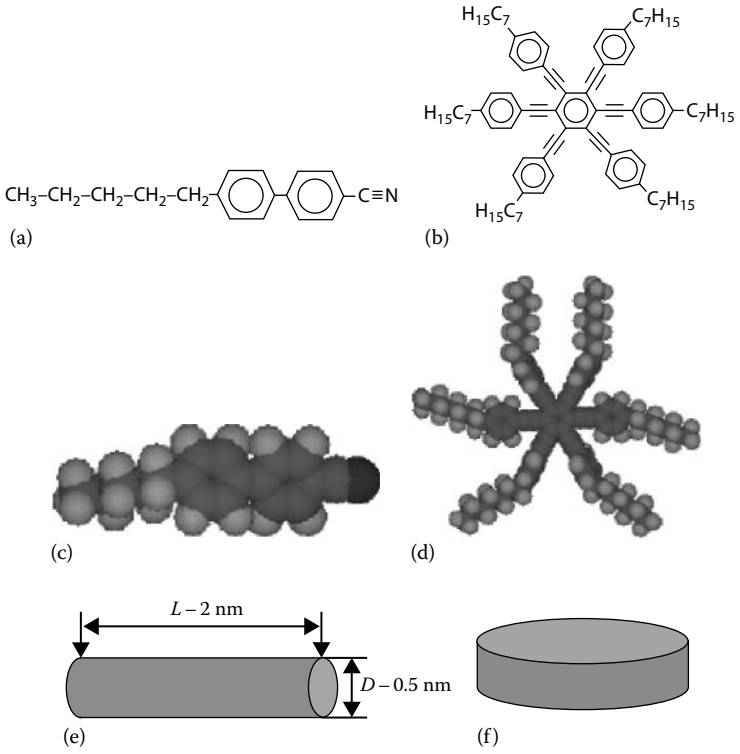


FIGURE 8.6 Rod-like and disk-like LC molecules: (a) Chemical structure, (c) space-filling model, and (e) physical model. Discostic liquid crystal: (b) chemical structure, (d) space-filling model, and (f) physical model. (From Wu, S.-T. and Yang, D.-K., *Fundamentals of Liquid Crystal Devices*, John Wiley & Sons, New York, 2006.)

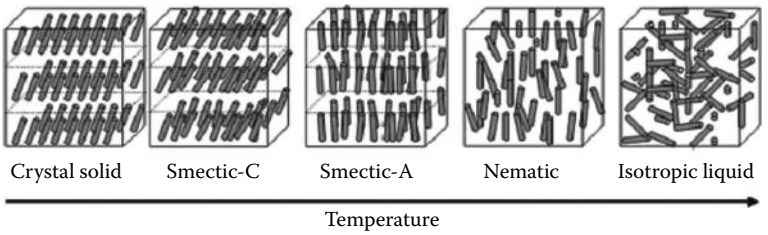


FIGURE 8.7 Phase change of rodlike molecules versus temperature. (From Wu, S.-T. and Yang, D.-K., *Fundamentals of Liquid Crystal Devices*, John Wiley & Sons, New York, 2006.)

Note that an LC might not have all phases shown in Figure 8.7. A unit vector $\vec{n}$, the LC director axis, represents the direction of the long molecular axis [19]. In nematic phase, LC molecules have orientation order but do not have positional order. As the temperature decreases, they tend to have positional orders in addition to orientation order. The LCs that change their behavior due to temperature change are called thermotropic. Lyotropic is another type of LC that has molecular structure changes due to the concentration of the solvent. LCs have many applications in different fields of science, medicine, industry, and engineering. Due to the LC anisotropic behavior, they have been widely used in LC displays, optical devices, optical tunable filters, and sensors [20].

As mentioned before, thermotropic LC phase changes when temperature changes. The nematic phase of thermotropic LC is a phase that molecules have long-range orientation order. The average molecule orientation is denoted by director axis $\vec{n}$. In this phase, molecules are free to move and do not have positional orientation. Molecules move freely but keep their long-range orientation order. Nematic LCs have uniaxial symmetry in director axis direction; therefore, $\vec{n} = -\vec{n}$. Because of the anisotropic property of LC, it will exhibit different permittivities in different directions. A uniaxial anisotropic LC has parallel permittivity for electric field parallel to the director axis and perpendicular permittivity for electric field perpendicular to the director axis $\vec{n}$. The dielectric anisotropy is defined as $\Delta\varepsilon = \varepsilon_{\parallel} - \varepsilon_{\perp}$, which could be either positive or negative. For instance, 5CB $\Delta\varepsilon = 10.5$ and BL006 $\Delta\varepsilon = 17.5$ are positive, and MBBA $\Delta\varepsilon = -0.59$ shows negative permittivity. Because of the anisotropy of LC dielectrics, it is possible to use an electric field to detect changes in the orientation of the director axis.

8.3.2 Capacitive Transduction Technique

Capacitive sensors have been widely used in sensor applications. The principle of capacitive transduction is that the value of a capacitive sensor changes because of a change in the physical quantity. This feature has been used for measurement of humidity, temperature, distance, pressure, etc. Although capacitive transduction is not new to the sensor world, UAHuntsville has been the first to use the capacitive sensor for LC sensor application [16,17,21]. The uniaxial anisotropic nematic LC has been used as the dielectric material of a capacitor. A change in the director axis of the LC in the presence of a chemical or biological agent causes capacitance change that can be measured and monitored accurately.

Figure 8.8 shows a typical director axis of a uniaxial calamitic LC material in the presence of vertical and horizontal electric fields. The nematic LC shows two permittivities, $\varepsilon_{\parallel}$ in the direction of electric

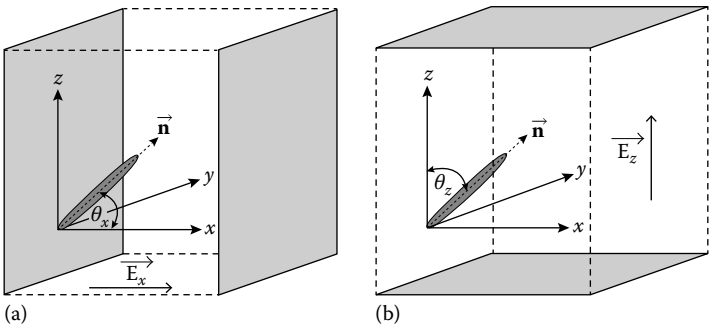


FIGURE 8.8 LC director axis and electric fields for horizontal and vertical parallel plate electrodes: (a) Vertical plates and (b) horizontal plates. (From Abu-Abed, A., Capacitive transduction for liquid crystal based sensors in ordered and partially disordered systems, PhD dissertation, University of Alabama in Huntsville, Huntsville, AL, 2007.)

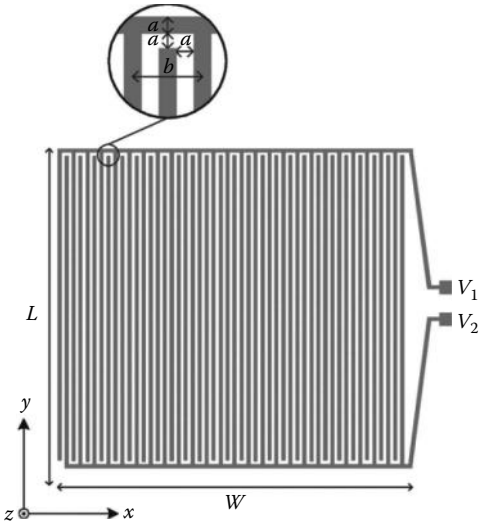


FIGURE 8.9 IDC structure. (From Abu-Abed, A., Capacitive transduction for liquid crystal based sensors in ordered and partially disordered systems, PhD dissertation, University of Alabama in Huntsville, Huntsville, AL, 2007.)

field parallel to molecule director axis and $\varepsilon_{\perp}$ for electric field perpendicular to the molecule director axis. In the molecule principal director axis coordinate, the permittivity tensor can be written as

$$\varepsilon_{\mathrm{m}} = \begin{bmatrix} \varepsilon_{\perp} & 0 & 0 \\ 0 & \varepsilon_{\perp} & 0 \\ 0 & 0 & \varepsilon_{\parallel} \end{bmatrix} \tag{8.1}$$

Considering homogeneous material, the director axis orientation can be determined using two capacitance measurements. Two parallel plate capacitors in the horizontal and vertical directions can provide enough information to determine the director axis angles as shown in Figure 8.8. The capacitance value can be calculated using a simple parallel plate capacitance equation:

$$C_i = \frac{\varepsilon_{\mathrm{eff},i} \cdot \varepsilon_0 \cdot A}{d} \tag{8.2}$$

Although vertical and horizontal electrodes provide enough information to determine director axis orientation, in order to have an open interface of the LC and chemical agents, an IDC can be used. Figure 8.9 shows the IDC structure.

An IDC consists of two latticed electrodes in the *XY* plane, and each electrode consists of *N* fingers. Each finger has a width of "*a*" and a spacing of "*a*" to the adjacent finger. The number of fingers can be large to provide a larger capacitance value. Electroplating can also be used to elevate electrodes of the finger and increase capacitance value and sensor sensitivity.

8.3.3 Excitation of the Liquid Crystal Sensor

Since the LC capacitive sensor is a passive component, an excitation voltage should be applied to the sensor in order to measure the capacitance. The amplitude and frequency of the ac excitation signal are important for a high-accuracy measurement. If the excitation voltage amplitude is

higher than Friedrich's voltage, the applied voltage can change the orientation of molecules and so the capacitance value. Excitation frequency of the LC sensor is very important for the determination of bandwidth and operating frequency of the charge preamplifier. One of the limiting factors in the resolution of the charge amplifier is the noise floor of the measurement system. In CMOS amplifiers, flicker noise or $1/f$ noise is dominant in low frequencies, and as the frequency of measurement decreases, this noise increases as shown in Equation 8.3. The first term in Equation 8.3 is the thermal noise, and the second term is the flicker noise that is inversely proportional to the frequency of operation:

$$v_N^2 = \left[\frac{8kT(1+\eta)}{3g_m} + \frac{K_F}{2fC_{OX}WLK'} \right] \Delta f \tag{8.3}$$

where
- k is Boltzmann constant
- T is temperature
- g_m is transistor transconductance
- η is a constant
- K_F is flicker noise coefficient
- f is frequency
- C_{OX} is oxide capacitance
- W is transistor width
- L is transistor length
- K' is a constant

Therefore, it is important to design the amplifier at a frequency that maximizes the resolution (lower noise) and accuracy of measurement. As a result, very low ppb agents can be detected with an integrated sensor microchip. Voltage investigation is important because low voltages are used for sensor capacitance measurement and high voltages are used to monitor anchoring energy for higher sensor sensitivity [22]. In order to measure voltage and frequency characteristic of the LC sensor, an experiment has been set up, and capacitance values under different voltages and frequencies of excitation have been measured. Excitation voltage and frequency have been changed from 0.5 to 20 V and 20 Hz to 2 MHz, respectively. Figure 8.10 shows the fabricated interdigitated sensor on a glass substrate.

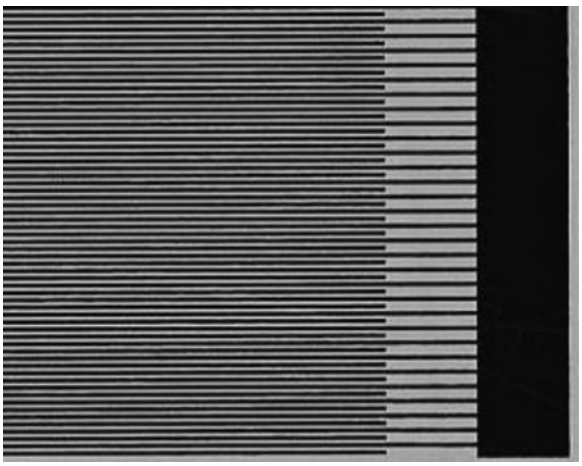


FIGURE 8.10 Fabricated interdigitated capacitive sensor.

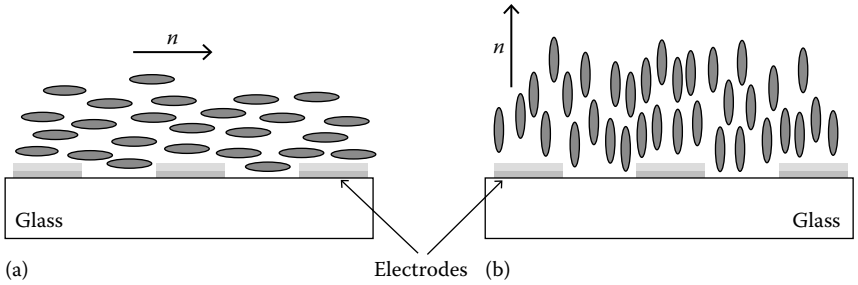


FIGURE 8.11 LC sensor molecule alignments: (a) Homogenous alignment and (b) homeotropic alignment.

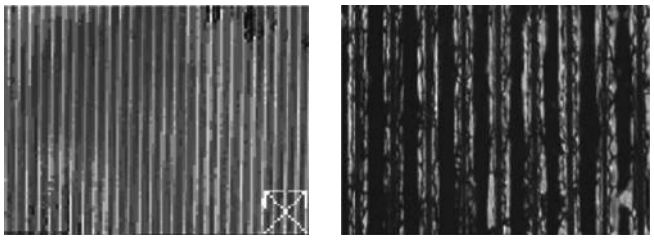


FIGURE 8.12 LC sensor optical observations: (a) Homogenous alignment and (b) homeotropic alignment.

Two different orientations for the LC molecules can be considered, one parallel to the substrate that is called homogenous and one perpendicular to the substrate or homeotropic orientation. These two orientations are shown in Figure 8.11.

In order to obtain homogenous alignment, 4% polyvinyl alcohol (PVA) in deionized water is spin coated on the interdigitated sensor surface and baked for 2 min at 90 °C. Then the surface is rubbed in one direction for a few times. After putting the LC on the surface, molecules tend to align in homogenous orientation. From optical observation, the homogenous alignment will look bright under a cross polarizer with 45° angle as shown in Figure 8.12a. For homeotropic alignment, 0.1% cetyltrimethylammonium bromide (CTAB) solution in deionized water is spin coated on the surface of the IDC. After baking at 90 °C for 5 min, the LC will be spin coated on the substrate. The resulting orientation will have homeotropic orientation and will always look dark regardless of the angle that it makes with the polarizer as shown in Figure 8.12b.

E7 LC has been used for the measurement because its clearing point is higher than room temperature and it remains anisotropic and has been used in LC chemical and biological sensing experiments.

It is has been shown that the sensor exhibits higher capacitance for low frequency rather than high frequency, because of frequency-dependent permittivity. Low-frequency excitation voltage causes LC molecules to vibrate. At high frequencies, LC loses its anisotropy that is not good for sensor applications. Although the higher frequency would provide less flicker noise for the amplifier, available bandwidth is also important for the rms input noise that should be considered for the amplifier design. The proper band for excitation is a flat band where the capacitance does not change very much versus frequency [ECE UAH]. Therefore, the proper range is the range where anisotropy is dominant [23–25].

8.3.4 LC Sensor Interface Circuit

The LC capacitive sensor is fabricated on a CMOS chip with the interface circuit for more sensitivity and accuracy in the detection. Different structures have been used for interfacing capacitive sensors in the literature. The most sensitive and accurate methods are the switched capacitor (SC) and linear differential charge amplifier.

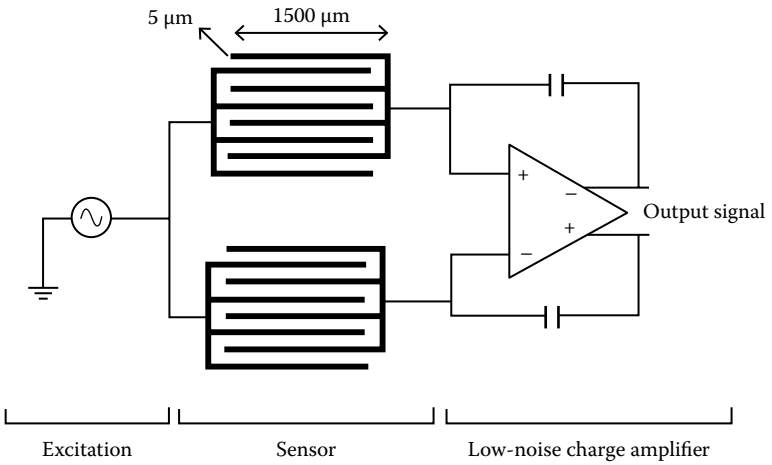


FIGURE 8.13 Block diagram of the differential measurement system.

Special requirements should be considered for interfacing the LC sensors. The LC excitation voltage should not contain dc components since it will cause charge migration in the film. Second, the amplitude of the ac excitation signal for measurement applications should be kept below Friedrich's voltage (a few tenth of a volt) in order not to make LC molecules reorient. These constraints should be considered in the design of the preamplifier. The block diagram of the sensor and the amplifier is shown in Figure 8.13.

The excitation voltage will be applied to the common node of the two sensors, and the signal output is amplified using a differential charge amplifier. Among different amplifier structures, a folded cascode amplifier has been used for this application. This amplifier has high output gain and wide input common mode range and can easily be compensated using loading capacitance. The amplifier has been used in many sensor and MEMS interface applications [26,27]. Furthermore, for reducing common disturbances and parasitic signals, a fully differential structure has been used. As discussed earlier, very low-frequency excitation causes the LC molecules to vibrate. Increasing the frequency of excitation decreases the sensor capacitance, which reduces measurement resolution. In order to have large capacitance and stable measurement, a 1 kHz excitation signal has been selected.

The input noise level of the low-frequency amplifier limits the measurement system resolution. Noise in a CMOS transistor can be modeled as a mean squared voltage noise at the transistor gate [28]. This noise consists of two components: thermal noise and $1/f$ or flicker noise as shown in Equation 8.3. Since thermal and flicker noises together affect the input noise level of the amplifier, they should be minimized simultaneously. A nonlinear optimization method has been used to design the preamplifier that minimizes the total input noise including thermal and flicker with a predefined constraint such as area limited design or power limited design. Note that power dissipation and area of the amplifier have a significant effect on the noise of the amplifier, but using an optimization algorithm can help in finding the best point with minimum loss of chip resources. A total input referred noise of the amplifier of 5.4 nV/$\sqrt{\text{Hz}}$ has been obtained, which corresponds to 15 aF capacitance resolution of the amplifier. The complete amplifier schematic is shown in Figure 8.14.

In order to compensate for the ac offset produced by sensor capacitor mismatch, CMOS switches have been used that change the number of fingers participating in the sensor as shown in Figure 8.15.

Two programmable ring counters provide sixteen-stage ac offset trimming for each interdigitated sensor with approximately 1% resolution. Furthermore, CMOS switches have been used for discharging feedback capacitors at a relatively long period of time to avoid saturation. The amplifier open-loop specifications are reported in Table 8.1.

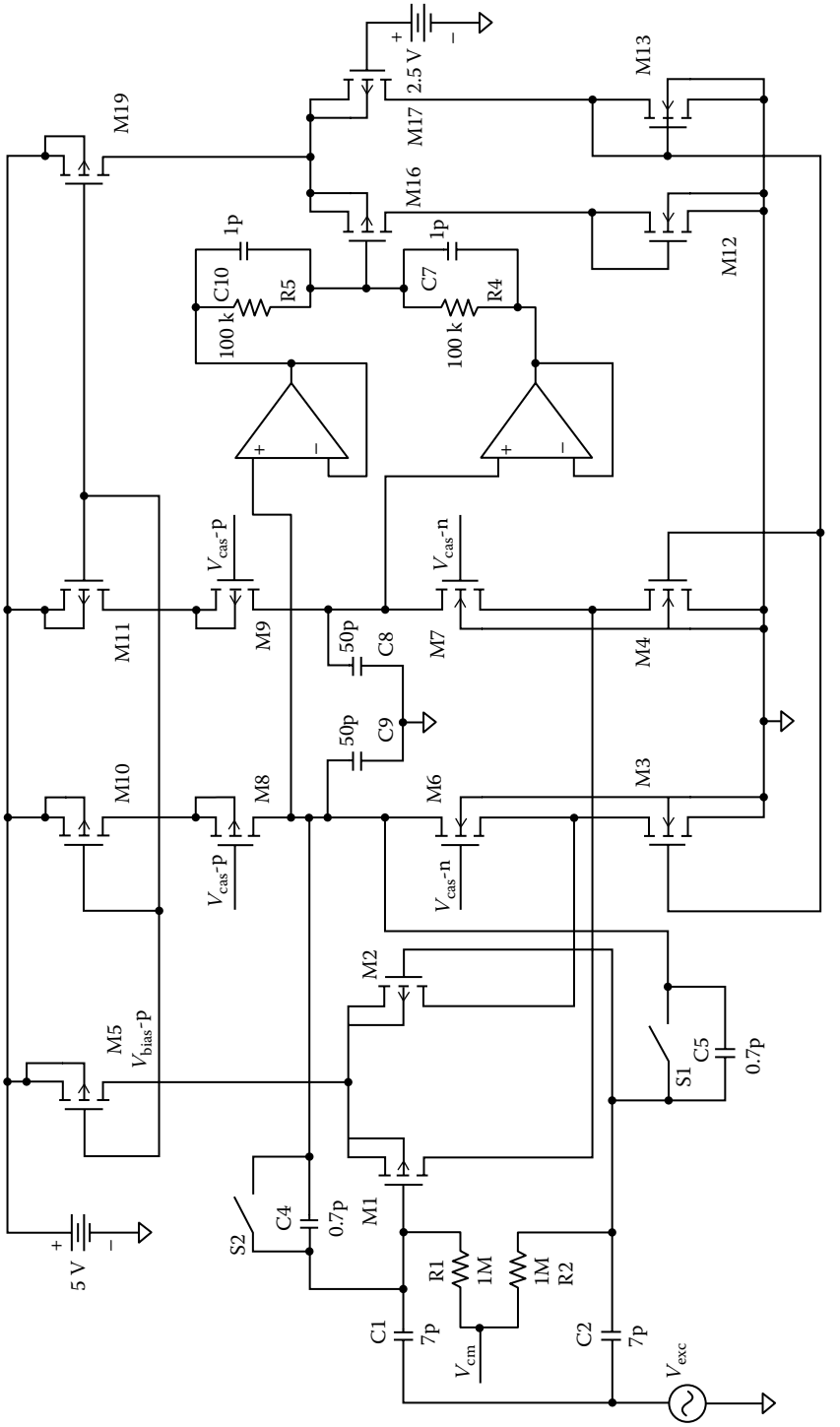


FIGURE 8.14 Fully differential charge amplifier that has been optimized for noise and performance.

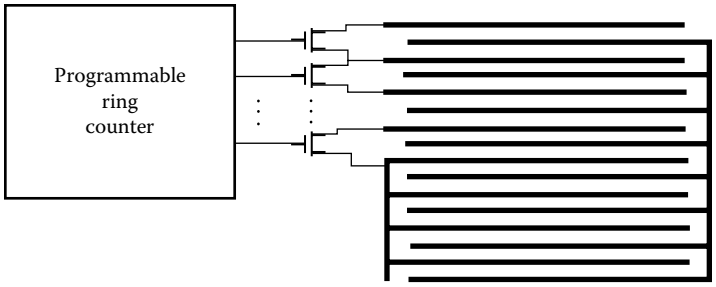


FIGURE 8.15 AC offset compensation using interdigitated sensor fingers and analog switches.

TABLE 8.1 Open-Loop Amplifier Specifications

A_{ol} = 74.5 dB	PM = 84.7°
PSRR = 99.4 dB	Output swing = 4.5 V
CMRR = 118 dB at 1 kHz	Pd = 1 mW

8.3.5 Gain Equation

For the differential amplifier gain in the midband frequency range, Figure 8.16 has been considered. Using charge conservation theory, Equations 8.4 and 8.5 can be written for the differential amplifier:

$$(v_{in} - v_A)C_{s1} = (v_A - v_{o1})C_f \tag{8.4}$$

$$(v_{in} - v_B)C_{s2} = (v_B - v_{o2})C_f \tag{8.5}$$

Subtracting Equation 8.5 from 8.4 and considering that $v_A = v_B$,

$$(v_{in} - v_B)(C_{s2} - C_{s1}) = -(v_{o2} - v_{o1})C_f \tag{8.6}$$

Considering matched feedback capacitors and small v_B, the gain equation can be written as

$$\Delta v_o \approx \frac{\Delta C_s}{C_f} v_{in} \tag{8.7}$$

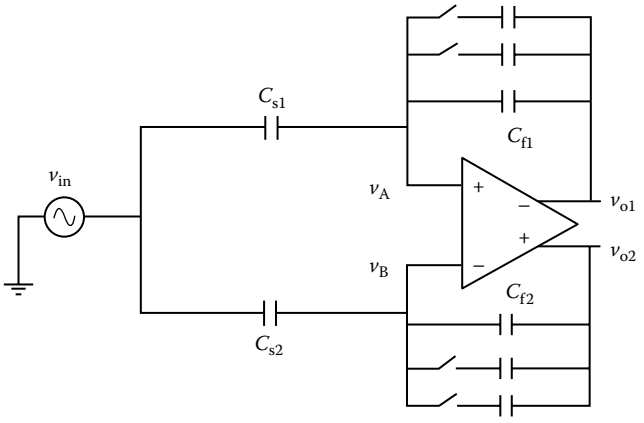


FIGURE 8.16 Sensor amplifier with variable gain.

The output differential signal is proportional to the sensor capacitance change divided by the feedback capacitor. Variable gain is achieved by using different capacitance values for the feedback capacitor. These capacitors are selected using CMOS analog switches as shown in Figure 8.16.

8.3.6 Switched Capacitor Implementation

Although a linear amplifier provides good noise and transfer characteristics without clock feedthrough problems, it has a few disadvantages including dc offset cancellation difficulties and large area-consuming resistors for input biasing. An SC technique is more complicated to design, but it does not need large resistors, and dc offset can be cancelled using auto-zero techniques. The same fully differential amplifier has been used as the core amplifier for the SC circuit. Figure 8.17 shows the schematic of the SC amplifier. Correlated double sampling (CDS) has been used to cancel the amplifier dc offset and low-frequency flicker noise [29]. Sensor capacitors are C2 and C5, the feedback capacitors are C7 and C8, and C9 and C10 are CDS capacitors. The SC circuit clock frequency should be higher than the Nyquist rate and in practice is selected to be more than 10 times of the input signal frequency. Very high clock frequency increases power dissipation of the amplifier.

Two nonoverlapping clocks S1 and S2 are used for the amplifier operation. The input signal is sampled on S1, and the feedback capacitor is discharged, while the dc offset is stored on CDS capacitors. On the amplification phase, clock S2, the differential signal is amplified at the output of the amplifier. High-frequency clock switching provides dc biasing for the input stage of the amplifier; therefore, separate biasing resistors are not required anymore.

Two main automatic offset cancellation techniques called chopper stabilization and CDS have been used for automatic offset cancellation. Chopper stabilization modulates the signal to a higher frequency; therefore, the dc offset and the signal is amplified in different bands and will be omitted. The CDS method is widely used in SC circuits for offset cancellation. The CDS method is a sampled data method that samples amplifier offset and stores it on a capacitor and subtracts it from the signal in the next clock cycle. Since CDS is a sampled data system, the sampling theorem applies to the offset cancellation. The flicker noise of

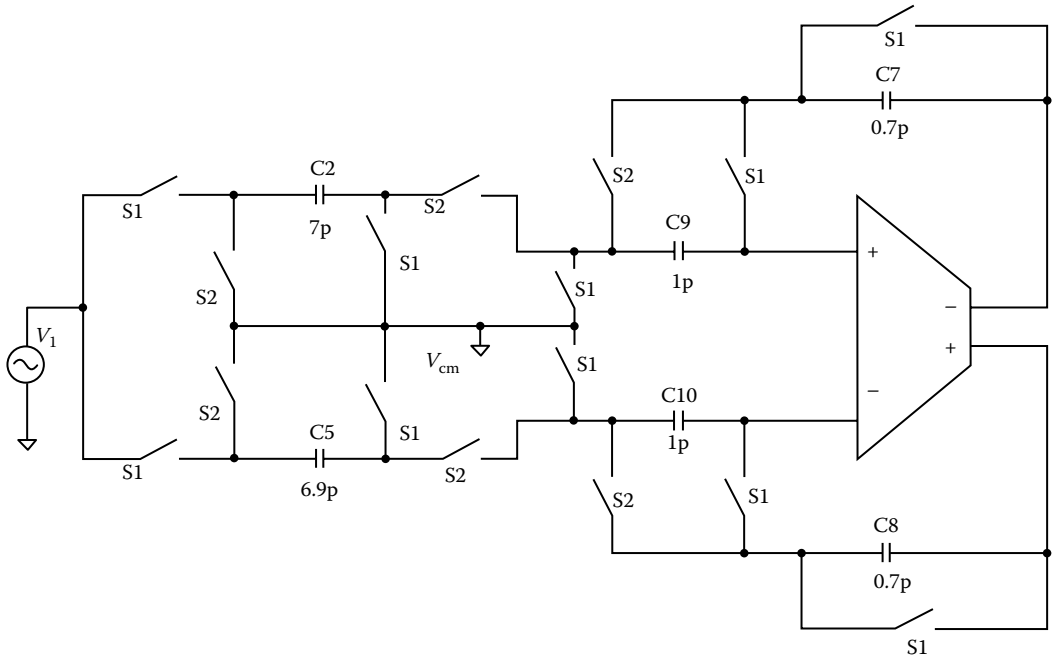


FIGURE 8.17 SC preamplifier with CDS.

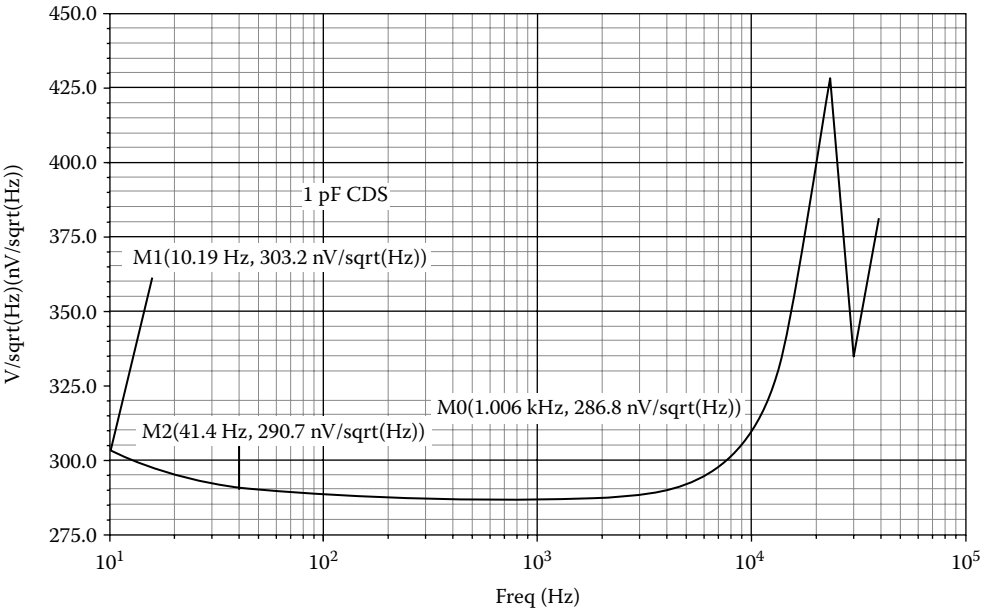


FIGURE 8.18 Input referred noise with a 1 pF CDS capacitor using PSS.

the CMOS amplifier is a low-frequency signal and is sampled by CDS and is decreased. Unfortunately, the amplifier thermal noise is under sampled by the CDS method and folds back to the amplifier signal band. Figure 8.18 shows the simulated input referred noise of the amplifier using Cadence periodic steady-state analysis and the CDS capacitor.

8.3.7 Demodulation and Filtering

After the sensor signal is amplified, demodulation and filtering are required to provide a dc signal for the analog-to-digital converter. The preamplifier output signal is buffered before entering the demodulation stage. The output signal of the demodulator goes to a low-pass filter to provide dc signal for the analog-to-digital converter. The demodulator used in this design is a switch-based multiplier that uses four analog switches as shown in Figure 8.19. The excitation of the LC sensor is amplified and transformed to a square wave after passing through multiple inverters and used for clocking the demodulator.

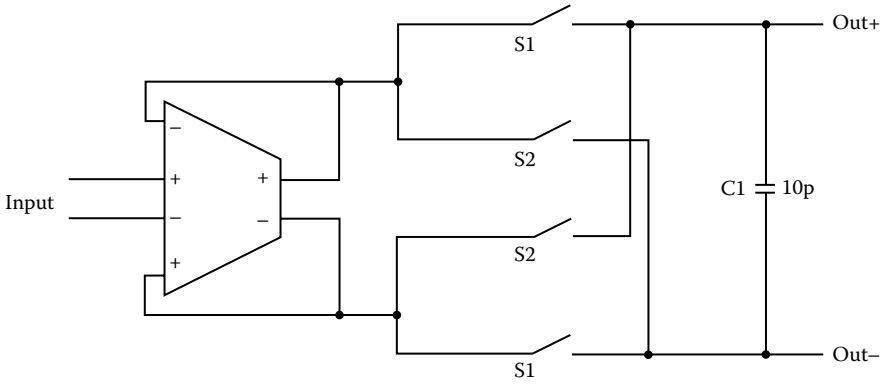


FIGURE 8.19 Buffer and switch demodulator.

Switches are transmission gates and because they do not have high currents, switch resistance is not critical. Furthermore, clock feedthrough and charge injections are highly reduced because of the differential structure. The preamplifier dc offset and out-of-band signals are modulated to higher frequencies, and the original signal that has the same frequency as the demodulator is down converted to a dc signal.

After the demodulation stage, a low-pass filter is required to filter the signal and prepare it for the analog-to-digital converter. Because the original signal has 1 kHz frequency, after demodulation, a filter with a much lower cutoff frequency is necessary to provide the dc output for the next stage. Considering a cutoff frequency of 100 Hz, the filter needs a very large time constant. If an *RC* filter is used, a 1 MΩ resistor with a 1.5 nF capacitor is required for the cutoff frequency of 100 Hz. These values are not practical for on-chip implementation. To achieve this low cutoff frequency, other integrated techniques like SC or *G*m–*C* filters can be used. In order to avoid switching and charge injection problems, the *G*m–*C* or (operational transconductance amplifier) OTA–*C* method has been employed.

The proposed *G*m–*C* filter structure is shown in Figure 8.20. This filter consists of a PMOS differential pair that uses source degenerative MOS transistors. The PMOS transistors have lower mobility, which is more useful in this structure than NMOS transistors. In order to increase output impedance and therefore gain of the OTA, cascade NMOS transistors have been used. The bias current has been selected to be 0.1 μA, and small transistor sizes for low *G*m have been selected. Degenerative transistors M11 and M22 (resistors) have large length for high resistivity that increases input linear range. Since a single-stage filter has negative feedback, it will further improve linearity.

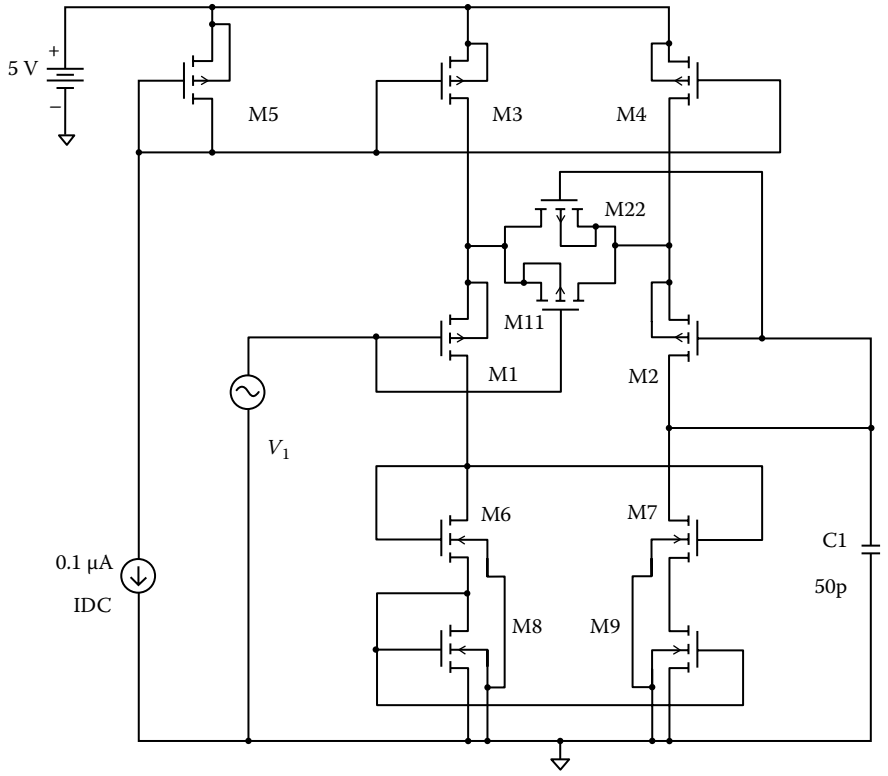


FIGURE 8.20 Schematic of the proposed *G*m–*C* filter.

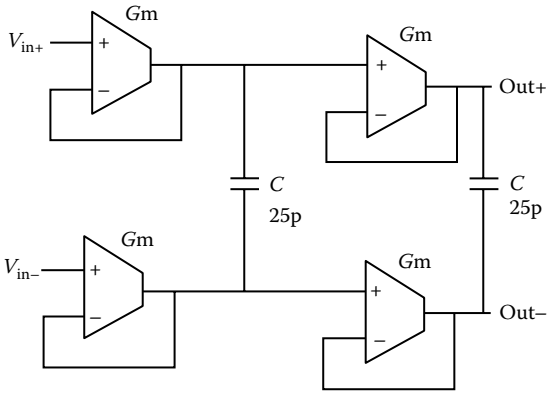


FIGURE 8.21 Differential low-pass Gm–*C* filter.

Figure 8.21 shows the Gm–*C* filter structure used in this design. The differential structure has been used to cancel common mode signals in the signal path.

Normally, the Gm of the Gm–*C* filters is tuned to get the proper cutoff frequency, but in this circuit, no extra tuning is necessary, because the corner frequency is well below 1 kHz. The dc output signal of this stage goes to the analog-to-digital converter section.

8.3.8 Analog-to-Digital Conversion

Among different techniques used for analog-to-digital conversion, sigma–delta is very attractive especially for low-frequency conversion. Other A/D conversion techniques like dual slope integrating and successive approximation need special analog fabrication processes such as laser trimming for high accuracy. The sigma–delta technique can achieve high accuracy (over 20 bits) in a digital process by using oversampling and noise shaping techniques without any postfabrication trimming or sophisticated analog component. This high accuracy can be achieved at the expense of more complicated digital filtering circuitry and time of conversion.

The sigma–delta converter is an oversampling technique where the input analog signal is usually sampled at a much higher frequency than the Nyquist rate. The oversampling rate (OSR) is defined as

$$\text{OSR} = \frac{f_s}{2f_o} \tag{8.8}$$

where

f_s is the sampling frequency

f_o is the signal highest frequency component

Figure 8.22 shows the block diagram of a sigma–delta converter. The input signal passes through an antialiasing filter to band limit the input signal. A sample and hold is used to sample the analog signal at

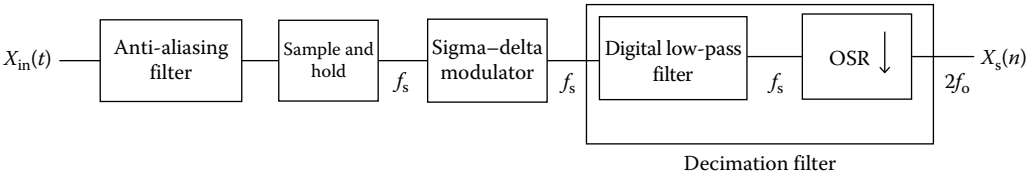


FIGURE 8.22 Block diagram of an oversampling A/D converter.

the first stage of signal processing. The sampled signal is modulated by a sigma–delta modulator stage. After modulation, a decimation filter implemented by a low-pass digital filter and down sampler provides the output digital data [30].

Antialiasing filters can be implemented using a simple first-order analog filter or a simple off-chip *RC* low-pass filter. Because sigma–delta conversion is usually implemented using an SC technique, the sample and hold is already in place by using the SC.

The sigma–delta modulator is a sampled data system and can be implemented using an SC technique. Because an SC samples the analog signal, there is no need for a separate sample and hold circuit. Since the noise of the first stage is important for the resolution of the A/D converter, the same low-noise folded cascode amplifier has been used as the main amplifier. The first-order sigma–delta A/D converter consists of an SC integrator, latch comparator, and feedback loop as shown in Figure 8.23. Two nonoverlapping clocks, S1 and S2, provide switching sequences. The CDS is used to cancel the dc offset and low-frequency flicker noise of the integrator. If a two-stage sigma–delta modulator is used, CDS is only applied for the first stage [31–34].

The output of the one bit quantizer is fed back to the integrator to select either of the reference voltages to integrate with the input signal. The gain of the integrator is set to 0.5 to avoid saturation, and reference voltages are 1 and 4 V. Therefore, the input signal range will be from 1 to 4 V. The differential signal from the previous stage is applied to the V_{in+} and V_{in-} inputs, and according to comparator output level, one reference voltage will be applied to oppose the charging direction of the integrator output (because of negative feedback). The time that comparator level remains high or low is proportional to the input signal level. The digital output pulse train goes to a decimation filter to convert to a baseband digital output. The decimation filter consists of averaging filters and low-pass filters that usually will consume a large part of the chip area.

In order to reduce the OSR and improve the noise transfer function of the sigma–delta modulator, the order of the modulator can be increased. Second-order modulators are common in sigma–delta A/D converters. Figure 8.24 shows the block diagram of a second-order sigma–delta modulator.

Switch capacitor implementation of the second-order sigma–delta modulator is shown in Figure 8.25.

Since the signals for the first and second stages have a phase difference, the feedback signal should be provided at the proper time for stability. Therefore, two flip-flops are used that latch the output signal on S1 and S2 clocks. Because the noise level and offset of the first stage is important, CDS offset cancellation has been used only for the first stage. The second-stage modulator is just a simple integrator.

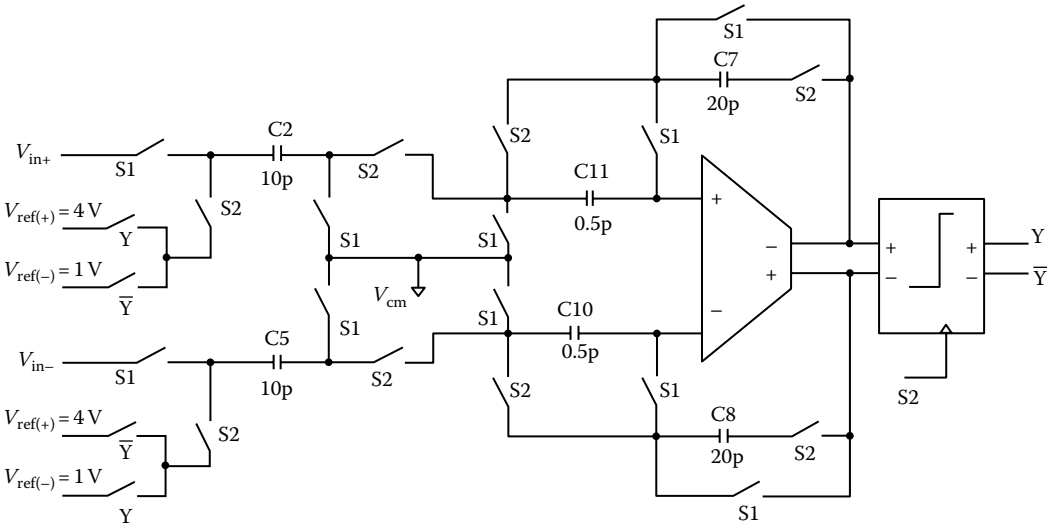


FIGURE 8.23 First-order SC sigma–delta A/D converter.

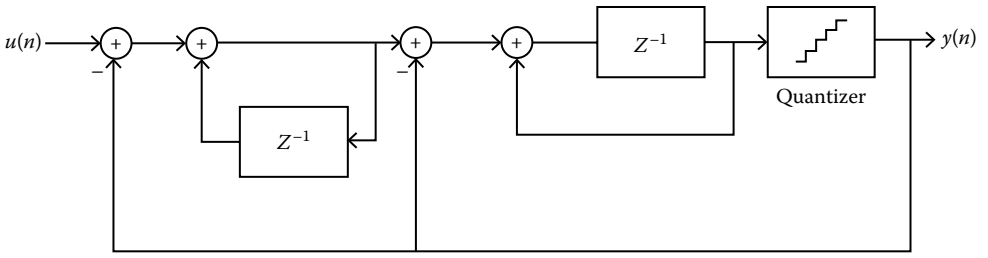


FIGURE 8.24 Second-order sigma–delta modulator block diagram.

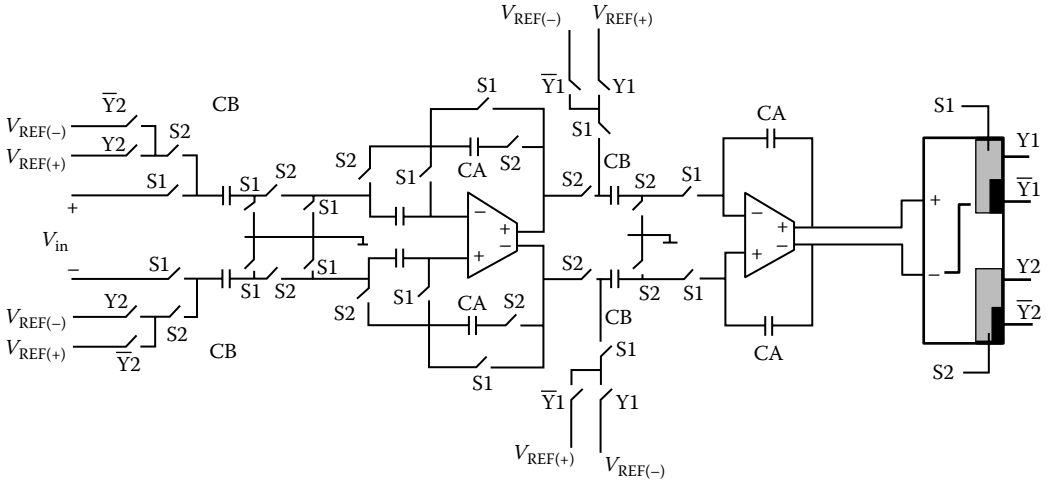


FIGURE 8.25 Second-order SC sigma–delta modulator.

The output signal of the second-stage integrator goes to the comparator. The clock frequency used for the second-order modulator is 51.2 kHz, which is much lower than 400 kHz selected for the first-order modulator and will provide more relaxing conditions for the main amplifier. The output signal of the integrators does not need a high slew rate OTA, because of the integrating function. The total microchip area including sensor and the interface circuit consumes 16 mm^2 of chip area using AMI0.5 technology.

8.4 Conclusion

The concept of single-chip chemical and biological sensor brings a lot of worthy advantages to chemical and biological detection world. Noise immunity, better accuracy, higher-resolution, mass production, and state-of-the-art CMOS fabrication technology bring a lot of opportunity for the researchers and engineers in the field. The challenge is to build the sensor system on an integrated circuit chip with the minimum postprocessing steps. In this case, the product will be very cheap and accurate if mass produced. Even though quite a number of methods have already been fabricated with the interface circuit on a chip, there are still many other methods that can be fabricated for better accuracy and mass production. Even though CMOS is currently the dominant technology for IC and MEMS fabrication, new technologies have also been reported for large array of sensors on other technologies such as amorphous silicon. The report in [35] uses LC chemical sensors on an a-Si:H chip that can be fabricated on glass and for large area electronics. A sensor array, as large as an iPad display, can be fabricated for handheld and deployable applications. The area of single-chip chemical and biological detectors will remain an active area of research as an interdisciplinary area for chemical and electric engineers for the years to come.

References

1. T. Vo-Dinh and B. Cullum, Biosensors and biochips: Advances in biological and medical diagnostics, *Fresen. J. Anal. Chem*., 366, 540–551, 2000.
2. C. Hagleitner, A. Hierlemann, D. Lange, and A. Kummer, Smart single chip gas sensor microsystem, *Nature*, 414, 293–296, 2001.
3. A. Hierlemann and H. Baltes, CMOS-based chemical microsensors, *R. Soc. Chem*., 128, 15–28, 2003.
4. F. Patolsky, G. Zheng, and C. M. Lieber, Nanowire-based biosensors, *Anal. Chem*., 78, 4260–4269, 2006.
5. J. C. Chen, J. C. Chou, T. P. Sun, and S. K. Hsiung, Portable urea biosensor based on the extended-gate field effect transistor, *Sens. Actuator B*, 91, 180–186, 2003.
6. T. S. Cho, K. Lee, J. Kong, and A. P. Chandrakasan, A low power carbon nano-tube chemical sensor system, *Proceedings of the IEEE Custom Integrated Circuit Conference*, San Jose, CA, pp. 181–184, 2007.
7. M. Schienle, C. Paulus, A. Frey, F. Hofmann, B. Holzapfl, P. Schindler-Bauer, and R. Thewes, A fully electronic DNA sensor with 128 positions and in-pixel A/D conversion, *IEEE J. Solid State Circuits*, 39(12), 2438–2445, 2004.
8. R. Shah and N. L. Abbott, Principles for measurement of chemical exposure based on recognition-driven anchoring transitions in liquid crystals, *Science*, 293(5533), 1296–1299, 2001.
9. J. Brake, M. Daschner, and N. L. Abbott, Biomolecular interactions at phospholipid decorated surfaces of thermotropic liquid crystals, *Science*, 302(5653), 2094–2097, 2003.
10. R. R. Shah and N. L. Abbott, Coupling of the orientations of liquid crystals and electrical double layers formed by the dissociation of surface immobilized salts, *J. Phys. Chem. B*, 105(21), 4936–4950, 2001.
11. R. R. Shah and N. L. Abbott, Using liquid crystals to image reactants and products of acid-base reactions on surfaces with micrometer resolution, *J. Am. Chem. Soc*., 121(49), 11300–11310, 1999.
12. V. K. Gupta and N. L. Abbott, Uniform anchoring of nematic liquid crystals on self-assembled monolayers formed from alkanethiols on obliquely deposited films of gold, *Am. Phys. Soc. Phys. Rev. E*, 54(5), 4540–4543, 1996.
13. V. K. Gupta, T. B. Dubrovsky, and N. L. Abbott, Optical amplification of ligand-receptor binding using liquid crystals, *Science*, 279(5359), 2077–2080, 1998.
14. J. A. Van Nelson, S. R. Kim, and N. L. Abbott, Amplification of specific binding events between biological species using lyotropic liquid crystals, *Langmuir*, 18(13), 5031–5035, 2002.
15. J. J. Skaife and N. L. Abbott, Quantitative interpretation of the optical textures of liquid crystals caused by binding of immunoglobulins to surface-bound antigens, *Langmuir*, 16(7), 3529–3536, 2000.
16. R. G. Lindquist, A. Abu-Abed, and C. Woo-Hyuck, Liquid crystal sensors with capacitive transduction, *Proceedings of the 5th IEEE Conference on Sensors*, Daegu, Korea, 2006.
17. A. Abu-Abed, R. G. lindquist, and W.-H. Choi, Capacitive transduction for liquid crystal based sensors. Part 1: Ordered system, *IEEE Sens. J*., 7(12), 1617–1624, 2007.
18. L. M. Bilnov, Electro-optical effects in liquid crystals, *Soviet Physics Uspekhi*, 17(5), 658, 1975.
19. S.-T. Wu and D.-K. Yang, *Fundamentals of Liquid Crystal Devices*, John Wiley & Sons, New York, 2006.
20. S. A. Alboon and R. G. Lindquist, Flat top liquid crystal tunable filter using coupled Fabry-Perot cavities, *Opt. Express*, 16(1), 231–236, 2008.
21. A. Abu-Abed, Capacitive transduction for liquid crystal based sensors in ordered and partially disordered systems, PhD dissertation, University of Alabama in Huntsville, Huntsville, AL, 2007.
22. J. Namkung, Y. Zou, A. Abu-Abed, and R. G. Lindquist, Capacitive techniques to monitor of anchoring energy for liquid crystal sensors, *IEEE Sens. J*., 10, 1479, 2010.
23. B. I. Leev, V. N. Sergienko, P. M. Tomchuk, and E. K. Frolova, Nematic liquid crystals in frequency and amplitude modulated electric fields, *Liq. Cryst. J*., 28(7), 973–982, 2001.
24. A. Golovin, Introduction to dielectric measurement of nematic liquid crystals, Kent State University, LCI & CPIP (online source, 2010).
25. A. Hassanzadeh and Y. Zou, Frequency and voltage characteristics of capacitive liquid crystal chemical and biological sensors, ECE Department, University of Alabama in Huntsville, Huntsville, AL, 2010.

26. B. V. Amini, R. Abdolvand, and F. Ayazi, A 4.5 mW closed loop sigma-delta micro-gravity CMOS-SOI accelerometer, *Proceedings of the IEEE International Solid State Circuit Conference*, San Francisco, CA, February 2006.
27. S. Lichun, M. Palaniapan, and T. W. Wan, A continuous-time capacitance to voltage converter for micro capacitive pressure sensors, *Proceedings of the International MEMS Conference*, Istanbul, Turkey, 2006.
28. P. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, Oxford University Press, Oxford, U.K., 2002.
29. C. C. Enz and G. C. Temes, Circuit techniques for reducing the effects of Op-Amp imperfections: Autozeroing, correlated double sampling and chopper stabilization, *Proc. IEEE*, 84(11), 1584–1614, 1996.
30. D. A. Johns and K. Martin, *Analog Integrated Circuit Design*, John Wiley & Sons, New York, 1997.
31. S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta Sigma Data Converters, Theory Design and Simulation*, IEEE Press, New York, 1997.
32. B. E. Boser and B. A. Wooley, The design of sigma-delta modulation analog to digital converters, *IEEE J. Solid State Circuits*, 23(6), 1298–1308, 1988.
33. V. F. Dias, G. Palmisano, P. O'Leary, and F. Maloberti, Fundamental limitations of switched-capacitor sigma-delta modulators, *IEEE Proc.*, 139, 27–32, 1992.
34. R. Gaggl, A. Wiesbauer, G. Fritz, C. Schranz, and P. Pessl, A 85-dB Dynamic range multibit delta sigma ADC for ADSL-CO applications in 0.18 um CMOS, *IEEE J. Solid State Circuits*, 38(7), 1105–1114, 2003.
35. A. Hassanzadeh and R. G. Lindquist, Design of an a-Si:H interface circuit for liquid crystal chemical and biological sensors, *IEEE Sens. J.*, 12(5), 1284–1288, 2012.

9

CMOS Integrated Gas Sensors

Prasanta K. Guha
Indian Institute of Technology, Kharagpur

Sumita Santra
Indian Institute of Technology, Kharagpur

Julian W. Gardner
University of Warwick

This chapter provides a comprehensive discussion on the necessity and importance of developing gas sensors on a low-cost complementary metal oxide semiconductor (CMOS) platform technology. The main focus is on the challenges of integrating microhotplate-based gas sensors into a standard CMOS process: microhotplate-based sensors typically operate at much higher temperatures of 200 °C–500 °C. This is beyond the normal limit of CMOS operation at +125 °C. The chapter focuses on both different on-chip circuit schemes to interface these microhotplate-based sensors and also the challenges associated with depositing/growing nanosensing materials (particularly carbon nanotubes [CNTs] and graphene) on a CMOS platform. Finally, there is a brief discussion on the existing gas sensor market and its future outlook.

9.1 Introduction

In the past decade, gas and chemical sensors have been increasingly important because of their desired application in industry, automobiles, health care, and environmental monitoring. Hence, considerable effort is being directed toward the development of miniaturized smart gas sensors and systems. More recently, research has been carried out on the selective detection of specific toxic gases (e.g., CO and NO_2) and combustible gases (e.g., H_2 and CH_4) and to develop innovative technologies to integrate such sensors onto a CMOS platform.

The current gas sensor market is dominated by electrochemical cells, resistive metal-oxide sensors, pellistors, and infrared (IR) sensors. Among them, electrochemical cells hold the largest market share. Most of the gas sensors available today are expensive costing on average $25; this is because the sensors are usually made in small volumes and a semi-automated manufacturing process. These sensors are discrete and hence require separate electronic board to drive and monitor the sensors, which makes them even more expensive (in excess of $100 as a module).

The present gas sensor market has been estimated to be worth around $500 M and is expanding at a rate of 10% [1]; this sustained market expansion is possible because of more awareness within the domestic and commercial world for safer and cleaner environment and also the introduction of stricter

legislation from different governments. Although the gas sensor market is small compared with the main stream electronic markets (e.g., PCs, tablets, and mobile phones), the market can open up drastically for chemical sensors in the near future; this is because there are recent tendencies from the blue chip companies to accommodate gas and other sensors (to give more value-added features) in their cutting-edge electronic gadgets. For example, the NASA Ames Research Center has recently developed a prototype (known as tricorder) that can convert an iPhone into a chemical sensor, which is claimed to be capable of different applications from detecting chemical attacks in future military conflicts to testing blood glucose indirectly by measuring acetone in exhaled breath [2]. The Japanese mobile giant NTT Docomo has announced that it is bringing a smartphone-ready device to measure acetone in breath for diet support [3]. Recently, Nokia has forecasted a novel mobile phone concept known as Scentsory; it evokes the human senses of sight, sound, touch, and even smell for a multisensory experience; the heart of the smell sensor consists of array of chemical sensors (known as e-nose) [4]. But this emerging mass market requires smaller smart gas sensors of much lower cost (each sensor should cost below $1) and much lower power (so that longer battery life can be maintained) compared to present-day available gas sensors.

This penetration of mass market will only be possible when gas sensors can be developed on a high-volume low-cost CMOS platform. This is because CMOS-based sensors do not only give miniaturized devices (<1 mm^2) but also offer lower power consumption and option of batch fabrication, which reduces the cost and also improves the reproducibility of the device performance. Furthermore, the on-chip electronics along with sensors always gives an added cost advantage compared to separate interface electronic chips and sensors on a printed circuit board—provided we can satisfy the process constraints of CMOS when fabricating the integrated or smart sensors.

Although the widely used electrochemical sensors consume very low power, it is not possible to integrate them on CMOS because they require significant volumes of liquid electrolytes and non-CMOS catalytic materials such as platinum or silver electrodes. In addition, electrolyte evaporation limits its life. A common gas sensor is based on nondispersive infrared detection (NDIR) technique; it requires high power consumption (0.5 W) for the infrared (IR) source, and they are difficult to integrate on CMOS because of the requirement of non-CMOS IR band-pass filter.

Research in the last few years has shown that it is possible to integrate microhotplate-based gas sensors on a CMOS platform, so the chapter will focus on the microhotplate-based resistive and calorimetric sensors and their interfacing electronics. The key issues of growing nanosensing materials on a CMOS platform are also discussed; particularly, the growth technique of recently discovered materials like CNTs or graphene and their response in the presence of gas is described.

9.2 Microhotplate-Based Gas Sensors

There are primarily two types of microhotplate-based gas sensors: (1) resistive and (2) calorimetric.

Resistive gas sensors show a considerable change in their electrical resistance upon exposure to reactive gases or vapors; these usually contain films of nonstoichiometric wide-bandgap semiconducting oxides that react with the analytes at elevated temperatures of 200 °C and 600 °C. The Taguchi gas sensor (TGS) is perhaps one of the earliest commercial resistive gas sensors, first patented and sold around 1970; even today they are sold by the millions by Figaro Inc., Japan [5], and other companies now that the patent has elapsed. The device consists of a platinum coil heater (to heat up the sensing material) within a ceramic tube, painted manually with a thick layer of tin oxide. Under the oxide layer is a pair of thin interdigitated gold electrodes to measure the change in resistance of the tin oxide in the presence of an oxidizing or reducing gas (e.g., NO_2 or CO).

Calorimeters (also known as pellistors, which is an abbreviation of the words pellet and resistor) are used to detect gases that are either combustible (catalytic) or that have a significant difference in thermal conductivity (noncatalytic) to that of air (e.g., catalytic pellistors are popular as a methane sensor). This consists of a small catalyst-coated bead (e.g., alumina) around a platinum heater wire and requires

a second pellistor (without the catalytic additive) as a reference arm in a potential divider. This clearly adds to the size and cost of the pellistor.

A new generation of microresistive/microcalorimetric gas sensors is being developed based on a microhotplate structure using CMOS/MEMS technology. A microhotplate can be defined as a miniaturized thin membrane that is thermally isolated from the rest of the substrate and usually consists of a microheater structure (for heating up the sensing material), a temperature sensor (to measure the temperature of the top membrane layer), and lastly contact electrodes for the sensitive layer. These types of thin microstructures are usually power efficient; hence, higher operating temperatures can be reached at lower power consumption (typically <50 mW at 500 °C). Microhotplates are much faster than the existing sensors (thermal time constants of the order of 1–50 ms) due to their lower thermal mass, which also enables a pulsed mode of operation, lowering the power consumption even further. In the case of the latest microcalorimeters, membrane technology is again used and like microresistive sensors except these do not require any electrodes. Two identical sensors (one having an active sensing material and the other playing the role of a reference) are again used, and the difference in temperature is detected between the two due to the combustion of the target gas at around 550 °C.

Microhotplate membrane structure can be realized by the front or back etching of a wafer—or even using a combination of both. If it is through front etching, then the structure is known as a suspended membrane (i.e., the heated area is suspended by, say, four arms with opening between the arms that acts as etch opening), and if the etching is from back side, then it is known as closed membrane.

Researchers have developed microhotplate structure using different materials. The key issue is that the heater material should be able to sustain a high operating temperature and that the membrane material should be a dielectric with a low thermal conductivity (compared to silicon). For example, platinum heaters and also platinum electrodes are one of the popular choices because of their reliable operation at high temperature, but platinum is not a CMOS material. The other used heater materials are doped silicon, aluminum, and doped polysilicon to name a few, though these are standard CMOS materials—but most of the devices were fabricated not using a commercial CMOS foundry; hence, monolithic circuit integration is not possible. It is not easy to develop a microhotplate structure on standard CMOS, because an additional etching step is required to form the membrane and also the heating area has to be properly isolated from the rest of the chip for a faithful operation of the interfacing circuits—however, there are some groups who took up this challenge and successfully demonstrated this integration. For example, a group at National Institute of Standards and Technology (NIST) [6,7], a group at Physical Electronics Laboratory (PEL) of ETH Zurich [8], and a group at High Voltage Microelectronics and Sensors (HVMS) of University of Cambridge in collaboration with Microsensors & Bioelectronics Laboratory (MBL) of University of Warwick [9] are perhaps some of the earliest to report microhotplate-based gas sensors on a commercial CMOS platform.

The microhotplate developed by the NIST group was by using front etch bulk micromachining technique (top view is shown in Figure 9.1a [6]). The hotplate has a polysilicon heater with aluminum metal (provided by the CMOS process) spreading plate for temperature uniformity. A separate gold electrode was deposited as a postprocess for sensing material contact. The outer perimeter of the microhotplate design has a p-type implant that was used as an etch stop. The device was power efficient (10 °C/mW) and fast (thermal time constant 1 ms).

The microhotplate developed in ETH Zurich was released by KOH back etching. There is a silicon island underneath the heater area, which serves as heat spreader that also mechanically stabilizes the membrane; this is obtained by an electrochemical etch stop technique. The formation of such an island can be found in papers published by Kloeck et al. [10] and also in Müller et al. [11]; the latter one discussed the formation of such islands at wafer level. The Zurich group had to use a postprocess Pt/TiW coating on an aluminum electrode to avoid a problem due to aluminum oxide formation.

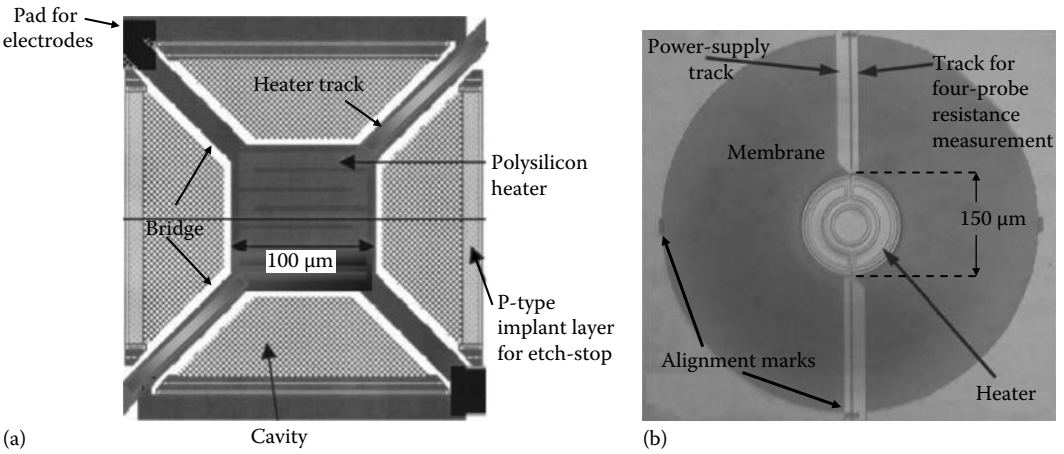


FIGURE 9.1 (a) The microhotplate developed by the NIST group using front etch bulk micromachining technique. (Adapted from Afridi, M.Y. et al., *IEEE Sensors Journal*, 2, 644, 2002.) (b) Cambridge CMOS Sensors Ltd smart ultralow power microhotplate that uses a high-temperature metal microheater. (Adapted from Ali, S. Z. et al., *J. Microelectromech. Syst*., 17, 1408, 2008.)

The microheater was made with polysilicon; later on, they also reported a metal-oxide semiconductor field-effect transistor (MOSFET) heater.

The Cambridge/Warwick microhotplate was developed using a silicon-on-insulator (SOI) CMOS process with an FET as the heater structure. The microhotplate was a closed membrane type made with a silicon oxide/nitride dielectric layer. The membrane was realized by a deep reactive-ion etching (DRIE) technique from the back side to remove the underlying handle silicon as a post-CMOS process at wafer level. Here, a buried oxide acts as an etch stop. However, the device works only up to 300 °C [9]. Later on, the operating temperature range was increased by using metals with a higher melting point (in standard CMOS) instead of aluminum [12,13]. The report of long-term reliability (continuous run over three months) showed that these devices are very stable and more stable than doped polysilicon. The group has launched a spin-off company (Cambridge CMOS Sensors Ltd) based on this patented smart ultralow power microhotplate [14]. The newly developed microhotplate from this group is shown in Figure 9.1b, which uses a high-temperature metal microheater.

9.3 Interface Electronics

The main advantage of CMOS integration is the on-chip interface electronics. The interface circuit can be integrated in two ways: (1) monolithic (sensor and circuit are on the same silicon die), and (2) hybrid (sensor and circuit are on separate dies). The advantage/disadvantage of one approach over another has been discussed in detail in the literature [15]. Briefly, in the case of the monolithic approach, sensors need to be fabricated with CMOS materials and follow strict foundry specifications, whereas it offers lower power consumption and fewer packaging issues. In the case of a hybrid approach, the main advantage is that the sensors can be fabricated with any required materials using much less stringent design rules. However, the parasitic capacitance/resistance, due to having bonding wires between the dies, is clearly larger and less predictable/repeatable than with the monolithic approach.

As far as the front end interfacing is concerned, a common problem with any microsensor is that they tend to generate smaller electrical signals at the sensor output, so the front end circuitry needs to be of very low noise, that is, a good analog design is required. Present submicron CMOS processes offer a very low rail-to-rail supply (say 0–1.1 V); this is good for low power, but this makes front end design much harder when the sensor drifts (i.e., a smaller operating range).

As microresistive sensors are based on a change in the electrical conductivity at high operating temperatures, the main circuit blocks required for interfacing these sensors are (1) driving circuits for microheaters, (2) temperature control circuits, and (3) interface circuits for the gas-sensing material. We will now consider these blocks in turn:

1. Microheaters can be driven with an accurate, stable voltage or current source, which is capable of driving large currents (tens of mA). Circuits can be designed to provide a constant signal (dc) or a modulated signal (ac or pulse signal) to the heater. In the case of a voltage source, it is necessary to use a current-limiting resistor in order to avoid any high current following through the heater. Current sources can be designed in CMOS employing a current mirror circuit. A four-wire (or two-wire) measurement option is often required to measure the change in resistance of the heater with temperature; this can be used to extract temperature information of the sensing layer.
2. The temperature of the sensing material plays a vital role in improving the selectivity of almost all sensing materials. Thus, it is very important to read and control the temperature of the sensing layer accurately. Bota et al. [16] in their work used the same resistor element as the heater and temperature sensor. They used a pulse width modulation (PWM) technique, where during the off mode of the microheater, a small current was driven through the heater to measure its temperature (as shown in Figure 9.2). Now, instead of using the microheater for temperature extraction, separate temperature sensors (e.g., silicon *p–n* junction diode, silicon/polysilicon resistor) can also be used to read the temperature of the sensing layer. A small current (∼μA) is usually used to drive the temperature sensor to avoid any self-heating effect. One of the commonly used interfacing techniques is to have one temperature sensor on the microhotplate and the other identical one outside the hotplate and use an accurate instrumentation amplifier to amplify the difference of temperature readings [13].

   The amount of power required to maintain a constant temperature depends upon the ambient temperature and on the structure of the microhotplate, particularly on the nature and thickness of the sensing material and membrane. Pulse mode heating can be used to reduce the average power consumption; this is achieved by using a switch in series with the microheater. The heating can be controlled by on–off mode (also known as bang–bang control), proportional mode, or proportional integral derivative (PID) mode. There have been some reports of an

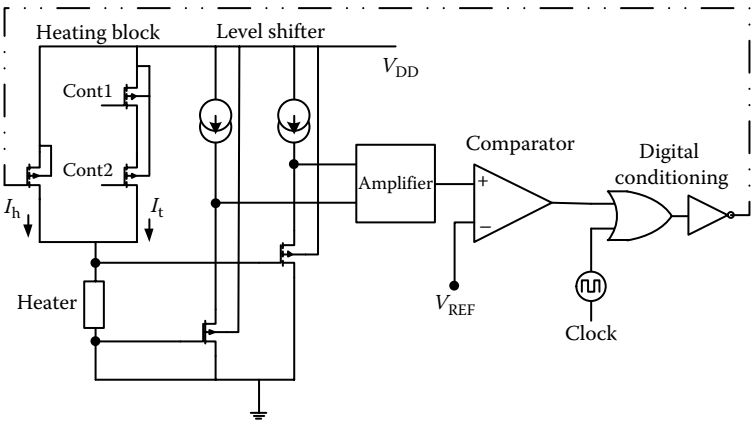


FIGURE 9.2 A PWM technique, where during the off mode of the microheater, a small current was driven through the heater to measure its temperature. I_h, heating current; I_t, temperature measuring current; Cont1 and Cont2, controlling switch for the measuring current; V_{DD}, supply voltage. (Adapted from Bota, S. A. et al., *Analog Integr. Circuits Signal Process.*, 40, 175, 2004.)

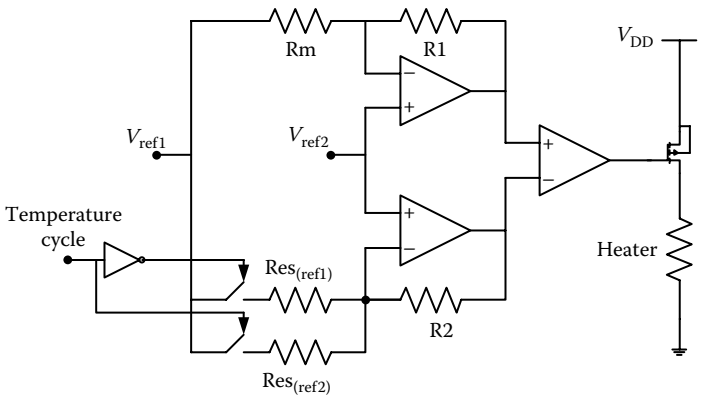


FIGURE 9.3 An on–off control circuit where if the microhotplate is cooler than a set-point temperature, the heater is turned on at maximum power, and once it is hotter than the set-point temperature, the heater is switched off completely. Res(V_{ref1} and V_{ref2}), resistances for setting temperature; V($_{ref1}$ and $_{ref2}$), reference voltages; V_{DD}, supply voltage; Rm, comparing resistor with Res_{ref}. (Adapted from Cardinali, G. C. et al., *Analog Integr. Circuits Signal Process.*, 14: 275, 1997.)

on–off control circuit in the literature (as shown in Figure 9.3) [17]. The on–off controller is the simplest form, where if the microhotplate is cooler than a set-point temperature, the heater is turned on at maximum power, and once it is hotter than the set-point temperature, the heater is switched off completely. However, bang–bang control can give rise to instability (unless controlled carefully), which in turn could damage the heater due to very rapid changes in voltage (hence temperature).

Barrettino et al. [8,18] proposed a simple proportional temperature controller circuit as shown in Figure 9.4. The operational amplifier (which contains an integral stabilization capacitor of 8 pF) drives a power transistor that provides current to the polysilicon heater. The inputs of the operation amplifier consist of the control voltage ($V_{control}$) and the voltage drop across the temperature sensor. The same group has also reported an on-chip digital PID controller using a 0.8 μm

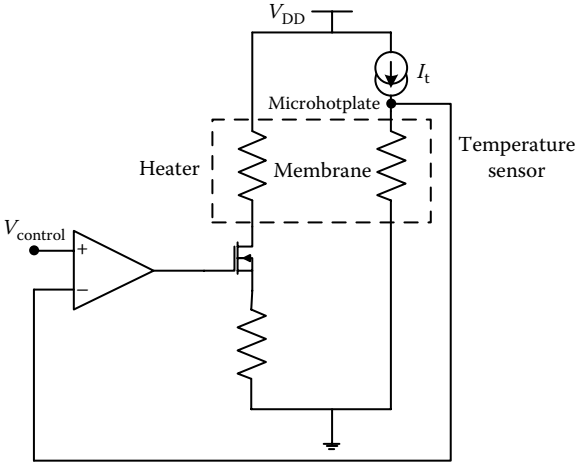


FIGURE 9.4 A simple proportional temperature controller circuit. The operational amplifier drives a power transistor that provides current to the polysilicon heater. I_t, current through the temperature sensor. (Adapted from Barrettino, D. et al., A smart single-chip micro-hotplate-based chemical sensor system in CMOS technology, in *Proceedings of the International Symposium on Circuits and Systems*, Phoenix, AZ, pp. 157–160, 2002.)

commercial CMOS process for controlling heater temperature smoothly and shows a micrograph of the integrated sensor chip with three PIDs and 10 bit ADC [19]:

3. The design of this interfacing circuit is perhaps the most challenging part. This is because
   a. The baseline resistance (baseline resistance means resistance in the absence of any gas) of different sensing materials varies over a very wide range (may be from 1 kΩ to 1 GΩ); this is equivalent to a range of 120 dB.
   b. The baseline resistance of nominally the same material also differs from one gas sensor to another (up to 50%), because resistance depends on grain size, morphology, and many other such parameters; that is, a synthesis mechanism is the key for controlling these factors.
   c. The sensing material resistance changes in the presence of a gas; the amount of change depends on the concentration of the gas, temperature of operation, and sensitivity of the material toward that particular gas.
   d. The sensing material resistance drifts with aging (through poisoning/stability issues) and also due to the incomplete release of analytes at the end of the measurements.

So the interface circuit must handle the precision and wide dynamic range that is demanded by the sensing element. The resolution of the circuit needs to be high so that it can amplify the change in resistance in the presence of a gas at PPM levels or lower. Along with these, a drift cancellation circuit will make the system smarter and more accurate, because drifting gives the impression of change in resistance due to the presence of a gas. The drift control circuit also reduces the possibility of output saturation in the subsequent signal conditioning amplifier stages. The problem here is to know the difference between drift and genuine signal without using calibration gases.

In the literature, there have been reports of several approaches to deal with the aforementioned challenges. However, simple voltage divider or Wheatstone bridge techniques that are popular to determine unknown resistances are not suitable at the CMOS level because of the need for trimming (or variable resistance) and also a large resistor value demands a large silicon area (prohibitively expensive). One of the earlier reports came from the Baltes group of *ETH Zurich*, where they proposed logarithmic interface circuit (shown in Figure 9.5) for the sensing material. Although lacking in resolution (~8 bit) and linearity, this simple circuit was able to cover a wide resistance range of sensing material due to compression [8].

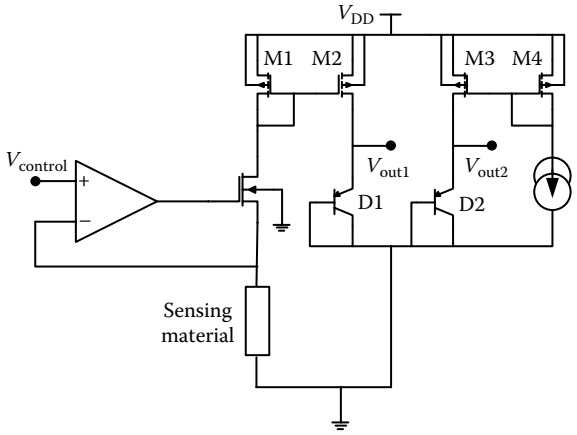


FIGURE 9.5 A logarithmic interface circuit for the sensing material. This simple circuit was able to cover a wide resistance range of sensing material due to compression. V_{DD}, supply voltage; M1, M2, M3, and M4, MOSFETs for current source; D1 and D2, diode temperature sensors. (Adapted from Barrettino, D. et al., A smart single-chip micro-hotplate-based chemical sensor system in CMOS technology, in *Proceedings of the International Symposium on Circuits and Systems*, Phoenix, AZ, pp. 157–160, 2002.)

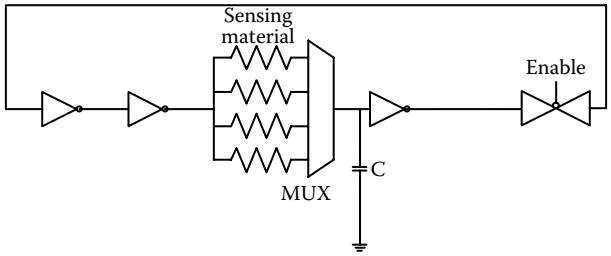


FIGURE 9.6 An array of sensors within the chain of a ring oscillator. This simple scheme can cover a wide resistance range. MUX, multiplexer; C, external capacitor. (Adapted from Merino, J. L. et al., *IEEE Trans. Instrum. Meas*., 53, 1173, 2004.)

The other key scheme found in the literature was resistance-to-frequency conversion. In [20], Merino et al. reported an interfacing circuit (shown in Figure 9.6) where they inserted an array of sensors within the chain of a ring oscillator. This simple scheme can cover a wide resistance range though oscillator accuracy can be affected by parasitic capacitances associated with sensing material. De Marcellis et al. [21,22] came up with a scheme that can also determine parasitic capacitance along with the change in resistance of the sensing material in the presence of a gas. The resistance-to-frequency conversion circuit is shown in Figure 9.7; this contains two comparators, an inverting integrator and an EX-OR digital logic block. The parasitic capacitor role can also be isolated if the resistance of the sensing material is not directly in the oscillating section of the circuit. Such circuit schemes were reported in [23–25]. The circuit structure for this can be achieved by using the sensing material resistor as a voltage-to-current converter or at the reference arm of a current mirror and then using that mirror current to charge and discharge a capacitor (as shown in Figure 9.8). The capacitor voltage is then fed into a Schmitt trigger or through two comparators having threshold voltage V_H and V_L; the resultant square wave then controls the charging and discharging switches. This square wave time period contains information of the value of the sensing material resistor.

There have been some reports to remove/compensate drift of sensing material resistance [26–28]. One popular baseline drift removal approach has been reported by Koickal et al. [27] and is shown in Figure 9.9. In this scheme, during the setup phase, each sensor is driven by a small value current source, and the voltage across the sensor is digitally stored using a simple counting A/D converter. This stored value is converted back to an analog signal (using a D/A converter) and then subtracted from the sensor signal, thus removing the baseline resistance.

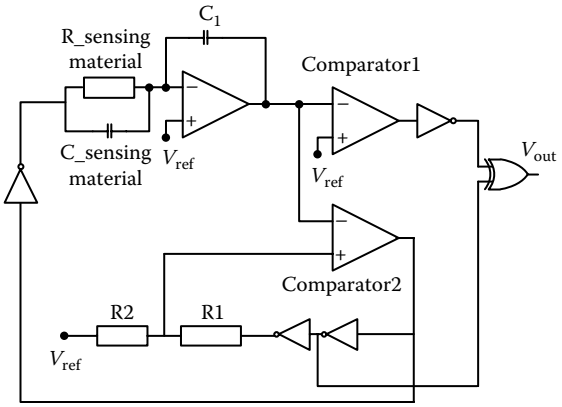


FIGURE 9.7 The resistance-to-frequency conversion circuit contains two comparators, an inverting integrator and an EX-OR digital logic block. V_{ref}, reference voltage generated through voltage divider and buffer; R1 and R2, resistors; V_{out}, output voltage. (Adapted from De Marcellis, A. et al., *Sens. Actuators B Chem*., 132, 477, 2008.)

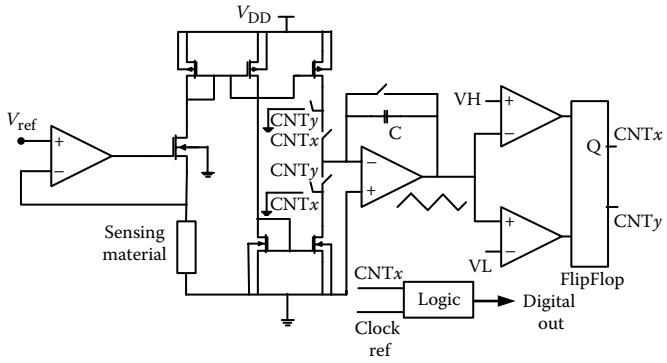


FIGURE 9.8 Using the sensing material resistor as a voltage-to-current converter and then using that mirror current to charge and discharge a capacitor. CNT*x* and CNT*y*, switches to charge (through current source) and discharge (through current sink) the capacitor C; V_{DD}, supply voltage. (Adapted from Grassi, M. et al., *IEEE J. Solid-State Circuits*, 42, 1543, 2007.)

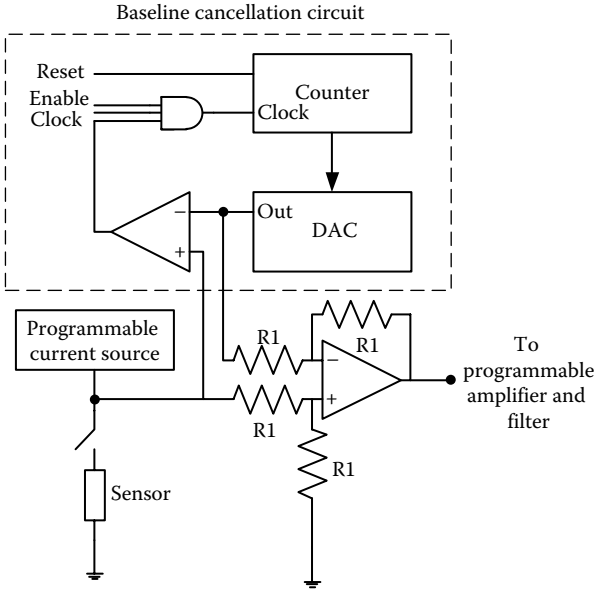


FIGURE 9.9 A drift removal circuit. Each sensor is driven by a small value current source, and the voltage across the sensor is digitally stored using a simple counting A/D converter. The stored value is converted back to an analog signal (using a D/A converter) and then subtracted from the sensor signal, thus removing the baseline resistance. (Adapted from Koickal, T. J. et al., *IEEE Trans. Circuits Syst. I Regul. Pap.*, 54, 60, 2007.)

There has been increasing demand to measure multiple gases simultaneously. For example, in a boiler combustion process, there could be a requirement to measure hydrogen or methane as well as CO, CO_2, and even oxygen. Alternatively, in the case of automotive gases, there is a desire to measure CO, NO_2, and unburnt hydrocarbons. In some ways, this is the same requirement faced when developing electronic nose technology. Dickson and Goodman [29] reported an array of 18 × 18 chemical sensors by combining polymer-based chemoresistors with a standard integrated circuit technology to classify different odors. The corresponding circuit scheme consists of row and column selection (M1–M4 transistor switches) to select sensor and readout circuit in the form of a transmission gate (M5–M9 transistor switches)—which passes the sensor voltage to a column output bus for amplification and off-chip processing (as shown in Figure 9.10).

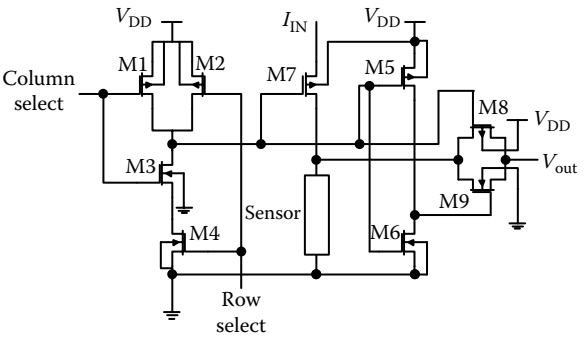


FIGURE 9.10 A typical circuit to measure multiple gases simultaneously. It consists of row and column selection (M1–M4 transistor switches) to select sensor and readout circuit in the form of a transmission gate (M5–M9 transistor switch)—which passes the sensor voltage to a column output bus for amplification and off-chip processing. I_{IN}, excitation current for sensor. (Adapted from Dickson, J. A. and R. M. Goodman, Integrated chemical sensors based on carbon black and polymer films using a standard CMOS process and post-processing, in *Proceedings of the International Symposium on Circuits and Systems*, Presses Polytech. Univ. Romandes, Geneva, Switzerland, pp. 341–344, 2000.)

9.4 Sensing Materials

Modern solid-state resistive gas sensors are generally based on wide-bandgap metal semiconducting oxides (e.g., tin oxide, tungsten oxide, and zinc oxide). The metal oxides are popular as a sensing layer because of their large response in the presence of a gas at part per million (ppm) level or below and their ability to respond to different reactive gases. Resistive sensors commercially available today are usually made of bulk materials (i.e., materials with large grains) as the chemical sensing layer. However, sensors developed on a CMOS platform are miniaturized in size, so large grain materials are not suitable for such sensors (because of their low surface area) to get proper sensitivity. Hence, nanosensing materials (large surface area gives better sensitivity) are the primary contender for a new generation of gas sensors. Researchers have already reported promising performances from various gas detection experiments.

The sensing material growth for CMOS-based gas sensors is very challenging. This is because of the constraints imposed by the CMOS platform, for example:

1. High temperature is often required to synthesize nanomaterials. This can be detrimental for on-chip electronics, particularly for metal tracks (made with aluminum) because of the electromigration issue. However, this metal electromigration effect can be reduced, if one can use metals with higher melting point (e.g., tungsten) rather than aluminum, or voltage switching.
2. Harsh environments (e.g., plasma treatment) and harsh chemicals are also frequently required for nanomaterial synthesis; these can damage semiconductor devices. Particularly, this is detrimental for any fragile structure (e.g., thin membrane-based microhotplates) present in the die.

Hence, novel CMOS friendly methods are being developed for nanomaterial synthesis [30,31]. These new methods use lower temperature (<400 °C) techniques and also less harsh chemicals.

Instead of using conventional techniques (e.g., chemical vapor deposition, aerosol-assisted chemical vapor deposition, atomic layer deposition, flame spray pyrolysis to name a few) to synthesize nanomaterials, techniques like spray coating and even ink-jet printing can be very useful for depositing materials (like metal oxides, CNTs, and nanocomposites) on specified areas. This is usually done with commercially available nanomaterials, so one can avoid a harsh environment that is often necessary to grow nanomaterials. Nowadays, printers are available that can deposit a volume even as low as nanoliters with high precision, and these can be computer controlled [32,33]. Even though these methods are simple and CMOS friendly, it is necessary to dissolve the nanomaterials properly to avoid any clogging at the nozzle.

Although metal oxides are popular as the sensing layer, a well-known problem with these sensing materials is their lack of selectivity. This situation is usually dealt with by using an array of sensors covered with different sensing layers, multicomponent or pattern recognition algorithms (e.g., principal component regression (PCR)), and multiway analysis or by doping the sensing materials (e.g., palladium doping makes metal oxide more selective toward hydrogen). This doping can be done by either physical (e.g., coating through sputtering) or chemical (e.g., with proper chemical solution) means. This functionalization is usually achieved using metal or metal oxide or even with polymer and its composites. Researchers have also been developing mixed metal oxides (two or more metal oxides together) to improve the selectivity. The selectivity can also be improved by the analysis of transient sensor responses to changes in analyte concentration or sensor temperature. The last method that has been extensively investigated [34–37] involves controlling the temperature of the sensing material surface, whether by selecting a fixed temperature to maximize sensitivity to a particular analyte or by programming or modulating the temperature—through these, a single sensor device can provide the type of selectivity that would otherwise require arrays of various doped fixed temperature sensors [15].

Apart from metal oxides, other new-generation materials that have recently been investigated as sensing materials are CNTs (multiwall and single wall) and graphene. As already mentioned, metal oxides react with analytes at elevated temperature (200 °C–500 °C), but CNTs can respond to gases at much lower temperature (even at room temperature). So CNT/graphene-based sensors will be more power efficient compared to their metal-oxide counterpart.

CNTs are usually synthesized using CVD technique at a very high temperature (usually >700 °C). Recently, a unique method of growing CNTs on the CMOS microhotplate was demonstrated by Santra et al. [30] using local growth technique. This is also primarily a CVD technique where instead of heating the whole chip, a microhotplate was used to form small catalytic islands and then to form CNTs by dissociating gas (SEM CNT picture is shown in Figure 9.11). This way, the whole chip will not be exposed to high temperature, and on-chip electronic circuits will remain intact.

The operational principle of graphene as a gas-sensing layer is believed to be similar to that of metal oxides where conductivity changes due to adsorbed gas molecules at the graphene surface (which act as donors or acceptors). However, graphene has extra edge over metal oxides due to its unique properties, for example,

1. As graphene is a 2D material, its whole volume is exposed to the analytes, which increases the sensitivity.
2. It has high conductivity (which implies very low Johnson noise), so even a few extra electrons will give a notable change in carrier concentration.

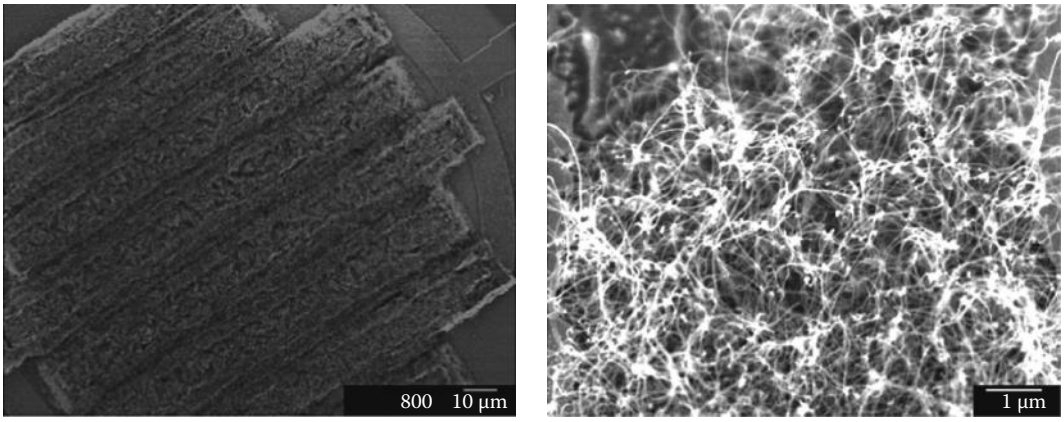


FIGURE 9.11 SEM CNT picture of a CVD technique where instead of heating the whole chip, a microhotplate was used to form small catalytic islands and then to form CNTs by dissociating gas. (Adapted from Santra, S. et al., *Nanotechnology*, 21, 485301, 2010.)

Unfortunately, until now, graphene synthesis is tedious and unreliable; this was usually performed by the mechanical peeling of graphite layers. Recently, researchers have developed chemical methods to generate device quality graphene flakes coated on a wafer scale. Often, this involves the chemical oxidation and exfoliation of graphite to produce graphite oxide and its subsequent reduction to graphene. Recently, Fowler et al. [38] reported the development of chemically converted graphene using spin coating of hydrazine dispersions on interdigitated electrodes, which is integrated on microhotplates. The graphene showed good response in the presence of NH_3 and NO_2. Schedin et al. [39] have demonstrated the ability of detecting individual gas molecule by graphene layer.

Instead of using nanosensing materials in chemoresistive mode, nanomaterials particularly nanowires/nanotubes can be used as FET devices for gas sensor applications [40]. In such devices, the Fermi level within the bandgap of the nanowire can be varied and used to control surface process electronically. The nanowire acts as a conductive channel that joins source and drain electrode. Two methods are common to produce such an FET structure, the first one consisted of nanowire dispersion in ethanol by ultrasonication and drying on a silicon dioxide–silicon substrate and then in electrode deposition by electron beam lithography. The second procedure consisted of deposition of dispersed nanowires on predefined electrode arrays. However, such an approach has to be more automated so that devices on a wafer level are possible to develop.

9.5 Conclusion and Future Outlook

Most commercial gas sensors today are made in a batch process, have an average cost around $25, and are about 1 cm^3 in volume or consume 100 mW of power. Consequently, they are not compatible with the mass market (e.g., 500 million mobile phones that were sold in 2012) that demands a small size (ideally the size of a microSD package) and have low cost (much less than $1). In this chapter, we have discussed the development of CMOS integrated gas sensors that offer the potential to replace existing commercial gas sensors, such as the electrochemical cell, pellistor, NDIR, and Taguchi type. This technology offers the potential not only to make low-cost CMOS gas-sensing elements but also to integrate associated circuitry to make smart sensing modules. The full integration of the sensing element, drive circuitry, and communication interface enables the cost per unit to fall in volume (millions) to a few tens of cents and thus satisfies the emerging market of mobile phones, tablets, and other electronic goods. We believe that the integration of CMOS technology together with a new generation of nanomaterials (such as nanowires, nanotubes, and nanolayers) will ultimately satisfy the mass market and thus make CMOS integrated sensors a ubiquitous commodity just like PCs, mobile phones, and cars!

References

1. Gas Sensor Market-Global Industry size, Share, Trends, Analysis, and Forecast, 2012-2018, PR Newswire, http://www.prnewswire.com/news-releases/gas-sensors-market—global-industry-size-share-trends-analysis-and-forecast-2012—2018-219628201.html, accessed on August 19, 2013.
2. NASA's Smartphone Attachement Smells, Identifies Chemical Compounds, Regator, http://regator.com/p/254799922/nasas_smartphone_attachment_smells_identifies_chemical_compounds/, accessed on August 19, 2013.
3. Technology Reports, Breath Acetone Analyzer to Achieve "Biochip Mobile Terminal",, NTT DOCOMO, Inc., http://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/rd/technical_journal/bn/vol14_1/vol14_1_051en.pdf#page=1, accessed on August 19, 2013.
4. Multi-Sensor remote Communication, Nokia Scentsory Phone, TrendHunter, http://www.trendhunter.com/trends/nokia-scentsory-concept, accessed on August 19, 2013.
5. Operating Principles of Figaro Gas Sensors, Technical Information, Figaro, http://www.figarosensor.com/, accessed on August 19, 2013.

6. Afridi, M.Y. et al., A monolithic CMOS microhotplate-based gas sensor system. *IEEE Sensors Journal*, 2: 644–655, 2002.
7. Afridi, M.Y. et al., A monolithic implementation of interface circuitry for CMOS compatible gas-sensor system, in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Phoenix, AZ, 2002.
8. Barrettino, D. et al., A smart single-chip micro-hotplate-based chemical sensor system in CMOS-technology, in *Proceedings of the International Symposium on Circuits and Systems*, Phoenix, AZ, pp. 157–160, 2002.
9. Covington, J.A., F. Udrea, and J.W. Gardner, Resistive gas sensor with integrated MOSFET micro hot-plate based on an analogue SOI CMOS process, in *Proceedings of the IEEE Conference on Sensors*, Orlando, FL, 2002.
10. Kloeck, B. et al., Study of electrochemical etch-stop for high-precision thickness control of silicon membranes. *IEEE Transactions on Electron Devices*, 36(4): 663–669, 1989.
11. Müller, T. et al., An industrial CMOS process family adapted for the fabrication of smart silicon sensors. *Sensors and Actuators A*, 84(1–2): 126–133, 2000.
12. Ali, S.Z. et al., Tungsten-based SOI microhotplates for smart gas sensors. *Journal of Microelectromechanical Systems*, 17(6): 1408–1417, 2008.
13. Guha, P.K. et al., Novel design and characterisation of SOI CMOS micro-hotplates for high temperature gas sensors. *Sensors and Actuators B*, 127(1): 260–266, 2007.
14. Products, Micro-hoyplates, Cambridge CMOS Sensors, http://www.ccmoss.com/microhotplates.html, accessed on August 19, 2013.
15. Gardner, J.W. et al., CMOS interfacing for integrated gas sensors: A review. *IEEE Sensors Journal*, 10: 1833–1848, 2010.
16. Bota, S.A. et al., A monolithic interface circuit for gas sensor arrays: Control and measurement. *Analog Integrated Circuits and Signal Processing*, 40: 175–184, 2004.
17. Cardinali, G.C. et al., A smart sensor system for carbon monoxide detection. *Analog Integrated Circuits and Signal Processing*, 14(3): 275–296, 1997.
18. Barrettino, D. et al., CMOS-based monolithic controllers for smart sensors comprising micromembranes and microcantilevers. *IEEE Transactions on Circuits and Systems I*, 54(1): 141–152, 2007.
19. Barrettino, D. et al., A single-chip CMOS micro-hotplate array for hazardous-gas detection and material characterization, in *Proceedings of the International Solid-State Circuit Conference*, L.C. Fujino, Ed., San Francisco, CA, pp. 312–313, 2004.
20. Merino, J.L. et al., A reusable smart interface for gas sensor resistance measurement. *IEEE Transactions on Instrumentation and Measurement*, 53(4): 1173–1178, 2004.
21. De Marcellis, A. et al., Uncalibrated integrable wide-range single-supply portable interface for resistance and parasitic capacitance determination. *Sensors and Actuators B*, 132(2): 477–484, 2008.
22. De Marcellis, A. et al., A CMOS integrable oscillator-based front end for high-dynamic-range resistive sensors. *IEEE Transactions on Instrumentation and Measurement*, 57(8): 1596–1604, 2008.
23. Grassi, M., P. Malcovati, and A. Baschirotto, A 141-dB dynamic range CMOS gas-sensor interface circuit without calibration with 16-bit digital output word. *IEEE Journal of Solid-State Circuits*, 42(7): 1543–1554, 2007.
24. Leung, C.K. and D.M. Wilson, Integrated interface circuits for chemiresistor arrays, in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Kobe, Japan, pp. 5914–5917, 2005.
25. Jayaraman, B. and N. Bhat, High precision 16-bit readout gas sensor interface in 0.13 μm CMOS, in *Proceedings of the IEEE International Symposium on Circuits and Systems*, New Orleans, LA, pp. 3071–3074, 2007.
26. Koickal, T.J. et al., Smart interface circuit to ameliorate loss of measurement range in chemical microsensor arrays, in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference*, Ottawa, Ontario, Canada, pp. 548–550, 2005.

27. Koickal, T.J. et al., Analog VLSI circuit implementation of an adaptive neuromorphic olfaction chip. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 54(1): 60–73, 2007.
28. Rairigh, D. et al., Baseline resistance cancellation circuit for high resolution thiolate-monolayer-protected gold nanoparticle vapor sensor arrays, in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Seattle, WA, pp. 2002–2005, 2008.
29. Dickson, J.A. and R.M. Goodman, Integrated chemical sensors based on carbon black and polymer films using a standard CMOS process and post-processing, in *Proceedings of the International Symposium on Circuits and Systems*, Presses Polytech. Univ. Romandes, Geneva, Switzerland, pp. 341–344, 2000.
30. Santra, S. et al., Post-CMOS wafer level growth of carbon nanotubes for low-cost microsensors—A proof of concept. *Nanotechnology*, 21: 485301(7pp.), 2010.
31. Santra, S. et al., ZnO nanowires grown on SOI CMOS substrate for ethanol sensing. *Sensors and Actuators B*, 146: 559–565, 2010.
32. NonoPrinting, Altatech, http://www.altatech-sc.com/en/products/altadrop.php, accessed on August 19, 2013.
33. Benchtop molecular printer, Nano technology and mems, Tesscorn, http://www.tesscorn.com/nanotechnology.htm, accessed on August 19, 2013.
34. Heilig, A. et al., Gas identification by modulating temperatures of SnO_2-based thick film sensors. *Sensors and Actuators B*, 43: 45–51, 1997.
35. Kunt, T.A. et al., Optimization of temperature programmed sensing for gas identification using micro-hotplate sensors. *Sensors and Actuators B*, 53(1–2): 24–43, 1998.
36. Meier, D.C. et al., The potential for and challenges of detecting chemical hazards with temperature-programmed microsensors. *Sensors and Actuators B*, 121(1): 282–294, 2007.
37. Lee, A.P. and B.J. Reedy, Temperature modulation in semiconductor gas sensing. *Sensors and Actuators B*, 60(1): 35–42, 1999.
38. Fowler, J.D. et al., Practical chemical sensors from chemically derived graphene. *ACS Nano*, 3(2): 301–306, 2009.
39. Schedin, F. et al., Detection of individual gas molecules adsorbed on graphene. *Nature Materials*, 6: 652–655, 2007.
40. Comini, E., Metal oxide nano-crystals for gas sensing. *Analytica Chimica Acta*, 568: 28–40, 2006.

10

Energy Harvesting for Sensors: DC Harvesters

Oscar Lopez-Lapeña
Universitat Politècnica de Catalunya

Maria Teresa Penella
Urbiotica S.L.

Manel Gasulla
Universitat Politècnica de Catalunya

Nowadays, photovoltaic (PV) panels and thermoelectric generators (TEGs) constitute the most important direct current (dc) energy transducers that can be used to harvest energy from the environment and power an autonomous sensor node. They generate a direct current and voltage for constant operating conditions. A current–voltage (I–V) curve is defined for each operating condition in which an optimum bias point exists where maximum power is achieved. The voltage at the maximum power point (V_{MPP}) depends on the operating conditions and it changes as these conditions vary with time. Several energy conditioners have been proposed and commercialized to transfer the energy to a storage unit or a load. The small magnitudes of the currents and voltages of such transducers constitute the main challenge in the design of these circuits. Ultralow-power consumption energy conditioners have to be designed to make feasible the use of these power sources.

This chapter presents a description of these energy transducers and their associate energy conditioning circuits.

10.1 Photovoltaic Energy Harvesting

PV technology is the most widespread form of energy harvesting because of its availability, low cost, and power density. PV panels generate electric energy when exposed to light irradiance. Consisting of semiconductor junctions (P–N), their operation is based on the photons' impact ionization effect. Each one of these P–N junctions is known as a *PV cell* and can be associated in series and parallel to form a PV panel. The generated current (I) depends on the bias voltage (V), the temperature (T), and the light irradiance (G). Figure 10.1 shows the symbolic representation of a PV panel, its equivalent electrical model, and a couple of graphs showing the current dependency on the bias voltage and on environmental

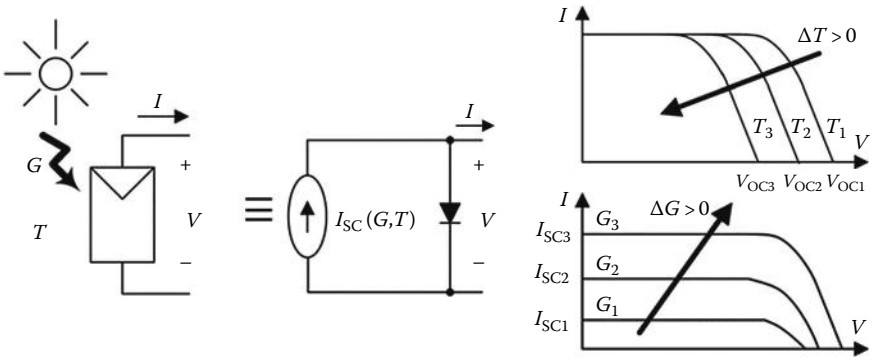


FIGURE 10.1 On the left, the symbolic representation of a PV panel and its equivalent electrical model. On the right, *I*–*V* curves of a PV panel for several temperatures (*T*) and irradiances (*G*) with operating conditions ($G_3 > G_2 > G_1$ and $T_3 > T_2 > T_1$).

conditions (*G* and *T*). For a given *T*, the *short circuit current* (I_{SC}) is proportional to the light irradiance. The current *I* is slightly reduced when the voltage *V* is below a threshold value and is abruptly reduced above it. Considering a single PV cell, this voltage limit corresponds to the diode threshold voltage in the equivalent model of Figure 10.1. The *open-circuit voltage* (V_{OC}) delimits the maximum voltage range in which the PV transducer provides energy and works as a power source.

As shown in Figure 10.1, the values of I_{SC} and V_{OC} depend on the temperature and irradiance operating conditions. PV panel manufacturers conventionally specify the panel parameters under standard test conditions (STCs) ($T = 25$ °C, $G = 1000$ W/m² and light spectrum AM1.5). They usually indicate I_{SC}, V_{OC}, and *the nominal power. Nominal power* is defined as the maximum power achieved from the *I*–*V* curve under STCs.

PV panels can be connected in series or parallel to achieve, respectively, higher voltage or current levels. If all of them have identical operating conditions and performance, the resulting power is scaled to the number of PV panels. Otherwise, for instance, in a shadow projection that produces differences in the light irradiance conditions, the current in a series association is limited by the I_{SC} of the less illuminated PV panel. In the case of a parallel association, the less illuminated PV panel could work as a load reducing the global generated power. To reduce these power losses, a diode can be connected in series with each of the parallel connected PV panels. In the series configuration, an antiparallel diode must be connected instead. Although both solutions can reduce these power losses, identical PV panels and physical orientation are recommended whenever they are connected in series or parallel.

Light efficiency (η_{PV}) indicates how much light energy is converted to electric energy. The efficiency depends on the PV panel technology and on the light spectrum. Although new technologies such as dye-sensitized solar cells are becoming commercially available, silicon PV panels are still the most used because of their low cost and high efficiency. Basically, there are three technologies based on silicon: amorphous, polycrystalline, and monocrystalline PV panels. Each is designed to satisfy a different trade-off between price and efficiency. While amorphous panels have the lowest efficiency (~5%), they are the cheapest ones and are more suited to indoor applications where an artificial light source is used. On the other hand, monocrystalline panels are the most efficient (~20%) for sunlight but they are also the most expensive.

PV energy harvesters can be used to power outdoor or indoor autonomous sensors. Figure 10.2 shows typical irradiance conditions in both environments. In a well-lit office, around 500 lux, the surface of a PV panel would be submitted to a *G* around 2 W/m² (for an incandescent light). When using an amorphous panel, the resulting output power density is about 100 mW/m². These power levels are much lower than the power of an outdoor monocrystalline panel. In this case, the irradiance varies with the time and can reach 1000 W/m². Therefore, the output power density could be as high as 200 W/m². So, the power availability, and thus the design constraints for PV systems, is very different indoors and outdoors.

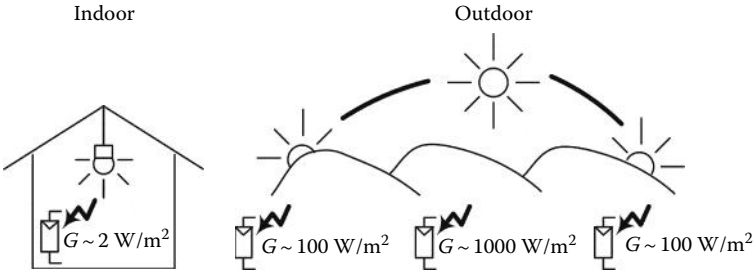


FIGURE 10.2 Typical irradiance (G) conditions in an indoor and an outdoor application.

Design engineers must size PV panels to overcome the average power consumed by the overall sensor node ($\bar{P}_L$), including the energy harvester. If the PV panel works at its maximum power point (MPP), the minimum nominal power under STC (P_{PV}) is determined by

$$P_{PV} > \frac{\bar{P}_L \times 1000\ \text{W/m}^2}{\eta_{Tot} \times \bar{G} \times (1 - \delta_P \times (T_{max} - 25\ °\text{C}))} \tag{10.1}$$

where $\bar{G}$, η_{Tot}, and δ_P are, respectively, the average light irradiance on the surface of the PV panel, the overall efficiency of the energy harvesting system, and the temperature degradation factor of the PV power. The temperature T_{max} is the maximum temperature of the PV panel, which will be higher than the ambient temperature. Both external parameters, such as wind speed, and internal parameters of the PV panels, such as the surface protection material, determine the temperature increment. At G = 800 W/m², 20 °C ambient temperature, and wind speeds of 1 m/s, temperature increases up to 42 °C and 48 °C typically.

The optimal size and orientation of an outdoor PV panel is not evident because the Sun position varies with time. The panel surface must be perpendicular to the light rays to achieve maximum irradiance (G_{Opt}). Figure 10.3 shows the calculation of the irradiance for a surface tilted by angle β. If β is different from the optimum tilt angle (β_{Opt}), the resulting irradiance and generated power is lower. Hence, the sunlight direction must be taken into account to select the panel orientation. On the Earth surface, the sunlight can be split into three components: direct sunlight, diffuse sunlight, and albedo. Diffuse sunlight and albedo are, respectively, the light scattered by molecules and particles in the atmosphere and the light reflected from the Earth surface. Both of them are usually neglected with respect to the direct sunlight to determine the optimal panel orientation. The direction of direct sunlight is described by the Sun position.

The orientation of a PV panel is defined by the azimuth (α) and the tilt (β) angles as shown in Figure 10.4. The azimuth is the angle between the axis of the PV panel and the geographic north direction. The tilt angle is measured between the PV panel surface and the horizontal plane. It can also be defined as the angle between an imaginary line to the center of the Earth and the normal vector of the panel surface.

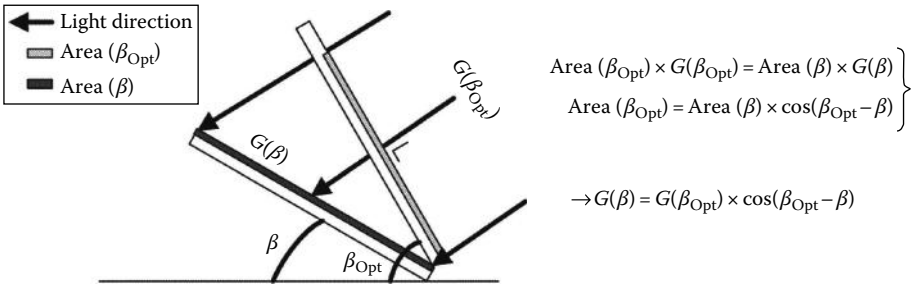


FIGURE 10.3 Calculation of the light irradiance on a surface tilted at angle β.

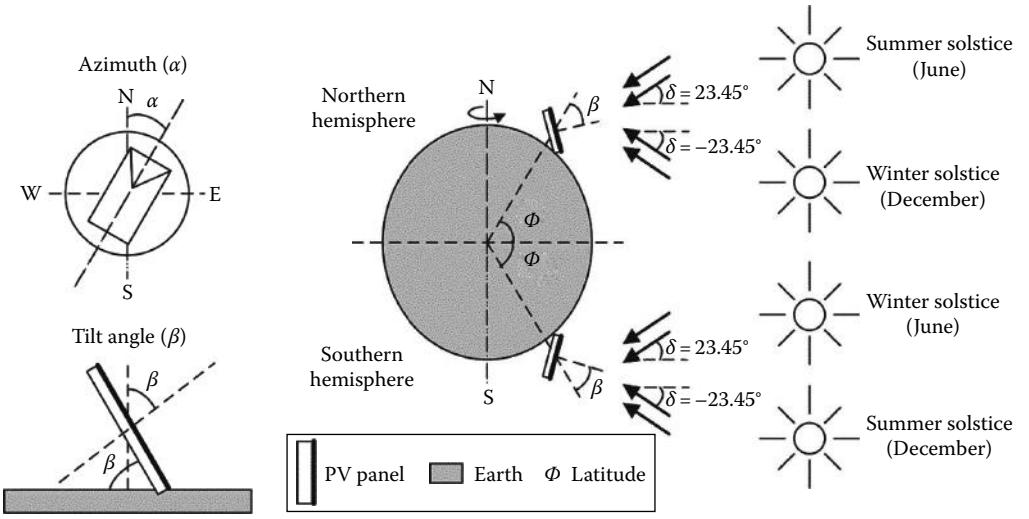


FIGURE 10.4 PV panel orientation and sun declination angle.

The PV system of an autonomous sensor is fixed at a given orientation and tilt because the energy benefit from an automatic orientation system does not compensate for the additional power consumption and cost. The optimal fixed azimuths are respectively 0° (toward the south) and 180° (toward the north) in the northern and southern hemispheres. On the other hand, the optimal fixed tilt angle depends not only on the hemisphere where the PV panel is placed but also on the latitude, the season of the year, and the distribution of the power consumption of the autonomous sensor along the time.

Figure 10.4 shows the variation of the direct sunlight direction during the year, described by the declination angle (δ). δ is the angle between the sun's rays and the equatorial plane. The optimal tilt for the winter solstice is latitude (Φ) minus 23.45° at that solstice (i.e., for the winter solstice and in the northern hemisphere, the optimal tilt will be Φ + 23.45°). Take into account that both the Summer and Winter solstices take place on opposite dates in both hemispheres and thus, δ has a different sign for the same solstice. Φ also has a different sign in both hemispheres.

For long operation periods of the autonomous sensor, the proper tilt angle for the worst-case scenario must be considered. The required P_{PV} for any available tilt angle must be evaluated for each month, using Equation 10.1 and determining the best angle. In this way, an optimum angle and a minimum P_{PV} is determined for each month. The optimum angle for the overall period will be the optimum angle of the month in which maximum P_{PV} is required. In the case of a constant value of $\overline{P_L}$ along the year, the worst case will be the month with the lowest irradiation, which typically is December and June for the northern and southern hemispheres, respectively. Thus, the proper tilt angle will be $\Phi - \delta$ (+23.45° in the northern hemisphere and Φ − 23.45° in the southern hemisphere). Even so, it is not recommended to use tilt angles lower than 15° to let the water from rainfalls slide and avoid the accumulation of dust.

An energy conditioning circuit is used to transfer energy from a PV panel to a storage unit, that is, a battery or a (super) capacitor. Two types of energy conditioners can be distinguished: the *direct coupled* and the *maximum power point trackers* (MPPTs). Figure 10.5 shows these two alternatives by using a battery as the storage unit.

In order to obtain the average light irradiance for a tilt angle and a period of time, meteorological records kept by weather stations all over the world can be used. Although the measurements are mostly limited to a horizontal surface ($G_{0°}$), the irradiance (G_β) on a surface tilted an angle β can be deduced from β_{Opt} and the relationship shown in Figure 10.3. The ratio between G_β and $G_{0°}$ is called the tilt factor and depend on β and on the latitude of the PV panel. Tilt factor tables [1] or interactive databases such as SoDa [2] can be used to estimate the average light irradiance on a PV surface for any tilt angle and period of time.

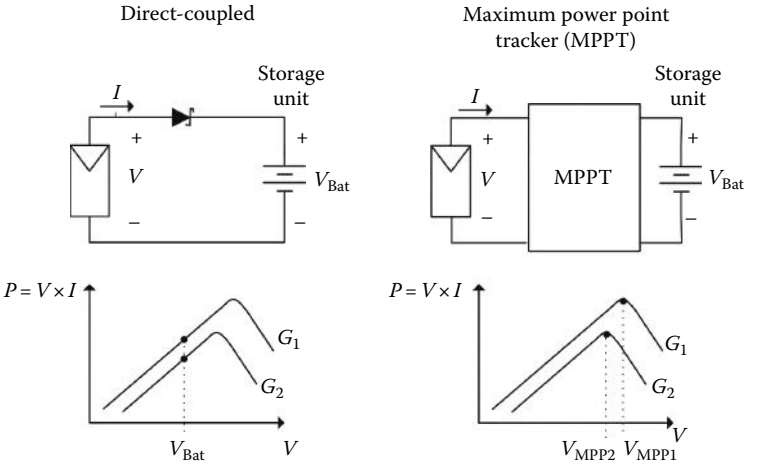


FIGURE 10.5 Energy conditioner schemes.

The *direct coupled* connection is the simplest solution because it simply uses a Schottky diode. The diode permits charging the storage unit from the PV panel and prevents its discharge when V_{OC}, mostly at low irradiance condition, is below the storage unit voltage (V_{Bat}). It is a proper solution for indoor PV panels, where the generated power is very low, because the power loss is limited to that of the diode. Unfortunately, the transducer voltage is fixed by the storage voltage and so, the transducer works outside its MPP. Therefore, Equation 10.1 cannot be used to size the PV panel. Instead, I_{SC} at STC is calculated as:

$$I_{SC} > \frac{\overline{P}_L \times 1000\ \text{W/m}^2}{\eta_{Tot} \times \overline{G} \times V_{Bat} \times (1 - \delta_I \times (T_{min} - 25\ ^\circ\text{C}))} \tag{10.2}$$

where δ_I and T_{min} are, respectively, the temperature degradation factor of the I_{SC} and the minimum temperature of the PV panel. In this design condition, $V_{Bat} < V_{MPP}$ was assumed to approximate the current I by I_{SC}. To account for variations of V_{Bat} with temperature and of V_{MPP} with temperature and irradiance, a conservative (low) value of V_{Bat} ($<V_{MPP}$) is chosen, which leads to oversizing the PV panel. As $V_{Bat} < V_{MPP}$, I_{SC} can be used instead of I in Equation 10.2.

In applications where V_{Bat} is so small that significant power losses are caused by the diode voltage drop, a transistor can be used instead. The SPV1001 of STMicroelectronics is an implementation example of such alternative solution [3].

MPPTs are more sophisticated energy conditioners that bias PV panels at their MPP. Although they have been widely used in high-power PV systems [4], their use in low-power panels is recent. To be feasible, the power benefit caused by following the MPP must outperform the power consumption of the power conditioning circuit. Although this design goal can be easily reached for high-power PV panels, it is more challenging for low-power PV panels (<1 W) used in low-power autonomous sensors. High-performance controllers that were used for high-power PV systems, such as digital signal processors (DSPs), are not feasible for small autonomous sensors because of their high cost and power consumption.

An MPPT is composed of a dc/dc switching converter and a tracking controller (see Figure 10.6). The converter biases the PV panel to a reference voltage given by the tracking controller. The controller measures the appropriate electrical variables of the PV panel (e.g., I and V) to determine the optimal reference voltage that makes the panel work at its MPP.

Switching converters are chosen because of their high efficiencies. Figure 10.7 shows a boost converter that can be used whenever the storage unit voltage is higher than the voltage of the PV panel. The state of

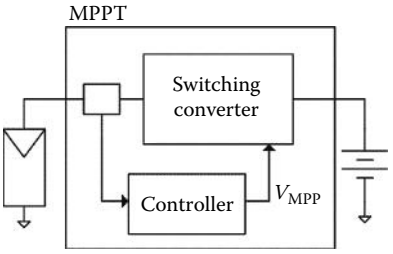


FIGURE 10.6 Block diagram of an MPPT.

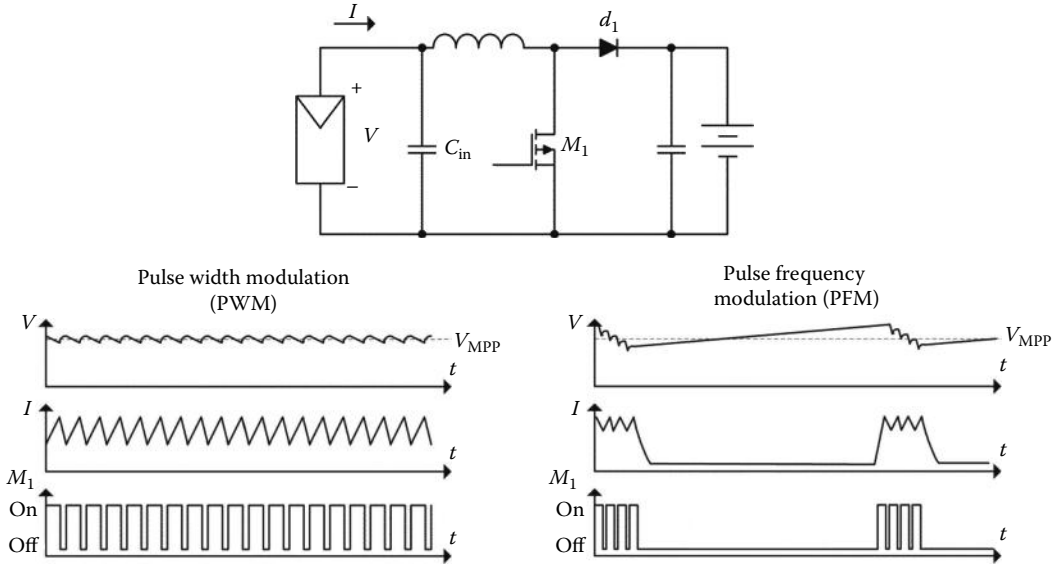


FIGURE 10.7 Switching converter modulation schemes.

the transistor can be controlled by a pulse width modulator (PWM) or by a pulse frequency modulator (PFM). The traditional PWM used in switching converters is not the most suitable solution in low-power applications. A periodic switching sequence holds the transistor in on state during a portion of the sequence period, called *duty cycle*. The voltage of the PV panel is adjusted by means of the *duty cycle*. As the transistor is continuously switching, the resulting power consumption is too high for low-power PV panels (<1 W). Instead, PFM can be used to further reduce power consumption. The converter remains inactive, with low-power consumption, during a long period of time while the capacitor C_{in} is charged by the PV panel. When the voltage V reaches a higher threshold value, the converter is activated to discharge the capacitor to a lower threshold value. In this way, V is held inside a small hysteresis cycle that is centered in V_{MPP}. In [5,6], the authors propose implementations of the PFM switching converter using commercial voltage regulators. There are also commercial MPPT ICs, such as the BQ25504 from Texas Instruments, that uses a PFM switching converter and achieves a current consumption as low as 330 nA [7].

The most popular MPP tracking controllers are *the fractional open-circuit voltage* (FOCV) and the *perturbation and observe* (P&O). The FOCV is an open-loop controller that is based on the empirical relationship between V_{OC} and V_{MPP} of the PV panel. The value of V_{MPP}, which varies with the light irradiance and the temperature, is nearly proportional to V_{OC}. Figure 10.8 shows the functional scheme of this controller and the factor of proportionality, K. The PV panel is periodically disconnected and V_{OC} is captured by a *sample and hold* (S&H) to obtain V_{MPP} through the empirical relationship. This is the most usual MPPT controller for low-power PV panels because of its simplicity [7–9].

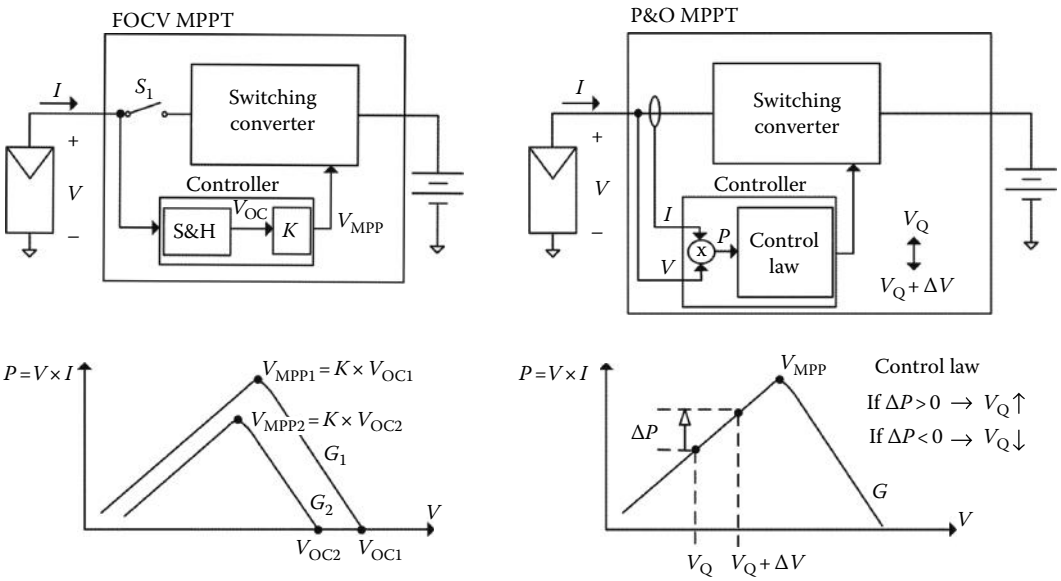


FIGURE 10.8 Low-power MPP tracking controllers.

On the other hand, the P&O is a more sophisticated and accurate tracking controller. It is a closed-loop controller that reaches the MPP by comparing the PV power at two bias points. Figure 10.8 shows the control law that determines if the bias voltage (V_Q) must be increased or decreased to reach the V_{MPP}. The variation of the PV power (ΔP) is measured between one bias voltage (V_Q) and another bias voltage ($V_Q + \Delta V$). Considering a positive perturbation in the voltage (ΔV), the working voltage will be increased if ΔP is higher than zero. Otherwise, it will be decreased. The SPV1040 from STMicroelectronics is an example of a commercial MPPT IC that is based on a P&O and is intended for low-power applications, such as small autonomous sensors [10]. Discrete solutions that implement the control law in a microcontroller [6,11] or using discrete analog circuits [5] can also be found in the literature. In all of them, the main design goal is to reduce the power consumption of the controller. A possible solution consists of reducing the activity of the microcontroller by using an ultralow frequency clock [11]. Nevertheless, such slow microcontrollers cannot be used to implement other tasks of the sensor node. Prediction algorithms, such as those proposed in [6], allow the application of *dynamic power management* techniques that reduce the power consumption, making feasible the use of high-speed microcontrollers.

Another alternative energy conditioning solution is provided by the LT3105 from Linear Technology [12]. The switching converter biases the PV panel to a fixed voltage that is configured in the design process. This is a proper solution if small variations of V_{MPP} are expected.

10.2 Thermoelectric Energy Harvesting

TEGs are used to provide electric energy from a temperature difference between a heat source (hot side) and a cold ambient (cold side). The source can be, for instance, a machine, a warm-blooded animal, or a human being whose temperature is different from the environment. Potential applications could be animal monitoring, fire detection, and cars, ships, and aircrafts industry. Besides its applicability in sensor nodes, it can also power small electronic circuits such as the *Seiko Thermic*, which is a wristwatch driven by body heat [13].

As shown in Figure 10.9, a TEG is implemented with a *Peltier cell*, which is physically based on the *Seebeck effect*. The hot side of TEG is thermally connected to the heat source, and the cold side is held

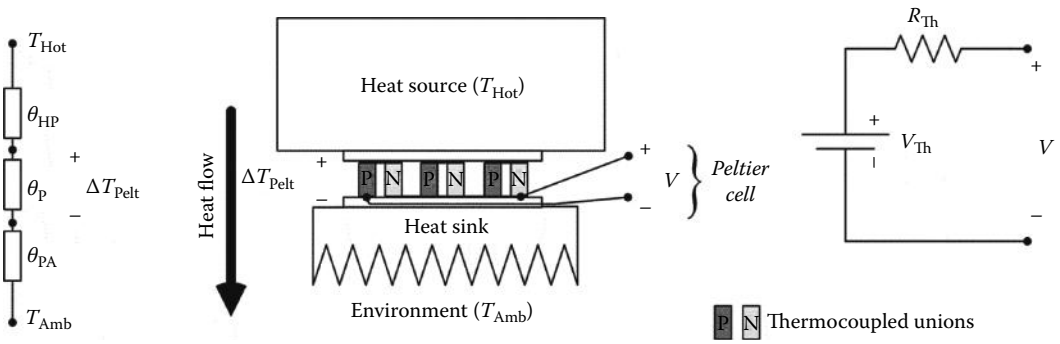


FIGURE 10.9 TEG implemented with a *Peltier cell* connected between a heat source and a heat sink. Equivalent thermal and electrical models are shown, respectively, on the left and on the right.

close to the ambient temperature (T_{Amb}) using a heat sink. The temperature gradient produces a heat flow and a temperature drop (ΔT_{Pelt}) between both faces of the TEG, given by

$$\Delta T_{Pelt} = \frac{\theta_P}{\theta_{HP} + \theta_P + \theta_{PA}} (T_{Hot} - T_{Amb}) \tag{10.3}$$

where θ_{HP}, θ_P, and θ_{PA} are, respectively, the thermal resistances between the heat source and the TEG, between both faces of the TEG, and between the TEG and the environment.

The electrical behavior can be approximated by a Thevenin equivalent circuit. The Thevenin voltage (V_{Th}) is proportional to the TEG temperature drop (ΔT_{Pelt}). The proportionality factor depends on the number (N) of the thermocoupled PN unions of the *Peltier cell* and on the *Seebeck* coefficients of the P-side (S_P) and the N-side (S_N). As the thermocoupled unions are electrically connected in series and thermally in parallel, the value of V_{Th} is

$$V_{Th} = N \times (S_P - S_N) \times \Delta T_{Pelt} \tag{10.4}$$

The Thevenin resistance (R_{Th}) depends on the cross area and length of the thermocoupled unions and on their electrical resistivity.

Nowadays, the thermocoupled unions are built with *bismuth telluride* (BiTe) that has a *Seebeck* coefficient of 0.2 mV/°C ($S_P - S_N$). An example of a commercial *Peltier cell* that could be used as a TEG is the MPG-D751 from Micropelt GmbH [14]. It is formed by 540 thermocoupled unions (N) in an area around 14 mm^2. Despite its small area, the thermal resistance is as small as 12.5 °C/W. The equivalent Thevenin resistance is around 300 Ω and the voltage V_{Th} increases 140 mV/°C at 23 °C. A practical way to increase the generated voltage and power per area consists of using thinner thermocoupled legs. Unfortunately, the industrial technology of BiTe TEG is near its miniaturization limits and it is not possible to decrease the cross section without decreasing their length [15]. An alternative feasible solution consists of using multistage *Peltier cells*. They consist of using two or more stages of *Peltier cells* that are thermally and electrically connected in series. For the same number of thermocoupled unions, the area will be divided by the number of stages and so, higher V_{Th} will be achieved. Figure 10.10 shows the Thevenin voltages for two stages of TEGs thermally connected in series. Considering θ_P much lower than $\theta_{HP} + \theta_{PA}$, the resulting voltage is doubled from a single stage. Nevertheless, the resulting voltage is not doubled if both TEGs are connected thermally in parallel and electrically in series. As the thermal resistances are connected in parallel, the temperature drop is reduced by half and V_{Th} is hardly incremented.

The size and cost design constraints of TEG usually lead to low generated voltages and power levels. In [16], voltages ranging from 150 to 250 mV with a maximum current of 18 mA are achieved with a TEG coupled to a human hand. Such low voltage levels constitute the main challenge when designing a

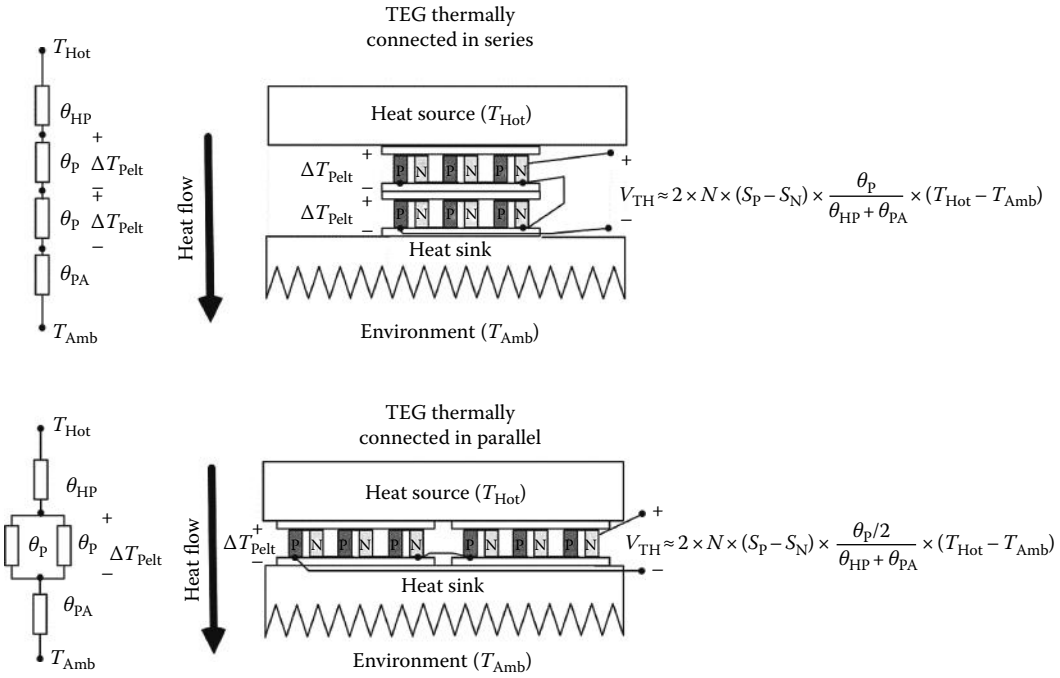


FIGURE 10.10 Comparison between two TEGs thermally connected in series with a parallel connected configuration. To calculate the generated voltage V_{TH}, it was taken into account that both TEGs are electrically series connected and $\theta_P \ll \theta_{HP} + \theta_{PA}$.

power conditioning circuit for the TEG. The minimum start-up voltage of conventional boost (step-up) dc/dc converters is limited to above 0.7 V by the threshold voltage of transistors. If lower input voltages have to be processed, special low-threshold transistors or special power conditioning architectures must be used. An example is the charge pump S-882Z from Seiko that employs silicon on isolation (SOI) transistors with low threshold voltages [17]. The minimum start-up voltage is 300 mV and the quiescent current is 0.6 μA but the power efficiency is below 20%. The manufacturer recommends using S-882Z with a boost switching converter to improve the overall efficiency. The S-882Z is used during the start-up to charge a capacitor until the minimum start-up voltage of the switching converter is reached. The capacitor powers the control circuit of the boost and is charged by the output of the converter through a diode once starts working. Figure 10.11 shows a scheme of the power conditioning circuit and the evolution of the main electrical variables.

An alternative solution to step-up of such low input voltage consists of using an electromagnetic transformer [18]. The induced voltage on the secondary is equal to the voltage in the primary multiplied by the turns ratio. Therefore, the primary voltage will be stepped-up if high turns ratio is selected. Unfortunately, electromagnetic transformers can just step-up an ac voltage source and TEG provides a dc voltage. To overcome this problem, an ac voltage is generated by an oscillator. The oscillator is formed by a JFET transistor that is fed back by a coupled solenoid of the transformer (see Figure 10.12). A JFET transistor is selected instead of bipolar or a MOSFET enhancement-mode channel transistor because it provides a modulated conduction channel from zero input voltage (V_{IN}). Therefore, the oscillator can start working from very low input voltages and a sinusoidal voltage is induced in the secondary (V_S). The output capacitor is charged from V_S through a rectifier diode and powers the control circuit of a boost converter (V_{DD}).

Similar starter circuits are also used by commercial power conditioning ICs such as the LTC3108 from Linear Technology [19]. It is operative from 20 mV input voltage for a 1:100 turns ratio.

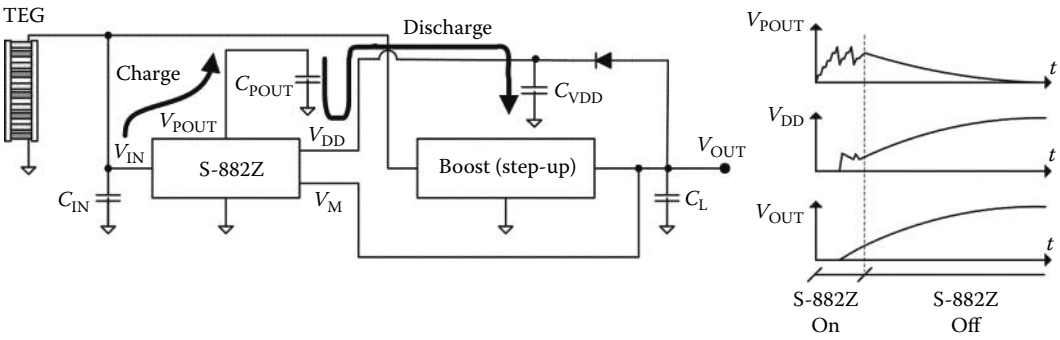


FIGURE 10.11 Scheme of a TEG power conditioner based on the charge pump S-882Z from Seiko.

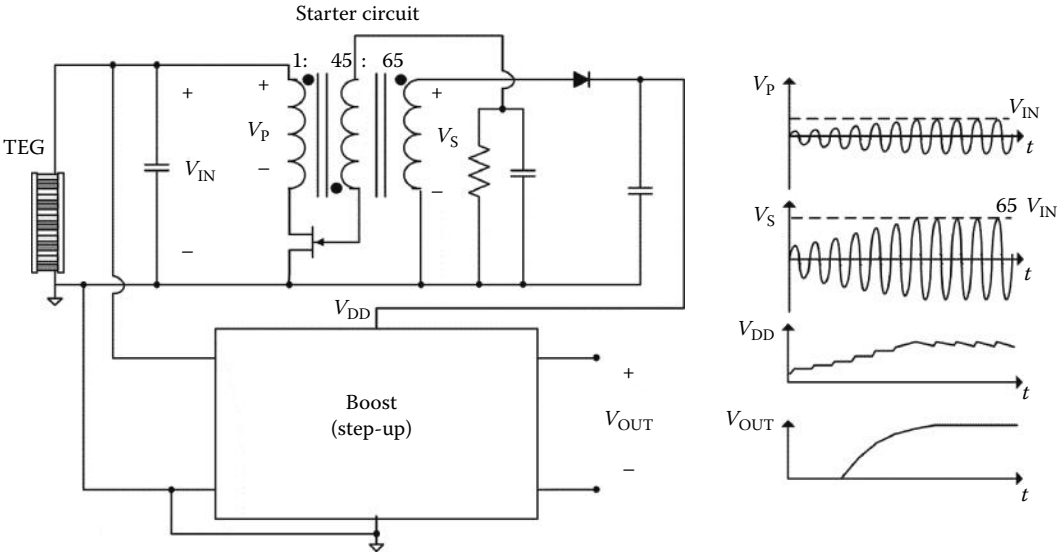


FIGURE 10.12 Scheme of a TEG power conditioner based on a transformer.

Another commercial IC converter, specifically the BQ25504 from Texas Instruments, tracks the MPP of TEGs. As it is based on PFM converter controlled by a simple FOCV, it has an ultralow current consumption that makes feasible to work for such low-power sources. TEGs, as well as PV panels, hold an empirical relationship between the V_{OC} and MPP, and so, the FOCV technique can also be applied. Taking into account that the electrical behavior can be modeled by an equivalent Thevenin source, V_{MPP} will be the half value of V_{OC} ($K = 0.5$).

References

1. S. Roberts, *Solar Electricity. A Practical Guide to Designing and Installing Small Photovoltaic Systems*, Prentice Hall, Upper Saddle River, NJ, 1991.
2. SoDa, SOlar radiation DAta, SoDa homepage, accessed date: July 2012, http://www.soda-is.com, accessed on July 2012.
3. ST, SPV1001. Cool bypass switch for photovoltaic applications, accessed date: July 2012, http://www.st.com, accessed on July 2012.

4. T. Esram and P. L. Chapman, Comparison of photovoltaic array maximum power point tracking techniques, *IEEE Trans. Energy Convers*., 22(2), 439–449, 2007.
5. O. Lopez-Lapeña, M. T. Penella, and M. Gasulla, A new MPPT method for low-power solar energy harvesting, *IEEE Trans. Ind. Electron*., 57(9), 3129–3138, 2010.
6. O. Lopez-Lapena, M. T. Penella, and M. Gasulla, A closed-loop maximum power point tracker for subwatt photovoltaic panels, *IEEE Trans. Ind. Electron*., 59(3), 1588–1596, 2012.
7. Texas Instruments, Ultra low power boost converter with battery management for energy harvester applications. BQ25504, accessed date: July 2012, http://www.ti.com, accessed in July 2012.
8. F. I. Simjee and P. H. Chou, Efficient charging of supercapacitors for extended lifetime of wireless sensor nodes, *IEEE Trans. Power Electron*., 23(3), 1526–1536, 2008.
9. A. Chini and F. Soci, Boost-converter-based solar harvester for low power applications, *Electron. Lett*., 46(4), 296–298, 2010.
10. ST Microelectronics, SPV1040. High efficiency solar battery charger with embedded MPPT, accessed date: July 2012, http://www.st.com, accessed in July 2012.
11. C. Alippi and C. Galperti, An adaptive system for optimal solar energy harvesting in wireless sensor network nodes, *IEEE Trans. Circuits Syst. I, Reg. Papers*, 55(6), 1742–1750, 2008.
12. Linear Technology, LT3105. 400 mA step-up DC/DC converter with maximum power point control and 250 mV start-up, accessed date: July 2012, http://cds.linear.com, accessed in July 2012.
13. M. Kishi, H. Nemoto, T. Hamao, M. Yamamoto, S. Sudou, M. Mandai, and S. Yamamoto, Micro-thermoelectric modules and their application to wristwatches as an energy source, in *Proceedings of the 18th International Conference on Thermoelectrics (ICT'99)*, pp. 301–307, 1999.
14. Micropelt, Thin film thermogenerators and sensing devices. MPG-D602. MPG-D751, accessed date: July 2012, http://www.micropelt.com, accessed in July 2012.
15. V. Leonov, T. Torfs, P. Fiorini, and C. Van Hoof, Thermoelectric converters of human warmth for self-powered wireless sensor nodes, *IEEE Sens. J*., 7(5), 650–657, 2007.
16. L. Mateu and N. L. Cosmin Codrea, Human body energy harvesting thermogenerator for sensing applications, in *Proceedings of IEEE International Conference on Sensor Technologies and Applications*, pp. 366–372, 2007.
17. Seiko Instruments, Inc., S-882Z series. Ultra-low voltage operation charge pump IC for step-up DC-DC converter startup, accessed date: July 2012, http://www.sii-ic.com, accessed in July 2012.
18. J. M. Damaschke, Design of a low-input-voltage converter for thermoelectric generator, *IEEE Trans. Ind. Appl*., 33(5), 1203–1207, 1997.
19. Linear Technology, LTC3108. Ultralow voltage step-up converter and power manager, accessed date: July 2012, http://cds.linear.com, accessed in July 2012.

11

Energy Harvesting for Sensors: AC Harvesters

Maria Teresa Penella
Urbiotica S.L.

Oscar Lopez-Lapeña
Universitat Politècnica de Catalunya

Manel Gasulla
Universitat Politècnica de Catalunya

11.1 Introduction

Radiofrequency (RF) and mechanical energy transducers generate an alternating (ac) voltage and current that is proportional to the available power in the environment. There are other transducers (of different environmental energies) that also generate an ac signal. However, we have chosen these energies as the most representative in terms of the variety of energy conditioning circuits used. Other ac harvesters can use similar circuits and techniques to that presented in this chapter. Indeed, RF and mechanical energy sometimes benefit of the same energy conditioning circuit structures. However, note that each transducer has its own peculiarities that must be accounted for.

Signals generated by these transducers must be rectified in order to charge the battery or the (super) capacitor that powers the load (autonomous sensor). When using just a rectifying stage, the load fixes the working point of the harvester and the system is not optimized under all environmental conditions. Additionally, under some circumstances, the voltage coming from RF or mechanical energy transducers is quite low and needs to be stepped-up. One advantage of having an ac signal is that step-up transformers or rectifiers (among other step-up circuits) can readily be used to obtain the output voltage needed to power the autonomous sensor. Energy transducers and rectifiers can be modeled with a Thevenin equivalent circuit that delivers maximum power to a matched resistive load. Consequently, several works have proposed maximum power point trackers (MPPTs) that bias the transducer, or the rectifier, at its maximum power point (MPP). MPPTs are more complex circuits that, if well designed, can highly improve efficiency at very low input powers.

This chapter describes two important ac energies (RF and mechanical) and the most used energy harvesting circuits.

11.2 Radiofrequency Energy Harvesting

RF signals are present in the environment as they are deliberately radiated by broadcasting stations and cellular phone antennas. The ambient available RF power must be harvested by means of broadband and/or circularly polarized antennas [1], because in most cases, there is either a broad range of frequencies and different polarizations or there is lack of detailed information on the ambient RF waves. Except for locations near transmitters, the available ambient power density (S) is in the nW/cm^2 range, which is too low to be useful. However, near transmitters, the power is much higher, and this is why deliberate RF power transmission on the industrial, scientific, and medical (ISM) application frequency bands has also become an alternative to power autonomous sensors [2]. In this case, S can be obtained from Equation 11.1:

$$S = P_{\text{EIRP}} \frac{1}{4\pi d^2} \tag{11.1}$$

where

P_{EIRP} is the effective radiated power, which is related to that of an isotropic radiator (an ideal antenna)
d is the distance from the transmitter

This formula is for free-space propagation; for other scenarios, the power decays more rapidly with distance, unless the location features a structure that acts as a waveguide (e.g., in corridors). Additionally, there are other factors that affect S, including multipath propagation and reflections; thus, calculating the real S requires accurate simulations or field measurements. Furthermore, the maximum P_{EIRP} for ISM frequency bands is regulated (for European regulations, see Annex 11.1 in [3]). The available power that the antenna can deliver to a matched load (P_{AV}) depends on S and on the effective aperture (area) of the antenna (A_e), as shown in Equation 11.2 (known as the *Friis relation*):

$$P_{\text{AV}} = S \cdot A_e = S \frac{\lambda_{\text{RF}}^2}{4\pi} G_r = P_{\text{EIRP}} G_r \frac{\lambda_{\text{RF}}^2}{(4\pi d)^2} \tag{11.2}$$

where

$A_e = \lambda_{\text{RF}}^2 G_r / 4\pi$
G_r is the antenna gain
λ_{RF} is the wavelength

Equation 11.2 clearly demonstrates that the higher the frequency (λ_{RF} is inversely proportional to frequency), the lower the power range. In Europe, the most popular frequency bands for ISM applications are centered at 433 MHz, 868 MHz, and 2.4 GHz.

Figure 11.1 shows the building blocks of an RF energy harvester, encompassing at least an antenna, an impedance matching block, and a rectifier. In some cases, a low-pass filter can be used between the impedance matching and the rectifier in order to prevent harmonics from flowing back to the antenna and reducing voltage amplitude at the input of the rectifier. If a filter is used, an additional second impedance matching block is needed to match the filter output impedance to the input impedance of the rectifier. However, as these filters are not conventionally used in RF energy harvesters, we will not deal with these two blocks.

An RF transducer (antenna) can be modeled as a voltage ac source in series with an output impedance (Figure 11.1). Antenna shapes and dimensions vary widely, and several distinct antenna designs have been employed in RF and radiofrequency identification (RFID) harvesting applications. The series impedance basically comprises the radiation resistance (R_S), the loss resistance (R_{loss}), and a reactive part (X_{ant}). Depending on the antenna design, X_{ant}, can be positive, thus presenting an inductive behavior

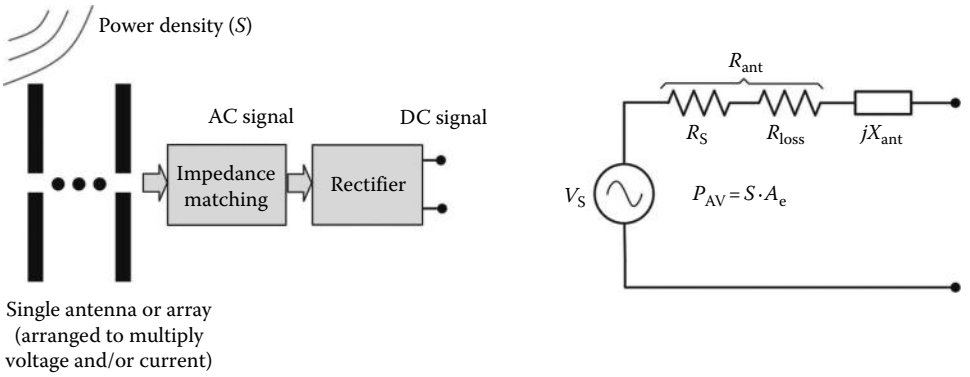


FIGURE 11.1 General block diagram of an RF harvester and the electrical equivalent model for an antenna.

(e.g., in loop-shape antennas) or negative, thus presenting a capacitive behavior (e.g., in patch antennas). The size of the antenna is inversely proportional to the desired resonant frequency. Despite this frequency dependency, some miniaturization techniques, such as high-dielectric substrates or meandered, bent, or tip-loaded structures, can be employed to minimize the final size of the system. The amplitude of the voltage generated on the antenna when matched ($\hat{v}_S$) depends on P_{AV} and R_S as

$$\hat{v}_S = 2\sqrt{2R_S P_{AV}} \tag{11.3}$$

Thus, at a given P_{AV}, $\hat{v}_S$ increases for a larger R_S. $\hat{v}_S$ must be relatively high in order to overcome the voltage drop of the diodes and reduce the losses from the rectifier. P_{AV} depends on A_e and S and, therefore, depends on the antenna characteristics. If greater power is sought, voltages or currents can be added by connecting several antennas (in series and/or in parallel) to form an array; in this case, the antenna's A_e is increased at the expense of a larger physical area [1]. A single rectifying circuit for the whole array reduces the number of rectifying elements but can complicate array design. An antenna with rectifying and matching elements is called a *rectenna. Rectennas* can also be connected in series or parallel in order to add current or voltage; again, at the expense of greater effective area (and physical dimensions). The efficiency of the *rectenna* depends on input power, and for microwatt-level inputs is 20% maximum [1].

The transmitted power reaches its maximum when the antenna sees at its output an impedance that is the conjugate of its own impedance: $Z^*_{ant} = R_S - jX_{ant}$. In Figure 11.2, Z_{in} (R_{in} and C_{in}) is the equivalent input impedance of the rectifier and ensuing load. Z_{in} should be adapted through a matching network (gray components in Figure 11.2) to present Z^*_{ant} at the antenna (point A in Figure 11.2). This procedure is known as *impedance matching*. A properly matched design is as important as a well-designed antenna or rectifier. Numerous matching configurations are available. The choice of the specific configuration depends on the number of elements used for the matching and the way in which they are positioned. Nonetheless, to date, three main circuits have been proposed for RFID or RF harvesting [4]: *a transformer* (Figure 11.2a), *a shunt inductor* (Figure 11.2b), and an *L network* (Figure 11.2c). A shunt inductor matching does not boost the antenna voltage and can be used for high R_S antenna for which $\hat{v}_S$ is relatively high (see Equation 11.3). As this is not always easy, a transformer or an *L* network can be used to boost the voltage generated on the antenna.

Rectifier circuits together with an output low-pass filter (in general a single capacitor) provide a direct current (dc) output voltage at the ensuing load. There are three main options for the rectifier:

1. A diode (which, together with the antenna, forms a rectenna) [1] Figure 11.3a
2. A bridge of diodes (or diode-connected transistors) (Figure 11.3b)
3. A voltage rectifier multiplier [5] (Figure 11.4 left shows a single stage circuit)

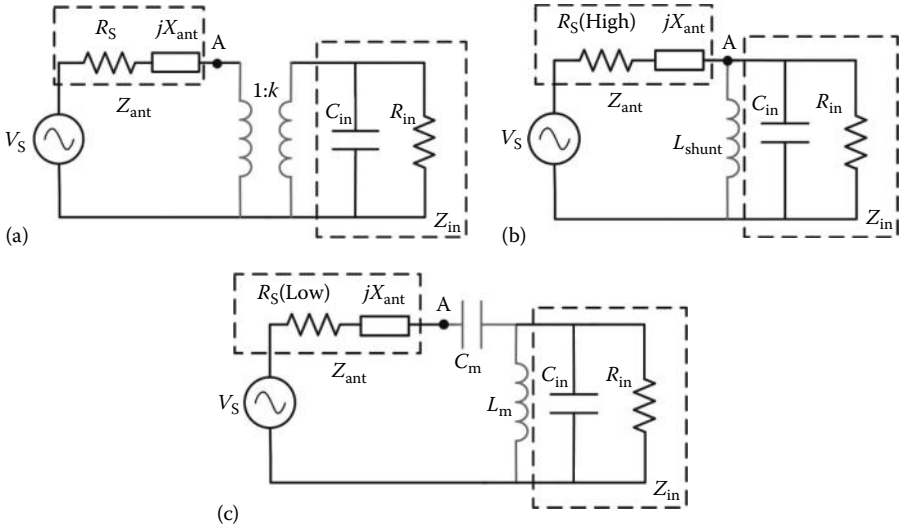


FIGURE 11.2 Typical matching network circuits for RFID or RF harvesting circuits: (a) transformer, (b) shunt inductor, and (c) *L* network. The matching network elements are shown in gray.

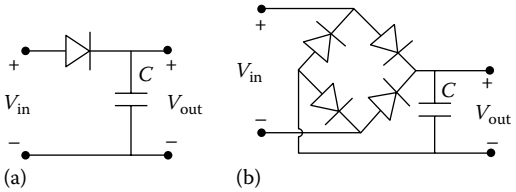


FIGURE 11.3 (a) Single diode and (b) diode bridge rectifiers.

All of these circuits are broadband (i.e., they exhibit the same rectifying characteristics in a broad band of frequencies). For kilohertz or megahertz frequencies, p–n diodes and transistors are used as rectifiers, whereas for microwave frequencies (>1 GHz), Schottky diodes (GaAs or Si type) are preferred, because they have shorter transit times. The rectification performance of the diode depends on its saturation current, its junction capacitance, and its conduction resistance. A large saturation current is sought, because it leads to a low forward voltage drop.

The diode (Figure 11.3a) and the diode bridge (Figure 11.3b) provide an output dc voltage to the load (V_{out}) whose value is lower than the peak value of the incoming signal. The *voltage rectifier multiplier*, as its name indicates, multiplies the peak amplitude of the incoming signal, Figure 11.4. At long distances (low P_{AV} and, consequently, low $\hat{v}_S$; see Equation 11.3), the dc voltage level from diode or diode bridge rectifiers is not high enough to power an electronic circuit, so the voltage rectifier multiplier appears to be the best solution [5]. The output voltage of a *voltage rectifier multiplier* can be further multiplied by cascading several stages as in Figure 11.5.

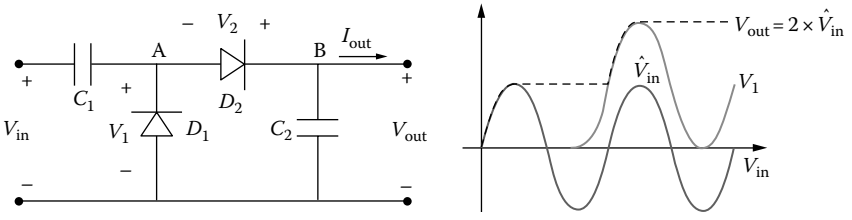


FIGURE 11.4 A single stage of a voltage multiplier rectifier and waveforms during the transient.

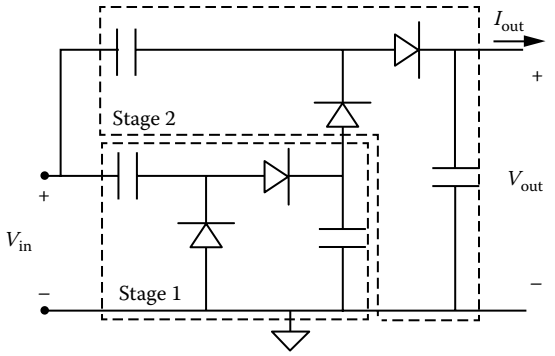


FIGURE 11.5 Two-stage voltage rectifier multiplier.

The diodes of the rectifier can also be implemented using CMOS transistors connected as diodes, which is the case for most RFID tags [5]. Using rectifying elements with low threshold voltage and low reverse current will increase the sensitivity of the rectifier (i.e., the minimum input voltage at which the rectifier can work under given load requirements). To overcome the technological issues of traditional CMOS, researchers have proposed several rectifier designs. For example, to obtain transistors with very low threshold voltage, Curty et al. [6] employed the silicon-on-sapphire (SOS) CMOS process, and Karthaus and Fischer [7] used silicon–titanium Schottky diodes. Others have avoided using these relatively expensive processes by biasing the voltage at the gate of the transistor, which results in almost zero threshold voltage transistors [8,9]. RFID tags usually include the required number of stages that give the desired output voltage at the expected minimum input power, $\bar{P}_{in}$.

P2110 and P1110 from Powercast [2] are two chips that integrate the rectifier and are designed for 50 Ω antennas and for the 850–950 MHz frequency band. The first one can deliver a regulated voltage from 1.8 to 5.25 V at input powers down to –11 dBm. Efficiency is around 10% at this power level and goes up to around 50% at higher power levels. The other circuit can more efficiently harvest input powers from –5 to 20 dBm.

Seen from the load, the rectenna (or the antenna plus any of the previously explained rectifiers) can be modeled with a Thevenin equivalent as a dc voltage source (V_{th}) in series with a Thevenin resistance (R_{th}). V_{th} depends on environmental conditions and R_{th} is constant for a certain range of input powers and generated voltages. Consequently, the system will deliver maximum power to a matched resistive load equal to R_{th}.

To boost the output voltage of the rectifier whenever it is too low, step-up switching converters can be used. Additionally, they can also be used to implement an MPPT [10,11]. Considering the Thevenin equivalent circuit of the *rectenna*, the simplest MPPT technique emulates a fixed resistance equal to R_{Th}. In this way, maximum power is extracted from the *rectenna* for a wide range of environmental conditions. A discontinuous conduction mode boost converter (Figure 11.6) with a simple open-loop controller emulates a resistor (R_{eq}) if its input voltage (V_{in}) is much lower than the output voltage (V_{out}). The average input current is proportional to V_{in} with a proportional constant equal to $1/R_{eq}$ that depends on the duty cycle of the converter's transistor, *D*. *D* depends on the converter timing parameters T_{hf}, t_{on}, *k*, and T_{lf} from Figure 11.6. Using a fixed *D* that leads to $R_{eq} = R_{Th}$, the input voltage the voltage at the maximum power point, reaches V_{MPP} ($V_{Th}/2$). Following this control scheme, the reported overall efficiency in [10] was 16.7% for a power density of 70 μW/cm^2 at the antenna (P_{AV} was around 2.5 mW) when connecting a thin-film lithium battery (4.15 V) at the output of the dc/dc converter.

In addition to the method previously explained, some of the techniques used for photovoltaic energy harvesting, such as the fractional open-circuit voltage (FOCV) MPPT, can also be applied to RF harvesting. In RF energy harvesters, the proportional constant, *k*, of FOCV will be 0.5. In any case, the power consumption of the energy conditioning circuit must be in the microwatt level or below in order to achieve a net power gain.

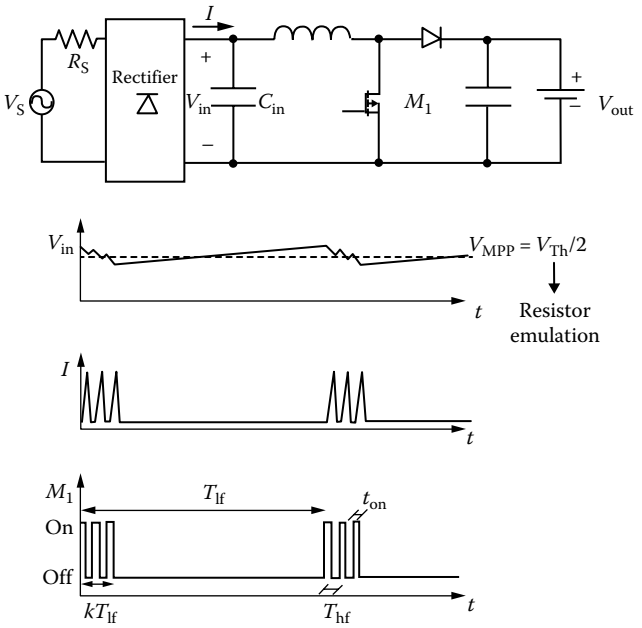


FIGURE 11.6 Boost converter working in discontinuous conduction mode acting as a resistor emulator.

11.3 Mechanical Energy Harvesting

Mechanical energy harvesting is a very active research topic that has spawned numerous reviews (e.g., [12]). It is based on kinetic energy [13], the sources of which include liquid or gas flow, vibrations, human activity, and pressure variations (e.g., acoustic noise and atmospheric pressure). Available power in flow energy increases cubically with an increase in the speed of the liquid or gas. Low-level vibrations occur in machinery, outdoor windows, and transport vehicles; they produce frequencies between 50 and 200 Hz and acceleration amplitudes between 1 and 10 m/s^2. Human activity can actively or passively generate kinetic energy. Active human power requires deliberate movement, whereas passive human power exploits common daily activities (e.g., heel strike while walking).

Mechanical energy can be coupled by one of the following conversion principles [14]: electrostatic, piezoelectric, and electromagnetic. Reported efficiencies are 0.32% [12], 0.5% (for polyvinylidene fluoride [PVDF]) to 20% (for lead zirconate titanate [PZT]) [15], and 6% [12], respectively. The generated signals are ac and must be rectified in order to power autonomous sensors. Figure 11.7 shows a general block diagram for the three types of conversion principles and the equivalent electrical models of the respective harvesters.

Electrostatic converters are based on variable capacitors. One plate is fixed and the other changes with the mechanical force, thereby changing the value of the capacitance. They are IC compatible as microelectric variable capacitors can be fabricated via silicon micromachining techniques. Energy can be extracted via charge-constrained or voltage-constrained approaches. Charge and voltage are related through capacitance ($Q = C \cdot V$); therefore, a change in capacitance produces variations in either voltage (charge constraint [16], Figure 11.8) or charge (voltage constraint [17], Figure 11.9). In both approaches, the extraction cycles are basically divided into three phases. In the first phase, the variable capacitor (C_{VAR}) is charged from an energy reservoir (can be a battery or a capacitor). Then, the harvest phase starts. In the charge-constrained circuit of Figure 11.8, the voltage of C_{VAR} increases as the distance between the plates increases (capacitance decreases) and charges a temporary capacitor (C_S). If the diodes are replaced by transistors, less power will be lost during this phase; however, the control signals will need to be synchronized with the mechanical movement. In the voltage-constrained circuit of Figure 11.9,

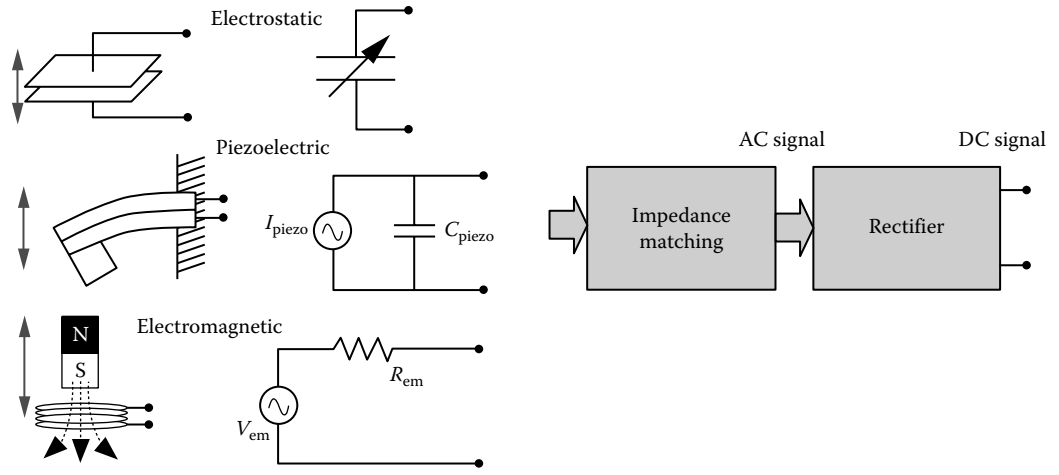


FIGURE 11.7 General block diagram of a mechanical harvester. Includes electrical models of the three different harvesters used for mechanical energy.

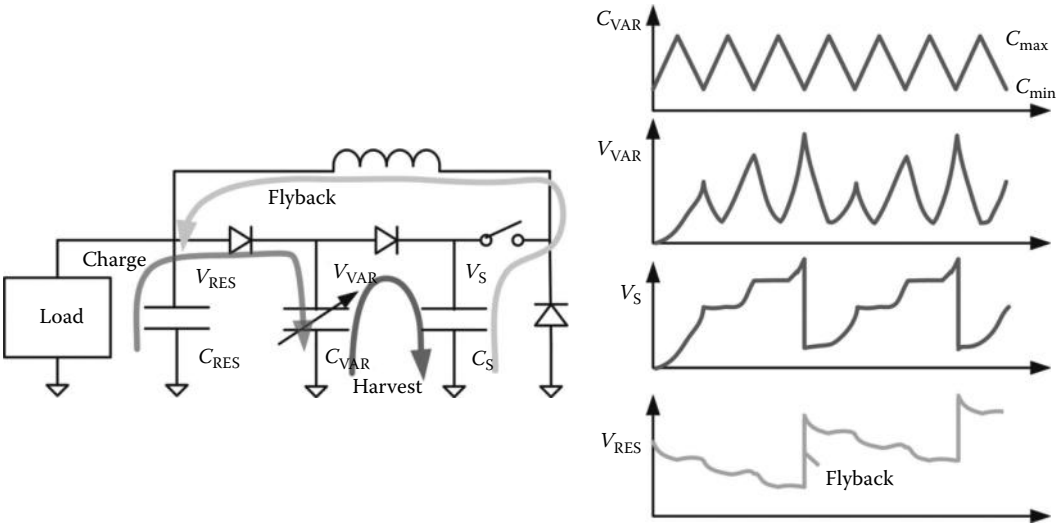


FIGURE 11.8 Charge-constrained energy harvesting circuit for electrostatic mechanical harvesters.

the energy reservoir is directly connected to the harvester and $V_C = V_{BAT}$ during the harvest phase. Thus, current flows to the energy reservoir as the capacity of the harvester decreases (distance between the plates increase). For the charge-constrained scheme, the third phase is performed after several harvest phases. In this phase, the energy in C_S is transferred to the main energy reservoir unit (C_{RES}) through a buck converter (flyback phase), and afterward, the process is started again. In the voltage-constrained scheme, after the harvest phase, the remaining energy in the variable capacitor can be recovered (recovery phase). Nonetheless, this energy is very small and not worthy to extract.

However, these converters have several drawbacks: They need a power source reservoir for starting the extraction cycles, and in the charge constraint scheme, they operate at high voltages, which limit their implementation to more expensive integration processes [18].

Reversible electric polarization in response to strain from mechanical stress is called the *piezoelectric effect*. Piezoelectric materials are anisotropic: their properties vary according to the direction of force and the orientation of the polarization and the electrodes. Several operational modes can be employed

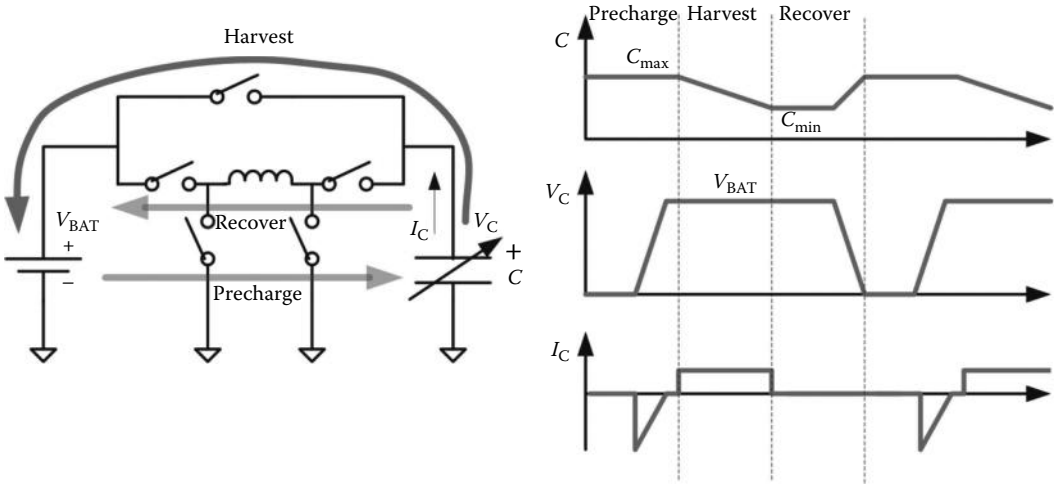


FIGURE 11.9 Voltage-constrained energy harvesting circuit and related curves for electrostatic mechanical harvesters.

for piezoelectric harvesters, the most widely used of which are modes 33 and 31. Mode 33 is used to extract energy from impact or when harvesting energy from passive human power. Mode 31 is more suited for cantilever structures with a proof mass at the free end. Cantilevers are usually bimorphic structures comprising two piezoelectric materials bound together, with a shim in between them. For power extraction, cantilevers typically operate at their resonant frequency. Piezoelectric transducers can be electrically modeled as an ac current source (I_{piezo}) in parallel with a capacitor (C_{piezo}) [19] (Figure 11.7). The main challenge with piezoelectric harvesters is integrating them into an IC.

The maximum voltage at which C_{piezo} will be charged (piezoelectric open-circuit voltage, V_{POC}) is generally in the volt level and can be directly connected to conventional bridge rectifiers (Figure 11.3b) without significant power losses. V_{POC} depends on I_{piezo} (determined by the displacement achieved), the angular frequency ω of the movement, and C_{piezo} (see Equation 11.4). In this case, V_{POC} is usually much higher than the diodes voltage drop (V_D) and must at least be $2V_D$ in order to extract energy from the transducer. For example, a piezoelectric energy harvester such as Volture [20] from Mide (tuned to work at vibrations frequencies between 80 and 175 Hz) can produce several milliwatts at 9.8 m/s^2 vibration with a conventional rectifier. V_{POC} depends on I_{piezo} (determined by the displacement achieved), the angular frequency of the movement and C_{piezo}:

$$V_{POC} = \frac{I_{piezo}}{\omega \cdot C_{piezo}} \tag{11.4}$$

Nonetheless, for microscale implementations or under certain environmental conditions, the output voltages can also be relatively small. Furthermore, more power can be generated if some charge is injected to the piezoelectric element before force is done against it in order to increase the force it posses to the movement. Thus, other circuits have been proposed [21]:

- Synchronous switched extraction circuits that extract all the generated energy from C_{piezo} at lower voltages than the bridge rectifier
- Circuits that inject some charge to the piezoelectric element before force is done against it

In the synchronous switched extraction circuit (Figure 11.10), the charge is accumulated in C_{piezo} until I_{piezo} crosses zero (at the maximum displacement of the piezoelectric transducer), then the switch (S) is closed and charge is extracted with a resonant LC circuit. The switch S is closed for a period that

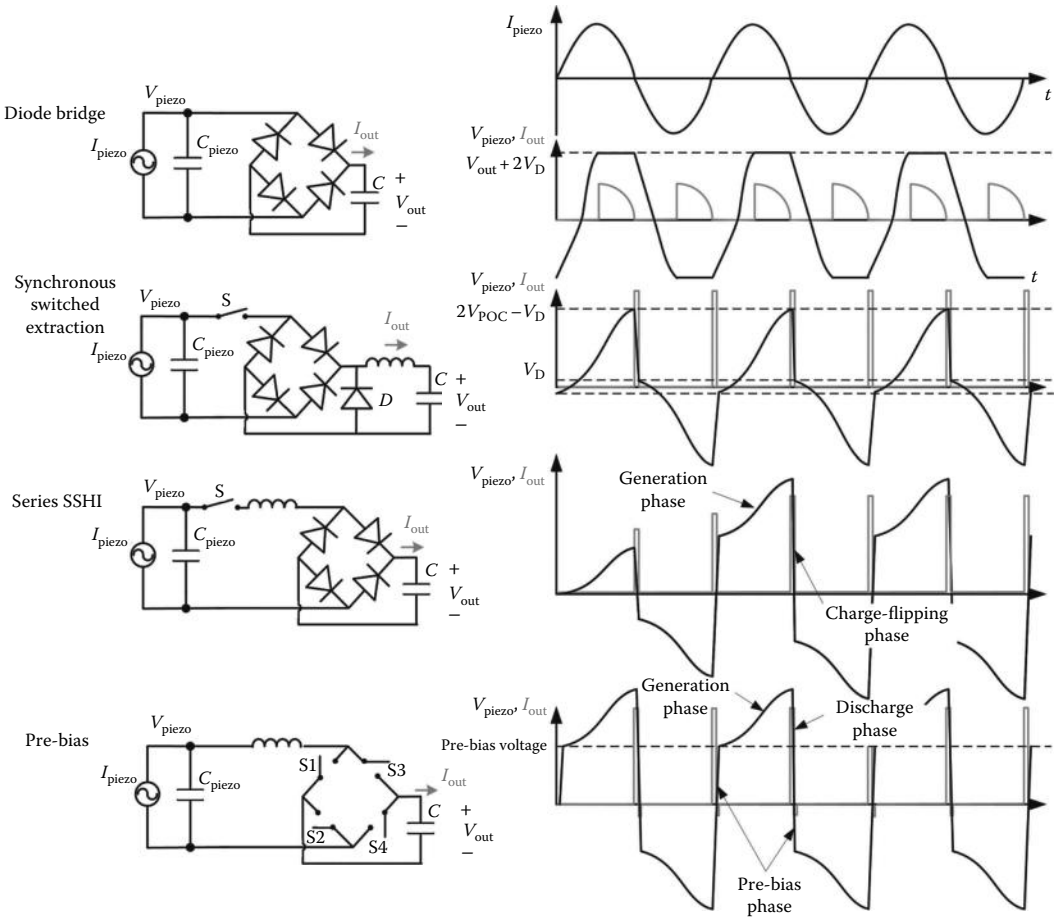


FIGURE 11.10 Different extraction circuits for piezoelectric energy harvesters.

corresponds to half the cycle of the resonant frequency of the *LC* circuit allowing the extraction of all the charge in C_{piezo} whenever the output voltage (V_{out}) is carefully chosen. The discharge time of C_{piezo} should be very small related to the period of the movement. Diode *D* is used to allow the current in *L* (if any) freewheel to *C* once C_{piezo} has been fully discharged. This circuit can handle lower V_{POC}, and contrary to the bridge rectifier circuit, the voltage in C_{piezo} is zero at the end of the charge extraction; thus, more power is harvested. This circuit is feasible whenever $V_{POC} > V_D$.

There are two main types of circuits that inject charge to the piezoelectric element: synchronous, switched harvesting with inductor (SSHI) and prebias circuits.

In SSHI circuits, there is an initial generation phase that stores charge in C_{piezo}, and afterward an inductor, in parallel or series (Figure 11.10) with the piezoelectric transducer, is switched at the zero crosses of I_{piezo} in order to invert the polarity of C_{piezo} (charge flipping phase). In parallel SSHI configuration, V_{POC} can be lower than V_D.

There are several implementations of prebias circuits, the simplest uses switches in an H bridge configuration (Figure 11.10). There are three phases for each semiperiod of I_{piezo}: Firstly, S1 and S4 are closed and C_{piezo} is precharged to the reservoir voltage; then, they are opened and the harvest phase starts; and afterward, when I_{piezo} reaches zero, S1 and S4 are closed again and C_{piezo} is discharged to 0 V through an *LC* resonant circuit. In the other semiperiod, the same phases are repeated with diodes S2 and S3. These circuits allow extracting more power from the harvester than SSHI circuits at the cost of more control overhead because the rectification bridge is implemented with switches.

In any case, there is an MPP at which extracted power is maximal. As happened with RF harvesting circuits, this MPP corresponds to an optimal resistance load in parallel to the output capacitor of the rectifier. Thus, a resistor emulator circuit as that described in Figure 11.6 has also been described for piezoelectric harvesters [22].

In some case, resonant matching networks have also been used to boost the voltage at the input of the rectifier [23]. This work uses a commercial piezoelectric energy harvesting circuit, LTC3588 from Linear Technology, which integrates a full bridge rectifier. Linear has indeed several energy harvesting circuits for piezoelectric energy harvesting that work with input voltages between 2.7 and 20 V and can output voltages between 1.8 and 5 V with current consumptions below one microampere.

Electromagnetic induction, based on Faraday's law (see Equation 11.5), refers to generation of a voltage (V_{em}) in a conductor (typically, a coil) located within a magnetic flux ϕ, usually generated by a permanent magnet. The generated voltage depends on various factors, including the strength of the magnetic field, the velocity of the relative motion between the coil and the magnet, and the number of turns in the coil. Electromagnetic energy transducers can be electrically modeled as low-level ac voltage sources with low series impedance (Figure 11.7):

$$v = -\frac{d\phi}{dt} \tag{11.5}$$

Since magnets are quite difficult to integrate into ICs, integrated prototypes of these transducers show very low output powers. This type of harvester is best suited for high-frequency and low-amplitude vibrations. Some commercially available energy harvesters at macroscale (e.g., the Perpetuum Ltd. [24]) are based on this principle.

Even for macroscale implementations, the generated voltage can usually be around hundreds of millivolts [19]. Bridges or single diodes could be used as the rectifier (Figure 11.3). In that case, the rectifying elements can be implemented with diode-connected transistors that will present lower threshold voltages. However, the output voltage of the rectifier will be lower than the peak value of the input signal (see Equation 11.2), and thus, it will still need to be stepped-up by using, for example, a boost dc/dc converter.

Voltage can also be boosted by using transformer matching circuits before rectification. In Figure 11.11, the voltage of the harvester is boosted by the turns ratio of the transformer, and afterward, each semiperiod of the ac signal is rectified with one of the diodes. One of the major challenges of this circuit is IC integration as transformers are quite difficult to integrate into ICs.

The voltage multiplier rectifier (Figure 11.4) or modified versions of this circuit can also be used to rectify and boost the output voltage of the harvester. The lack of magnetic components makes the voltage multiplier rectifier a good option for IC integration.

More complex circuits such as dual polarity boost converters (Figure 11.12) have also been proposed [19]. There, each converter is used to boost and rectify one of the two semiperiods of the generated

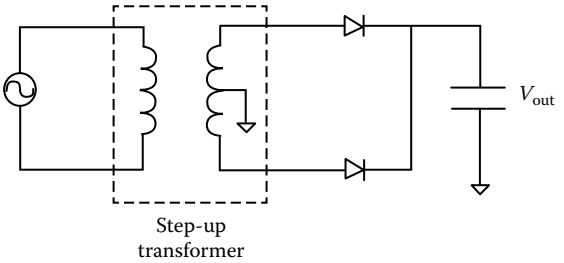


FIGURE 11.11 Circuit to boost the output of an inductive mechanical harvester by using a step-up transformer prior to rectification.

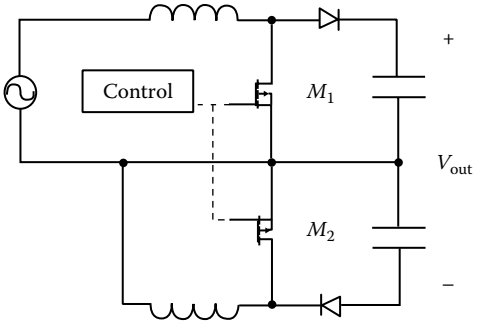


FIGURE 11.12 Dual polarity boost circuit for electromagnetic mechanical energy harvesting.

signal. Rectification is achieved by the alternate activation of one of the two converters. This circuit requires some control overhead.

Finally, MPPT circuits that make the harvester work as if it had an optimal resistor at its output, through resistor emulation as in RF (Figure 11.6) or by setting the working voltage to half the harvester open-circuit voltage (V_{OC}) [25], have also been proposed. In [25], they used switched capacitors to fix the harvester's working voltage at 0.5 V_{OC}, where V_{OC} needs to be periodically sampled.

The trend in energy harvesters is to integrate several energy conditioning circuits into the same chip in order to extract energy from multiple energy sources. Several commercial products and scientific works have recently appeared. Nevertheless, these circuits use the most suitable energy conditioning circuit for each energy source because, as can be deduced from this work, each transducer has its peculiarities. A commercial product from Cymbet, CBC915 [26], can be used indistinctly with different energy harvesters and make them work at their MPP. Furthermore, in [27], they propose a circuit that accepts energy from optical, thermal, and mechanical energy harvesters at the same time and makes all the harvesters work at their MPP for a wide range of environmental conditions. For the mechanical harvester, they used a conventional rectifier and then a buck–boost converter makes this system work at its MPP.

References

1. J. A. Hagerty, F. B. Helmbrecht, W. H. McCalpin, R. Zane, and Z. B. Popovic, Recycling ambient microwave energy with broad-band rectenna arrays, *IEEE Transactions on Microwave Theory and Techniques*, 52(3), 1014–1024, 2004.
2. Powercast, Powercast homepage, accessed date: 2010, http://www.powercastco.com/, accessed in July 2012.
3. European Radiocommunications Office Recommendation, ERC Recommendation 70-03 (Tromsø 1997 and subsequent amendments) relating to the use of short range devices (SRD), accessed date: http://www.erodocdb.dk/docs/doc98/official/pdf/REC7003E.PDF, accessed in July 2012.
4. M. T. Penella-López and M. Gasulla-Forner, *Powering Autonomous Sensors*. New York: Springer, 2011.
5. J.-P. Curty, M. Declercq, C. Dehollain, and N. Joehl, *Design and Optimization of Passive UHF RFID Systems*. New York: Springer Science+Business Media, 2007.
6. J. P. Curty, N. Joehl, F. Krummenacher, C. Dehollain, and M. J. Declercq, A model for u-power rectifier analysis and design, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(12), 2771–2779, 2005.
7. U. Karthaus and M. Fischer, Fully integrated passive UHF RFID transpoder IC with 16.7 uW minimum RF input power, *IEEE Journal of Solid-State Circuits*, 38(10), 1602–1608, 2003.
8. T. Le, K. Mayaram, and T. Fiez, Efficient far-field radio frequency energy harvesting for passively powered sensor networks, *IEEE Journal of Solid-State Circuits*, 43(5), 1287–1302, 2008.

9. T. Umeda, H. Yoshida, S. Sekine, Y. Fujita, T. Suzuki, and S. Otaka, A 950-MHz rectifier circuit for sensor network tags with 10-m distance, *IEEE Journal of Solid-State Circuits*, 41(1), 35–41, 2006.
10. T. Paing, J. Shin, R. Zane, and Z. Popovic, Resistor emulation approach to low-power RF energy harvesting, *IEEE Transactions on Power Electronics*, 23(3), 1494–1501, 2008.
11. A. Dolgov, R. Zane, and Z. Popovic, Power management system for online low power RF energy harvesting optimization, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 57(7), 1802–1811, 2010.
12. P. D. Mitcheson, E. M. Yeatman, G. K. Rao, A. S. Holmes, and T. C. Green, Energy harvesting from human and machine motion for wireless electronic devices, *Proceedings of the IEEE*, 96(9), 1457–1486, 2008.
13. S. Roundy, J. M. Rabaey, and P. K. Wright, *Energy Scavenging for Wireless Sensor Networks, with Special Focus on Vibrations*, 1st edn. Norwell, MA: Kluwer Academic Publishers, 2004.
14. S. P. Beeby, M. J. Tudor, and N. M. White, Energy harvesting vibration sources for microsystems applications, *Measurement Science and Technology*, 17(12), R175, 2006.
15. N. S. Shenck and J. A. Paradiso, Energy scavenging with shoe-mounted piezoelectrics, *IEEE Micro*, 21(3), 30–42, 2001.
16. A. Kempitiya, D. A. Borca-Tasciuc, and M. M. Hella, Analysis and optimization of asynchronously controlled electrostatic energy harvesters, *IEEE Transactions on Industrial Electronics*, 59(1), 456–463, 2012.
17. E. O. Torres and G. A. Rincon-Mora, Electrostatic energy harvester and Li-Ion charger circuit for micro-scale applications, *Proceedings of 49th IEEE International Midwest Symposium on Circuits and Systems, MWSCAS*, San Juan, Puerto Rico, pp. 65–69, 2006.
18. E. O. Torres and G. A. Rincon-Mora, Energy-harvesting system-in-package microsystem, *Journal of Energy Engineering*, 134(4), 121–129, 2008.
19. P. Mitcheson, T. Green, and E. Yeatman, Power processing circuits for electromagnetic, electrostatic and piezoelectric inertial energy scavengers, *Microsystem Technologies*, 13(11), 1629–1635, 2007.
20. Mide, Volture piezoelectric vibration energy harvester, accessed date: http://www.mide.com/products/volture/peh20w.php, accessed in July 2012.
21. J. Dicken, P. D. Mitcheson, I. Stoianov, and E. M. Yeatman, Power-extraction circuits for piezoelectric energy harvesters in miniature and low-power applications, *IEEE Transactions on Power Electronics*, 27(11), 4514–4529, 2012.
22. G. K. Ottman, H. F. Hofmann, and G. A. Lesieutre, Optimized piezoelectric energy harvesting circuit using step-down converter in discontinuous conduction mode, *IEEE Transactions on Power Electronics*, 18(2), 696–703, 2003.
23. N. Krihely and S. Ben-Yaakov, Piezoelectric harvesting circuit with extended input voltage range, *Proceedings of the IEEE 26th Convention of Electrical and Electronics Engineers in Israel*, Eilat, Israel, pp. 670–674, 2010.
24. Perpetuum, Perpetuum Ltd., accessed on: October 2, 2006, http://www.perpetuum.com/
25. D. Maurath and Y. Manoli, A self-adaptive switched-capacitor voltage converter with dynamic input load control for energy harvesting, *Proceedings of ESSCIRC '09*, Athens, Greece, pp. 284–287, 2009.
26. Cymbet Corporation, EnerChip energy processor for energy harvesting applications, access date: July 2012, http://www.cymbet.com/pdfs/DS-72-15.pdf, accessed in July 2012.
27. S. Bandyopadhyay and A. P. Chandrakasan, Platform architecture for solar, thermal, and vibration energy combining with MPPT and single inductor, *IEEE Journal of Solid-State Circuits*, 99, 1–17, 2012.

12

Hydrogen Gas Sensors: Flow

H. El Matbouly
LMST-Université du Québec à Trois-Rivières

F. Domingue
LMST-Université du Québec à Trois-Rivières

V. Palmisano
European Commission-JRC Institute for Energy

L. Boon-Brett
European Commission-JRC Institute for Energy

12.1 Introduction

The demand for hydrogen as a process gas and as a renewable energy carrier to reduce dependence on fossil fuels and to mitigate the problems associated with greenhouse gas emissions is increasing [1,2].

Hydrogen has a very low density (0.0899 kg/m^3) and boiling point (20.39 K) and high diffusion coefficient (0.61 cm^2/s in air). It has a low minimum ignition energy (0.017 mJ), high heat of combustion (142 kJ/g H_2), and wide flammable range (4–75 vol% in air), as well as a high burning velocity and an ignition temperature of 560 °C. Hydrogen also acts as a strong reducing agent for many elements and has a high permeability through many materials [3]. In addition, hydrogen is a colorless, odorless, and tasteless flammable gas that makes it undetectable by human senses.

Due to the unique physical and chemical properties of hydrogen, hydrogen-based applications require rigorous safety precautions in comparison to other industrial gases. Since the use of odorants in many hydrogen applications is not possible, dedicated detection devices or sensors are essential to warn of the presence of unwanted hydrogen leaks. Hydrogen sensors play an essential role in ensuring the safety of hydrogen-based applications. As such, sensors boost end user confidence and therefore facilitate the safe use and commercialization of hydrogen as an alternative fuel. Some applications require not only detecting the presence of hydrogen but also measuring the concentration of the gas in the ambient [3].

This chapter reviews the available commercial hydrogen-sensing technologies and their operating principles. The chapter discusses the performance characteristics of these sensor types, such as measuring range, sensitivity, selectivity, and response time. Available emergent technologies and their impact on improved hydrogen gas sensing are also presented.

Hydrogen sensors are transducer devices that detect hydrogen gas and produce an electric signal with a magnitude proportional to the hydrogen gas concentration.

There are several different types of hydrogen sensors and the most common commercial types are catalytic (CAT), electrochemical (EC), thermal conductivity (TCD), semiconductive metal–oxide (MOX), and metal–oxide–semiconductor (MOS) sensors. The following section presents an overview of the main sensor types available on the market.

12.2 Commercial Sensor Technologies and Their Detection Principles of Catalytic Hydrogen Gas Sensor

The detection principle of CAT hydrogen gas sensors is based on heat generation due to hydrogen reaction with oxygen on the surface of a CAT material. There are two types of CAT hydrogen sensors: pellistor type and thermoelectric type.

A pellistor-type sensor is made of two platinum wires embedded in separate beads: a sensing bead and a reference bead. The sensing bead is impregnated with catalyst (usually palladium) to accelerate the oxidation of hydrogen. The reference bead compensates for environmental effects such as changes in humidity or temperature. These two beads are mounted in a Wheatstone bridge as shown in Figure 12.1.

When the target gas is introduced, CAT combustion of the gas occurs on the surface of the sensing bead, releasing heat and causing the temperature of the bead to rise. As a result, the resistance of the embedded platinum wire changes, which imbalances the Wheatstone bridge and provides an output signal that varies with target gas concentration. CAT-type sensors were first suggested by Jonson in 1923 and used in mines for methane detection [4,5]. CAT pellistor-type sensors are well-developed and widely available.

On the other hand, thermoelectric-type CAT sensors use the thermoelectric Seebeck effect to transform the heat generated from the exothermic hydrogen oxidation reaction into an electric signal. The thermoelectric voltage generated is related to the temperature difference by the following expression:

$$U = S\Delta T \tag{12.1}$$

where

U is the thermoelectric voltage generated

S is the Seebeck coefficient

ΔT is the temperature difference due to heat generation

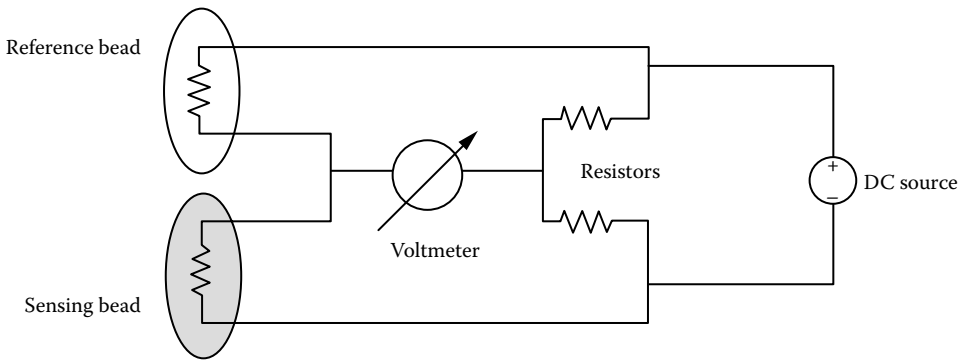


FIGURE 12.1 Wheatstone bridge circuit connected to a CAT hydrogen sensor. The sensing bead represents the variable resistor that imbalances the bridge in the presence of hydrogen.

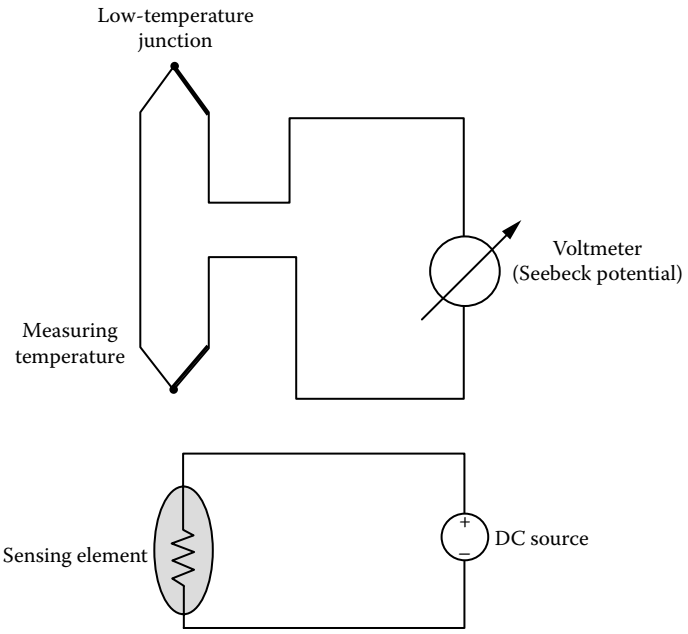


FIGURE 12.2 Schematic diagram representing hydrogen sensing by thermoelectric effect.

Thermoelectric hydrogen gas sensors are fabricated as a film of a thermoelectric material deposited on an insulating substrate material. One-half of the thermoelectric material is coated with a catalyst, and oxidation of hydrogen on the catalyst surface generates heat. This generates a temperature gradient over the thermoelectric film that produces a voltage signal measurable by an electric circuit. One advantage of thermoelectric hydrogen sensors over pellistor sensors is that they can operate at room temperature or slightly elevated temperatures (<100 °C) [6]. Figure 12.2 represents the CAT thermoelectric gas sensor. The thermoelectric effect was first used in sensing application in 1985 and reported by McAleer et al. [7].

12.2.1 Thermal Conductivity Hydrogen Gas Sensor

TCD hydrogen sensors exploit the exceptionally high TCD of hydrogen (0.168 W/m·K at 25 °C), seven times higher than air (0.024 W/m·K at 25 °C). The hydrogen concentration is inferred by the rate at which a sensing thermal element releases heat compared to a reference element. In the classical configuration the two elements are placed into two separate cells connected in a Wheatstone bridge circuit: when the gas is introduced at the sensing element, the difference in thermal conduction from the hot to the cold elements leads to a change in temperature that affects the resistance of the sensor and thus imbalances the bridge. This imbalance can be measured and can be used to indicate the hydrogen concentration. The main difference between the CAT bead sensor and the TCD sensor is the absence of a chemical reaction [8]. Figure 12.3 represents the circuit diagram of the TCD sensor.

12.2.2 Electrochemical Hydrogen Gas Sensor

EC sensors detect the change in charge transport or electrical properties due to EC reactions of hydrogen gas at the sensing electrode. The sensor structure consists of three electrodes stacked parallel and separate by an electrolytic solution as shown in Figure 12.4. There are two types of EC gas detection: amperometric and potentiometric [4].

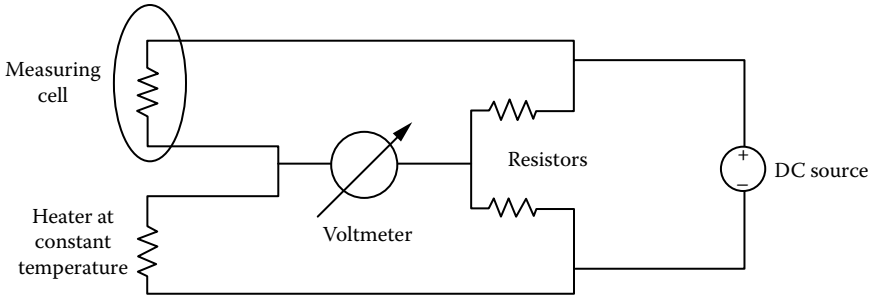


FIGURE 12.3 TCD sensor measuring circuit using a Wheatstone bridge. The measuring cell represents the variable resistor that imbalances the bridge in the presence of hydrogen.

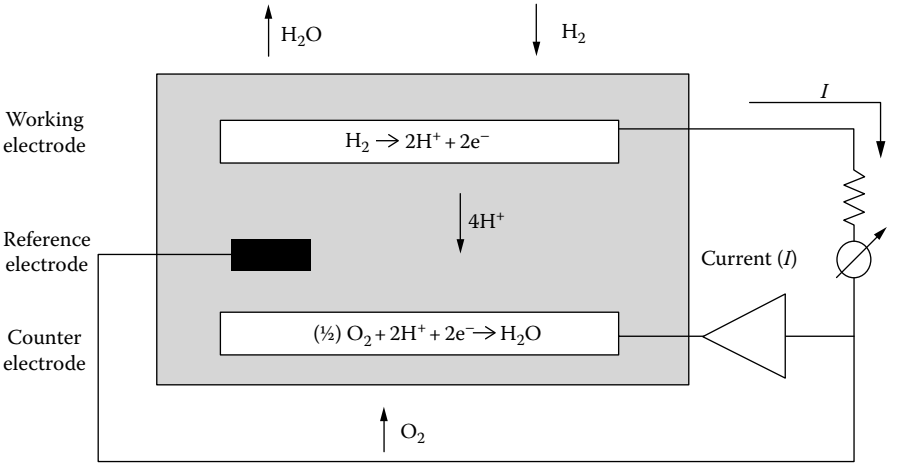


FIGURE 12.4 EC cell for hydrogen detection.

Amperometric sensors work at a constant applied voltage where the sensor signal is a diffusion-limited current. When hydrogen penetrates the gas diffusion barrier, it reaches the surface of the working electrode (anode) where the following oxidation reaction takes place:

$$H_2 \rightarrow 2H^+ + 2e^-$$

The electrons released in this reaction flow to the cathode (counter electrode) via an external circuit; this electric current constitutes the sensor signal. Protons released from the oxidation reaction cross the electrolyte and reach the counter electrode where the reduction of oxygen takes place as follows:

$$(1/2)O_2 + 2H^+ + 2e^- \rightarrow H_2O$$

The flow of electric current is proportional to the hydrogen gas concentration according to Faraday's law:

$$i = zQF$$

where

F is the Faraday constant, 96486.7 A · s/mol
z is the number of exchanged electrons per molecule
Q is the conversion rate of hydrogen in mol/s

The reference electrode maintains a constant voltage at the working electrode.

Potentiometric sensors have the same structure but work at zero current, and the measured quantity is the potential difference or electromotive force between the sensing electrode and the reference electrode. The electrode potential is related to the hydrogen gas concentration and can be written according to the Nernst equation as follows:

$$E = E^0 + \frac{RT}{zF}\ln\left(\frac{a}{a_0}\right) \tag{12.2}$$

where

E is the electrode potential
E^0 is the standard electrode potential
R is the universal gas constant
T is the absolute temperature
F is the Faraday constant
z is the number of electrons taking part in the reaction
a is the chemical activity of the analytes (proportional to hydrogen concentration)
a_0 is the activity of the reference

The processes of diffusion and chemical reaction must be at equilibrium before an accurate signal can be obtained from these sensors [6,9].

12.2.3 Semiconductive Metal–Oxide Hydrogen Sensors

MOX sensors consist of two metal electrodes connected to an active material. The electrodes allow the measurement of the resistance change of the active layer that depends on the target gas concentration. The active material is a semiconducting MOX, such as SnO_2, ZnO, TiO_2, FeO, Fe_2O_3, NiO, Ga_2O_3, In_2O_3, Sb_2O_5, MoO_3, and WO_3, whose electric resistance varies in the presence of hydrogen gas.

When the sensor reaches its operating temperature (generally between 180 °C and 450 °C), oxygen chemisorbs on the surface of the MOX material and forms negatively charged adsorbed oxygen ions that create a surface potential that acts as a barrier to electron flow. Thus, oxygen is necessary to operate this type of sensor. The electric resistance of the active layer is attributed to this barrier potential. In the presence of hydrogen, the surface density of negatively charged oxygen and, consequently, the barrier potential decrease, thus reducing the electric resistance of the active layer. The resistance change can be related to the hydrogen concentration. Typically the MOX film is heated using a resistive metal film [9,10]. Figure 12.5 represents MOX sensor structure.

12.2.4 Metal–Oxide–Semiconductor Hydrogen Sensors

MOS gas sensors are characterized by a three-layer structure: a CAT metal layer is deposited over an oxide layer on a semiconductor substrate. Their operation is based on work function (measured in electron volts) variation due to the presence of hydrogen. For hydrogen detection a hydrogen-sensitive CAT metal, such as platinum or palladium (the gate metal), is deposited on the oxide layer. In the presence of hydrogen, the current–voltage (I–V) characteristics of the MOS transistor change [11]; this change can be detected indicating the presence of the gas as Figure 12.6 illustrates.

In addition to MOS transistor hydrogen sensors, there are metal–semiconductor (MS) Schottky diodes and MOS capacitors structures that can be also used to detect the hydrogen.

In Schottky diodes the CAT metal, typically palladium or platinum for hydrogen detection, is directly in contact with a semiconductive material (Figure 12.7). Hydrogen molecules adsorb onto the surface of the CAT metal and dissociate into hydrogen atoms inducing a change in the work function of the metal

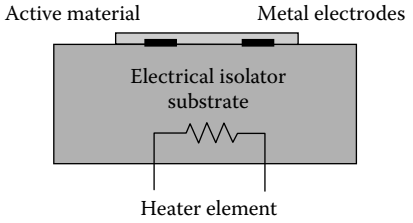


FIGURE 12.5 Typical MOX hydrogen gas sensor structure.

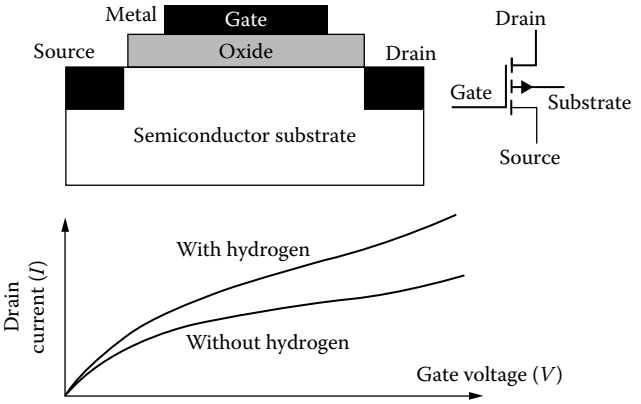


FIGURE 12.6 MOS transistor structure and the *I–V* for hydrogen detection.

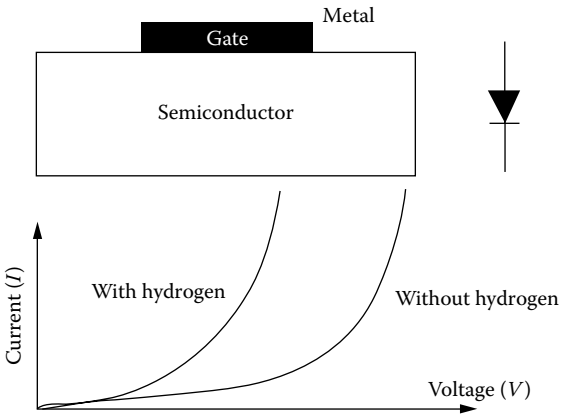


FIGURE 12.7 Schottky diode hydrogen gas sensor structure and the *I–V* curve presenting its principle of operation.

and hence the Schottky barrier height. This change in the Schottky barrier height causes a shift in the *I–V* characteristics. The response can be measured as the change in voltage when the diode is operated at constant bias current [6].

A MOS capacitor is the simplest MOS device and a cross-sectional view of a typical MOS capacitor is shown in Figure 12.8. A thin oxide layer is sandwiched between a metal layer (palladium or platinum) and a silicon substrate layer. The absorption of hydrogen and the chemical reactions at the MOX interface cause a change in the *C–V* characteristics of the capacitor and shift the flat band voltage by an amount proportional to the hydrogen gas concentration [9]. The first MOS capacitor hydrogen sensor was reported by Steele et al. in 1976 [12].

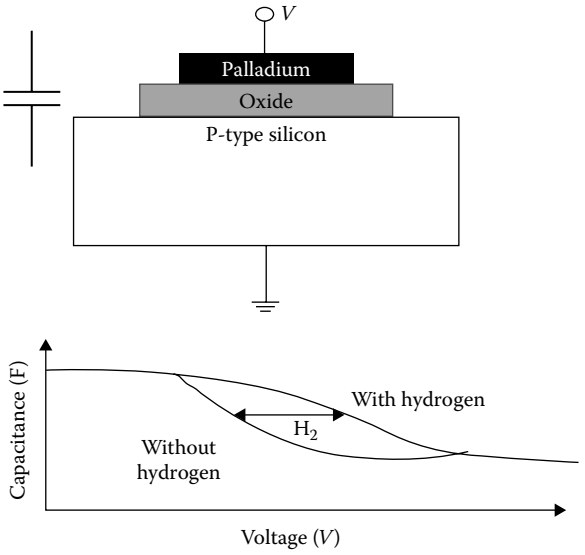


FIGURE 12.8 Typical MOS capacitor for hydrogen detection.

12.2.5 Pd-Based Hydrogen Sensors

Another class of hydrogen sensor is based on the interaction of palladium with hydrogen: the hydrogen molecule is split at the thin palladium film interface and atomic hydrogen can be accommodated in the metal lattice resulting in a change of its physical properties. The hydrogen concentration can be related to the change in the electric resistance (resistance-based sensors), to the change in metal lattice expansion (mechanical sensors), or to the change in dielectric function (optical sensors) of the sensing material. For a review of those sensors, see [6].

12.3 Hydrogen Sensor Specifications

The performance of commercial hydrogen sensors has been evaluated [6,9,13]. Table 12.1 compares the performance specifications of the most common commercial sensor types discussed in the previous sections.

12.3.1 Partial List of Hydrogen Sensor Suppliers and Manufacturers

Tables 12.2 through 12.5 list some known hydrogen sensor suppliers and manufacturers, with links to their websites.

TABLE 12.1 Performance Specifications of Commercial Hydrogen Sensor Types

	Temperature Range (°C)		Humidity (RH%)		Measuring Range (vol%)	Pressure (kPa)		Response	Recovery	Power	Life Time
Sensor	Min	Max	Min	Max	Max	Min	Max	Time (s)	Time (s)	(mW)	(Years)
CAT	−20	70	5	95	4	70	130	<30	10	1000	5
TCD	0	50	0	95	100	80	120	<10	7.5	<500	5
EC	−20	55	5	95	4	80	110	<90	2	2–700	2
MOS	−40	110	5	95	4.4	70	130	<2	10	700	10
MOX	−20	70	10	95	2	80	120	<20	10	<800	>2

Source: Hubert, T. et al., *Sens. Actuators B Chem.*, 2011.

TABLE 12.2 Partial List of Hydrogen CAT Sensor Suppliers

Sensor Supplier or Manufacturers	Websites
Neodym	http://www.neodymsystems.com/
Applied Nanotech	www.appliednanotech.net
ACME Engineering Prod.	http://www.acmeprod.com
RKI Instruments	http://rkinstruments.com/
General Monitors	http://www.generalmonitors.com/
City Technology Ltd	http://www.citytech.com/

TABLE 12.3 Partial List of Hydrogen TCD Sensor Suppliers

Sensor Supplier or Manufacturers	Websites
J. Dittrich	http://www.dittrich-systeme.de/
Neroxis	http://www.neroxis.ch/
BlueSens	http://www.bluesens.com
Advanced Gasmitter	http://www.sensors-inc.com/
MST IT	http://www.mst-it.com/
AppliedSensor	http://www.appliedsensor.com/

TABLE 12.4 Partial List of Hydrogen EC Sensor Suppliers

Sensor Supplier or Manufacturers	Websites
City Technology Ltd	http://www.citytech.com/
Solid Sense	http://www.solidsense.de/
Nemoto & Co., Ltd	http://www.nemoto.co.jp
Euro-Gas	http://www.euro-gasman.com/Sensors_intro.htm
Membrapor	http://www.membrapor.ch/
Hanwei Electronics Co., Ltd	http://english.hwsensor.com/
Synkera Technologies	http://www.synkera.com/

TABLE 12.5 Partial List of MOS and MOX Hydrogen Sensor Suppliers

Sensor Supplier or Manufacturers	Websites
Hanwei Electronics Co., Ltd	http://english.hwsensor.com/
Figaro Engineering	http://www.figarosensor.com/
e2v Technologies	http://www.e2v.com/products-and-services/instrumentation-solutions/gas-sensors/our-gas-sensor-technology/
Winsen	http://www.winsensor.com/english/news.asp
AppliedSensor	http://www.appliedsensor.com/
UST Umweltsensortechnik GmbH	http://www.umweltsensortechnik.de/index3.htm
Synkera Technologies	http://www.synkera.com/

12.4 Requirements of Hydrogen Gas Sensors

To use hydrogen gas sensors in industrial applications, general requirements described in EN IEC 60079-29-1 have to be followed [14]. The previously discussed hydrogen sensors were developed to meet these requirements as it is shown in Table 12.6, and they are suitable for applications such as the production of ammonia or methanol or for safety monitoring in nuclear power plants, where hydrogen is produced via the thermochemical splitting of water as well as stationary systems.

TABLE 12.6 Performance Requirements for Industrial Applications

Parameters	Performance Requirements
Measuring range	Up to 1 vol% H_2 in air
Lower detection limit	<0.1 vol%
Response time	<30 s
Recover time	<30 s
Life time	3–5 years

Source: Boon-Brett, L. et al., *Int. J. Hydrogen Energy*, 35, 371, 2010.

TABLE 12.7 Performance Requirements for Automotive Applications

Parameters	Performance Requirements
Measuring range	0%–4% H_2 in air
Response time	<3 s
Recovery time	<3 s
Ambient temperature	−40 °C to +125 °C
Humidity	0%–100% Relative humidity

Source: Boon-Brett, L. et al., *Int. J. Hydrogen Energy*, 35, 371, 2010.

However, for new hydrogen-based applications, the development of new technologies for the detection and measurement of hydrogen concentration requires new performance specifications for which some standards already exist [15].

The use of hydrogen as an automotive fuel represents a new application for which stringent performance specifications have been cited [16]. Table 12.7 reports the requirements for automotive applications in terms of measuring range, response time, recovery time, and operating range of temperature and humidity.

Other requirements for hydrogen sensors in automotive applications are high reliability, long lifetime, high selectivity, and adaptability to the application interface. Low-cost, small dimensions, and

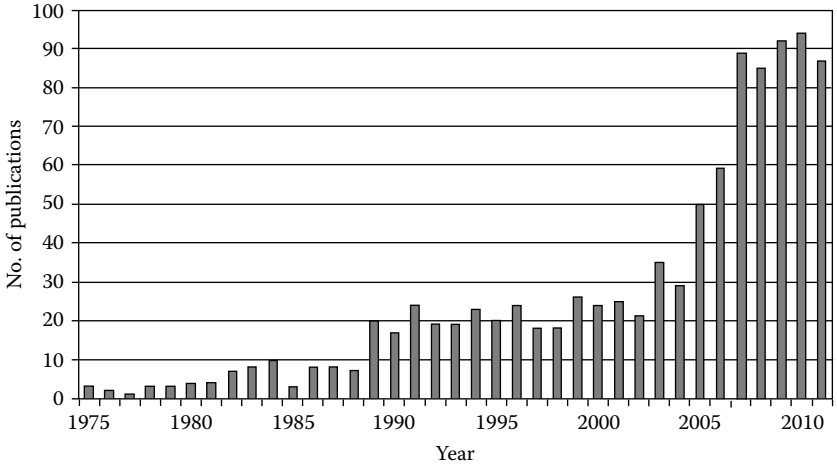


FIGURE 12.9 The number of publications in the area of hydrogen sensing according to an inquiry (June 2012) in SciVerse Scopus (www.scopus.com).

low power consumption are also important concerns for the automotive industry. From the analysis of the performance of hydrogen sensors available on the market (see Table 12.1), it is possible to conclude that there are some weaknesses concerning the performance requirements for automotive industry: although some of the automotive requirements are satisfied by the reported MOS sensor, dimensions, power consumption, and lower detection limit still need to be improved. The use of innovative technologies may enhance hydrogen sensor performance and add extra specifications allowing for integration with different systems/applications [6,7]. In fact research on novel approaches for hydrogen detection is increasing as reflected in the growing number of relevant publications in this area. Figure 12.9 shows the number of publications from 1975 until 2011.

12.5 Emerging Technologies in Hydrogen Gas Sensors

Emerging nanofabrication technologies enable the fabrication of submicron and nanometric devices working according to new hydrogen detection principles that add further performance improvements for gas sensing [17,18].

This section reviews and analyzes various emerging hydrogen-sensing techniques and discusses their advantages and weaknesses. Research and development work being performed to compensate these weaknesses is also presented.

12.5.1 Acoustic Hydrogen Gas Sensors

Acoustic technology offers a novel way of hydrogen sensing since it has the potential for passive wireless gas sensing that facilities the integration in sensor network applications. An acoustic sensor is constructed from two transducers on a piezoelectric substrate such as $LiNbO_3$ or $LiTaO_3$: one of these transducers is responsible to convert the input signal into an acoustic wave, while the second converts the acoustic wave back into an electric output signal as shown in Figure 12.10. This technique claims more robustness under different environment conditions [6].

The operating principle of acoustic gas sensor is based on the propagation of an elastic wave in the device by piezoelectric effect. The gas detection layer (usually palladium) is placed in the path of the acoustic wave between the input and the output transducers. When the sensitive layer absorbs hydrogen, a change in its mechanical (mass density), electrical properties (electric conductivity), or both will occur that produces a detectable change in the frequency of the acoustic wave. The resonant frequency at which the elastic wave propagates depends on the width of the transducer, while the mode of operation depends on elastic wave propagation mode, device geometry, and adopted technology [9]. Figure 12.10 shows a top view of the acoustic gas sensor structure.

Surface acoustic wave (SAW) mode is the most commonly used mode in hydrogen gas detection [19]. In this mode, the use of palladium or a polymer as the gas detection layer permits a mass density change (mass loading) due to the absorption of hydrogen. In this mode, the mass-loading effect dominates since the conductivity of the sensing material is too low to modify the propagation of the

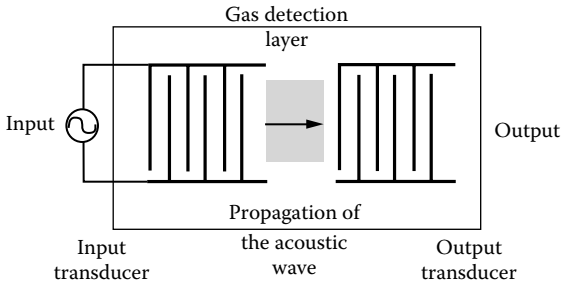


FIGURE 12.10 Top view of an acoustic gas sensor with the input and output transducer.

acoustic wave [20]. To modify the propagation of SAW by changing the electric conductivity, semiconducting MOXs such as WO_3 [21], InO_x [22], and SiN_x [23] are used. In order to improve the sensitivity of the sensitive layer in SAW devices, the gas detection layer is nanostructured to increase the surface-to-volume ratio and hence having more surface contact with the hydrogen that permits a fast response to the gas [24]. Nanostructured ZnO as well as palladium nanowires have been investigated [25,26] showing an enhancement in the sensitivity compared to film structures of the same materials. To improve the selectivity to hydrogen, SAW sensors incorporating a hydrogen catalyst with the semiconducting MOXs have been reported [27].

However, this propagation mode shows some weaknesses, especially a poor stability for temperature and humidity [8]. Therefore, the other propagation modes must be exploited for automotive industry requirements. For example, use of shear horizontal SAW (SH-SAW) sensors instead of SAW sensors reduces the effect of humidity thanks to the absence of vertical component of the wave in the first mode. Other wave propagation modes can be investigated for robustness as well as to facilitate their generation by the transducers [9].

12.5.2 Micromachined Hydrogen Gas Sensors

The advances in the field of microfabrication have permitted fabrication of three dimensional structures from a silicon substrate using micromachining techniques such as bulk micromachining, surface micromachining, dissolved wafer process, LIGA, and electrodischarge machining (EDM) [28]. The available commercial micromachined hydrogen sensors apply the same detection principles discussed previously but with microfabricated sensing elements. MOX, TCD, and CAT micromachined sensors have the advantage of improved performance compared with their conventionally fabricated counterparts.

Silicon wafers coated with silicon dioxide (SiO_2) and silicon nitride (Si_2N_3) are usually used for integration with other components or for micromachining as a sacrificial layer. The structure of the micromachined sensor is almost the same for all these detection principles, which consists of a suspended membrane on which the sensing elements are placed. The membrane is fabricated by either bulk or surface micromachining. Figure 12.11 shows the structure of the micromachined (TCD) gas sensor.

In case of CAT micromachined gas sensor, there are two suspended membranes: one for the active element and the other for the reference element that replace the beads. This can be considered as miniaturization of the classical bead structure that is usually in the order of centimeters [29]. Figure 12.12 is a cross section of a micromachined CAT sensor.

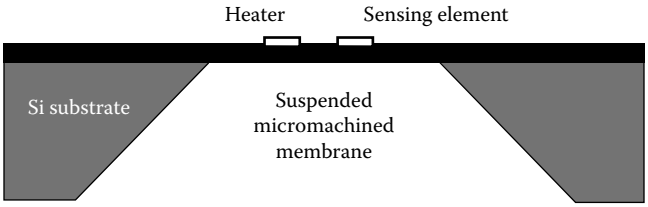


FIGURE 12.11 Micromachined structures for TCD H_2 commercial sensors.

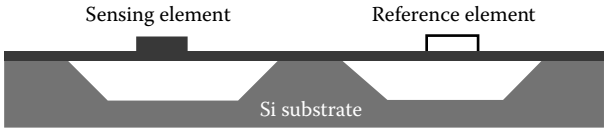


FIGURE 12.12 The structure of micromachined CAT gas sensor on Si substrate.

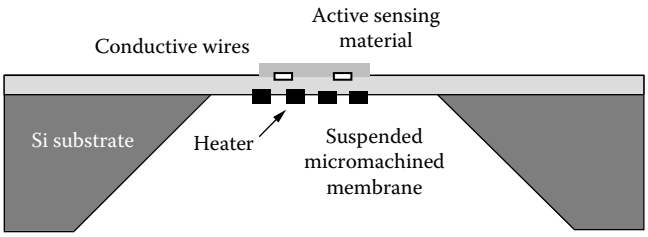


FIGURE 12.13 Micromachined structure of MOX hydrogen gas sensor.

Semiconductive MOX hydrogen sensors can also be fabricated using micromachining techniques whereby the MOX material is deposited on the micromachined suspended membrane. Some MOX structures have the heating element on the same surface as the active material [30], but it is also possible for the heater to be deposited on the underside of the membrane (Figure 12.13).

Micromachined hydrogen sensors show improved performance over conventionally designed and manufactured hydrogen sensors. They are faster and more stable under different environmental conditions. Moreover, they are cheaper and smaller and consume less power. However, there are some challenges facing these sensors and their low cost can only be realized when mass produced.

12.5.3 Nano-/Microelectromechanical System Hydrogen Gas Sensors

Nanolithography allows fabricating nano-/microelectromechanical system (N-MEMS) with high surface-to-volume ratio and with dimensions in the order of a few atomic distances. N-MEMS devices are used in several applications such as mechanical resonators, RF switches, actuators, and sensors. N-MEMS resonators have a quality factor in the order of ~10^3 and an operation frequency in the MHz range. Their functionality is based on the mechanical motion of their components due to an external stimulus. Designing an N-MEMS resonator implies choosing the material and the geometry of the structure. For example, a double-clamped N-MEMS beam resonator has a resonant frequency f_0 given by [31–33]

$$f_0 = 1.03 \frac{w}{L^2} \sqrt{\frac{E}{\rho}} \tag{12.3}$$

where

w is the width

L is the length of the beam

E and ρ are the Young modulus and the density of the material, respectively

The N-MEMS resonator is usually actuated by an external source to reach its mechanical resonance frequency.

NEMS/MEMS resonator devices allow for gas detection by introducing a gas-sensitive layer. For example, N-MEMS can be used for hydrogen sensing in MHz range using gold–palladium alloy as detection material [31]. The sensing mechanism of this sensor is based on resonant frequency shift due to the presence of the hydrogen gas. Another N-MEMS hydrogen sensor has been presented [34], showing the advantage of low power consumption and high sensitivity under ambient conditions [35–37].

12.6 Conclusion

This chapter has reviewed commercially available hydrogen sensor technologies and has discussed the physical principles of hydrogen detection in each of these technologies. The chapter has summarized the performances and operating conditions of each sensor type and compared them to the current industrial

requirements and to the emergent hydrogen applications requirements. Gaps still exist between the sensor performance requirements for emerging applications and the actual performance of commercial hydrogen sensors. Novel hydrogen detection techniques have the potential to enhance hydrogen sensor performance in order to meet the needs of these emerging markets.

References

1. S. Blanchette Jr., A hydrogen economy and its impact on the world as we know it, *Energy Policy*, 36(4), 522–530, 2008.
2. ENEA-consulting, Application des systèmes Hydrogène pour les besoins du développement à horizon 2020–2025, Mai 2011, http://www.enea-consulting.com/wp-content/uploads/Rapport-H2-Dvlpmt-Final.pdf, accessed in August 2013.
3. W.M. Haynes (Ed.), *CRC Handbook of Chemistry and Physics*, 92th edn., 2011–2012.
4. J.G. Firth, A. Jones, T.A. Jones, The principles of the detection of flammable atmospheres by catalytic devices, *Combustion and Flame*, 20, 303–311, 1973.
5. E. Jones, The pellistor catalytic gas detection, in: P. Moseley, B.C. Tofield (Eds.), *Solid State Gas Sensors*, Adam Hilger, Bristol, England, pp. 17–31, 1987.
6. T. Hubert, L. Boon-Brett, G. Black, U. Banach, Hydrogen sensors—A review, *Sensors and Actuators B: Chemical*, B157, 329–352, 2011.
7. J.F. McAleer et al., Tin dioxide gas sensor: Use of the Seebeck effect, *Sensors and Actuators*, 8, 251–257, 1985.
8. M. Watanabe, R. Inoue, D. Ichikawa, K. Furusaki, Development of thermal conductivity type hydrogen sensor, *ECS Transactions*, 28(20), 31–42, 2010.
9. I. Kerroum, H. El Matbouly, F. Domingue, Survey of commercial sensors and emerging miniaturized technologies for safety applications in hydrogen vehicles, *IEEE Sensors Applications Symposium (SAS)*, Brescia, Italy, 2012.
10. B. Lutz, R. Wind, C. Kostelecky, D. Routkevitch, D. Deininger, Development of ultra-low power metal oxide sensors and arrays for embedded applications, *AIP Conference Proceedings*, New York, Vol. 1362, pp. 62–63, 2011.
11. T. Usagawa, Y. Kikuchi, Device characteristics for Pt–Ti–O gate Si–MISFETs hydrogen gas sensors, *Sensors and Actuators B*, 160, 105–114, 2011.
12. M.C. Steele, J.W. Hile, B.A. MacIver, Hydrogen-sensitive palladium gate MOS capacitors, *Journal of Applied Physics*, 47, 2537–2538, 1976.
13. L. Boon-Brett, J. Bousek, P. Moretto, Reliability of commercially available hydrogen sensors for detection of hydrogen at critical concentrations: Part II—Selected sensor test results, *International Journal of Hydrogen Energy*, 35, 562–571, 2009.
14. IEC 60079-29-1. Ed. 1.0, Explosive atmospheres—Part 29-1: Gas detectors—Performance requirements of detectors for flammable gases, 2007, http://webstore.iec.ch/preview/info_iec60079-29-1%7Bed1.0%7Db.pdf, accessed on August 9, 2013.
15. ISO 26142, Hydrogen detection apparatus, 2010, http://www.iso.org/iso/catalogue_detail.htm?csnumber=52319, accessed on August 8, 2013.
16. L. Boon-Brett, J. Bousek, G. Black, P. Moretto, P. Castello, T. Hübert, U. Banach, Identifying performance gaps in hydrogen safety sensor technology for automotive and stationary applications, *International Journal of Hydrogen Energy*, 35, 371–384, 2010.
17. M. Dragoman, A. Muller, D. Neculoiu, G. Konstantinidis, K. Grenier, D. Dubuc, L. Bary, R. Plana, H. Hartnagel, E. Fourn, E. Flahaut, Carbon nanotubes-based microwave and millimeter wave sensors, *Proceedings of the 37th European Microwave Conference*, Munich, Germany, October 2007.
18. P. Zhang, Design and fabrication of chemiresistor type micro/nano hydrogen gas sensors using integrated electrodes, PhD thesis, University of Central Florida, Orlando, FL, 2008.

19. A. D'Amico, A. Palma, E. Verona, Surface acoustic wave hydrogen sensor, *Sensors and Actuators*, 3, 31–39, 1982.
20. S.J. Ippolito, S. Kandasamy, K. Kalantar-Zadeh, W. Wlodarski, A. Holland, Comparison between conductometric and layered surface acoustic wave hydrogen gas sensors, *Smart Materials and Structures*, 15, 131–136, 2006.
21. W.P. Jakubik, Investigations of thin film structures of WO_3 and WO_3 with Pd for hydrogen detection in a surface acoustic wave sensor system, *Thin Solid Films*, 515, 8345–8350, 2007.
22. S.J. Ippolito, S. Kandasamy, K. Kalantar-Zadeh, W. Wlodarski, K. Galatsis, G. Kiriakidis, N. Katsarakis, M. Suchea, Highly sensitive layered ZnO/$LiNbO_3$ SAW device with InO_x selective layer for NO_2 and H_2 gas sensing, *Sensors and Actuators B*, 111–112, 207–212, 2005.
23. A.C. Fechete, W. Wlodarski, K. Kalantar-Zadeh, A.S. Holland, J. Antoszewski, S. Kaciulis, L. Pandolfi, SAW-based gas sensors with rf sputtered InO_x and PECVD SiN_x films: Response to H_2 and O_2 gases, *Sensors and Actuators B*, 118, 362–367, 2006.
24. C. Drake, S. Deshpande, D. Bera, S. Seal, Metallic nanostructured materials based sensors, *International Material Reviews*, 52, 259–317, 2007.
25. F.-C. Huang, Y.-Y. Chen, T.-T. Wu, A room temperature surface acoustic wave hydrogen sensor with Pt coated ZnO nanorods, *Nanotechnology*, 20, 065501, 2009.
26. C. Caliendo, I. Fratoddi, M.V. Russo, Nanostructured organometallic polymer and palladium/polymer hybrid: Surface investigation and sensitivity to relative humidity and hydrogen in surface acoustic wave sensors, *Nanotechnology*, 18, 125504, 2007.
27. P.J. Shaver, Activated tungsten oxide gas detector, *Applied Physics Letters*, 11, 255, 1997.
28. S.A. Tadigadapa, N. Najafi, Developments in microelectromechanical systems (MEMS): A manufacturing perspective, *Transactions of the ASME*, 125, 816–823, 2003.
29. E.-B. Lee, I.-S. Hwang, J.-H. Cha, H.-J. Lee, W.-B. Lee, J.J. Pak, J.-H. Lee, B.-K. Ju, Micromachined catalytic combustible hydrogen gas sensor, *Sensors and Actuators B*, 153, 392–397, 2011.
30. I. Simon, N. Bârsan, M. Bauer, U. Weimar, Micromachined metal oxide gas sensors: Opportunities to improve sensor performance, *Sensors and Actuators B*, 73, 1–26, 2001.
31. X.M.H. Huang et al., Nanomechanical hydrogen sensing, *Applied Physics Letters*, 86, 143104, 2005.
32. C. Durand, Développement de résonateurs électromécaniques en technologie Silicon On Nothing, à détection capacitive et amplifiée par transistor MOS, en vue d'une co-intégration permettant d'adresser une application de référence de temps, PhD thesis, Université des Sciences et Technologies de Lille, Lille, France, N d'ordre: 4337, janvier 2009.
33. N.V. Lavrik, M.J. Sepaniak, P.G. Datskos, Cantilever transducers as a platform for chemical and biological sensors, *Review of Scientific Instruments*, 75(7), 2229, 2004.
34. J. Henrriksson, L.G. Villanueva, and J. Brugger, Ultra-low power palladium-coated MEMS resonators for hydrogen detection under ambient conditions, *Transducers'11*, Beijing, China, June 5–9, 2011.
35. X.L. Feng, C.J. White, A. Hajimiri, M.L. Roukes, A self-sustaining ultrahigh-frequency nanoelectromechanical oscillator, *Nature*, 3, June 2008.
36. M. Dragoman et al., Carbon nanotubes-based microwave and millimeter wave sensors, *Proceedings of the 37th European Microwave Conference*, Munich, Germany, October 2007.
37. M. Li, B. Myers, X. Tang, S.J. Aldridge, C. McCaig, J.J. Whiting, R.J. Simonson, N.S. Lewis, M.L. Roukes, Nanoelectromechanical resonator arrays for ultrafast, gas-phase chromatographic chemical analysis, *Nanoletters*, 10, 3899–3903, 2010.
38. G. Tortissier, Étude et développement d'une plateforme de détection chimique à ondes acoustiques de surface pour environnement sévère haute température, PhD thesis, University Bordeaux 1, Bordeaux, France, 2009.
39. M. Kimura, N. Takashima, MEMS hydrogen gas sensor for the entire concentration range of hydrogen gas, *Sensors and Materials*, 23(7), 201, 2011.

13

Thermal Sensors: Flow

Nam-Trung Nguyen
The University of California, Berkeley

13.1 Introduction

This chapter deals with fundamentals of thermal methods for sensing and measuring mass fluid flows. The flow can be defined as the motion of a continuum (fluid) in a closed structure (channel, orifice), and it is the measured object. The associated physical quantity sensed is the mass flux that flows through a unit cross section. The equation for the volume flow rate Φ_v is given by

$$\Phi_v = \frac{dV}{dt} = vA \tag{13.1}$$

where

V is the volume through in the time t

v is the average velocity over the cross-sectional area A of the channel

with the relation between the volume V, mass M, and the density of fluid ρ

$$V = \frac{M}{\rho} \tag{13.2}$$

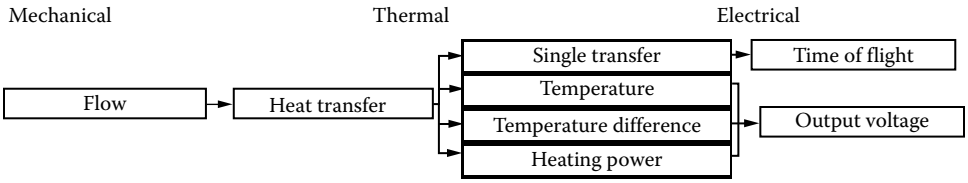


FIGURE 13.1 The three signal domains and the signal transfer process of a thermal flow sensor.

the mass flow rate Φ_m can be derived by using Equations 13.1 and 13.2 to obtain

$$\Phi_m = (\Phi_v \rho) + \left(V \frac{d\rho}{dt} \right) \tag{13.3}$$

For time invariable fluid density, one obtains

$$\Phi_m = \Phi_v \rho = Av\rho \tag{13.4}$$

A thermal mass flow sensor will generally output a signal related to the mass flux

$$\phi_m = \frac{\Phi_m}{A} = v\rho \tag{13.5}$$

and convert the mechanical variable (mass flow) via a thermal variable (heat transfer) into an electrical signal (current or voltage) that can be processed by, for example, a microcontroller. Figure 13.1 illustrates this working principle. The working range for any mass flux sensor is somewhat dependent on the fluid properties, such as thermal conductivity, specific heat, and density, but not on the physical state (gas or liquid) of the fluid.

13.2 Principles of Conventional Thermal Mass Flow Sensors

With two heater control modes and two evaluation modes, there are six operational modes shown in Table 13.1 and three types of thermal mass sensors:

- Thermal mass sensors that sense the effect of the flowing fluid on a hot body (increase of heating power with constant heater temperature, decrease of heater temperature with constant heating power). They are usually called hot-wire, hot-film sensors, or hot-element sensors.
- Thermal mass sensors that respond to the displacement of temperature profile around the heater, which is modulated by the fluid flow. These are called calorimetric sensors.
- Thermal mass sensors that sense the passage time of a heat pulse over a known distance. They are usually called time-of-flight sensors.

TABLE 13.1 Operational Modes of Thermal Mass Flow Sensors

Heater Controls	Constant Heating Power		Constant Heater Temperature	
Evaluation	Heater temperature	Temperature difference	Heating power	Temperature difference
Operational modes	Hot-wire and hot-film types	Calorimetric type	Hot-wire and hot-film types	Calorimetric type
	Time-of-flight type		Time-of-flight type	

13.2.1 Hot-Wire and Hot-Film Sensors

The dependence of the heat loss between a fine wire as well as a thin film and the surrounding fluid has traditionally been the most accepted method for measuring a fluid flow: the hot-wire method. The hot-film method uses film sensors for detecting the flow. The basic elements of this sensor type are discussed.

Flow channel: In contrast to the thermal anemometer described in Chapter 62 of Volume 1 of this book for point measurement, thermal mass flow sensors of the hot-wire and hot-film type have a flow channel defining the mass flow. Figure 13.2 shows two typical arrangements.

Sensor element: Hot-wire sensors are fabricated from platinum, platinum-coated tungsten, or a platinum–iridium alloy. Since the wire sensor is extremely fragile, hot-wire sensors are usually used only for clean air or gas applications. On the other hand, hot-film sensors are extremely rugged; therefore, they can be used in both liquid and contaminated-gas environments. In the hot-film sensor, the high-purity platinum film is bonded to the rod. The thin film is protected by a thin coating of alumina if the sensor will be used in a gas or of quartz if the sensor will be used in a liquid. The alumina coatings have a high abrasion resistance and high thermal conductivity. Quartz coatings are less porous and can be used in heavier layers for electrical insulation. Typical hot-wire and hot-film sensors are shown in Figure 13.3.

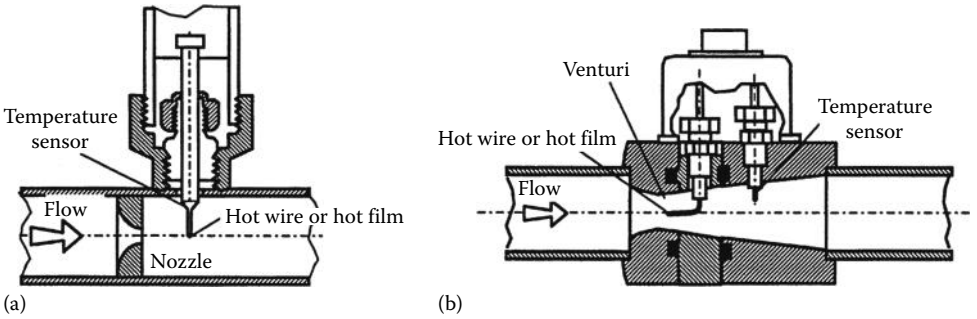


FIGURE 13.2 Typical arrangements of flow channel: (a) nozzle type and (b) venturi type.

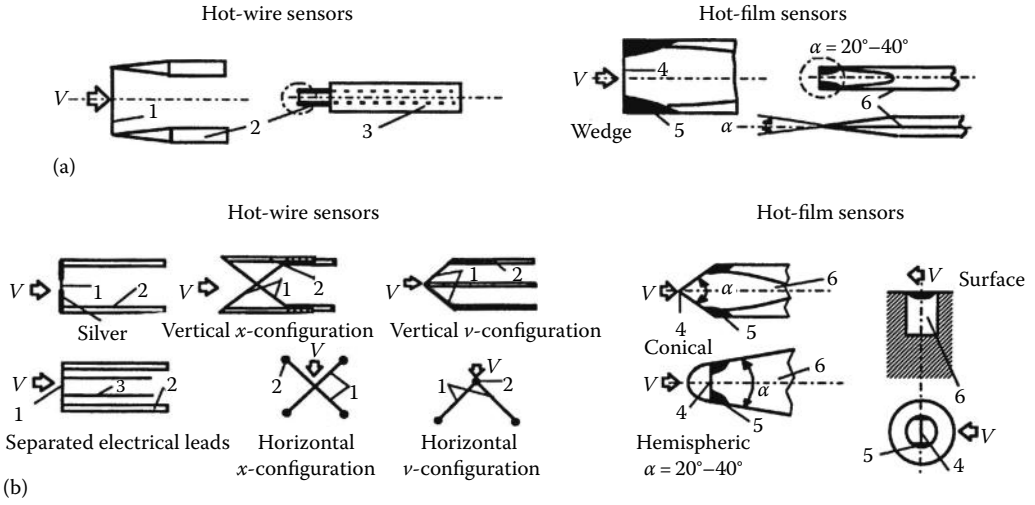


FIGURE 13.3 Typical hot-wire and hot-film sensors: (a) basic shape of the sensor element and (b) further shapes of the sensor element. (1, hot-wire; 2, sensor supports; 3, electric leads; 4, hot-film; 5, contact caps; 6, quartz rod.)

The sensor element, whether it is a wire or a film, should be a resistor that has a resistance with a high temperature coefficient α. For most sensor materials, the temperature dependence can simply be expressed by a first-order function:

$$R = R_r[1 + \alpha(T - T_r)] \tag{13.6}$$

where

R is the resistance at operating temperature T
R_r is the resistance at reference temperature T_r
α is the temperature coefficient

For research applications, cylindrical sensors are most common, either a fine wire (typical diameters from 1 to 15 μm) or a cylindrical film (typical diameters from 25 to 150 μm). Industrial sensors are often a resistance wire wrapped around a ceramic substrate that has typical diameters from 0.02 to 2 mm.

Table 13.2 shows some typical parameters of industrial hot-wire and hot-film sensors.

Control and evaluation circuit: The constant-current and constant-temperature bridge are conventional circuits for control and evaluation of heat-wire or heat-film sensors.

A constant-current Wheatstone bridge with a hot-wire sensor is shown schematically in Figure 13.4a. In this circuit, resistors R_3 and R_4 are much larger than sensor resistor R_1. Therefore, current through R_1 is essentially independent of changes in the sensor resistor R_1. Any flow in the channel cools the hot wire, decreases its resistance as given by Equation 13.6, and unbalances the bridge. The unbalanced bridge produces an output voltage V_o, which is related to the mass flow. Because the output voltage V_o from the bridge is small, it must be amplified before it is recorded. The value of the thermal coefficient α for R_1 and R_2 should be equal in order to eliminate signal errors due to changes in ambient temperature. Similarly, thermal coefficients for R_3 and R_4 should also be equal.

TABLE 13.2 Typical Parameters of Hot-Wire and Hot-Film Sensors

Parameter	Hot Wire	Hot Film
Sensor element	Platinum hot wire (diameter of 70 μm)	Platinum hot film (alumina coated)
Operational mode	Constant heater temperature in air	
Working temperature range	−30°C to 200°C	
Characteristics	Nonlinear	
Accuracy in %	±4	±2
Time response in ms	<5	12
Sensitivity in mV kg^{-1} h^{-1}	1	5

Source: Schnell, G., *Sensoren in der Automatisierungstechnik (Sensors in the Automation Techniques)*, Verlag Viehweg, Braunschweig, Germany, 1993.

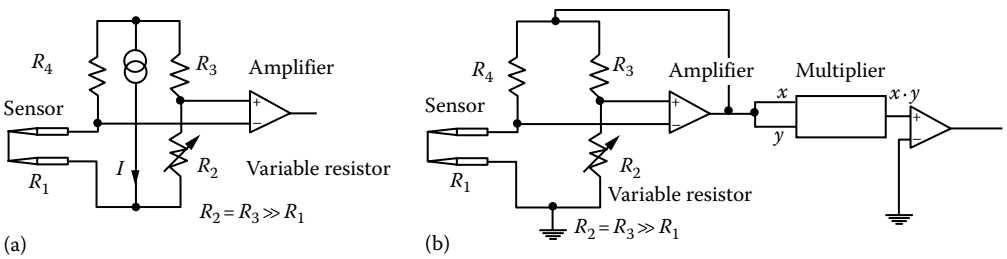


FIGURE 13.4 Control and evaluation circuit of heat-wire and heat-film sensors: (a) constant-current bridge and (b) constant-temperature bridge.

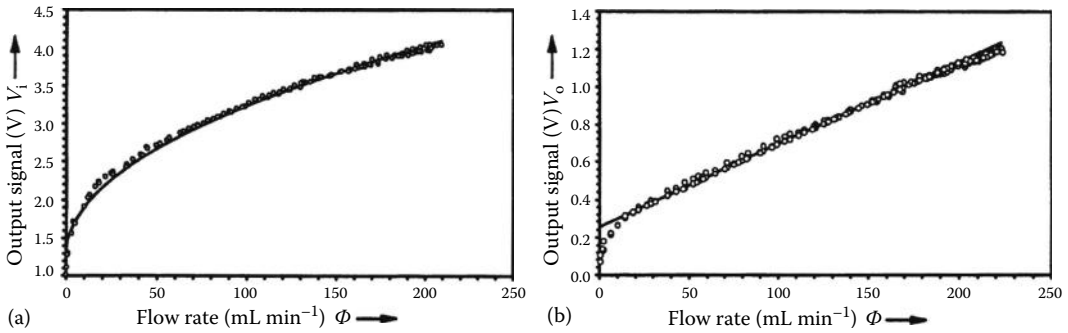


FIGURE 13.5 Sensor characteristics of the constant-temperature mode before (a) and after (b) linearization.

The constant-temperature Wheatstone bridge is shown in Figure 13.4b. The bridge is balanced under no-flow conditions with the variable resistor R_2. The flow cools the hot wire, and its resistance decreases and unbalances the bridge. A differential amplifier balances the bridge with the feedback voltage. The output signal can be linearized before recording.

Because the output signal has a square-root-like characteristic, the linearizer can be realized easily using a multiplier (i.e., AD534 of analog devices) with two equal input signals (a squarer). Figure 13.5 illustrates the results. With this method, there is linearization error in the low-flow range.

13.2.2 Calorimetric Sensors

The displacement of the temperature profile caused by the fluid flow around a heating element can be used for measuring very small mass flow. Depending on the location of the heating and sensing elements, there are two types of calorimetric sensors: the intrusive sensors that lie in the fluid and the nonintrusive sensors that are located outside the flow. Table 13.3 illustrates the two typical calorimetric sensors.

The intrusive type has many limitations. The heater and the temperature sensors must protrude into the fluid. Therefore, corrosion and erosion damage these elements easily. Furthermore, the integrity of the piping is sacrificed by the protrusions into the flow, thus increasing the danger of leakage.

In the nonintrusive sensor type, the heater and the temperature sensors essentially surround the flow by being located on the outside of the tube that contains the flow. The major advantage of this sensor type is the fact that no sensor is exposed to the flowing fluid, which can be very corrosive. This technique is generally applied to flows in the range of 1–500 L min^{-1}. The larger flows are measured using the bypass arrangement. Figure 13.6a shows the measured shift of the temperature distribution around the heater. The asymmetricity of the temperature profile increases with flow.

TABLE 13.3 Typical Parameters of Calorimetric Sensors

Parameter	Gases	Liquids
Working temperature	0 °C–70 °C	0 °C–70 °C
Accuracy in %	±1%	±1%
Linearity	±0.2%	±0.2%
Flow range 1:50	min. 5 mL min^{-1}	min. 5 g h^{-1}
	max. 100 L min^{-1}	max. 1000 g h^{-1}

Source: Mass Flow and Pressure Meters/Controllers, Bronkhorst Hi-Tec B.V., Ruurlo, the Netherlands, 1994.

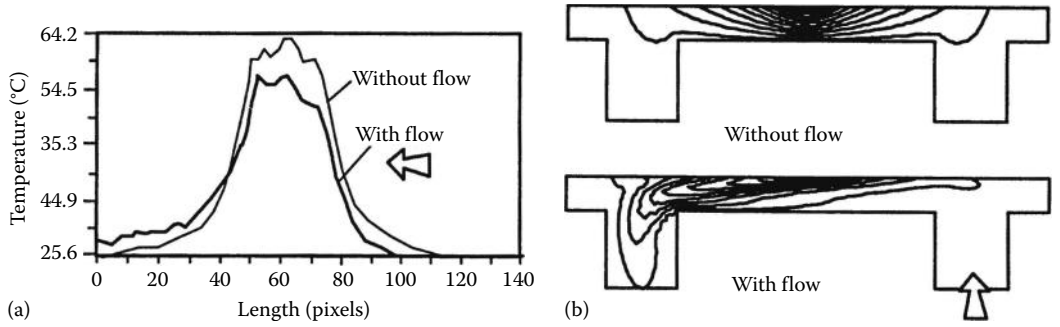


FIGURE 13.6 Temperature distribution around the heater: (a) measurement and (b) numerical simulation.

The measurement was carried out using a thermography system [1]. To understand the working principle, the effect of fluid mechanics and heat transfer should be reviewed. The mathematical theory for this problem is discussed later in this chapter. Figure 13.6b illustrates the influence of the flow over the temperature distribution.

Because the calorimetric mass flow sensors are sensitive in low-flow ranges, bypass designs have been introduced in order to make the sensors suitable for the measurement of larger flow ranges. The sensor element is a small capillary tube (usually less than 3 mm in diameter). They ensure laminar flow over the full measurement range. The laminar flow elements are located parallel to the sensor element as a bypass (Figure 13.7a). They are usually a small tube bundle, a stack of disks with etched capillary channels [2] (Figure 13.7b), or a machined annular channel.

Figure 13.8 shows typical parameters of calorimetric flow sensors. Compared to the hot-wire or hot-film sensors, this sensor type has good linearity and is only limited by signal noise at low flows and saturation at high flows. While the linear range may exceed a 100:1 ratio, the measurable range may be as large as 10,000:1 [3]. The small size of the capillary sensor tube is advantageous in minimizing the electric power requirement and also in increasing the time of response. Because of the small size of the tube, it necessitates the use of upstream filters to protect against plugging of dust particles. With the bypass arrangement, a relatively wide flow range is possible.

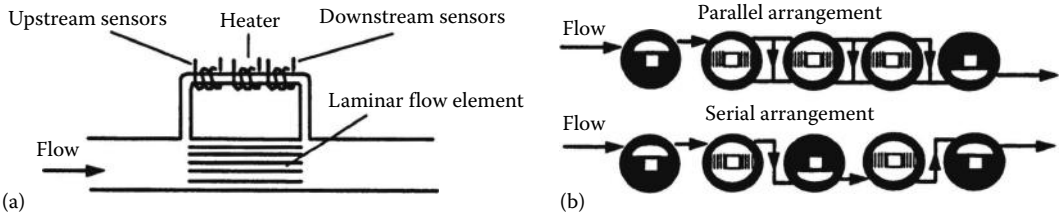


FIGURE 13.7 Bypass arrangement for large flow range: (a) principle and (b) a solution for the laminar flow element. (From *Mass Flow and Pressure Meters/Controllers*, Bronkhorst Hi-Tec B.V., Ruurlo, the Netherlands, 1994.)

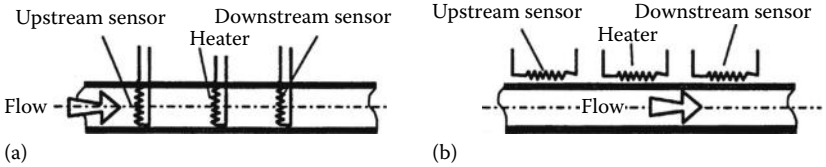


FIGURE 13.8 Typical calorimetric sensors: (a) intrusive type and (b) nonintrusive type.

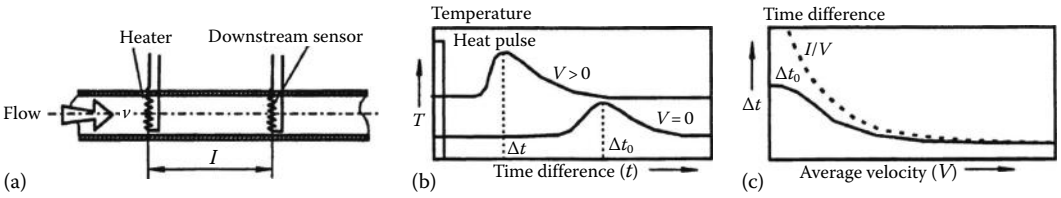


FIGURE 13.9 Time-of-flight sensors: (a) principle, (b) temperature at downstream sensor, and (c) sensor characteristic.

13.2.3 Time-of-Flight Sensors

The time-of-flight sensor consists of a heater and one or more temperature sensors downstream, Figure 13.9a. The heater is activated by a current pulse. The transport of the generated heat is a combination of diffusion and forced convection. The resulting temperature field can be detected by temperature sensors located downstream. The detected temperature output signal of the temperature sensor is a function of time and flow velocity. The sensor output is the time difference between the starting point of the generated heat pulse and the point in time at which a maximum temperature at the downstream sensor is reached, Figure 13.9b. At the relatively low flow rates, the time difference depends mainly on the diffusivity of the fluid medium. At relatively high flow rates, the time difference tends to relate to the ratio of the heater–sensor distance and the average flow velocity [4].

Because of the arrangement shown in Figure 13.9a, the time-of-flight sensors have the same limitations as the intrusive type of calorimetric sensors: corrosion, erosion, and leakage. Since the signal processing needs a while to measure the time difference, this sensor type is not suitable for dynamic measurement. The advantage of this type of volumetric flow sensor is the independence of fluid properties as well as fluid temperature in the higher flow range. The influence of fluid properties on the mass flow sensor output is described in [3], as well as an approach to compensate for changes in these properties, which is valid for both hot-element and calorimetric sensors.

13.2.4 Mass and Heat Transfer

The most important signal of the transfer process shown in Figure 13.1 is the thermal signal. There are different kinds of thermal signals: temperature, heat, heat capacity, and thermal resistance. In the following, the transfer of heat and the interaction between heat and temperature will be explained by three mean heat transfer processes: conduction, convection, and radiation. The first two processes can be described by the general equation of a transfer process. The transfer variables in the equation can be the momentum (momentum equation), the temperature (energy equation), or the mass (mass equation).

A transfer process consists of four elements: accumulation, conduction, induction, and convection. The *accumulation process* describes the time dependence of the transfer variable. The *conduction* presents the molecular transfer. The *convection* is the result of the interaction between the flow field and the field of the transfer variable. The *induction* describes the influence of external fields and sources.

13.2.5 Conduction

When there is a temperature gradient in a substance, the heat will flow from the hotter to the colder region, and this heat flow q (in W m^{-2}) will be directly proportional to the value of the temperature gradient:

$$q = -\lambda \frac{dT}{dx} \tag{13.7}$$

where T is the temperature. The previous expression is called Fourier's law of heat conduction and defines the material constant λ (in W K^{-1} m^{-1}), the thermal conductivity. Figure 13.10 shows the order of the thermal conductivity λ. of different materials.

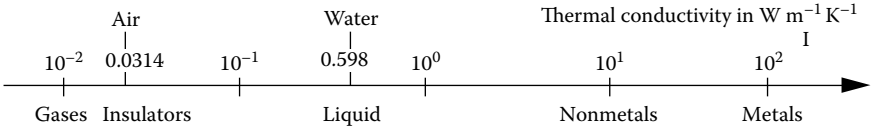


FIGURE 13.10 The order of thermal conductivity of different materials.

The differential form of the heat-conduction equation is a special case of the energy equation (see Section 13.2.6). The transfer equation only consists of the accumulative, conductive, and inductive terms:

$$\frac{\partial T}{\partial t} = \frac{\lambda}{\rho c}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) + \frac{q'}{\rho c} \tag{13.8}$$

where

ρ is the density

c is the specific heat at constant pressure

q' (in W m^{-3}) is the amount of heat (in joules) per unit of volume and time that can be generated inside the material itself, either through the action of a separate heat source or through a change in phase of matter

In the steady state without internal heat sources, the equation of conduction reduces to

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = 0 \tag{13.9}$$

13.2.6 Convection

In general, there are two kinds of convection: forced convection and natural convection. The first one is caused by a fluid flow and the other one by itself because of the temperature dependency of fluid density and the buoyancy forces. To describe convection, three conservation equations are required:

Conservation of mass: continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla(\rho v) = 0 \tag{13.10}$$

Conservation of momentum: Navier–Stokes equation

$$\frac{\rho \partial v}{\partial t} + v\nabla v = -\nabla p + \eta \nabla^2 v + \rho g \tag{13.11}$$

Conservation of energy: energy equation

$$\frac{\partial T}{\partial t} + v\nabla T = \left(\frac{\lambda}{\rho c}\right)\nabla^2 T + \frac{q'}{\rho c} \tag{13.12}$$

where η is the dynamic viscosity of the fluid. The temperature field and the heat power can be found by solving these three equations. For designing and understanding the thermal flow sensor, the convective heat transfer can be expressed in the simplest form:

$$Q = \varepsilon A \Delta T \tag{13.13}$$

where

Q (in W) is the heat transfer rate (or the heat power)
ε is the heat transfer coefficient between the heated surface A and the fluid
ΔT is the temperature difference between the heated body and ambient

The dimensionless Nusselt number describes the heat transfer. The relationship between the heat transfer coefficient ε and the Nusselt number Nu can be expressed as follows:

$$\varepsilon = \mathrm{Nu}\frac{\lambda}{L} \tag{13.14}$$

where L is the characteristic length (the length L of a flat plate, the hydraulic diameter D_h of a tube, and the half of the perimeter of a wire, Figure 13.11). The hydraulic diameter D_h can be calculated using the wetted perimeter U and the cross-sectional area A of the tube:

$$D_h = \frac{4A}{U} \tag{13.15}$$

The relevant dimensionless number that describes the flow is the Reynolds number

$$\mathrm{Re} = \frac{vL}{\nu} \tag{13.16}$$

where

v is the average flow velocity
ν the kinematic viscosity of the fluid, which is defined by the density ρ and the dynamic viscosity η:

$$\nu = \frac{\eta}{\rho} \tag{13.17}$$

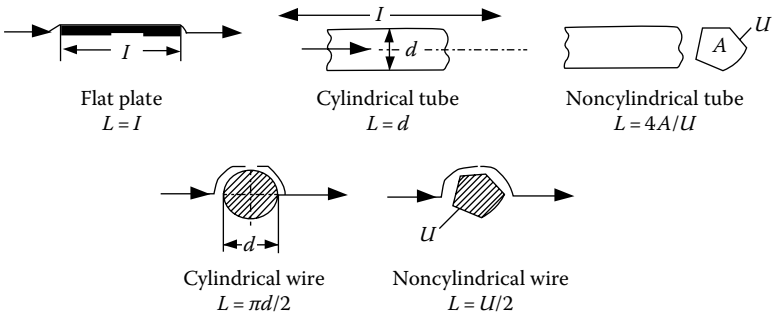


FIGURE 13.11 Definition of the characteristic length of different objects. (From VDI-Wärmeatlas, VDI-Verlag, 1994.)

TABLE 13.4 Nusselt Number (Nu) of Forced Convection

Object	Nu_{lam} for Laminar Regime	Nu_{turb} for Turbulent Regime	Average Nusselt Number Nu
Flat plate	$Nu_{lam} = 0.664\sqrt{Re^3}\sqrt{Pr}$	$Nu_{turb} = \dfrac{0.037\ Re^{0.8}\ Pr}{1 + 2.443\ \ Re^{-0.1}(Pr^{2/3} - 1)}$	$Nu = \sqrt{Nu_{lam}^2 + Nu_{turb}^2}$
	$Re < 10^5$; $0.6 < Pr < 2000$	$5.10^5 < Re < 10^7$; $0.6 < Pr < 2000$	$10 < Re < 10^7$; $0.6 < Pr < 2000$
Cylindrical tube	$Nu_{lam} = 3.65 + \dfrac{0.19(Re\ Pr\ d/l)^{0.8}}{1 + 0.117(Re\ Pr\ d/l)^{0.467}}$	$Nu_{turb} = \dfrac{\xi/8(Re - 1000)\,Pr}{1 + 12.7\sqrt{\xi/8}\,(Pr^{2/3} - 1)}\left[1 + \left(\dfrac{d}{l}\right)^{2/3}\right]$	
	$Re < 2300$; $0.1 < (Re\ Pr\ d/l) < 10^4$	where $\xi = (1.28\ \log_{10}Re - 1.64)^{-2}$	
Short cylindrical tube $0.1 < d/l < 1$	$Nu_{lam} = 0.664^3\sqrt{Pr}\sqrt{Re\,d/l}$	Thermal entrance, fully developed flow, $2300 < Re < 10^4$	
	$Re < 2300$; $0.1 < (Re\ Pr\ d/l) < 10^4$	$Nu_{turb} = \sqrt[3]{3.66^3 + 1.61^3\ Re\,Pr\,d/l}$	
		Thermal entrance, developing flow, $2300 < Re < 10^6$	
Wire	$Nu_{lam} = 0.664\sqrt{Re^3}\sqrt{Pr}$	$Nu_{turb} = \dfrac{0.037\,Re^{0.8}\ Pr}{1 + 2.443\ \ Re^{-0.1}(Pr^{2/3} - 1)}$	$Nu = 0.3 + \sqrt{Nu_{lam}^2 + Nu_{turb}^2}$
	$10 < Re < 10^7$; $0.6 < Pr < 1000$		

Source: VDI-Wärmeatlas, VDI-Verlag, 1994.

TABLE 13.5 Average Nusselt Number of Free Convection in Some Special Cases

Cases	Equation	References
Vertical flat plate or wire	$Nu = 0.55(Gr\ Pr)^{1/4}$	[8]
Horizontal flat plate	For the upper surface: $Nu = 0.76(Gr\ Pr)^{1/4}$	[9]
	For the lower surface: $Nu = 0.38(Gr\ Pr)^{1/4}$	

Table 13.4 shows a collection of formulae for calculating the Nusselt number. The fluid properties (kinematic viscosity ν and Prandtl number Pr) should be chosen at the average temperature T_{av} between the heater temperature $T + \Delta T$ and the fluid temperature T:

$$T_{av} = T + \frac{\Delta T}{2} \tag{13.18}$$

In the case of natural convection, the Nusselt number depends on the Grashof number Gr, which describes the influence of buoyancy forces:

$$Gr = \frac{g\beta L^3 \Delta T}{\nu^2} \tag{13.19}$$

where

g is the acceleration due to the gravity (9.81 m s^{-2})
β the thermal expansion coefficient of the fluid
ΔT is the temperature difference between the hot fluid and the ambient

The average Nusselt number can be calculated for a laminar flow ($10^4 < Gr\ Pr < 10^8$) in Table 13.5.

13.2.7 Radiation

A body can either emit or absorb thermal radiation. Radiation is not important for the operational principle of thermal mass flow because of its relatively low magnitude.

Following, the physical and mathematical backgrounds of these three sensor types are discussed in detail. The working principle and the influence of fluid properties on the sensor signal can be determined using these mathematical models. However, the mathematical models only describe the relationship between thermal variables (heat power, temperature) and the average velocity. Further relationships between mass flow, mass flux, thermal, and electrical variables can be derived using Equations 13.1 through 13.6.

13.3 Analytical Models for Calorimetric Flow Sensing

13.3.1 Model for the Intrusive Type of Calorimetric Sensors

In a quasi-static situation, the incoming heat at a certain point in the fluid must be equal to the outgoing heat. The heat is transported either by conduction in the fluid and/or supporting beams or by convection through the thermal mass of the fluid. Ultimately, the heat is transported to the walls of the tube; see Figures 13.8 and 13.12. A heat balance equation results in a differential equation for T in x. The temperature profile in the y and z directions is assumed to be constant and linear, respectively [5].

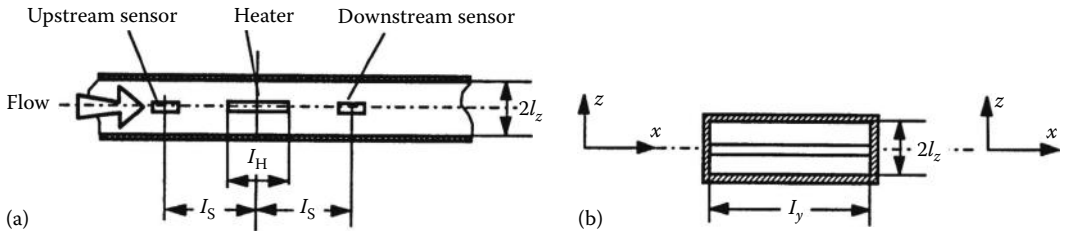


FIGURE 13.12 Analytical model for the intrusive type of calorimetric sensors: (a) length cut and (b) cross section.

Referring to Figure 13.12 and using A as the cross-sectional area of the flow channel ($A = l_y 2l_z$), ρ as the fluid density, c as the fluid heat capacity (at constant pressure), v as the average fluid velocity, and λ as the fluid thermal conductivity, one finds that

$$\frac{\partial^2 T}{\partial x^2} - v\left(\frac{\rho c}{\lambda}\right)\frac{\partial T}{\partial x} - \left(\frac{T}{l_z^2}\right) = 0 \tag{13.20}$$

or

$$\frac{\partial^2 T}{\partial x^2} - \left(\frac{v}{a}\right)\frac{\partial T}{\partial x} - \frac{T}{l_z^2} = 0 \tag{13.21}$$

where $a = \lambda/\rho c$ is the thermal diffusivity of the fluid. Equation 13.21 is linear in T. Solving the differential equation using a heater length l_H, a heater power Q, and the boundary condition

$$\lim_{x \to \pm\infty} T(x) = 0 \tag{13.22}$$

the following temperature distribution results

$$x < \frac{-l_H}{2} \quad \text{for } T(x) = T_0 \exp\left[\gamma_1\left(x + \frac{l_H}{2}\right)\right] \tag{13.23}$$

$$\frac{l_H}{2} < x < \frac{l_H}{2} \quad \text{for } T(x) = T_0 \tag{13.24}$$

$$x > \frac{l_H}{2} \quad \text{for } T(x) = T_0 \exp\left[\gamma_2\left(x - \frac{l_H}{2}\right)\right] \tag{13.25}$$

where

$$\gamma_{1,2} = \frac{\left(v \pm \sqrt{v^2 + 2a^2/l_z^2}\right)}{(2a)} \tag{13.26}$$

$$T_0 = \frac{P}{\left[\left(2\lambda l_y l_H / l_z\right) + A\lambda\left(\gamma_1 - \gamma_2\right)\right]} \tag{13.27}$$

The temperature difference between the two sides, upstream (at $x = l_s$) and downstream (at $x = -l_s$), can be then calculated as

$$\Delta T(v) = T_0\left\{\exp\left[\gamma_2\left(\frac{l_s - l_H}{2}\right)\right] - \exp\left[\gamma_1\left(\frac{-l_s + l_H}{2}\right)\right]\right\} \tag{13.28}$$

13.3.2 Model for the Nonintrusive Type of Calorimetric Sensors

A simple, 1D model is used to show the working principle of the nonintrusive type with capillary-tube and heater wire winding around it. Geometric parameters and assumptions are given in Figure 13.13. Because of the symmetry, only half of the capillary tube will be considered for the calculation model. The conservation of thermal energy in a lumped element (Figure 13.13c) can be given in the following equation:

$$Q_{\text{cond.},x,\text{fluid}} + Q_{\text{conv.},x,\text{fluid}} + Q_{\text{cond.},x,\text{wall}} = Q_{\text{cond.},y,\text{fluid}} \tag{13.29}$$

The indices define the conduction or convection in the *x*- or *y*-axis in the fluid as well as in the heated wall. Defining the temperature along the *x*-axis as $T(x)$, the average flow velocity as v, the thermal conductivities of wall material as λ_w and of fluid as λ, the thermal diffusivity of fluid $a = \lambda/(\rho c)$,and the thickness of the average thermal boundary layer as δ, one finds the heat balance equation:

$$\left[\frac{1}{2} + \left(\frac{\lambda_W d_W}{\lambda \delta}\right)\right]\frac{\partial^2 T}{\partial x^2} - \frac{v}{(2a)}\frac{\partial T}{\partial x} - \frac{T(x)}{\partial^2} = 0 \tag{13.30}$$

The thickness of the average thermal boundary layer δ depends on the flow velocity [6]. For gases with a small Prandtl number (Pr < 1) or liquids with a low Reynolds number, one can assume that

$$\delta = r \tag{13.31}$$

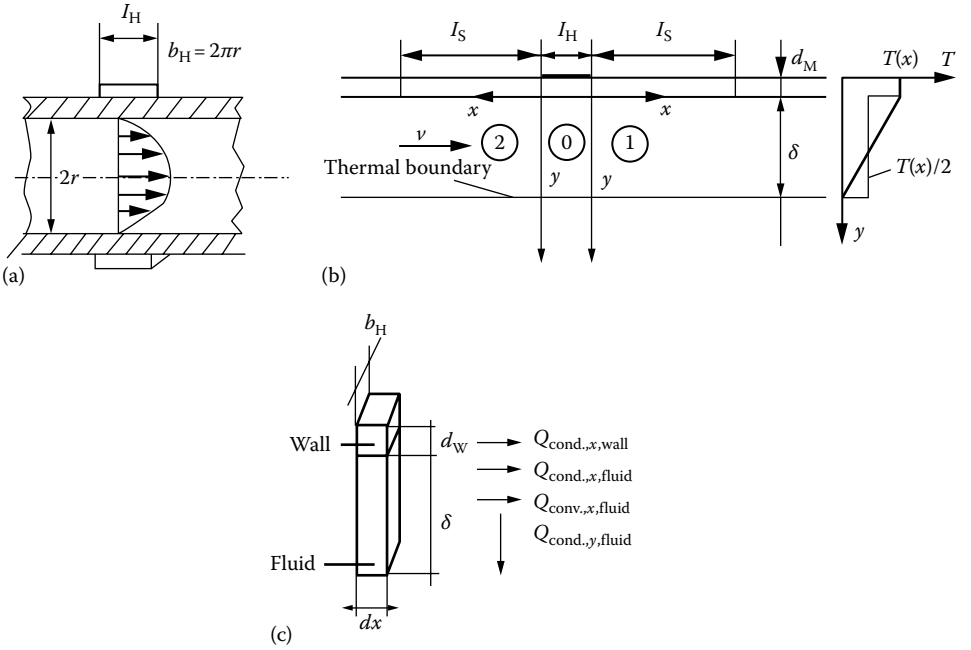


FIGURE 13.13 Analytical model for the nonintrusive type of calorimetric sensors: (a) heater and channel geometry, (b) model geometry, and (c) model of a lumped element.

After solving Equation 13.30 in the local coordinate systems 1 and 2 (Figure 13.13b), one obtains the temperature difference $\Delta T(v)$ between the temperature sensors:

$$\Delta T(v) = \vartheta_0 \left[\exp(\gamma_2 l_s) - \exp(\gamma_1 l_s) \right] \tag{13.32}$$

with

$$\gamma_{1,2} = \frac{\left(v \pm \sqrt{v^2 + 16a^2\kappa/\delta^2} \right)}{(4a\kappa)} \tag{13.33}$$

The dimensionless factor

$$\kappa = \frac{1}{2} + \frac{(\lambda_w d_w)}{(\lambda \delta)} \tag{13.34}$$

describes the influence of the wall on the heat balance. If the wall is neglected, we get $\kappa = 1/2$ as in the similar case of Equation 13.21. The heater temperature T_0 can be calculated for the constant heat power Q:

$$T_0 = \frac{Q}{\left\{ 2\pi r\lambda \left[(l_H/\delta) + \sqrt{(v^2\delta^2)/((4a^2) + 4\kappa)} \right] \right\}} \tag{13.35}$$

13.3.3 Model for the Time-of-Flight Type

The transport of the heat generated in a line source through a fluid is governed by the energy Equation 13.12. The analytical solution of this differential equation for a pulse signal with input strength $q_0' (\mathrm{W\,m^{-1}})$ is given in [7] as

$$T(x,y,t) = \left(\frac{q_0'}{4\pi\lambda t} \right) \exp\left\{ -\frac{[(x - vt)^2 + y^2]}{4at} \right\} \tag{13.36}$$

where a denotes the thermal diffusivity. By measuring the top time τ at which the signal passes the detection element ($y = 0$), in other words differential Equation 13.36 with respect to time, one can obtain the basic equation for the so-called time of flight of the heat pulse:

$$v = \frac{x}{t} \tag{13.37}$$

For Equation 13.36 to be valid, the term $4at$ must be much smaller than the heater–sensor distance x. This assumes that forced convection by the flow is dominating over the diffusive component. In other words, Equation 13.36 is true at high flow velocities. When the diffusive effect is taken into account, the time of flight is given by

$$\Delta t = \tau = \frac{[-2a + (4a^2 + v^2x^2)^{1/2}]}{v^2} \quad v \neq 0 \tag{13.38}$$

$$\tau = \frac{x^2}{4a} \quad v = 0 \tag{13.39}$$

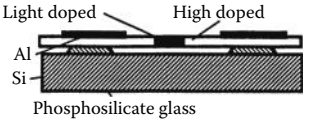
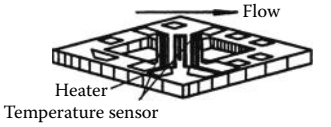
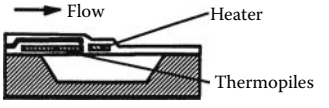
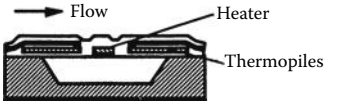
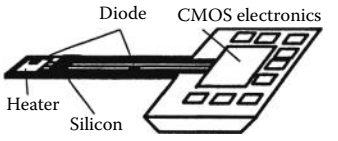
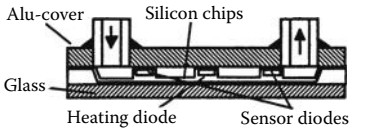
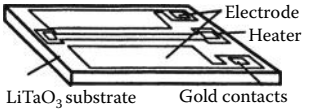
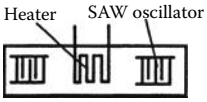
13.3.4 Principles of Microflow Sensors

In research papers, the first reference to thermal mass flow sensor normally cited is that of King in 1914. Since then, microsystems technology has been developed. The development of thermal flow sensor can be realized in micron size using the three current technologies: bulk micromachining, surface micromachining, epimicromachining and LIGA techniques (LIGA: German description of "Lithographie, Galvanoformung, Abformung"). These fabrication techniques (except LIGA) are compatible with conventional microelectronic processing technology. Thermal flow sensors developed using these technologies will be called "microflow sensors" in this section. The operational modes are similar to the conventional thermal flow sensor. In Table 13.6, the microflow sensors are classified after their transducing principle. With these new sensors, very small flows in the nanoliter and microliter range can be measured. Table 13.7 shows some realized examples of microflow sensors.

TABLE 13.6 Transducing Principle of Microflow Sensors

Transducing Principle	Realization	Application
Thermoresistive	Metal film (platinum), polycrystalline silicon, single crystalline silicon or metal alloys	Measurement of temperature, temperature difference, and heat power
Thermoelectric	*p*Si–Al (bipolar technology), polySi–Al (CMOS technology), or *p*PolySi–*n*PolySi thermopiles	Measurement of temperature and temperature difference
Thermoelectronic	Transistors, diodes	Measurement of temperature and temperature difference
Pyroelectric	Pyroelectric materials ($LiTaO_3$) with metal or silicon resistors as heater and electrodes	Measurement of heat power
Frequency analog	SAW oscillators	Measurement of temperature

TABLE 13.7 Examples of Microthermal Mass Flow Sensors

Transducing Principle	Hot Wire or Hot Film	Calorimetric
Thermoresistive [11]		
Thermoelectric [6,12]		
Thermoelectronic [13]		
Pyroelectric [14]		
Frequency analog [15]		

13.3.5 Smart Thermal Flow Sensors and Advanced Evaluation Circuits

Conventional sensors usually have separate electronics, which causes high cost and prevents large serial production. An integrated smart thermal flow sensor is defined as a chip that contains one or more sensors, signal conditioning, A/D conversion, and a bus output [6]. Therefore, there is a need for advanced evaluation methods that convert the thermal signal directly into a frequency and duty-cycle output.

13.3.6 Duty-Cycle Modulation for the Hot-Wire Sensors

The constant heater temperature can be controlled by modulation of the amplitude of the heat voltage (conventional principle) or by modulation of the duty cycle. The heater is activated when the output signal is high. This heater controlling output goes low when the temperature level $T_0 + \Delta T$ is reached. During low output, the heating temperature decreases to the temperature level $T_0 + \Delta T$, where the output goes high again and restarts the heating cycle. The temperature level is determined by a reference resistor and a variable resistor. An increase in the flow rate increases the convective cooling of the heater, and it needs a longer time to reach the temperature level. That results in the higher output time t_H. Figure 13.14 shows the working principle of the duty-cycle modulation, where t_d is the time delay in the switching action. Defining the maximum heating power as Q_{max} (the output signal is always high), one obtains the relation:

$$Q = Q_{max}\left(\frac{t_H}{t_{total}}\right) \tag{13.40}$$

13.3.7 Electrical Sigma–Delta Modulation for Calorimetric Sensors

Conventional signal-conditioning circuits use an analog-to-digital converter (ADC) to get digital sensor readout, which can be regarded as an amplifier or a voltmeter. They read the information signal from the sensor, but do not interact with the sensor. In contrast, sigma–delta converters are a part of the sensor function since they act as a feedback amplifier for the sensor output. Hence, sigma–delta conversion normally results in a much more robust sensor signal than is provided by conventional ADCs. Following, the principle of sigma–delta conversion applied to calorimetric flow sensors is explained.

The transistors T_1 and T_2 represent two switchers that feed the constant current I_0 into the *RC* network. It is assumed in this example that the temperature on the resistor R_{s2} is higher than the

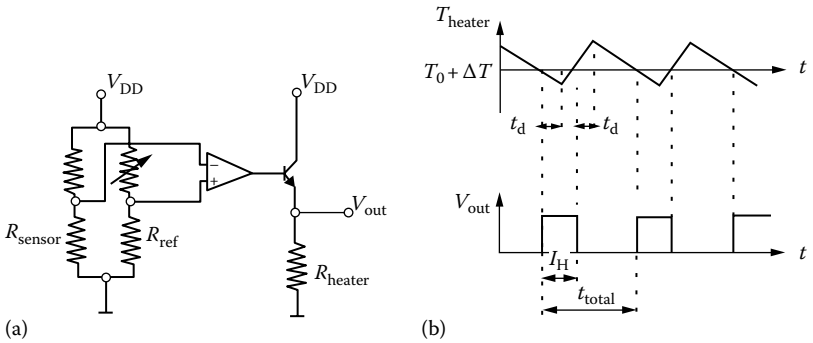


FIGURE 13.14 The duty-cycle modulation for hot-wire sensors: (a) basic circuit and (b) the heater temperature detected by R_{sensor} and the output voltage V_{out} vs. the time.

temperature on R_{s1}. Therefore, the resistance of R_{s2} is larger than R_{s1}. It results in a larger time constant for the charging of C_2. With the help of the comparators and the D flip-flop, the constant current I_0 can be switched on and off. The switching signals $f1$ and $f2$ have, in the same time period, different pulse numbers $N + S$ and N:

$$\frac{R_{s2}}{R_{s1}} = \frac{(N+S)}{N} = 1 + \frac{\Delta R}{R_{s1}} \tag{13.41}$$

Thus

$$\frac{\Delta R}{R_{s1}} = \frac{S}{N} \tag{13.42}$$

with

$$R_{s1} = \frac{(V_{ref} N)}{(I_0 G)} \tag{13.43}$$

The resistance difference as well as the temperature difference can be calculated:

$$\Delta R = \frac{(S V_{ref})}{G I_0} \tag{13.44}$$

The counting and recording of S and G are shown in Figure 13.15.

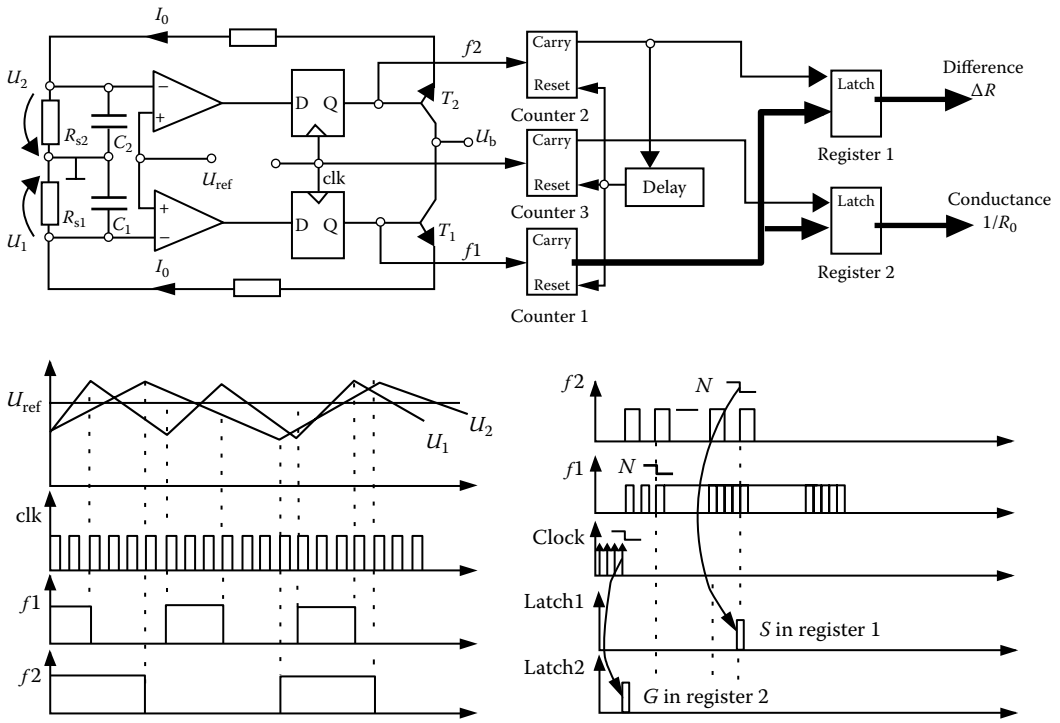


FIGURE 13.15 Sigma–delta conversions for calorimetric sensors.

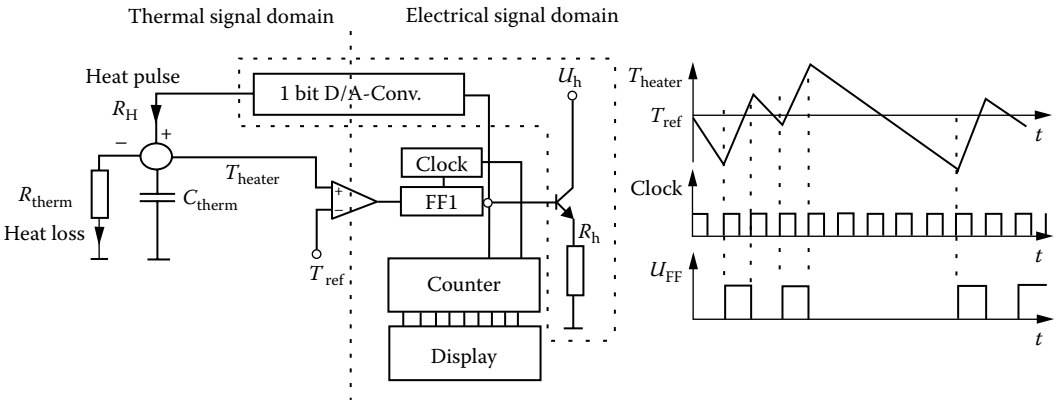


FIGURE 13.16 Principle of thermal sigma–delta modulation.

13.3.8 Thermal Sigma–Delta Modulation

The principle of thermal sigma–delta modulation is based on the conventional electrical sigma–delta: the thermal sigma–delta converter uses a thermal integrator (thermal *R*/*C* network) instead of an electric integrator (electric *R*/*C* network). Figure 13.16 shows the principle of thermal sigma–delta modulation. The comparator output modulates the flip-flop. The flip-flop synchronizes the heating signal by its clock. Therefore, the heating periods are chopped into further small pulses that depend on the clock frequency. The heater is actuated step by step until the comparator switches again, and one obtains a frequency analog output at the flip-flop [10].

13.4 Calibration Conditions

13.4.1 Calibration of Hot-Wire and Hot-Film Sensors

The hot-wire and hot-film sensors are based on the point velocity measurement (see Figure 13.2). Therefore, the measurement results depend on the velocity distribution inside the flow channel. Achieving a high signal-to-noise ratio can thus require spatial arrays of hot-wire and hot-film sensors that give more information about the velocity field and thus more accurate results of the mass flow in channel. With the use of a nozzle, a Venturi, or a flow conditioner, the flow profile is preconditioned, which leads to an acceptable accuracy.

13.4.2 Temperature Dependence of Fluid Properties

Most fluid properties depend on the working temperature. The heat transfer process depends on the fluid properties. The measurement of fluid temperature (see Figure 13.2) keeps the heater on a constant temperature difference to the fluid and can also be used for compensation of variations in temperature. For these reasons, accurate thermal mass flow sensors require both flow calibration and calibration for temperature compensation. Furthermore, the influence of temperature and/or fluid composition can be derived utilizing the relationships developed in the sections on modeling (Equations 13.20 through 13.35).

13.4.3 Instrumentation and Components

Table 13.8 lists some companies that manufacture and market thermal mass flowmeters.

TABLE 13.8 Manufactures of Thermal Mass Flow Sensors

Manufacturers	Data
KOBOLD Instruments Inc. 1801 Parkway View Drive Pittsburg, PA 15205	Calorimetric type: MAS-Series; air; min. range 0–10 mL min; max. range 0–40 L min. T_{max}. 50 °C; max. pressure 10 bar; accuracy ±2% Hot-wire type: ANE-Series; air; range 0–20 m/s; working temperature 20 °C–70 °C
Hontzsch GmbH Box 1324, Robert-Bosch Str. 8 D-7050 Waiblingen, Germany	Hot-wire type: range 0.05–20 m/s
Bronkhorst Hi-Tech Nijverheidstraat 1A 7261 AK Ruurlo, Netherlands	Calorimetric type: gases and liquids; min. range (gas) 0–5 mL min; max. range (liquid) 0–1000 mL/min; max. pressure 400 bar
Honeywell MicroSwitch, Freeport, IL	Calorimetric type: gases; range 0–1000 mL min^{-1}; max. pressure 1.75 bar
Sierra Instruments Inc. 5 Harris Ct., Bldg. L Monterey, CA 93924	Calorimetric type: all gases from 1 mL/min to 10,000 L/min; –40 °C to 100 °C; 30 bar max.; 1% accuracy Hot-wire type in stainless-steel sheath: gases from 0 to 100 m/s; –40 °C to 400 °C; 100 bar max.; 2% accuracy
Brooks Instrument Division Emerson Electric Co. 407 W. Vine Street Hatfield, PA 19440	Calorimetric type: all gases from 1 mL/min to 10,000 L/min; –40 °C to 100 °C; 100 bar max.; 1% accuracy

References

1. N. T. Nguyen and W. Dötzel, A novel method for designing multi-range electrocaloric mass flow sensors: Asymmetrical locating with heater- and sensor arrays, *Sensors and Actuators A*, 62, 506–512, 1997.
2. *Mass Flow and Pressure Meters/Controllers*. Ruurlo, the Netherlands: Bronkhorst Hi-Tec B.V., 1994.
3. U. Bonne, Fully compensated flow microsensor for electronic gas metering, *Proceedings of the International Gas Research Conference*, November 16–19, 1992, Orlando, FL, Vol. 3, p. 859, 1992.
4. T. S. J. Lammerink, F. Dijkstra, Z. Houkes, and J. van Kuijk, Intelligent gas-mixture flow sensor, *Sensors and Actuators A*, 46/47, 380–384, 1995.
5. T. S. J. Lammerink, N. R. Tas, M. Elwenspoek, and J. H. J. Fluitman, Micro-liquid flow sensor, *Sensors and Actuators A*, 37/38, 45–50, 1993.
6. H. J. Verhoeven and J. H. Huijsing, An integrated gas flow sensor with high sensitivity, low response time and pulse-rate output, *Sensors and Actuators A*, 41/42, 217–220, 1994.
7. J. van Kuijk, T. S. T. Lammerink, H.-E. de Bree, M. Elwenspoek, and J. H. J. Fluitman, Multi-parameter detection in fluid flows, *Sensors and Actuators A*, 46/47, 380–384, 1995.
8. H. Schlichting, *Boundary Layer Theory*, 7th edn. New York: McGraw-Hill, 1979.
9. A. J. Chapman, *Heat Transfer*, 4th edn. New York: Macmillan, 1984.
10. J. H. Huijsing, F. R. Riedijk, and G. van der Horn, Developments in integrated smart sensors, *Proceedings of Transducer 93*, Yokohama, Japan, 1993, pp. 320–326.
11. R. G. Johnson and R. E. Higashi, A highly sensitive silicon chip microtransducer for air flow and differential pressure sensing applications, *Sensors and Actuators A*, 11, 63–67, 1987.
12. F. Mayer, G. Salis, J. Funk, O. Paul, and H. Baltes, Scaling of thermal CMOS gas flow microsensors experiment and simulation, *MEMS '96*, San Diego, CA, 1996, pp. 116–121.
13. C. Yang and H. Soeberg, Monolithic flow sensor for measuring millilitre per minute liquid flow, *Sensors and Actuators A*, 33, 143–153, 1992.
14. D. Yu, H. Y. Hsieh, and J. N. Zemel, Microchannel pyroelectric anemometer, *Sensors and Actuators A*, 39, 29–35, 1993.
15. S. G. Joshi, Flow sensors based on surface acoustic waves, *Sensors and Actuators A*, 44, 63–72, 1994.

14

SQUID Magnetometers

Robert L. Fagaly
Quasar Federal Systems

Superconducting quantum interference devices (SQUIDs) have been a key factor in the development and commercialization of ultrasensitive magnetic measurement systems [1]. In many cases, SQUID magnetometers offer the ability to make measurements where no other methodology is possible.

Because of their superconducting nature, SQUID sensors operate at cryogenic temperatures. Although SQUID electronics have the capability to operate well above 1 MHz, most applications tend to be at lower frequencies. The ability of a SQUID sensor to measure changes in magnetic fields and currents is based on four effects: superconductivity, the Meissner effect, flux quantization, and the Josephson effect. Magnetic flux penetrating a superconducting loop containing one or more Josephson junctions (the SQUID loop) results in an output voltage that is proportional to the magnetic flux.

Typical SQUID input circuits may include (Figure 14.1) superconducting detection coils with superconducting connections and an "input coil" that is inductively coupled to the SQUID loop. Input circuits with such "flux transformers" have applications for measuring magnetic fields and for the sensitive measurement of magnetic properties in physics, chemistry, geology, biology, and medicine as well as numerous other applications.

SQUID sensors are normally fabricated from two types of materials: The first, commonly referred to as a low-temperature superconductor (LTS), is based on niobium metal, whose ductility can allow the winding of small-diameter loops. The second, referred to as a high-temperature superconductor (HTS), uses $YBa_2Cu_3O_{7-\delta}$, which is a brittle ceramic material produced as a 2-D structure.

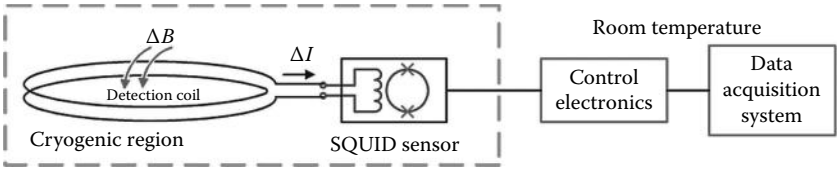


FIGURE 14.1 Schematic diagram of a typical SQUID input circuit.

14.1 Cryogenic Requirements

The superconducting nature of SQUIDs requires them to operate well below their superconducting transition temperature (9.3 K for niobium and 93 K for $YBa_2Cu_3O_{7-\delta}$). The thermal environment for the SQUID sensor and detection coil has typically been, respectively, liquid helium or liquid nitrogen contained in a vacuum-insulated vessel known as a dewar.

If the measurement is of an external field, the dewar must be magnetically transparent and metallic construction is not appropriate. Dewars for external field measurements (e.g., geophysical or biomagnetic) are normally constructed of nonmetallic, low-susceptibility materials to minimize their magnetic interactions with the SQUID sensors and detection coils. Materials used are typically fiberglass/epoxy composites such as G-10 (see Figure 14.2).

To get the detection coil(s) as close as possible to the object being measured, a "tailed" design is often used. This decreases the forces on the bottom of the dewar and allows the use of thinner end pieces (closer tail spacing). Dewars for biomagnetic measurements often have curved tails to get closer to the head, chest, or abdomen. Some SQUID magnetometers (primarily those for rock magnetometry) use metallic dewars cooled by closed cycle refrigeration [2].

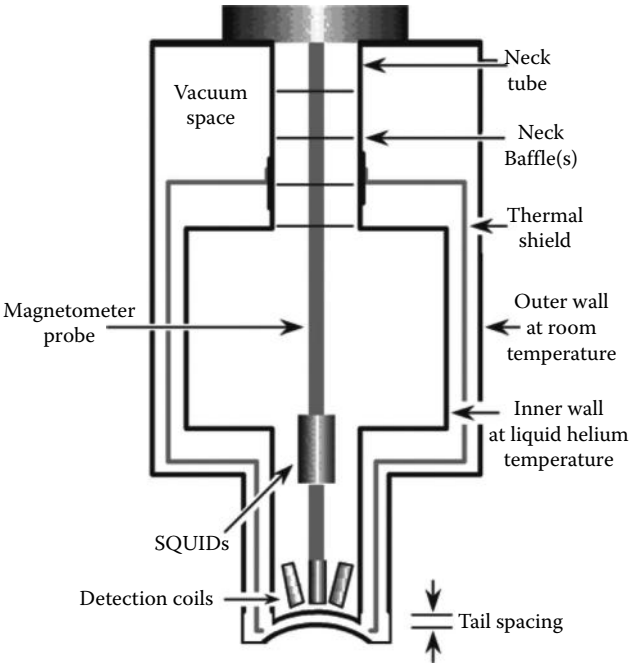


FIGURE 14.2 Typical design of a fiberglass dewar used for biomagnetic measurements.

14.2 Magnetic Field Sensing

The SQUID can act as a magnetometer by allowing flux to directly penetrate the SQUID loop. Because of the small diameter (~50 μm) of the SQUID loop, this results in field sensitivities at the 50 pT/√Hz level. LTS SQUID loops are shielded to isolate the Josephson junctions from such interference, and the input signal is inductively coupled via an input coil for greater sensitivity (Figure 14.1).

Because HTS materials are planar devices and suffer from the inability to make zero-resistance superconducting joints, complex coil geometries (e.g., Figure 14.3b, f, and e) cannot be fabricated. Instead, HTS magnetometers use bare SQUID loops (for magnetic microscopy), integral detection coils, or inductively coupled (e.g., flip chip [3]) detection coils.

The most common input circuit used with LTS SQUIDs is a simple superconducting detection coil that is used to detect changes in magnetic fields. This circuit is connected to the SQUID input terminals as shown in Figure 14.1. Since the total flux in a superconducting loop is conserved, any change in external field through the detection coil will induce a current in the SQUID, which must satisfy

$$\Delta\Phi = NA\Delta B = (L_{\text{coil}} + L_{\text{input}})\Delta I \tag{14.1}$$

where

ΔB is the change in applied field
N, A, and L_{coil} are the number of turns, area, and inductance of the pickup coil
ΔI is the change in current

To calculate the sensitivity and noise levels of a simple detection coil system, the inductance of the detection coil must be known. The inductance of a flat, tightly wound, circular multiturn loop of superconducting wire is given (in MKS units) by [4]

$$L = \mu_0 N^2 r\left[\log_e\left(\frac{8r_{\text{coil}}}{r_{\text{wire}}}\right) - 2\right] \tag{14.2}$$

where

μ_0 is the magnetic permeability of free space
r_{coil} is the radius of the detection coil
r_{wire} is the radius of the (superconducting) wire

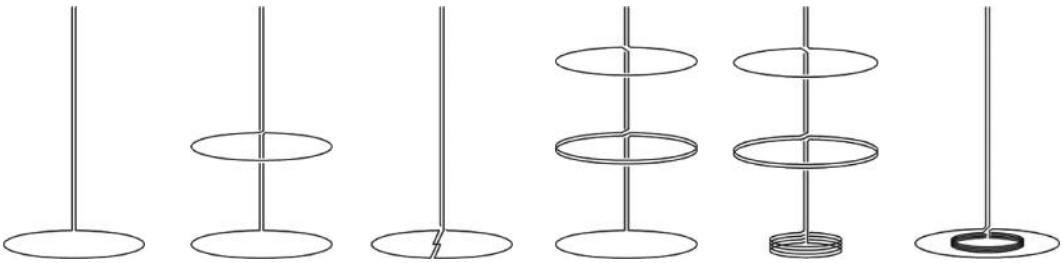


FIGURE 14.3 (a) Magnetometer [B_z], (b) first-derivative axial gradiometer [dB_z/dz], (c) first-derivative planar gradiometer [dB_z/dx], (d) second-derivative axial gradiometer [d^2B_z/dz^2], (e) second-derivative asymmetric axial gradiometer [d^2B_z/dz^2], and (f) first-derivative radial gradiometer [dB_z/dr].

Knowing the coil inductance L_{coil}, we can rewrite Equation 14.3 as

$$\Delta B = (L_{coil} + L_{input})\frac{\Delta I}{NA} \tag{14.3}$$

Since the SQUID system has an output proportional to the input current, maximum sensitivity is obtained by using the input circuit that provides the maximum current into the SQUID and satisfies all other constraints of the experimental apparatus. For a pure magnetometer of a given diameter, the maximum sensitivity will occur when the impedance of the detection coil matches that of the SQUID sensor ($L_{coil} = L_{input}$).

As an example, a single turn coil of 0.13 mm superconducting wire with a diameter of 4 cm when used with commercially available SQUIDs can have a sensitivity better than 20 fT/√Hz (1 fT = 10^{-15} T = 10^{-11} Gauss). For maximum sensitivity, L_{coil} and L_{input} should be matched.

Much higher sensitivity (10^{-17} T/√Hz) in a limited bandwidth (100 Hz in this case) has been achieved by incorporating rf SQUID into a parametric amplifier circuit [5]. Substituting a present-day direct current (dc) SQUID sensor could yield a sensitivity of $<3 \times 10^{-18}$ T/√Hz. The highest sensitivity reported for a commercially available HTS SQUID magnetometer is 3.5×10^{-15} T/√Hz [6].

14.2.1 Environmental Noise

One problem with pure magnetometers (and virtually all HTS SQUIDs) is that they are extremely sensitive to the outside world (Figure 14.5). This may be acceptable if one is measuring external fields, but if what is needed to be measured is very close to the detection coil and very weak, outside interference, for example, the Earth's magnetic field, 50 or 60 Hz power line frequencies and/or radio-frequency interference (rfi), may prevent measurements at SQUID sensitivities. If the experiment is the measurement of a magnetic field due to an object a few cm away from the detection coil, a gradiometer may offer a solution.

Since the magnetic field of a dipole source is proportional to the cube of the distance,

$$B_{dipole} \propto \frac{1}{z^3} \tag{14.4}$$

where z is the distance beneath the coil, it follows that the field from a distant source is relatively uniform in direction and magnitude at the sensor. If we connect in series two identical and exactly parallel loops wound in opposite senses, separated by a distance b (the baseline), we obtain a coil (Figure 14.3b) that will reject uniform fields.

Using the previous example, with two 4 cm diameter coils separated by 10 cm (the baseline), one obtains a gradient sensitivity of 2 fT/cm·√Hz. For objects that are less than a few baselines away from a gradiometer, this coil system will act as a magnetometer for the source of interest ($\propto z^{-3}$) and as a (Nth order) gradiometer for distant sources ($\propto z^{-3-N}$) [7]. Figure 14.4 shows the relative response of first- and second-order gradiometers to distant sources. As can be seen, for sources relatively close to the signal coil, the effective response is that of a pure magnetometer, while distant sources have significantly reduced signal strengths.

14.2.2 Gradiometer Balance

If the gradiometer is perfectly made with the area turns of both coils being identical (i.e., balanced), it will reject uniform fields. However, should one coil have a larger effective diameter than the other, the

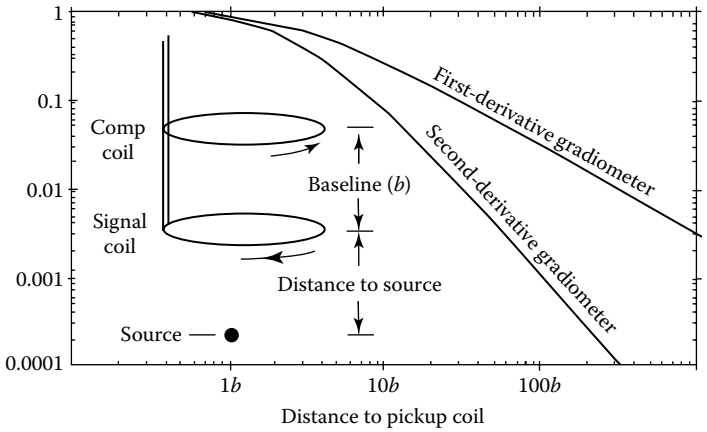


FIGURE 14.4 Response of gradient coils relative to magnetometer response ($1/z^3$ suppressed).

response of the coil will not be that of a perfect gradiometer but that of a gradiometer in series with a magnetometer. Mathematically, the balance, β, can be defined as

$$V_{\text{out}} \propto \mathbf{G} + \beta \cdot \mathbf{H} \tag{14.5}$$

where

V_{out} is the system response
G is the coil's response to a gradient field
H is the applied uniform field

Second-derivative coils ($N = 2$) have been constructed with balances better than one part per million [8]. That is, the effects of uniform magnetic fields on the detection coils are reduced by more than a million. This can allow a SQUID gradiometer to operate in relatively large (mT) ambient fields while maintaining sensitivities in the tens of fT. This has importance in measurements of magnetic fields of biological origin and nondestructive test and evaluation.

While nearly all HTS gradiometers are planar devices, it is possible to use a thick-film-patterned (as a planar gradiometer) HTS tape and bend it to form an axial gradiometer [9]. Note that there will be substantial gradiometer imbalance.

14.2.3 Magnetic Shields and Shielding Rooms [10]

To reduce the effect of environmental noise—assuming that the measurement is not of external fields and size permitting—the experiment can be placed in a magnetic shield (e.g., magnetic microscopy) or a magnetically shielded room (MSR) (e.g., biomagnetism). Attenuation of external magnetic fields can be accomplished by induced current shielding or flux shunting. Faraday's law ($\nabla \times \mathrm{E} = -\partial B/\partial t$) states that an electric field is induced in a material when it is exposed to a time-dependent magnetic field. If the shielding material is conducting ($\rho \ll \infty$), the electric field induces a current that generates a magnetic flux to oppose the externally applied field. This effect, which is frequency dependent, can be used to create an eddy current shield to attenuate electromagnetic interference. Eddy current shields can be constructed from sheet metal (e.g., aluminum) or wire (e.g., copper screen). While not capable of attenuating dc fields, the shielding factor of eddy current shields increases with frequency. Screened rooms are relatively inexpensive and properly constructed and are good at attenuating MHz and GHz noise sources.

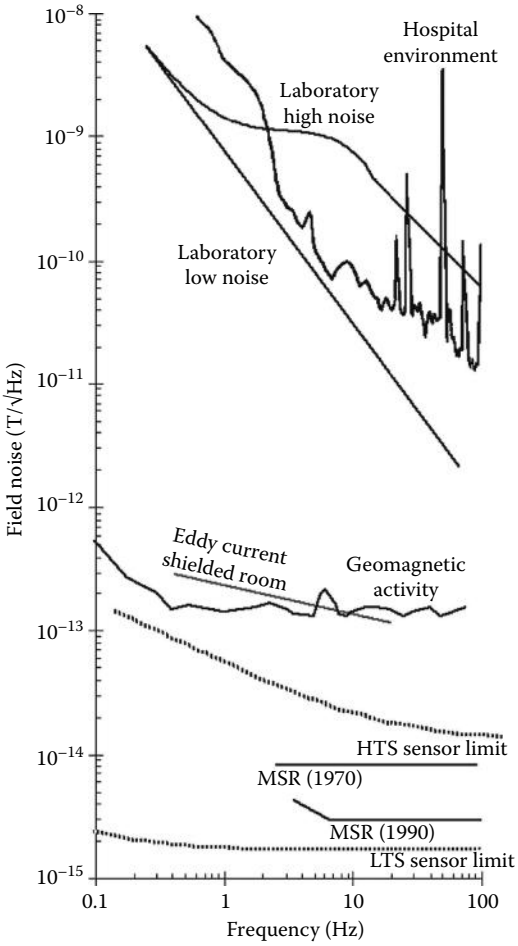


FIGURE 14.5 rms field noise spectra in various environments as a function of frequency.

The need for shielding at lower frequencies down to dc has led to the use of high-permeability materials (e.g., mu-metal) for shunting magnetic fields external to the experimental region. For biomagnetic measurements, it may be necessary to enclose not only the sensing device but also the subject in an MSR with large interior dimensions (e.g., 4 m × 3 m × 2 m). The use of multiple shields can act to further shield the interior of an MSR with commercial MSRs offering shielding factors >40 dB at dc, increasing to 100+ dB above 100 Hz.

The effective use of magnetic shields can reduce the background magnetic noise to levels that allow measurements below pT/√Hz levels (Figure 14.5).

A unique advantage of SQUIDs is that they can be used to make magnetic field sensors with dc response and sensitivities well below environmental noise levels. This has allowed noninvasive measurements of electrophysiological activity and a number of other magnetic sensing methodologies not possible with conventional electromagnetic sensing systems.

14.3 Geophysical Applications

The geomagnetic field at the surface of the Earth (Figure 14.6) is generated by a number of sources. There exists a background field of ~50 μT with a daily variation of ±0.1 μT. In addition, there is a $1/f^{3/2}$ contribution (below 1 Hz) from ultralow electromagnetic energy reaching the Earth from solar-generated

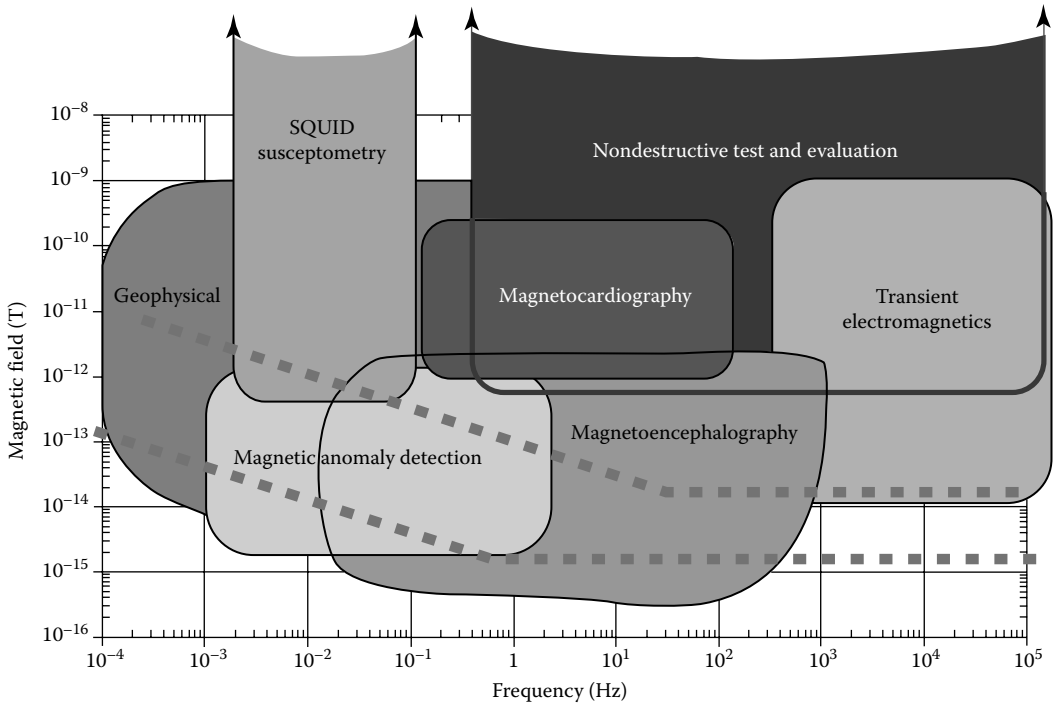


FIGURE 14.6 Field sensitivities and bandwidths typical of various applications. The dashed lines indicate the sensitivity of commercially available SQUIDs (the lower line is a typical LTS SQUID, and the upper line is a typical LTS SQUID).

disturbances in the ionosphere [10]. Other contributions include lightning strikes and man-made contributions such as structural steel and other localized magnetic materials such as furniture and instruments that distort the Earth's field and result in field gradients; moving vehicles that generate transient fields; electric motors; elevators; radio, television, and microwave transmitters; and the ever-present power line electromagnetic field and its harmonics.

A technique known as magnetotellurics (MT) [11] can be used to determine the electric conductivity distribution of the Earth's crust by measuring the Earth's electric and magnetic field. Because the Earth is a good electrical conductor compared to the air, the electric field generated in the ionosphere (due to the solar wind) is reflected at the Earth's surface, with components of both the electric and magnetic field decaying as they penetrate into the Earth. The decay length or skin depth (δ) can be expressed as

$$\delta \approx 0.5\sqrt{\rho\tau} \text{ (km)} \tag{14.6}$$

where

ρ is the electrical resistivity of the Earth
τ is the period of the electromagnetic wave

In magnetotellurics, the electric field (as a function of frequency) is related to the magnetic field via an impedance tensor where $\mathbf{E}(\omega) = \mathbb{Z}\mathbf{H}(\omega)$. The impedance tensor ($\mathbb{Z}$) contains four complex elements Z_{xx}, Z_{xy}, Z_{yx}, and Z_{yy} and is related to the resistivity ρ by the relation $\rho_{ij} \approx 0.2|Z_{ij}(\omega)|^2\tau$ where τ is the period in seconds and ρ is measured in $\Omega \cdot$m and $\mathbb{Z}$ in mV/km $\cdot$ nT.

SQUID magnetometers have been used to measure the Earth's magnetic field at frequencies ranging between 1 kHz and 10^{-4} Hz. MT has been effective in areas of thick sediments and on near-surface

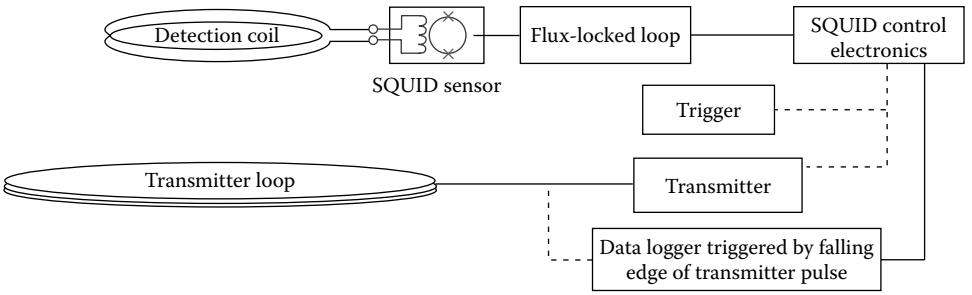


FIGURE 14.7 Block diagram of controlled source electromagnetic system.

metamorphic or volcanic rocks. The increased conductivity of hot saline regions associated with geothermal sites makes it well suited for location of hydrothermal reservoirs.

Rather than using the electromagnetic field generated by the solar wind, active methods create a time-dependent magnetic field generated by a loop driven by an electric current, with the magnetometer detecting the induced field. Measurements can be in either the time (e.g., transient electromagnetics [TEM]) or frequency (e.g., controlled source audio magnetotelluric [CSAMT]) domains. Figure 14.7 shows the key elements of a TEM measurement using relatively small field excitation coils [12]. Targets were centered beneath the transmitter coil, which had three turns with a current of 0.01 A (65 nT effective field).

14.3.1 Rock Magnetometers

Geophysical rock magnetometers [13] for paleomagnetism typically use three orthogonal detection coils to measure all three components of the remnant magnetic moment of rock and core samples with sensitivities approaching 10^{-10} emu/√Hz levels. They do not use a bias field magnet; in fact, the detection coil region is magnetically shielded to reduce the ambient magnetic fields in the detection coil region to well below the Earth's magnetic field. Another difference between rock magnetometers and SQUID susceptometers is that the sample tubes are made from fiberglass/epoxy rather than stainless steel primarily because metallic tubes conduct rfi into the pickup coil region and can cause SQUID noise. This also allows for time constants that approach microseconds. Typically, the sample chamber of rock magnetometers does not have variable temperature capabilities but are fixed at room (300 K) temperature. SQUID microscopes have also been used for paleomagnetic investigations.

A variation on the traditional rock magnetometer is the spinner magnetometer. In this case, the sample is rotated beneath a SQUID magnetometer. In this situation, the sample is not tightly coupled to the sensing magnetometer; thus, sensitivity will be significantly reduced. Rotating the sample on two axes permits the orientation and magnitude of the rock's magnetization to be determined. Sensitivities of 4×10^{-5} A/m have been achieved [14] using 100 fT/√Hz HTS SQUID magnetometers.

SQUIDs can also be used for magnetic anomaly detection (MAD) [15]. Knowledge of the total magnetic field gradient ($\nabla\mathbf{B}$) of an object can allow determination of its magnitude and direction [16]. From Maxwell's equations, four of the gradients are redundant (e.g., $\partial B_y/\partial x = -\partial B_x/\partial y$). Thus, only five gradient components (dB_x/dx, dB_y/dy, dB_x/dz, dB_y/dz, dB_z/dz) are necessary to determine $\nabla\mathbf{B}$. The three field (B_x, B_y, B_z) components are normally used to construct an 8-element array for MAD HTS tensor arrays. An 8-element vector/tensor array can also be made from pure (e.g., HTS) magnetometer elements [1].

14.4 Nondestructive Test and Evaluation

Magnetic sensing techniques such as eddy current testing (Figure 14.8e) have been used for many years to detect flaws in structures. A major limitation on their sensitivity is the skin depth, λ (the distance where the field is attenuated by a factor $1/e$) of metallic materials. For a sinusoidal varying wave,

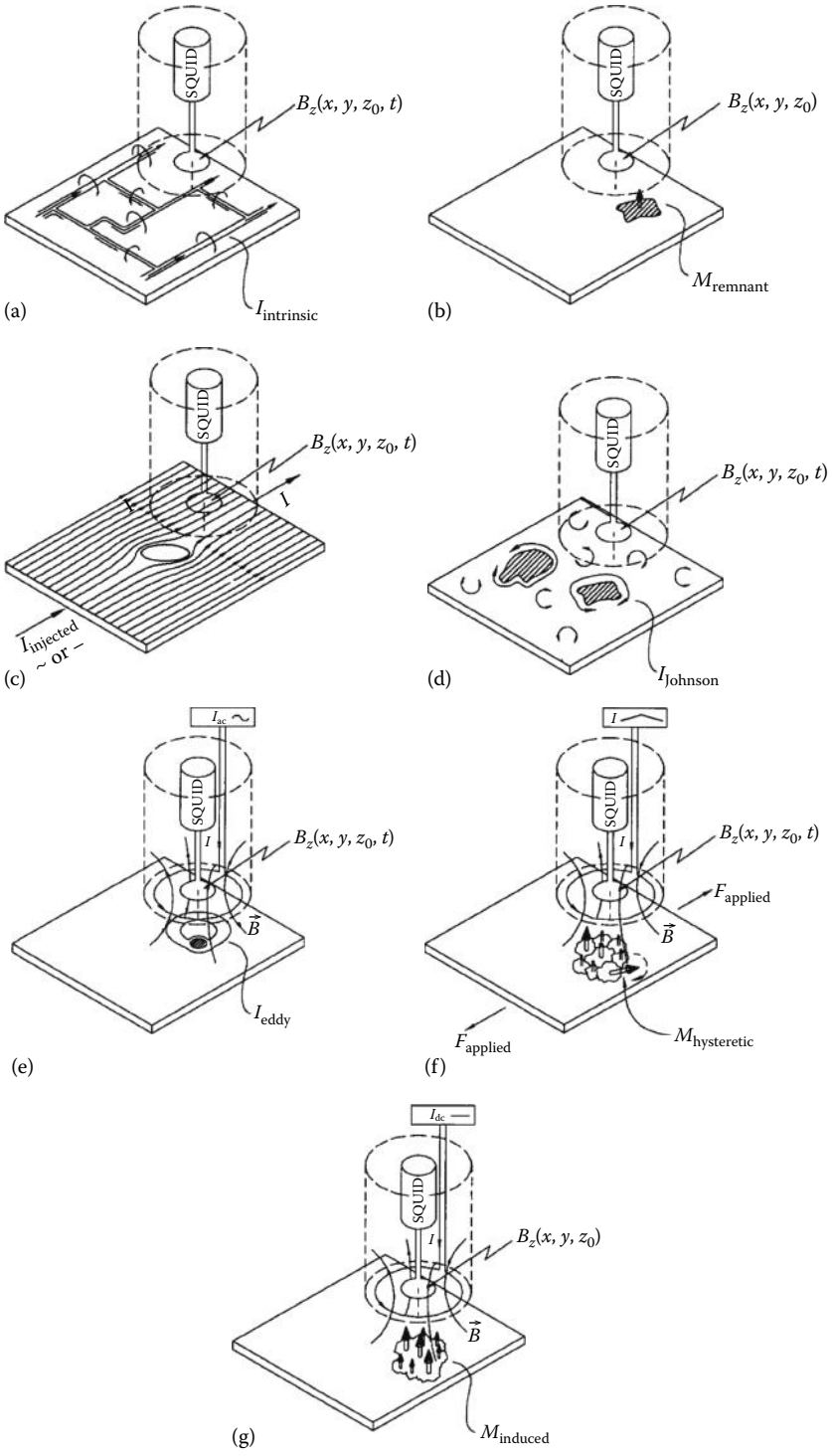


FIGURE 14.8 Measurement configurations for SQUID NDE: (a) intrinsic currents, (b) remnant magnetization, (c) flaw-induced perturbations in applied currents, (d) Johnson noise or corrosion activity in conductors, (e) eddy currents induced by an applied ac magnetic field, (f) hysteretic magnetization by application of stress or an applied field, and (g) diamagnetic and/or paramagnetic materials in an applied field. (Photo courtesy of J. P. Wikswo.)

$\lambda = \sqrt{\rho/\pi\mu_0 f}$ where f is the frequency of the applied field, ρ is the electrical resistivity, and μ_o is the magnetic permeability of free space. Because SQUID sensors have true dc response and superior sensitivity, they can see "deeper" into metallic structures. DC response also means that they can detect remnant magnetization—without the need for externally applied magnetic fields. Their flat frequency response and zero-phase distortion allow for a wide range of applications. Reference [17] gives an excellent overview of SQUID NDE research.

Magnetic susceptibility (Figure 14.8g) can be measured by applying an external magnetic field in the same orientation as the detection coil(s). If the material being studied is isotropic, the response of the material to the applied field will be sufficient to determine the material's magnetic properties.

A particular advantage of magnetic sensing is in examining defects beneath insulating barriers. For example, a cladded pipe with asbestos lagging can be quickly examined because the insulating barriers are virtually invisible to magnetic detection—the asbestos need not be removed or cut into. A similar use would be to detect corrosion underneath paint or other opaque coverings. Unlike eddy current measurements, which have a skin depth that is frequency dependent, the true dc response of a SQUID magnetometer allows measurements with little, if any, effect from intervening materials. This is true even if the intervening material is metallic.

14.4.1 SQUID Microscopes

Since electric currents create magnetic fields, the ability to measure small magnetic fields offers the potential to locate the causes of semiconductor failures. In failed integrated circuits (ICs), a short circuit would appear as a small area of intense magnetic flux. By overlaying a magnetic map (created by scanning a SQUID sensor over an IC—Figure 14.8a) onto a computer-aided design (CAD) map of an IC's features, it is possible to locate where faults occur in a device. SQUID microscopes have been used to make noncontact measurements of electronic circuits [18]—one instrument has better than 10 μm resolution [19]. Another [20] can detect 10 nA currents flowing in a conductor 100 μm from the sensor.

14.5 Medical Applications of SQUID

As mentioned, SQUID magnetometers may provide diagnostic capabilities in areas where there is no present measurement technique. The use of bioelectric signals as a diagnostic tool is well known in medicine, for example, the electrocardiogram (ECG) for the heart and the electroencephalogram (EEG) for the brain. The electrical activity that produces the surface electrical activity that is measured by EEG and ECG also produces magnetic fields [21]. The analogous magnetic measurements are known as the magnetocardiogram (MCG) and the magnetoencephalogram (MEG).

Magnetic fields of biological origin (Figure 14.9) are quite small compared to the Earth's magnetic field. The heart produces the strongest magnetic field with a strength on the order of 100 pT [22]. Brain activity can be ten thousand times weaker (ten billion times weaker than the Earth's magnetic field) [23]. In addition, magnetic measurements can be made for which there are no electrical analogs. These are measurements of static magnetic fields produced by ferromagnetic materials inhaled, injected, or ingested into the body and measurements of the magnetic susceptibility of materials in the body. Examples of these are magnetic dusts inhaled and retained in the lungs and the paramagnetic iron storage compounds ferritin and hemosiderin that occur naturally in the body.

The relative strengths of these various biomagnetic fields vary by more than 10,000. The strongest is associated with magnetic contaminants in the lung. An example would be dust inhaled by a machinist or an arc welder. The external magnetic field after exposure to a magnetizing field of a 0.3 mT (the equivalent of a toy magnet) can be as great as 1 nT. Evoked neurological responses, such as the response to visual or auditory stimuli, have field strengths at the surface of the skull of a few tens of fT. The sources of these fields vary from localized groups of neurons in the brain to the more dispersed muscle tissues in the heart, to magnetic particles distributed through the lungs.

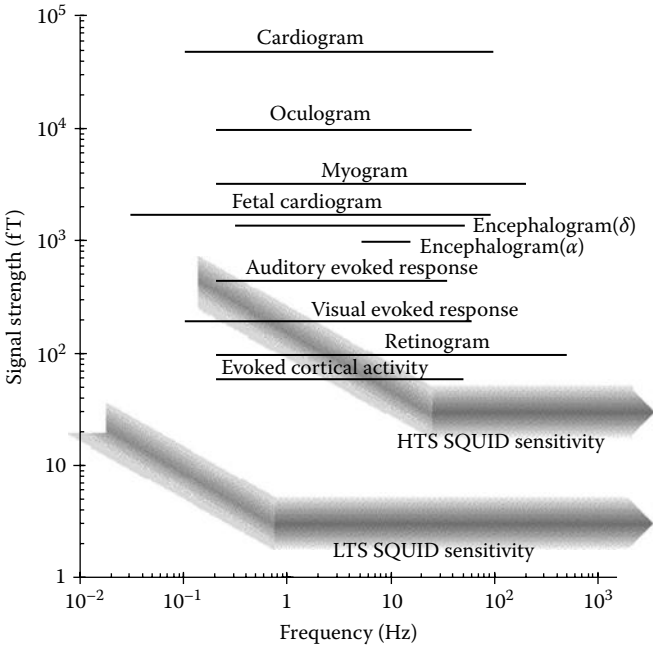


FIGURE 14.9 Typical amplitudes and frequency ranges for various biomagnetic effects.

Magnetic fields from active electrical sources in the body can be measured passively and externally to the body (i.e., noninvasively) by placing the magnetometer (actually first- or second-order gradiometers) in close proximity to the body's surface. As mentioned, the use of gradiometers can allow measurements to be made in unshielded environments at sensitivities below 5 fT/√Hz. This technique is applicable to measurement of the heart, brain, and skeletal muscles.

14.5.1 Magnetoencephalography

The magnetic analog to electroencephalography (EEG), MEG, has the advantage of being (in the case of depth electrodes) noninvasive and unambiguous in its ability to locate neurological activity. Within the brain, a population of neurons can be modeled as an equivalent current dipole (Figure 14.10) that generates well-defined magnetic field profile. Mapping of these field profiles can be used to infer the location

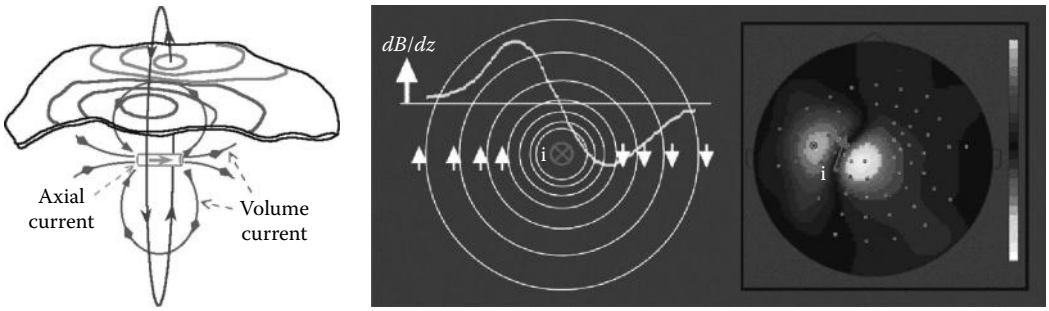


FIGURE 14.10 (a) Magnetic field generated by a current dipole. For a sphere, the dipole is located at midpoint of the maxima and minima, and (b) at a depth = distance $D/\sqrt{2}$, where D is the separation between the maxima and minima on the surface of the sphere.

of the equivalent active dipole site region to within millimeters [24]. Using evoked response techniques, the location of signal pathways and information processing centers in the brain can be mapped at different delay times (latencies) following the stimulus. MEG systems for clinical use can have channel counts in the hundreds [25].

14.5.2 Magnetocardiography

Measurement of the heart's conduction system is important in the diagnosis of diseases in which the heart has an abnormal rhythm, reduced mechanical performance, and higher susceptibility to sudden failure. DC MCG measurements have shown that a "current of injury" flows in the heart as a result of myocardial infarction. Although a current of injury signal can also be seen in the alternating current (ac)-coupled ECG, the magnitude of the injury and the location of the injured region cannot be determined with precision. Using MCG, it may be possible to determine both the location and extent of the injured heart tissue. Clinical studies are underway to test this hypothesis [26].

The SQUID magnetometer is sensitive enough to noninvasively detect the fetal MCG (fMCG). The fetal ECG (fECG) is often used by physicians as a measure of fetal distress, but detection of this fetal heart beat is often unreliable, especially between 25 and 34 weeks of gestation. Unlike fECG, it is possible to monitor the fetal heart (via fMCG) through all phases of gestation. Another advantage of magnetic (as compared to electrical) methods is that the spatial resolution of MCG allows it to better isolate the fetal heart in the presence of the maternal heartbeat [27].

14.5.3 Other Biomagnetic Applications

The location and quantity of ferrimagnetic materials in the body, typically due to contaminants in the lungs, can be determined by measuring the magnetic field of the body before and after a low-intensity (~10 mT) field is applied to magnetize the ferrimagnetic materials. The distribution of these materials can be determined by applying the field to only a small region of the body at a time [28].

Fields from paramagnetic or diamagnetic materials (such as iron stored in the liver) can also be measured. By comparing measurements made while a low-intensity magnetic field (~0.3 mT) is applied to a local region of the body with those of the background field when the subject is lowered beneath the detection coil(s), actual quantities of the (paramagnetic or diamagnetic) materials can be determined. Information on the depth of these materials can be obtained by using magnetization and detection coils of differing sizes in the same instrument or by measuring the field as a function of distance from the body. This technique is already being used to monitor patients suffering from iron overload diseases such as thalassemia and hemochromatosis [29].

It is only since the development of the SQUID that measurements of biomagnetic fields have become a reality. Current SQUID magnetometers for MEG utilize hundreds of channels to provide whole-head coverage, sensitivities at the 5 fT/√Hz level, and bandwidths up to 5 kHz.

Defining Terms

Josephson junction critical current (I_c): The current at which a Josephson junction transitions from superconducting to normal (resistive) behavior.

Magnetic flux (Φ): The product of the magnetic field (B) penetrating an area times the area ($\Phi = BA$). Flux is quantized in units of $h/2e = 2.068 \times 10^{-15}$ Wb (T · m^2).

SQUID: Superconducting quantum interference device that acts as a current- or flux-to-voltage amplifier that operates at cryogenic temperatures. The principles of SQUIDs are discussed in Chapter 5.

Transition temperature (T_c): The temperature in which a superconductor transitions from superconducting ($R = 0$) to normal ($R > 0$) behavior.

References

1. R.L. Fagaly, Superconducting quantum interference device instruments and applications, *Rev. Sci. Instrum*., 77, 10110, 2006.
2. G. Walker, *Miniature Refrigerators for Cryogenic Sensors and Cold Electronics*, Oxford, U.K.: Clarendon Press, 1989.
3. M.I. Faley, U. Poppe, K. Urban, E. Zimmerman, W. Glass, H. Halling, M. Bick, H.-J. Krause, D.N. Paulson, T. Starr, and R.L. Fagaly, Operation of HTS dc-SQUID sensors in high magnetic fields, *IEEE Trans. Appl. Supercond*., 9, 3386–3391, 1999.
4. F.W. Grover, *Inductance Calculations, Working Formulas and Tables*, New York: Dover, 1962.
5. M.B. Simmonds, W.A. Fertig, and R.P. Giffard, Performance of a resonant input SQUID amplifier system, *IEEE Trans. Magn*., 15, 478–481, 1979.
6. M.I. Faley, U. Poppe, K. Urban, D.N. Paulson, and R.L. Fagaly, A new generation of the HTS multilayer DC-SQUID magnetometers and gradiometers, *J. Phys.: Conf. Ser*., 43, 1199–1202, 2006.
7. G.-L. Romani, S.J. Williamson, and L. Kaufman, Biomagnetic instrumentation, *Rev. Sci. Instrum*., 53, 1815–1845, 1982.
8. F.B. Jaworski and D.B. Crum, Sources of gradiometer imbalance and useful balancing techniques, In: *SQUID Applications to Geophysics*, H. Weinstock and W.C. Overton, Jr., Eds., Tulsa, OK: Society of Exploration Geophysicists, 1981, p. 19.
9. D.E. Farrell, C.J. Allen, P.N. Arendt, S.R. Foltyn, D.N. Paulson, R.L. Fagaly, and G.M. Brittenham, High-Tc SQUID susceptometer, *Bull. Am. Phys. Soc*., 44, 1553, 1999. See also: M. Bick, K.E. Leslie, R.A. Binks, D.L. Tilbrook, S.K.H. Lam, R. Gnanarajan, J. Du, and C.P. Foley, Axial high-temperature superconducting gradiometer with a flexible flux transformer, *Appl. Phys. Lett*., 84, 5347–5349, 2004.
10. K. Vozoff, The magnetotelluric method in the exploration of sedimentary basins, *Geophysics*, 37, 98, 1972.
11. A.A. Kaufman and G.V. Keller, *The Magnetotelluric Sounding Method*, New York: Elsevier, 1981.
12. R. Wold, P. Weichman, M. Tondra, E. Lange, C. Nordman, D.N. Paulson, T. Starr, J. Johnson, and D. Wilkinson, Development of a standoff UXO detection system using SDT and high temperature SQUID sensor arrays, *UXO Forum'2000*, Anaheim, May 2000.
13. 2-G Enterprises, Series SRM Rock Magnetometers, http://www.2genterprises.com/index.html, last accessed August 5, 2013.
14. K.E. Leslie, R.A. Binks, C.J. Lewis, M.D. Scott, D.L. Tilbrook, and J. Du, Three component spinner magnetometer featuring rapid measurement times, *IEEE Trans. Appl. Supercond*., 11, 252–255, 2001.
15. T.R. Clem, C.P. Foley, and M.N. Keene, SQUIDs for geophysical and magnetic anomaly detection, In: *The SQUID Handbook: Volume II: Applications of SQUIDs and SQUID Systems*, J.J. Clarke and A.I. Braginski, Eds., New York: Wiley, 2004, Chapter 14.
16. W. Wynn, C. Frahm, P. Carroll, R. Clark, J. Welhoner, and M. Wynn, Advanced superconducting gradiometer/magnetometer arrays and a novel signal processing technique, *IEEE Trans. Magn*., 11, 701–707, 1975.
17. W.G. Jenks, S.S.H. Sadeghi, and J.P. Wikswo, SQUIDs for nondestructive evaluation, *J. Phys. D: Appl. Phys*., 30, 293–323, 1997.
18. R.L. Fagaly, SQUID detection of electronic circuits, *IEEE Trans. Magn*., 25, 1216–1218, 1989.
19. J. Kirtley, Imaging magnetic fields, *IEEE Spectrum*, 33(12), 40–48, 1996.
20. Neocera, Inc., Magma C-30 HTS SQUID microscope, http://www.neocera.com/index.html, last accessed August 5, 2013.
21. S.J. Williamson, G.-L. Romani, L. Kaufman, and I. Modena, Eds., *Biomagnetism, An Interdisciplinary Approach*, New York: Plenum Press, 1983.
22. D. Cohen, E.A. Edelsack, and J.E. Zimmerman, Magnetocardiograms taken inside a shielded room with a superconducting point-contact magnetometer, *Appl. Phys. Lett*., 16, 278–280, 1970.

23. S.J. Williamson and L. Kaufman, Biomagnetism, *J. Magn. Magn. Mater.*, 22, 129, 1981.
24. G.-L. Romani, S.J. Williamson, and L. Kaufman, Tonotopic organization of the human auditory cortex, *Science*, 216, 1339, 1982.
25. J. Vrba and S.E. Robinson, SQUID sensor array configurations for magneto-encephalographic applications, *Supercond. Sci. Technol.*, 15, R51–R89, 2002.
26. K. Tolstrup, B.E. Madsen, J.A. Ruiz, S.D. Greenwood, J. Canacho, R.J. Siegel, H.C. Gertzen, J.W. Park, and P.A. Smars, Non-invasive resting magnetocardiographic imaging for the rapid detection of ischemia in subjects presenting with chest pain, *Cardiology*, 29(106), 270–276, 2006.
27. J.F. Strasburger and R.T. Wakai, Fetal cardiac arrhythmia detection and in utero therapy, *Nat. Rev. Cardiol.*, 7, 277–290, 2010.
28. D.N. Paulson, R. Toussaint, R.L. Fagaly, and S.E. Robinson, Superconducting magnetometer system for detecting lung contaminants, *IEEE Trans. Magn.*, 23, 1315–1318, 1987.
29. G.M. Brittenham, D.E. Farrell, J.W. Harrris, E.S. Feldman, E.H. Danish, W.A. Muir, J.H. Tripp, and E.M. Bellon, Magnetic-susceptibility measurement of human iron stores, *N. Engl. J. Med.*, 307, 1671, 1982.

Further Information

Commercial Vendors of SQUID Measurement Systems

2-G Enterprises, Sand City, CA, http://www.2genterprises.com/, accessed on August 17, 2012.
ATB, Pescara, Italy, http://www.atbiomagnetics.com/home.php, accessed on August 17, 2012.
CardioMag Imaging, NY, http://www.cardiomag.com/, accessed on August 17, 2012.
Cryogenics Ltd., London, England, http://www.cryogenic.co.uk/, accessed on August 19, 2012.
CTF, Coquitlam, BC, Canada, http://www.ctf.com/, accessed on August 19, 2012.
Elekta Neuromag, Helsinki, Finland, http://ecatalog.elekta.com/functional-mapping.aspx, accessed on August 18, 2012.
Neocera, Beltsville, MD, http://www.neocera.com/, accessed on August 24, 2012.
Quantum Design, San Diego, CA, http://www.qdusa.com/, accessed on August 17, 2012.
Tristan Technologies, San Diego, CA, http://www.tristantech.com/, accessed on August 18, 2012.

15

Sensor Networks and Communication

Robert M. Crovella
NVIDIA Corporation

15.1 Introduction

15.1.1 What Is a Communication Network?

A communication network provides a system by which multiple users may share a single communication path (or medium) to exchange information. The telephone system is an example of a system containing many communication networks, which can be considered to be a single communication network as an abstract example. Communication networks are commonly used in various industries and applications to provide an economical means to allow multiple, geographically separated users to exchange information.

15.1.2 Ordinary Sensors versus Networked Sensors

A definition of the function of a sensor is to map or convert one measured variable (e.g., spatial, mechanical, electromagnetic) into another—usually electric—variable or signal. This signal may then be passed to a measurement or processing system for capture and analysis, or as a direct input to some controlled process. In this case, the measured variable is represented as an electric signal. This signal must be handled individually by the measurement system, and it may also be subject to corruption from a variety of sources, such as electromagnetic interference (EMI) in the case of an electric signal.

In applications where a number of sensing devices are needed, and/or where the sensing devices are distributed geographically (or are distant from the measurement and analysis system), the application designer may wish to use a communication network to transmit sensor data from the measurement point to the measurement and analysis system. Such applications typically involve some sort of digital computing machinery at the measurement and analysis point, or as part of the control system. Figure 15.1 depicts a representative block diagram of a sensor/network system showing the relationship of the various components.

Networked sensors can be distinguished into two components: those performing the measurement function and those components performing the communication function. In some cases, these two functions may be designed as a single unit, such that the "sensor" intrinsically includes communication capability. In other cases, an ordinary sensor may be connected to a conversion unit, which converts the output signal of the ordinary sensor into a form suitable for the network and manages the delivery of this information on the network.

15.1.3 Why Use Networked Sensors?

Network communication combined with sensor technology can provide several benefits to an application, as well as to the sensor designer. The most obvious benefit of a network is the simplification of the wiring for the transmission of the signals from one place to another. For system containing N users, the number of wires or cables T required to individually connect each user with each other user is given by Equation 15.1:

$$T = 2^{(N-1)} - 1 \tag{15.1}$$

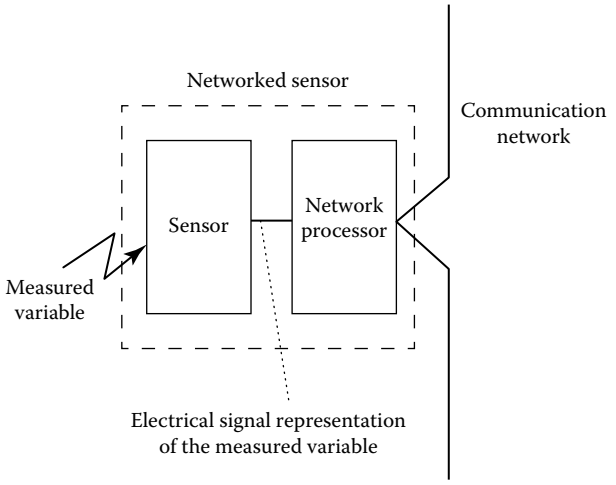


FIGURE 15.1 A networked sensor is an ordinary sensor with network communication components added.

assuming each wire or cable can carry information in both directions between the two users connected by that cable. For more than a few users, the number of cables required (T) to provide an individual connection between each pair of users is large. Sensors are often connected to a central measurement and analysis system and may only need to communicate with the central system. In this case, the number of individual wires or cables needed is equal to the number of sensors ($N - 1$). Even with this smaller number of cables, the wiring for a large number of sensors in some applications can be quite complex. A network may be able to reduce the total number of cables required to a much smaller number. In fact, in a sensor network, all of the sensors and the central measurement and analysis system can be connected to a single cable.

An indirect benefit of networking may be in its handling of the sensor signal. Because most modern networks are digital in nature, an analog sensor signal typically must be digitized before it can be transmitted on a network. With a networked sensor, the digitization will typically be carried out by circuitry in relatively close proximity to the sensor. As a result, the analog signal will have traveled a short distance before being converted to a digital signal. This can be a benefit in two ways. The first is that the analog signal will not suffer as much attenuation or degradation due to electric losses associated with carrying a signal over a great distance. The second is that once in digital form, the "signal" can be made relatively immune to the effects of distortion or degradation due to EMI. Although digital transmission of signals is still subject to EMI, modern protocols and transmission systems can be designed to be very robust, using signaling that is resistant to EMI as well as using error control techniques. As a result, the effect of attenuation and disturbances can be essentially eliminated by digital transmission of the signal.

Another benefit of networking is the ability to communicate a much wider range of information—in both directions—when compared with a single cable carrying a sensor signal. With many modern networks suitable for networked sensing applications, a microprocessor is used at the sensor to manage the handling of the sensor signal and its transmission on the network. But there is generally no need to limit the microprocessor to this one function alone. The combination of the network and the microprocessor provides a platform upon which many additional functions and features can be incorporated into the networked sensor. For example, the signal of a sensor may need a certain calibration or correction function applied to it before it can be used in calculations. It may be beneficial to load into the networked sensor (through the network) a set of correction parameters or coefficients and then have the microprocessor correct or calibrate the output of the sensor before transmitting it to the network. Sensors can be easily designed to have multiple sensing functions, such as temperature and pressure. Each signal can be handled separately and transmitted separately on the network, with no need for additional connections. Sensors may be designed to store certain types of information, such as the name of the manufacturer, or certain calibration parameters determined by the manufacturer at the time of manufacture. This information can then be read out over the network and used for a variety of purposes. A sensor can even be designed to have "intelligent" functions, such as the ability to sense its environment and determine when certain parameters have been exceeded (such as operating temperature range) or report a special message containing an "alarm" when the sensor signal level exceeds a certain threshold. The combination of the network and the microprocessor leads to an endless variety of functions and features that can be added to the basic sensor technology.

15.1.4 Potential Problems with Networked Sensors

Networked sensors will generally require more complex circuitry than equivalent, nonnetworked sensors. A drawback of analog-to-digital (A/D) conversion and digital transmission of signals is the time and level quantization effect that A/D conversion can have on the analog signal. These effects can be mitigated with modern, high-speed A/D converters (to minimize the effect of time quantization or the sampling effect) with the ability to convert in high resolution (i.e., using a large number of digital bits to represent the analog signal level). These drawbacks are not unique to networked sensors but rather to digitized sensor values and digital control whether or not it uses a network. Finally, the

capacity of the network to carry information (the bandwidth) must be considered in any communication system. Putting a large number of sensors on a single network may overload the information-carrying capability of the network, resulting in queuing delays in the reception of sensor signals and, in some cases, lost data.

15.2 Communication and Networking Concepts

In order to be able to select an appropriate network technology, it is necessary to understand some basic terminology so that the features and capabilities of various networks and technologies can be categorized and compared.

15.2.1 Station

A station represents a single communicating element on a network system. Each user of the network must access the communication capability of the network via a station. Each station will typically have some implementation of the open systems interconnection (OSI) network reference model as the means of utilizing the network system.

15.2.2 Media Access

Media access is the method by which individual stations determine when they are permitted to transmit or "use" the media. Media access control (MAC) is a function that is usually performed in the data link layer of the OSI reference model. Some well-known methods of MAC include carrier sense multiple access with collision detection (CSMA/CD) and token passing. CSMA/CD systems (such as Ethernet) allow all stations on a network equal access. Each station must "listen" to the network to determine periods of inactivity before transmitting. Any station wishing to use the network may begin transmitting provided that the network is not busy when it checks the network status. If multiple stations attempt to transmit simultaneously, a collision occurs. This is detected by all transmitting stations, which all must immediately stop transmitting and each wait a randomly determined period of time, before attempting to use the network again. Controller area network (CAN), for example, uses a variant of CSMA/CD for media access. Token-passing systems have a logical "token" that is exchanged among stations via network messaging. The station that holds the token has permission to transmit. All other stations are only permitted to receive messages. Stations wishing to transmit but not having the token must wait until the station holding the token passes it on. Another commonly used method of MAC is master–slave. In this method, one station on the network (designated the master) is generally in charge of, and originates, all communications. Slaves only respond to the master and only respond when the master initiates communications with them via sending a message to the slave. Profibus fieldbus messaging specification (FMS) (see later) is an example of a protocol that uses both token passing (in some cases) and master–slave (in some cases) to control media access.

15.2.3 Bandwidth

Bandwidth may have several different definitions. For digital communication systems, bandwidth describes the capacity of the system to transport digital data from one place to another. This term may be applied to the raw capability of the physical and data link layers to transport message data (*raw bandwidth*, closely related to the bit rate concept), or it may be applied to the effective rate at which user-meaningful information is transported (*effective bandwidth*). The bandwidth of a given system is generally inversely proportional to the worst-case node-to-node distance. The smaller the network span, the higher its bandwidth can be.

15.2.4 Addressing

Addressing is a concept that assigns generally unique identifiers to each station in a network system. This identifier (the address) can then be used by the network for a variety of purposes, including identifying the origin and/or destination of messages, or arbitrating access to a shared communications medium. Another addressing or identifier concept assigns unique identifiers not to stations but to unique pieces of data or signals that will be carried by the network. Stations then use an identifier according to what type of data they will be transmitting. Many but not all networking methods require establishment of an explicit address for each network station.

15.2.5 Arbitration

Arbitration is a function closely related to MAC. Arbitration is used by some networks to define the procedure followed when multiple stations wish to use the network simultaneously.

15.2.6 Signaling

Signaling refers to the actual physical (e.g., electrical, optical, or other) representation of data as they are carried on the media. For example, in some networks, data elements may be represented by certain voltage levels or waveforms in the media. In other networks, data elements may be represented by the presence of certain wavelengths of light in the media. The association of all the representable data elements (e.g., 0/1 or on/off) with the corresponding signal representations in the media is the signaling scheme or method. An important signaling method where electric wires are used as the medium is differential signaling. Differential signaling represents a particular data element (1 or 0) as two different states on a pair of wires. Determining the data element requires measuring the voltage difference between the two wires, not the absolute level of the voltage on either wire. Different data elements are then represented by the (signed) voltage difference between the two wires. For example, RS-485 represents a digital 1 data element as a 5 V signal level on the first wire and a 0 V signal level on the second wire, and a digital 0 as a 0 V signal level on the first wire and 5 V signal level on the second wire. One of the principal benefits of differential signaling is that it is possible to determine the data being transmitted without knowing the ground reference potential of the transmitter. This allows the transmitter and receiver to operate reliably, even when they have different ground potentials (within limits), which is a common occurrence in communication systems.

15.2.7 Encoding

Encoding refers to the process of translating user-meaningful information into data elements or groups of data elements to be transported by the network system. A code book refers to the set of all relationships between user-meaningful information and data carried by the network. Encoding may occur at several levels within the OSI reference model, as user-meaningful information is transformed successively until it becomes an actual network message, produced by the data link layer. Decoding is the reverse process, whereby a network message is successively translated back into user-meaningful information.

15.2.8 Modulation

Modulation in a classical sense refers to a signaling technique by which data or information is used to control some combination of the frequency, phase, and/or amplitude of a carrier signal. The carrier signal carries the information to a remote receiver where it will be demodulated to retrieve the information. Modulated network systems will briefly be explained in this chapter.

15.2.9 Message

A message is the fundamental, indivisible unit of information that is exchanged between stations. User-meaningful information will be grouped into one or more messages by the OSI network reference model.

15.2.10 Multiplexing

Multiplexing refers to the ability to use the media in a network to carry multiple messages or information streams "simultaneously." Multiplexed systems allow several communication channels to use the same physical wire or media. Each message or information stream may have different sources and destinations. Multiplexing may be accomplished using a variety of means. Time division multiplexing (TDM) involves breaking access to the media into a series of time quanta. During each time quantum, the media carries a separate message or information stream. The close arrangement of time quanta allows the network media to carry multiple messages "simultaneously." Code division multiplexing (CDM) involves the separation of the code book (see Section 15.2.7) into sections. Each section of the code book provides all of the messages that will be used for a particular information stream. Therefore, a particular information stream within the network media is distinguished by all of the messages that belong to the section of the code book for that stream. Frequency division multiplexing (FDM) divides an available bandwidth of a communication channel into several frequency ranges and assigns one information stream to each frequency range.

15.2.11 Protocols

A protocol is a defined method of information exchange. Protocols typically are defined at several levels within the OSI network reference model, such as at the application layer and at the data link layer. Protocols are used to define how the services provided by a particular layer are to be exercised, and how the results of these services are to be interpreted.

15.2.12 Service

A service represents a specific function or operation that is supported by a particular layer in the OSI network reference model. For example, an application layer service might be provided for the reading of or writing to a data element contained in another device (or station) on the network. This service might make use of a data link layer service that might be provided for supporting the exchange of a message with another device (or station) on the network.

15.2.13 Topology

Topology refers to the physical or geographic layout or arrangement of a network. Certain types of canonical topologies are commonly discussed in the context of networks, such as trunk line–branchline, star (or hub), ring, and daisy chain.

15.2.14 Bit Rate

Bit rate refers to the speed at which binary pieces of information (bits) are transmitted on a particular network. The raw bit rate of a network generally refers to the actual speed of transmission of bits on the network. The effective bit rate—or throughput—generally refers to the speed at which user information is transmitted. This number is less than or equal to the raw bit rate, depending on what percentage of the bits transmitted is used for carrying user information. The bits not carrying user information are overhead, used to carry protocol, timing, or other network information.

15.2.15 Duplex (Half and Full Duplex)

Half duplex refers to a communication system in which a station can either transmit information or receive information but not both simultaneously. A full-duplex network allows a station to transmit information and receive information simultaneously.

15.2.16 Error Control

Many network systems provide mechanisms to control errors. Error control has four aspects: prevention, detection, correction, and isolation. Error prevention may simply be shielding for the media to minimize electromagnetic disturbances, or it may be more complicated, such as signal sampling control to optimize the probability that a signal will be in the correct state when sampled. Error detection generally depends on detecting violations of protocol rules at various network levels or violations of computed data added to a message for error control purposes. Some examples of error detection techniques are parity and cyclic redundancy check (CRC). Both methods involve the computation of additional bits of information based on the data that are contained in a message and appending these bits to the message. For example, a particular protocol may require that the data link layer compute and append a CRC to a message prior to transmission. The receiver of the message may then also compute the CRC and compare it to the CRC that has been appended to the message. If a mismatch exists, then it is assumed an error has occurred. Error correction may take on a variety of forms. One of the simplest methods of error correction is to require that the data link layer of the transmitter retransmits a message that has been detected to have an error during transmission. This method is based on the assumption that the error was caused by a disturbance that is unlikely to occur again. Another method of error correction involves transmission of additional bits of information along with the user information in a message. These additional bits of information are computed by the transmitter to provide redundant information in the message. When fewer than a certain number of bit-level errors have occurred during the transmission of the message, the receiver is able to reconstruct the original user information accurately using the redundant information (bits) supplied within the message. Error isolation is a capability of some networks to localize the source of errors and isolate the sections of the network or the stations at which the errors have been localized. Error isolation allows the fault-free portions of the network to continue communicating even when other portions of the network have degraded to the point of generating errors.

15.2.17 Internetworking

There are occasions when communications between two or more points are best handled by multiple networks. This may be the case when a single network has limitations that prevent it from tying the points together (e.g., distance limits) or when multiple networks are required for other reasons (e.g., to carry different types of data). When multiple networks are used to provide communications, there may be a need to pass messages or information directly from one network to another.

A repeater may be used when the networks to be joined are logically identical, and the purpose is simply to extend the length of the network or extend its capabilities in some way. A repeater generally has no effect on messages and simply carries all messages from one cable or port to another (i.e., a change of physical media). A repeater allows for connection of networks at the physical layer level.

A bridge is similar to a repeater but allows for connection of networks at the data link layer level. Generally, a bridge will pass all messages from one network to another, by passing messages at the data link layer level.

A router usually has the function of partitioning similar networks. Two networks may be based on the same technologies and protocol but may not be logically identical. In these cases, some, but not all, of the messages on one network may need to be carried or transported to the other network. The router has

the function of determining which messages to pass back and forth based on certain rules. Functions to enable efficient, automatic routing of messages may be included in layer 3 (the network layer) of the OSI network reference model, and a router allows for connection of networks at the network layer level.

A gateway may have a function similar to a router, or it may have the function of joining dissimilar networks, that is, networks based on dissimilar technologies and/or protocols. When functioning like a router, a gateway usually performs its discrimination at a higher protocol level than a router. When a gateway joins dissimilar networks, generally a more complex set of rules must be designed into the gateway so that message translation, mapping, and routing can occur within the gateway as it determines which messages to pass from one network to the other.

15.2.18 ISO/OSI Network Reference Model

The explosion in the use and types of communication networks over the last several decades has led to more precise descriptions and treatment of communication networks in general. The International Organization for Standardization (ISO) has recognized one such method of precise description of networks, called the OSI reference model [1]. As shown in Figure 15.2, this model decomposes an arbitrary communication network into a "stack" of seven "layers." At each layer, certain types of network communication functions are described. The user of the communication system—usually another system that needs to communicate on the network—interacts with layer 7, the highest layer. The actual transmission medium (e.g., copper cable, fiber optic, free space) is connected to layer 1, the lowest layer. Most communication networks do not implement all of the layers in the reference model. In this case, formal definition, treatment, or inclusion of certain layers of the model in the actual network design is omitted. Layers 1, 2, and 7 are typically present in all networks, but the other layers may only be explicitly

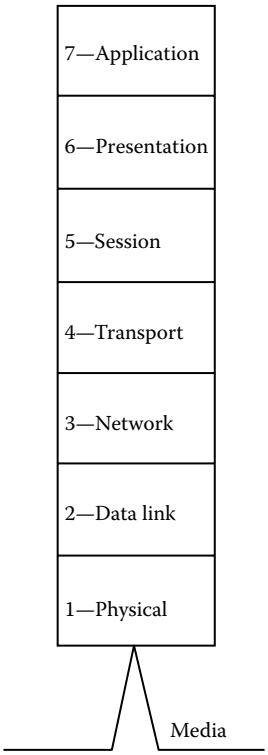


FIGURE 15.2 The ISO/OSI seven-layer model provides a method for segmenting communication functions.

included or identifiable when their function is an important part of the network communications. In many sensor communication networks, the functions performed by layers 3, 4, 5, and 6 are "collapsed" into vestigial additions to the functions of layer 7, the application layer.

15.2.19 Physical Layer

The physical layer is the lowest layer of the model. This layer is responsible for converting between the symbolic or data representation of the network messages and the actual physical representation of data in the network medium. This layer specifies the behavior of the electric circuits referred to as the transmitter and the receiver. It also defines physical structures for connectors.

15.2.20 Data Link Layer

The data link layer, or layer 2, is responsible for several functions. This layer manages access to the network medium (MAC), structures the bits of information into well-defined groups identified as "frames" or messages, handles identification of source and destination stations on the network, and provides for error-free transmission of a message from source to destination stations, all according to the data link layer protocol. A number of standard data link layer protocols exist, which act as the basis for many of the communication networks in wide use. Ethernet, or IEEE 802.3, for example, specifies a MAC sublayer that works with the IEEE 802.2 Logical Link Control layer to form the data link layer protocol used in the majority of office information networks [2].

15.2.21 Network Layer

The network layer encapsulates functions related to routing of messages, both within a single network and among multiple networks. This layer typically uses addressing in a variety of forms as a key part of the functions of directing and routing messages and the search and usage of the available communication paths.

15.2.22 Transport Layer

The transport layer provides any additional data transfer functions not directly provided by the data link layer for end-to-end reliable messaging. For example, some data transfer functions between stations may require the use of multiple data link layer messages to accomplish a reliable message transfer. The generation of multiple messages and the sequential disassembly, delivery, and assembly of data are accomplished by the transport layer. The transport layer also recovers from lost, duplicated, and misordered messages.

15.2.23 Session Layer

The session layer provides for a higher level of control and management of network usage and data flow than that provided at lower layers, including opening or building up a communication channel, maintaining the channel, and closing the channel. This layer is infrequently implemented in contemporary systems.

15.2.24 Presentation Layer

The presentation layer provides functions to transform data from formats that are transportable by the network to the user-accessible formats that are defined in the application layer and understood in the local station.

15.2.25 Application Layer

The application layer, or layer 7, provides communication services directly to the user application. The usage and formatting of these services is summarized in the application layer protocol. The user interacts with the network by invoking functions and services provided by the application layer and passing data to and from the network through these services.

15.3 Network Technologies

There is a wide range of technologies in various stages of development and standardization, which address virtually all levels or layers of the ISO/OSI network reference model. One or more of the available technologies will probably suit almost any networking need. An analysis of the available technologies and their limitations will also be beneficial if it is deemed that a networking method must be designed to meet a particular application. The selection and description of technologies is by no means complete or exhaustive. The technologies presented are selected from several industries that make common use of networking to communicate sensor data. Table 15.1 provides a comparison of selected parameters for a set of networks.

15.3.1 RS-232

RS-232 (ANSI/EIA/TIA-232-E-91) is a widely used method of communication, which has been standardized in a variety of places including the Electronics Industry Association [3]. RS-232 represents elements of layer 1 of the OSI model, for communicating between two (and only two) stations. RS-232 provides a separate wire for transmission of data in each direction between the two stations and gives the two stations different designations—data terminal equipment (DTE) and data communications equipment (DCE)—so that a method exists to distinguish which station will use which wire to transmit and receive. The signal levels for RS-232 represent a digital 1 bit as a voltage in the range of 5–12 V on the wire and a digital 0 bit as a voltage of negative 5–12 V on the wire. RS-232 is typically implemented in a full-duplex fashion, since each station can transmit to the other simultaneously using separate wires. RS-232 can be made to operate at a variety of bit rates but typically is used at bit rates from 300 bit/s up to 115,200 bit/s.

TABLE 15.1 Comparison of Selected Parameters (Maximum Values) for Various Network technologies[a]

	Length	Stations	Bit Rate	Wires	Media	Topology
RS-232	30 m	2	115 kb/s	2	TP	P–P
RS-485	1200 m	32	10 Mb/s	2	TP	D-C
Seriplex	1500 m	256	200 kb/s	4	2STP	D-C, Free
AS-i	100 m	32	167 kb/s	2	UP	T–B
Interbus-S	25.6 km[b]	64	500 kb/s	6	3STP	Ring
CAN	450 m	64	1 Mb/s	4	2STP	T–B
4-20 mA	1000 m	2	—	2	STP	P–P
HART	1000 m	2(15)	1200 b/s	2	STP	P–P (D-C)
Profibus	9600 m	126	12 Mb/s	2	STP	D-C
Found. Fieldbus	1900 m	32	2.5 Mb/s	2	STP	D-C
LonWorks	1400 m	64	1.2 Mb/s	2	STP	D-C, Free

Notes: P–P, point to point; D-C, daisy chain; T–B, trunk line–branchline; TP, twisted pair; STP, shielded twisted pair; UP, unshielded pair.

[a] Maximum parameters for networks are not achievable simultaneously and do not include repeaters, routers, or gateways. Maximum parameters are estimates based on available information.

[b] Maximum 400 m between stations.

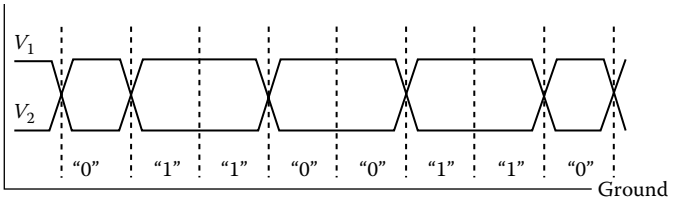


FIGURE 15.3 A sample RS-485 waveform showing voltages on differential wire pair (V_1, V_2) and superimposed bit intervals showing 0 and 1 bits. The ground reference is arbitrary within the defined signaling range.

15.3.2 RS-485

EIA RS-485 was made a standard in 1983, derived from the RS-422 standard. RS-485 provides for differential transmission of data on a pair of wires among 32 or more stations. Like RS-232, the standard is a layer 1 specification. RS-485 provides for half-duplex communication, since a station cannot simultaneously transmit and receive independent data streams. Each station in an RS-485 system can have either a transmitter or a receiver, or both (commonly called a transceiver). Most implementations provide a transceiver. When one transceiver is transmitting, all others should be receiving (i.e., not transmitting). Which station is allowed to transmit at which time is not specified in the standard and must be covered by a higher layer protocol (e.g., Interbus-S, Profibus distributed periphery [DP]). Figure 15.3 shows a sample RS-485 waveform, indicating the differential nature of the signaling.

15.3.3 Seriplex*

Seriplex® is a digital, serial multiplexing system developed by Automated Process Control, Inc., in Jackson, MS. Square D Corporation purchased Automated Process Control and the rights to Seriplex in 1995, and subsequently launched Seriplex Technology Organization (STO) to manage the protocol. Seriplex is designed to be particularly efficient at handling large numbers of digital or on/off input and output points. Seriplex provides three communication wires, one for a clock signal, one for a data signal, and a ground reference. The system can be operated in two different modes (peer-to-peer and master–slave). In master–slave mode, one station is designated the master. The master synchronizes all data transmission among stations by driving a digital waveform on the clock line that all stations listen to and use for timing of transmit and receive operations. The master generates a repetitive pattern on the clock line that causes all stations to transmit and/or receive data on each cycle or "scan" of the network. Each station is given an address and uses the address along with the clock signal to determine when to drive the data line (in the case of an input point) or when to monitor the data line for valid output data (in the case of an output point). There are variations possible in implementation that allow for various clock speeds and bit rates (16, 100, and 200 kHz). Other protocol details allow for the handling of analog or multibit input and output points (by combining several bits on sequential scans together), bus fault detection, input redundancy, and communication error control using multiple scans of the network. Implementing the protocol in a sensor or other device typically requires using a Seriplex application-specific integrated circuit (ASIC), which must be licensed from the STO [4].

15.3.4 AS-i

Actuator Sensor Interface (AS-i) was developed by a consortium of primarily European companies interested in developing a low-cost, flexible method for connecting sensors and actuators at the lowest levels of industrial control systems. The system is managed by an independent worldwide organization [5].

* Seriplex is a trademark of the STO.

The AS-i system provides a two-wire, nontwisted cable for interconnection of devices. Devices may draw current from the two wires (nominally at 24 V dc) for powering circuitry, and the data communications are modulated on top of the nominal dc level at a bit rate of 167 kHz, under the control of the master. A single parity bit per station is used for error detection. Similar to Seriplex, an AS-i device is typically implemented using a special ASIC that handles the communication.

15.3.5 Interbus-S

Interbus-S was developed by Phoenix Contact [6] and is controlled by the Interbus-S Club. The topology of the network is a ring, with data being sequentially shifted from point to point on the ring under the control of a network master. Each device in the ring acts as a shift register, transmitting and receiving data simultaneously at 500 kHz. The actual serial data transmission between stations conforms to RS-485. Interbus-S transmissions include a CRC for error detection. Interbus-S (Interbus-S remote bus) has also been extended to include a subprotocol called Interbus sensor loop (or Interbus-S local bus). This subprotocol provides an alternate physical layer, with a single twisted pair carrying power and data on the same lines and a reduction in the minimum size of the shift register in each station from 16 to 4 bits. Each Interbus sensor loop system can act as a single station on an Interbus-S network, or the sensor loop can be connected directly to a controller or master. Interbus-S devices are usually implemented with a special ASIC.

15.3.6 CAN

CAN is a data link layer (layer 2) network technology developed by Robert Bosch Corporation [7], with an application target of onboard automotive networking. The technology is standardized in ISO 11898 [8], licensed to all major integrated circuit manufacturers, and is widely available—both as separate CAN controllers and CAN controllers integrated with microprocessors. As a result, CAN has been used in a variety of industries. As a data link layer technology, it is not a complete network definition. A number of physical layer options are usable with CAN (e.g., twisted pair, fiber optic, radio-frequency [RF] wireless) and some have been subject to standardization (e.g., ISO 11898). Also, a number of application layer protocols have been developed for use with CAN, such as DeviceNet, Smart Distributed System (SDS), CANOpen [9], and SAE J1939 [10]. Both DeviceNet [11] and SDS [12] have developed systems for creating networks of industrial field devices for the factory floor, including sensors and actuators. SDS was originally developed by Honeywell, now it is a part of IEC 62062 Standards Part 5 Ed 1.0 B:2000 [12].

15.3.7 4–20 mA Current Loop

The 4–20 mA current loop is a widely used method for transferring information from one station (the transmitter) to another station (the receiver). Therefore, this system allows for only two stations. A typical current loop system assigns a sensing range (e.g., 0 °C–100 °C) to the current range between 4 and 20 mA. A loop exists (i.e., two wires) between the transmitter and receiver. The transmitter can impress a certain current in the loop (using a controlled current source) so that the receiver can measure the current in the loop (e.g., by placing a small resistor in series with the loop and measuring the voltage drop across the resistor). After measuring the current, the receiver can then determine the present level of the sensed signal within the defined sensing range. This method uses current signaling, instead of voltage signaling, and therefore is relatively unaffected by potential differences between the transmitter and the receiver. This is similar to the benefit of differential (voltage) signaling, which also requires two wires. Another characteristic of this method is that it is not primarily digital in nature, as many other sensor communication systems are. The measured value can vary continuously in the range of 4–20 mA and therefore can easily represent an analog sensing range, rather than a set of digital signals. Also, the signal is continuously variable and available. Another characteristic of this method is that the integrity of the loop can be verified. As long as the loop is unbroken and the transmitter is in good working order,

the current in the loop should never fall below 4 mA. If the current approaches 0 mA, then the receiver can determine that a fault exists—perhaps a broken cable. These systems are widely used in various process control industries (e.g., oil refining) for connecting sensors (transmitters) with control computers. Because one station is always the transmitter and one station is always the receiver, this is a unidirectional, half-duplex communication system.

15.3.8 HART*

HART® is a protocol that builds upon 4–20 mA communication systems. The basic idea is that additional data (beyond the basic sensor signal being carried in the current loop) can be transmitted by modulating a signal on top of the current flowing in the loop. The actual modulation method conforms closely to the Bell 202 standard for analog modem communications on telephone lines at 1200 bit/s. Because a 4–20 mA current loop carries a relatively slowly varying signal, it is easy to separate the 4–20 mA signal from the digital signal using filters. The Bell 202 standard uses continuous-phase frequency shift keying between two frequencies at up to 1200 shifts/s to modulate digital ones and zeros onto the 4–20 mA current loop. This method allows for bidirectional, full-duplex communication between the two stations, on top of the 4–20 mA signal. It is also possible to configure HART communications on a network that is not carrying a 4–20 mA signal, in which case up to 15 devices can be connected together on the network. HART was developed by Fisher-Rosemount Corporation and has been transferred to an independent foundation for management [13]. Because HART is compatible with US telephone systems, it can theoretically be run over the telephone line and is therefore capable of running over arbitrarily long distances.

15.3.9 Profibus

Profibus (PROcess FIeld BUS) is one of three networks standardized by a European standard [14]. Profibus is under the control of a global organization, PNO [15]. Profibus is an umbrella network standard that encompasses three subnetworks within the Profibus family. Profibus DP is the variant that is designed specifically for communication with field devices (sensors and actuators) at the device I/O level. Profibus process automation (PA) is a variant that has more capabilities designed to support the needs of device-level networking for process industries, such as oil refining. One of the capabilities of Profibus PA is its ability to be installed in an intrinsically safe (IS) way, thus providing a higher degree of safety in environments that may be explosive or otherwise hazardous. Profibus PA typically uses a special physical layer specification standardized under IEC 61158-2, which is used by several network systems for process automation applications. IEC 61158-2 specifies a two-wire twisted pair implementation carrying both power and data on the same two wires at 31.25 kbit/s. Profibus FMS represents the highest level implementation, which is used to link together controllers (not field or I/O devices) in a factory.

Profibus DP systems are typically master–slave systems, where usually a single network master (the host controller) communicates with a number of slave devices (remote I/O blocks and other I/O devices). The protocol provides for cyclic exchange of I/O information as well as on-demand exchange of other types of information. Profibus DP can be implemented on several different physical layers, including RS-485 and fiber optics, at various bit rates up to 12 Mbit/s. Profibus messages include a CRC for error detection.

15.3.10 Foundation Fieldbus

Foundation fieldbus (FF) is a networking standard that has grown out of an effort within industry standards organizations, especially ISA-SP50 [16] and IEC SC65C/WG6 [17], to provide a replacement for the 4–20 mA analog sensor communication standard. FF provides two basic levels of networking: H1 and H2. H1 is a

* HART is a trademark of the HART Communications Foundation.

lower-speed system that can provide IS operation and uses a single twisted pair to deliver both power and data communications to field devices, according to IEC 1158-2. Running at a bit rate of 31.25 kbit/s, H1 is very similar to Profibus PA, when run on the IEC 1158-2 physical layer standard. The H1 system is designed to be able to connect hierarchically "upward" to an H2 system, which acts as the host. FF H2 can be run at either 1 or 2.5 Mbit/s on twisted pair wires and also provides an IS option at the 1 Mbit/s rate. The H2 system can act as a network backbone in a factory environment, carrying data among various H1 systems.

15.3.11 WorldFIP

WorldFIP [18] is another technology of the three that were standardized in the European standard EN 50 170, running on the IEC 1158-2 physical layer. Many of the proponents of WorldFIP have embraced FF and contributed to the development of that standard. WorldFIP is a member of the FF, and FF has incorporated many of the capabilities of WorldFIP as a result. When run on the IEC 1158-2 physical layer, WorldFIP has similar capabilities to FF.

15.3.12 LonWorks*

LonWorks® is a networking technology developed and controlled by the Echelon Corporation [19]. LonWorks is designed to be a general-purpose networking technology suitable for a variety of industries. LonWorks has been applied extensively in the building automation and control industry, as well as a variety of other industries. The core LonWorks technology for devices is contained in special integrated circuits—called Neuron® chips—which combine several microprocessors to manage the network and communications and provide a general-purpose control environment. These chips are available from Motorola, Inc., and the Toshiba Corporation, which are licensees of the LonWorks technology. Echelon has also announced the possibility to license the LonTalk® protocol to other manufacturers for implementation in other microprocessors. LonWorks networks can be implemented on a variety of physical layers, including twisted pair at several bit rates and wireless options at 4800 bit/s, but the most common is a differential twisted pair system running at 78 kbit/s. Most of the networking details (the LonTalk protocol) are hidden from the user and are encapsulated as functions within the general-purpose control environment. The user programs (using a language like the C programming language) the Neuron chip for each station to behave in a certain way and communicate various data items to other stations. Then, specialized tools are used to tie all of the stations together (handling addressing and other network details) to yield a functioning network. The system combines flexibility with a certain amount of ease of implementation and can easily be applied to a variety of applications.

15.4 Applying Network Communications

15.4.1 Shielding

Many communication networks require shielding of the media (the cable). Shielding constitutes an electric conductor that completely encases the communication media (e.g., twisted pair) to provide protection against EMI. Shielding provides an electric conductive barrier to attenuate electromagnetic waves external to the shield and provides a conduction path by which induced currents can be circulated and returned to the source, typically via a ground reference connection. Shields in communication systems are often grounded at only a single point. This single point of ground prevents the shield from participating in a "ground loop," which is an alternative path for current to flow between two points of potential difference connected to a common ground. Ground loops can lead to noise problems and can be destructive if the stray currents are large enough, since a shield ground is usually not constructed to carry heavy currents.

* LonWorks, LonTalk, and Neuron are trademarks of the Echelon Corporation.

15.4.2 Media

The most common media types for network systems fall into three categories: electric, optical, and electromagnetic. Electric media are based on conductors (e.g., copper wire), whereas optical media are based on optical waveguides or fiber optics. Electromagnetic media consists of free space, or general electromagnetic wave-permeable materials, and are referred to as wireless systems. Within the category of electric media are a large variety of conductor configurations. The most common are unshielded pair, unshielded twisted pair (UTP), shielded twisted pair (STP), and coaxial (COAX). These conductor configurations have various properties that are significant for the transmission of electric signals, as well as varying degrees of immunity to EMI. As a rule of thumb, the quality of the transmission line characteristics (signal transmission and immunity to EMI) improves in the order listed. Twisted pair systems are generally easier to install, whereas coaxial and fiber-optic systems generally require more specialized tools and termination methods. Of course, wireless systems are easy to install, but attention must still be paid to the media. The characteristics of the free space such as distance and amount of EMI present must be considered for reliable operation of the network.

15.4.3 Bit Rate

Some networks provide only one choice of bit rate, whereas others provide user-selectable options for bit rate. Bit-rate options may be dependent on the type of media that is installed. As a rule of thumb, the bit rate chosen should be the lowest possible bit rate that still supports the application requirements for speed of data transfer and overall bandwidth. This generally results in more reliable operation and generally gives the network more immunity to minor degradations, specification violations, and EMI.

15.4.4 Topologies

There are a variety of network topologies that are commonly used. Topology refers to the physical arrangement and interconnection of stations by the media. Some networks can be run using several different topologies; some can only be run with a certain topology (e.g., Interbus-S requires a ring topology). The most common topologies are daisy chain, trunk line–branchline, ring, and star. Variations on these exist, and networks that incorporate or can be run on highly varied topologies are sometimes called free-form, tree, or free topology networks. Figure 15.4 depicts graphically several different types of topologies. In some cases, networks require certain topologies. Deviating from these can cause degradation in network behavior (e.g., corruption of messages) or network failure.

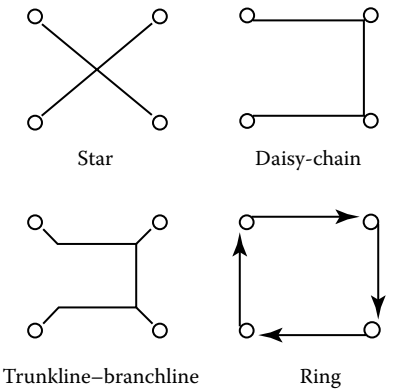


FIGURE 15.4 Some examples of the many possible network topologies using four stations.

15.4.5 Configuration

Most networks involve some sort of configuration. Configuration is the process of connecting stations together and assigning certain programmable parameters to each station required for proper operation of the network. The most common configurable parameter in many networks is the station address. Some networks may require other parameters to be preset, such as the communication speed or bit rate. Some networks have the capability to autoconfigure, which means to assign parameters automatically to stations as part of the network start-up process, without explicit user intervention (e.g., Interbus-S). Many networks define various tools, which may be computer based, to assign parameters to each station in order to configure the network. In other cases, the stations may incorporate switches or other manual means to configure the necessary parameters for network operation.

15.5 Recent Advances

15.5.1 Wireless Technologies

The need for networking is present even in environments where an electrical or optical cable cannot be easily distributed. This may be due to various limitations, such as difficulty in running a new cable from one building to another or connecting to sensors in motion or on vehicles. There are two general categories of wireless communications, based on electromagnetic frequency spectra. Various wireless technologies employ the infrared spectrum. These technologies generally have transmission limited to applications that have a direct line of sight between stations. Also, the distances are generally limited to 100 m or less. Because of these limitations, there are generally no legal restrictions in employing these frequency spectra, and infrared transceivers are now becoming available from a variety of manufacturers.

The other general category of wireless communications is based on RF communications. In most countries, the use of these spectra is tightly controlled by governmental agencies. As a result, employing wireless networking in most of these frequency ranges requires special licensing. However, a number of frequency ranges are reserved for low-power public communications. Within these frequency ranges, devices are allowed to communicate in an unlicensed fashion as long as they transmit according to certain rules about transmitted power output. RF-based wireless systems are generally not limited to line-of-sight applications and can be designed to cover greater distances than infrared-based systems.

Wireless technologies can be viewed as simply another choice for the physical layer media, that is, free space. As such, it is possible to consider, in some cases, a wireless media for implementation of a variety of protocols. For example, both CAN and LonWorks systems could be candidates for wireless networking.

15.5.2 Fiber Optics

Another physical layer media choice is fiber-optic media. Fiber-optic media employs pulses of light delivered along a tubular waveguide (glass or plastic fiber) to transmit information from one station to another. Fiber optics enjoy some benefits over traditional copper wiring. First, attenuation of light within fiber optics is generally about an order of magnitude less than attenuation of an electric signal within a copper wire. Second, fiber-optic transmission systems can be modulated (or pulsed) at much higher frequencies, yielding greater potential bandwidths and bit rates than copper media. Finally, fiber-optic systems are generally immune to the traditional sources of EMI that can cause trouble for copper media systems. There are also limitations in present implementations of fiber-optic systems. One of the limitations is that special tools and termination techniques must be used to connect a fiber-optic cable to a sensor or field device. Second, fiber-optic "taps" are not easily created. Therefore, most fiber-optic systems are implemented in a point-to-point fashion. When multiple devices are involved in a network, each device usually acts as an optical repeater, with a fiber-optic input and a fiber-optic output port.

15.5.3 Network Design Considerations

Designing a network communication system from the ground up can be a lengthy undertaking and should not be considered unless a careful review of available technologies has yielded no solutions to the particular requirements of the application. The designers must take into account a number of fundamental questions to shape the capability of the network. One topic mentioned frequently in the area of networking for control applications is the subject of determinism. This refers to the ability of the network to behave in a predictable fashion under any given set of stimuli or external conditions. Many networks do not exhibit this characteristic. Another question to be resolved is the subject of priority and media access. The designers must determine the conditions under which any particular station is allowed to transmit, and if multiple stations are attempting to transmit, how it will be determined which station will be given priority to transmit first. Media access methods often impact the ability of a network to behave in a deterministic fashion.

15.5.4 Integrating Sensors with Communications: IEEE P1451

A recent interesting development in the area of sensor networks is an effort being sponsored by the IEEE [20] out of its TC-9 committee, called IEEE P1451. This activity is working toward the development of a standard to define sensor (or transducer) interfaces to networks generically. The first part of the proposed standard, IEEE P1451.1, includes definitions for the interface between the device and the network (refer to Figure 15.1). The second part, IEEE P1451.2, includes definitions for the interface between the transducer (or sensor) and the network interface block within the device. P1451.2 includes a definition for a transducer electronic data sheet (TEDS), which defines a summary set of information pertinent to the sensor, allowing for standardized exchange of data on the network. The proposed standard has the potential benefits to make it easier to connect a sensor to a variety of networks and to allow similar sensors from different manufacturers to be handled in a similar fashion on the network.

References

1. ISO/IEC 7498-1:1994 Information technology—Open systems interconnection—Basic reference model: The basic model, International Organization for Standardization (ISO), Genève, Switzerland, http://www.iso.org/, accessed on August 24, 2013.
2. 8802-3: 1996 (ISO/IEC) [ANSI/IEEE Std 802.3, 1996 Edition] Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications, Institute of Electrical and Electronics Engineers, Piscataway, NJ, http://www.ieee.org/, accessed on August 24, 2013.
3. ANSI/EIA/TIA-232-E-91, Electronic Industries Association, Arlington, VA, http://www.eciaonline.org/eiastandards/, accessed on August 24, 2013.
4. Distributed, intelligent I/O for industrial control and data acquisition… The SERIPLEX Control Bus, Bulletin No. 8310PD9501R4/97, Seriplex Technical Organization, Raleigh, NC, http://static.schneider-electric.us/docs/Machine%20Control/Component%20Networks/SERIPLEX/8310PD9603.pdf, accessed on August 24, 2013.
5. AS-Interface U.S.A., Scottsdale, AZ, http://www.as-interface.com/, accessed on August 24, 2013.
6. Interbus-S protocol structure, Data Sheet 0005C, Phoenix Contact, Harrisburg, PA, http://www.profibus.com/technology/interbus/, accessed on August 24, 2013.
7. CAN Specification, Version 2.0, 1991, Robert Bosch GmbH, Stuttgart, Germany.
8. ISO 11898:1993 Road vehicles—Interchange of digital information—Controller area network (CAN) for high-speed communication, International Organization for Standardization (ISO), Genève, Switzerland, http://www.iso.org/, accessed on August 24, 2013.

9. CiA Draft Standard 301 (Version 3.0), CANopen communication profile for industrial systems, CiA Headquarters, Erlangen, Germany, http://www.can-cia.de/, accessed on August 24, 2013.
10. SAE J 1939—Recommended practice for serial control and communications vehicle network, Society of Automotive Engineers, Warrendale, PA, http://www.sae.org/, accessed on August 24, 2013.
11. DeviceNet Specification vl.4, Open DeviceNet Vendors Association, Coral Springs, FL, http://www.odva.org/, accessed on August 24, 2013.
12. Smart distributed system application layer protocol specification v2.0, 1996, Honeywell MICRO SWITCH Division, Freeport, IL, http://sensing.honeywell.com/index.php?ci_id=2717&la_id=1, Incorporated to IEC62026-5 ed 1.0 B:2000, Low voltage switchgear and controlgear - Controller device interfaces (CDIs) - Part 5 Smart Distributes Systems (SDS), available at http://www.techstreet.com/products/724176, accessed on August 23, 2013.
13. HART Communication Foundation, Austin, TX, http://www.hartcomm.org/, accessed on August 23, 2013.
14. EN 50 170—Volume 2, CENELEC Central Secretariat, Brussels, Belgium.
15. PROFIBUS Trade Organization U.S.A., Scottsdale, AZ, http://www.profibus.com/, accessed on August 24, 2013.
16. ISA, The International Society for Measurement & Control, Research Triangle Park, NC, http://www.isa.org/, accessed on August 24, 2013.
17. IEC 61158-2(1993-12), Fieldbus standard for use in industrial control systems—Part 2: Physical layer specification and service definition, International Electrotechnical Commission, Geneva, Switzerland, http://www.iec.ch/, accessed on August 24, 2013.
18. WorldFIP Headquarters, Antony, France, http://www.world-fip.org/, accessed on August 24, 2013.
19. 005-0017-01 Rev C, LonTalk Protocol, Echelon Corporation, Palo Alto, CA, http://www.echelon.com/, accessed on September 22, 2013.
20. P1451.2, Draft standard for a smart transducer interface for sensors and actuators—Transducer to microprocessor communication protocols and transducer electronic data sheet (TEDS) formats, Institute of Electrical and Electronics Engineers, Piscataway, NJ, http://www.worldfip.org/, accessed on September 22, 2013.

Further Information

Gibson, J. D., Ed., *The Communications Handbook*, Boca Raton, FL: CRC Press, 1997, includes recent material on communication techniques.

Peebles, P. Z., Jr., *Digital Communications Systems*, Englewood Cliffs, NJ: Prentice Hall, 1987, provides a good general text on communication.

Svacina, B., *Understanding Device Level Buses*, Minneapolis, MN: TURCK, Inc., 1996, is an in-depth study of the subject of communication networks for industrial field devices.

II

Electrical Variables

16

Electric Current Sensors*

Pavel Ripka
Czech Technical University in Prague

16.1 Introduction

Current sensors are required for numerous applications, including

- Industrial applications—large ac and dc currents
- Car applications: special requirements for hybrid cars
- Energy production, conversion, storage, and consumption
- Current leakage monitoring
- House automation
- Motor drives, power converters, and power modules
- Current sensors embedded in ICs

The basic literature on current sensors is a book by Iwansson [2] and comprehensive review articles [1,3]. Contactless current sensors are based on magnetic sensors that are covered by [4].

* This chapter is based on an abbreviated paper [1] with kind permission of IOP.

16.2 Basic Current-Sensing Parameters

The basic parameters of any sensor are *linearity*, *offset* (for dc sensors) and *sensitivity*, and also stability of offset and sensitivity with temperature and time. Also important is the sensor bandwidth (or frequency characteristics), as current waveforms often contain high-frequency component. Current sensors with ferromagnetic yoke or field concentrators also suffer from *hysteresis* and *perming* (change in offset after the sensor is overloaded).

Wide bandwidth: a large frequency range from dc to several MHz is required for applications in switching power converters.

High dynamic range and resistance against overload is another important parameter. Accuracy is often required even for a small fraction of the full scale. One example is the requirement to measure the current consumption of the house in the standby mode and during full operation.

Immunity against dV/dt is critical for use in modern power modules. Current sensors located close to fast-switching transistors can be subjected to transients of 10 kV/μs. Such transients can be coupled by parasitic capacitances to the output of the current sensor. Proper electrostatic shielding can efficiently suppress this effect, but it complicates the design of the sensor.

Low cost and high environmental resistance is a must for industrial sensors.

16.3 Geometrical Selectivity

Also important for contactless sensors is *geometrical selectivity*—the reading should be independent of the position of the measured conductor and the sensor should be resistant to external currents and magnetic fields [5].

The most common way to achieve geometrical selectivity is to use a closed magnetic circuit with a measured conductor inside. This is used in current transformers (CTs), fluxgate current sensors, and in most Hall current sensors. The high-permeability magnetic circuit (the yoke is often in the shape of a ring) concentrates all field lines from the measured current, so that the measured flux in the magnetic circuit does not depend on the actual position of the conductor. In this case, H in the yoke central line l is constant, and we can write

$$H \times l = N \times I \tag{16.1}$$

where

I is the measured current
N is the number of turns

The magnetic circuit also shields against external fields.

The best magnetic circuit for a current sensor is a toroid made from high-permeability material without an air gap. This is used in CTs and in some residual current sensors.

It is important to use a high cross-sectional area of the magnetic material to keep a high demagnetization factor against external fields.

The magnetic yoke should not be saturated, that is, $l > NI/H_{sat}$. This means that the magnetic circuit becomes very large for high measured currents. Air gaps can be introduced to avoid saturation, but these degrade the geometrical selectivity. Another approach is to compensate the measured current using a multiturn coil. However, generating the compensation current requires power.

Magnetic gradient techniques can be used if the power or space limitations do not allow the use of a yoke. Simple integrated current sensors use a folded conductor and a gradient field sensor. This suppresses the response from distant sources, giving a low gradient. For large currents,

the current bar should be kept straight, and circular arrays of typically four to eight sensors are used [6]. More complex weight sum field decoupling method allows using sensor arrays of different shapes [7].

16.4 Overview of Electric Current Measurement

16.4.1 Definitions and Standards

Electric current is caused by movement of charge, and in the older MKS system it was therefore naturally defined as a derived unit. In the present SI unit system, the Ampere is the fundamental unit that is realized using the current balance. A much more practical realization of the Ampere is via Ohm's law from the volt and the Ohm, which are easily reproducible using the Josephson junction and the quantum Hall effect.

The proposed new definition of the Ampere is a return to the natural concept based on elementary charge and time.

National metrology institutes offer calibration of current sensors with minimum expanded uncertainty of typically 10 μA/A for dc currents and 100 μA/A for ac currents. The ac current ratio is calibrated with minimum expanded uncertainty of 5 ppm in ratio error and 5 μrad in phase displacement.

16.4.2 Current Shunts and Other Contact Methods

The disadvantages of current shunts are obvious: the measured current has to be interrupted to introduce the sensor, shunts for large currents are bulky, they dissipate heat, and the output is galvanically connected with the measured circuit. However, current-sensing resistors are a robust and cheap solution for many applications. Self-made coaxial current shunt with minimum inductance and wide frequency bandwidth can easily be made from a set of metal resistors in parallel [8].

If precision is not a critical consideration, the current can be measured even through the voltage drop across a current-carrying copper trace or bus bar. In power converters, the current is often measured as a voltage drop across the inductance or less precisely using a voltage drop across the power switch or more precisely using a sensing MOSFET, which is matched with the power MOSFET, but has much lower channel width [9].

In order to measure very small currents, the value of the current-sensing resistor should be increased so that the measured drop is measurable. For measuring picoamps and lower currents, a current-to-voltage converter (Figure 16.1) with op-amp is preferred. The range is changed by output divider R_A, R_B rather than by changing the feedback resistor R_F.

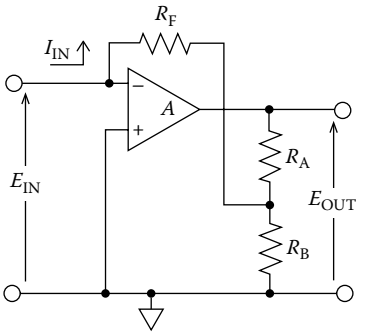


FIGURE 16.1 Current-to-voltage converter. (From Keithley, Inc., Low current measurements, AN 100, 2007, www.keithley.com.)

16.5 Contactless Sensor Principles

Current measurement using a shunt resistor is impractical or impossible in some cases. Contactless current sensors keep galvanic insulation between the measured current and their output voltage. A wide range of ac and dc contactless current sensors are produced by LEM, F.W. Bell, VAC, Honeywell, Telcon, and many other manufacturers. Magnetic field sensors can be used to measure currents in remote conductors at high potentials, in underground cables, and in building structures.

16.5.1 Instrument Current Transformers

A CT (or ICT) usually has a bulk ring core made of high-permeability material. For currents higher than 50 A, the primary winding is a single conductor through the core opening. The secondary winding should ideally be short-circuited; however, for most applications, it is connected to a small "burden" resistor, which serves for sensing the secondary current. The core is wound with high-permeability tape (for low-frequency devices) or is made from ferrite (for high-frequency current probes). In some devices, the secondary current is measured by active current-to-voltage converter in such a way that the burden is virtually zero. This leads to smaller error, as the transformer core works at lower flux density.

CTs are also made with an openable core, most often as ac current clamps for oscilloscopes or multimeters.

CTs are very popular devices:

- They are very simple and robust.
- They do not require external power.
- They have high galvanic insulation.
- They are cheap.
- They have a long lifetime with invariant parameters.

Figure 16.2 shows the equivalent circuit of the CT.

At low frequencies (such as 50/60 Hz), the dominant source of error is the magnetizing current I_L, which is inversely proportional to the frequency. To make this current small, one should make L_m large. This can be achieved by

- Using a core material with high permeability
- Increasing the core area
- Increasing the number of turns
- Virtually increasing the core permeability by a feedback [11]

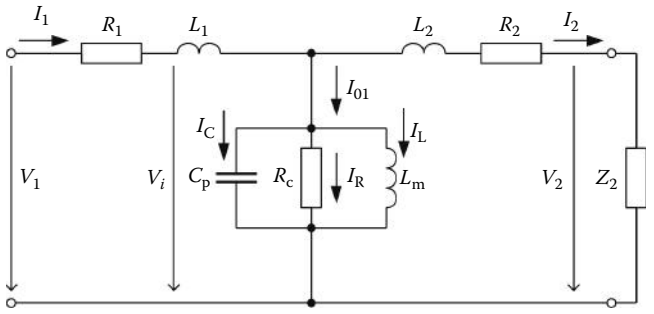


FIGURE 16.2 Equivalent circuit of the CT. R_1, R_2, L_1, and L_2 represent resistances and leakage inductances of primary and secondary windings; R_c is the resistance representing losses in the ferromagnetic core; L_m is the magnetizing inductance; C_p represents the parasitic capacitances of the winding; Z_2 is the burden; $I_L + I_R$ is the magnetization current. (From Ripka, P., *Meas. Sci. Technol.*, 21(11), 1, 2010.)

The magnetizing inductance L_m, together with secondary resistance R_2 and burden resistance R, forms a high-pass filter circuit. Simple approach to improve transformer accuracy at low frequencies is therefore to eliminate the burden resistance R by using a current-to-voltage converter [12].

Some CTs have almost a constant ratio error for a wide range of measured currents. In this case, the error can be compensated for by adding extra turns to the secondary winding. This is called turn error compensation. Transformers with such compensation show anomalous behavior—their error increases when the burden decreases. Magnetization of the CT core (e.g., due to a lightning strike) may also lead to increased errors when the measured current is small so that the CT core is not demagnetized [13].

At higher frequencies, parasitic capacitances between the turns and layers of the winding are the dominant source of error. Here, increasing the number of turns increases the error. For precise fast-CTs, amorphous cobalt-based cores or ferrite cores are used, which give high permeability at MHz frequencies. Such transformers are used to observe short beam pulses in particle accelerators. Bergoz manufactures transformers with cutoff frequencies up to 2 GHz.

CTs are often used in electronic watt and energy meters. For this application, it is very important to make them resistant against saturation—either from a strong permanent magnet or due to a dc component in the measured current [14]. Possible methods for avoiding such saturation are as follows:

1. Use flat-loop magnetic material, which has a very high saturation field H.
2. Use composite cores, consisting of a high-permeability ring for precision and a low-permeability ring for saturation immunity.
3. Detect saturation by other means.
4. Use a Rogowski coil instead of a CT, which has no magnetic core.

The composite high-permeability core gives low angular and amplitude error for most applications, but this is not guaranteed for all types of loads. Flat-loop materials do a better job, but these materials are expensive. It was recently shown that some cheap low-permeability rings can also be used for this application: they have a relatively large phase shift, but it is constant over the wide range of measured currents and can therefore be compensated [14].

Fe-based nanocrystalline alloys are promising materials for very precise small-size CTs.

16.5.2 Rogowski Coil

The Rogowski coil for measuring current is an air coil wound around the measured current conductor (Figure 16.3). The basic operating principle is given by the mutual inductance M between the primary (single turn) and the secondary (many turns). The output voltage is proportional to the derivative of the current:

$$V(t) = M\left(\frac{dI}{dt}\right) \tag{16.2}$$

The coil should be precisely manufactured with constant winding density and diameter. Ideally, a homogeneous coil has excellent geometrical selectivity, that is, it is insensitive to external fields and to the position of the measured conductor, as it follows Ampere's law in the open air:

$$\oint_C \vec{B} \cdot d\vec{l} = \mu_0 i_C \tag{16.3}$$

where the integration path C is the central line of the coil (usually a circle).

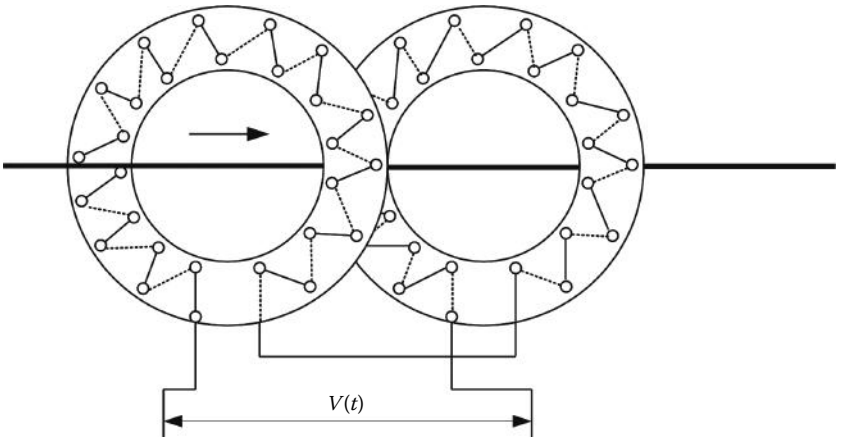


FIGURE 16.3 A Rogowski coil in PCB technology. Each layer is wound in the opposite direction to compensate for the loop created by turn advancement. (From Kojovic, L., High-precision Rogowski coils for improved relay protection, control and measurements, *The Line*, 7–8, 2002, http://www.cooperpower.com/Library/pdf/02049.pdf.)

In order to obtain the ac current waveform, a Rogowski coil is used together with an integrator. Single-chip digital integrators have been developed to process the signal of Rogowski coils (also known as *dI*/*dt* sensors) for energy meters. The use of an AD 7759 signal processor with built-in sigma–delta A/D converters gives 0.1% error from the measured value in the 1000:1 dynamic range [15].

The Rogowski coil contains no ferromagnetic material, and thus it has excellent linearity and an extremely large dynamic range. Users often rely on their linearity and use them to measure currents that are much higher than the currents used for calibration. Care should be taken with the geometry of the connecting bus bars and the return conductor to reduce the effect of magnetic coupling.

The Rogowski coil can be very simply temperature compensated [16].

Stationary Rogowski coils are used to measure ac or transient currents or changes in dc currents. For long-term measurements, the limiting factor is the offset drift of the integrator. It is also possible to measure dc current with an openable Rogowski coil: the output voltage is integrated while the coil is closed around the measured conductor.

Flexible Rogowski coils can be easily wound on top of a plastic cable with thick insulation, such as a coaxial cable with removed shielding. The inside conductor is used as a return loop, which compensates the perpendicular virtual loop created by the winding advance [17].

Rogowski coils are useful for measuring transient currents. The coil self-inductance and the parasitic capacitance form a resonant circuit. In order to increase the resonant frequency of this circuit, a lower number of turns should be used, which decreases the sensitivity. 1–3 MHz resonant frequency is typical for flexible coils, while multiturn solid coils may have resonance as low as 50 kHz.

Printed circuit board (PCB) technology was used to produce Rogowski coils with a more precise geometry and improved temperature stability. To compensate the virtual perpendicular loop, two PCB coils with opposite winding directions are connected in series. In order to achieve perfect compensation, it is important to design these two coils with the same diameter, as in Figure 16.3.

An improved twin-loop PCB Rogowski coil with a transfer ratio of 4 V/400 kA was used to measure the plasma current in the tokamak [19]. In this design, both coils are made in the same place, which means that one of them has a smaller diameter and thus the sensitivity to axial fields is not perfectly compensated.

High-frequency inductive current sensors are similar to Rogowski coils, but they use only a single turn as a secondary coil [20]. These sensors do not suppress external fields.

16.5.3 DC Current Transformers and Current Comparators

Classical CTs naturally measure only ac currents. DC CTs and dc current comparators use the fluxgate effect in a core that is periodically saturated by the excitation field. If no dc current is present, the core flux and also the voltage induced into the detection winding contain only odd harmonics. DC current shifts the magnetization characteristics, and the flux waveform is no longer symmetrical. This induces even harmonics proportional to the measured dc current. Usually phase-sensitive detectors are used to extract the second harmonics from the induced voltage. Dual core minimizes feedthrough of the excitation voltage into the measured circuit [21].

The accuracy of a typical commercial 40 A DC CT is 0.5%, linearity 0.1%, and current temperature drift <30 mA (–25 °C to 70 °C). Because of their low offset drift, fluxgate-based "DC CTs" are superior to current sensors that have a Hall sensor in the air gap. A disadvantage is the large power consumption of these devices. Top sensors of this type achieved linearity error of 10 ppm of FS, zero temperature drift of ±5 μA/°C, and resolution ±0.3 ppm of FS ±0.4 μA. Their frequency band is dc up to 100 kHz.

16.5.4 Hall Effect Current Sensors

Hall effect sensors are also known as Hall sensors. Many dc current sensors use a Hall element mounted in the air gap of a magnetic core (Figure 16.4). The yoke has two important effects: (1) increasing the sensitivity and (2) increasing the geometrical selectivity, that is, shielding the external fields and decreasing the influence of the position of the measured current.

Even when using a magnetic yoke, Hall current sensors are sensitive to external magnetic fields and nearby currents and also to the position of the measured conductor, due to the nonhomogeneity associated with the air gap.

A serious dc offset can be caused by the remanence of the magnetic core—only a few Hall current meters have an ac demagnetization circuit to erase perming after the sensor has been exposed to a large dc current or external field.

In general, the yoke material should have large saturation induction B_s and low coercivity H_c. A widely used material is a cheap grain-oriented FeSi, and for precise applications 50% Ni–Fe alloy is used. By proper annealing, the coercivity of 50% Ni–Fe can be decreased below 2 A/m. The nonhomogeneity of magnetic materials has an influence on the linearity of open-loop current sensors. By increasing the air gap, the influence of this nonlinearity is decreased, as is the influence of the remanence. The hysteresis error is approximately 0.2%–0.5% for a 1 mm air gap and different

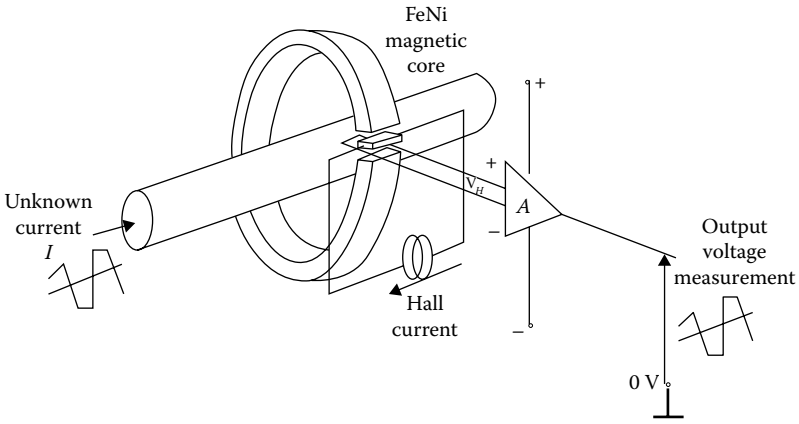


FIGURE 16.4 A Hall dc current sensor in open-loop configuration. (From *J. Magnet. Magnet. Mater.,* 304(2), Waeckerlé, T., Fraisse, H., Furnemont, Q., and Bloch, F., Upgrade Fe-50%Ni alloys for open-loop dc current sensor: Design and alloy-potential characteristics, e850, Copyright 2006. With permission from Elsevier.)

grades of 50% Ni–Fe [22]. However, a larger air gap leads to unwanted sensitivity to the position of the measured conductor within the core and also reduces the suppression of external currents and fields.

The sensor linearity can be increased by using the feedback principle. Feedback-compensated devices cancel the field in the yoke by using multiturn compensation winding. These sensors can achieve 0.02% error and a temperature coefficient of sensitivity of only 50 ppm/K. However, the main weak point is still the limited zero stability due to the Hall sensor offset: the typical offset drift of a 50 A sensor is 600 mA in the (0 °C–70 °C) range. This parameter is 20 times worse than that of fluxgate-type current sensor modules.

Some Hall current sensors use *field concentrators*, which increase the measured field but do not completely surround the measured current. In this case the danger of saturation is much lower, but the position of the measured conductor should be fixed, and suppression of the external fields and currents should be achieved by other means. A low-cost current sensor based on a highly sensitive Hall sensor with simple integrated flux concentrators is described in [23] (Figure 16.5). The field concentrators transform a lateral field locally into the vertical direction (Figure 16.6). The sensor is manufactured by

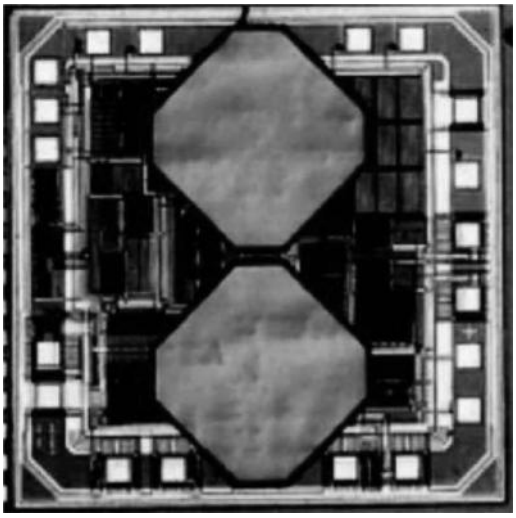


FIGURE 16.5 Internal structure of the CMOS integrated Hall magnetic sensor with twin integrated magnetic concentrators. (Courtesy of Sentron AG, Zug, Switzerland.)

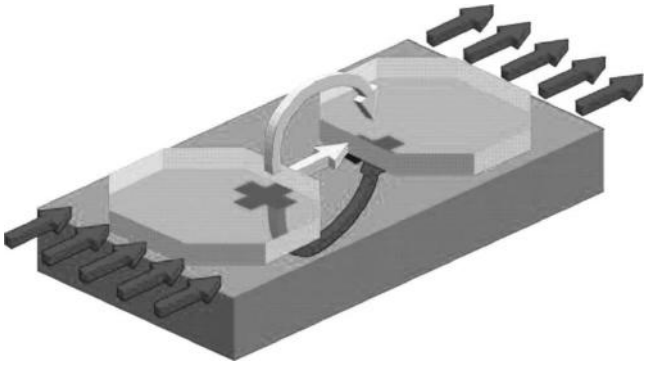


FIGURE 16.6 Field concentrators deflect the flux into the vertical direction. The two small crosses under the concentrators are Hall sensors. (Courtesy of Sentron AG, Zug, Switzerland.)

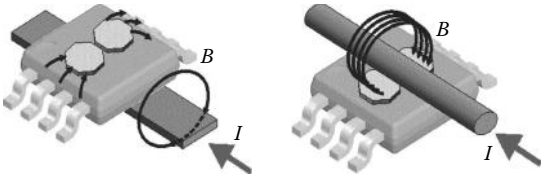


FIGURE 16.7 Current-sensing application of the CMOS integrated Hall magnetic sensor for a PCB conductor and a free-standing conductor. (Courtesy of Sentron AG, Zug, Switzerland.)

Sentron (Melexis group) with 1% accuracy in the ±12 A range. It can be used to measure the currents in both PCB conductors and free-standing conductors (Figure 16.7).

A *yokeless current transducer* with six Hall probes around the rectangular bus bar achieves 0.5% linearity and 0.2% temperature stability in the 100 kA range. Close conductors are typically suppressed by a factor of 100. An integrated current sensor based on two Hall sensors placed on both sides of the current strip was developed by Frick et al. [24]. The differentially connected Hall sensors reject the common-mode interference from the external magnetic fields. Proper autobalancing is achieved using built-in calibration coils. The achieved accuracy was 0.5% in the 5 A range with 1.5 kHz bandwidth.

Some current sensors use a combination of a Hall sensor (for the dc and low-frequency component) and a CT (for the high-frequency component), using a gapped ferrite core. A frequency range of 30 MHz was achieved for 40 A range [25].

The background calibration using integrated coils together with the spinning current technique can reduce the temperature sensitivity drift of the Hall sensor to less than 50 ppm/°C, while the bandwidth can be still as high as 500 kHz. Sensitivity variations due to mechanical stresses and ageing are also compensated [26]. This technique can significantly improve the performance of current sensors, even if feedback compensation is not possible.

16.5.5 AMR Current Sensors

These sensors employ anisotropic magnetoresistance effect. They have higher resolution than Hall sensors, but their sensing direction is in the chip plane, so they cannot be built into the narrow air gap in the magnetic yoke. Yokes with a large air gap suffer from field leakage sensitivity to external currents and magnetic fields. Thus, most current sensors based on AMR magnetoresistors are yokeless. The popular configuration is again a bridge measuring the magnetic gradient from the current, since it is resistant against homogeneous external magnetic fields (e.g., from distant currents). The bridge also suppresses the temperature variation of the electric resistivity of the magnetoresistive material (Figure 16.8). This type of sensor usually has the current conductor integrated with the sensors in a single device to ensure stable geometry. The bridge has all the barber poles in the same direction, which means it is made insensitive to a homogeneous external field but sensitive to measured current through the bus bar. The measured current can be compensated by a feedback current through a compensation conductor. A typical application is a galvanically isolated current sensing in a pulse width modulation (PWM) regulated brushless motor. These sensors are manufactured by Sensitec (also under the F.W. Bell label) with ranges from 5 to 220 A. The achieved linearity is 0.1%, temperature coefficient of sensitivity is 100 ppm/K, and the offset drift in the (−45 °C to +85 °C) range is 1.4% FS.

Similar sensors were developed with a giant magnetoresistance (GMR) detector (which employs GMR) and other types of magnetoresistors such as TMR, but they did not achieved the precision of AMR current sensors. The main obstacle is even characteristics of these sensors, which requires bias field and also hysteresis and nonlinearity of these sensors.

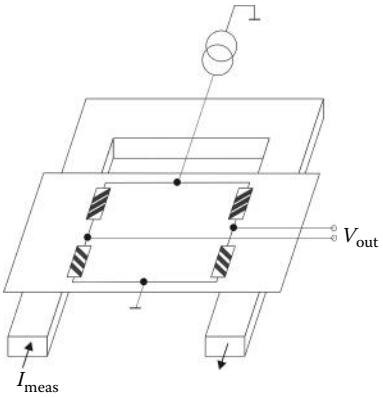


FIGURE 16.8 An AMR current sensor. (From Ripka, P., *Meas. Sci. Technol.*, 21(11), 1, 2010.)

16.5.6 Magneto-Optical Current Sensors (Including Fiber-Optic)

Optical current sensors have several advantages that are very attractive for power distribution applications:

1. Effective isolation from high potentials
2. Immunity against electromagnetic interferences
3. High dynamic range, no saturation effects
4. High bandwidth
5. Compact and lightweight design

These features offer a significant cost reduction in comparison to conventional high-voltage CTs. A good review of these sensors is [27].

Most optical current sensors are based on the Faraday effect—either in bulk material or in an optical fiber. The polarization plane of a linearly polarized light that travels through the magneto-optical material is rotated by angle α, which depends on magnetic field along the optical path. If the optical path is closed around the current I, α depends linearly on I and the sensor is insensitive to other currents.

16.5.6.1 Bulk Magneto-Optical Sensors

Magneto-optical point sensors (or unlinked sensors) use a piece of glass or a crystal rod placed in the neighborhood of the electric conductor [28]. The sensor is usually interrogated by optical fibers. These devices are robust, cheap, and sensitive. They belong to the class of "extrinsic fiber sensors," that is, sensors that use optical fibers for transmission, not for sensing. These sensors are employed in the first generation of the ABB magneto-optic current transducer (MOCT), which serves since 1995 in industrial applications. The ABB sensor achieves an accuracy class of 0.2 A in the 3000 A range.

The triangular prism bulk magneto-optical sensor has a closed sensing optical path around the measured conductor. This provides independence of the sensor output from the position of the measured conductor within the closed path and also resistance to external conductors and external homogeneous magnetic fields.

16.5.6.2 All-Fiber Sensors (or Intrinsic Fiber Sensors)

In wound fiber devices, too, the magneto-optical material encloses the electric conductor, and thus these sensors are not sensitive to external currents and magnetic fields.

Optical fiber is made from materials that have much lower sensitivity than the magneto-optic crystals used for bulk sensors, but their sensitivity can be increased by using a higher number of turns of the fiber wound around the measured conductor. Fiber-optic sensors are simple devices; they suffer from the spurious birefringence induced in bent fibers [29].

Commercially available all-fiber current sensors are manufactured by ABB since 2005. These devices also use two circularly polarized light waves travelling in the same direction. The phase difference caused by the magnetic field is measured by the interferometer. This configuration effectively suppresses the bending-induced linear birefringence in the fiber. The measuring range is 600 kA, corresponding to 360° phase difference, while the resolution is 0.25 A. The sensitivity can be increased by using multiple turns of the optical fiber around the measured conductor. Great care was taken to insulate the measuring fiber from mechanical stress and vibrations, as this is a major source of temperature sensitivity drift.

Similar sensors were developed for measuring large dc currents: the achieved accuracy is 0.1% for currents up to 600 kA [30].

16.5.6.3 Other Types of Extrinsic Fiber-Optic Current Sensors

These sensors utilize optical fibers for galvanically isolated connection to a nonoptical current sensor such as Rogowski coil [31]. The required electric supply energy can be converted from the incoming light.

16.6 Other Types of Current Sensors

16.6.1 Superconducting Current Sensors

Superconducting quantum interference detector (SQUID) magnetoencephalographs measure the fT fields produced by neural currents in the brain. SQUIDs can also be used to measure the small currents that pass through a coil inductively coupled to the SQUID loop. The SQUID should be shielded against external magnetic fields or made as a gradiometer with a very short baseline. SQUIDs normally measure only flux changes, but an array of SQUIDs can be used to build an absolute current sensor with 1 nA resolution. As the SQUID is a nonlinear device (the voltage-flux characteristic is similar to a sine wave), the flux should be compensated by a flux-locked loop. This is a technical problem at higher frequencies. A noise level of 7.4 pA/√Hz up to 10 MHz was achieved using a two-stage SQUID instead of an external op-amp [32].

16.6.2 Magnetometric Location and Measurement Hidden Currents

The field from a long straight conductor decreases with $1/r$, assuming that the return conductor is at a large distance. If this is not the case, the magnetic field is lower, and it should be calculated from the actual geometry. The other extreme is a small current loop, which creates a field decreasing with $1/r^3$ distance. Underground electric conductors can be located and their current can be remotely monitored by measuring their magnetic field at several points. In the case that the cable contains both forward and return currents, it can be detected only from a small distance, and the current value cannot be measured.

Sensitive magnetic sensors such as fluxgates should be used. It is an advantage if the sensor is vectorial.

16.7 Current Clamps

Current clamps usually have an openable magnetic yoke, so they can be mounted without interrupting the measured conductor. The role of the yoke is again to concentrate the field lines so that the sensor reading is not dependent on the actual position of the clamped conductor and the device is insensitive to unclamped conductors.

AC clamps are usually based on CT principles. The measured conductor forms the primary winding, and the multiturn secondary winding is terminated by a small frequency-independent resistor.

DC clamps are usually based on a Hall sensor in the air gap. These devices may have 10 mA resolution, but the maximum achieved accuracy is typically 30 mA, even if they are of the compensated type. Their main disadvantage is unwanted sensitivity to external fields, due to the air gap in the magnetic circuit that is necessary for the Hall sensor. Even the change of position with respect to the Earth's field causes significant error, and the offset should be manually nulled.

Not only traditional clamps but also flexible sensors can form an openable magnetic circuit. The most popular are flexible Rogowski coils as discussed earlier.

16.8 Discussion and Conclusions

Shunt resistors are still widely used, but they are very large and power consuming for high currents.

Traditional dc and ac contactless current sensors are available for ranges from mA to kA with precision from 3% (uncompensated Hall current sensors) to 0.1% (compensated Hall devices and magnetic amplifiers). Higher precision is easily achievable with current comparators. Current sensors based on the AMR effect are very promising. If the measured current does not have a dc component, CTs and *di*/*dt* sensors (Rogowski coils) are the best choice.

For kA and larger currents, Faraday effect current sensors are attractive, as they are smaller and lighter than traditional sensors and they also have significantly lower power consumption. For smaller currents, optical sensors are not competitive with other sensor types: they are expensive, and they would require many turns of the fiber cable to achieve the required sensitivity.

References

1. Ripka, P., Electric current sensors: A review, *Measurement Science and Technology*, 21(11), 1–23, 2010. http://dx.doi.org/10.1088/0957-0233/21/11/112001
2. Iwansson, K., Sinapius, G., and Hoornaert, W., *Measuring Current, Voltage and Power*, Elsevier, Amsterdam, the Netherlands, 1999.
3. Ziegler, S., Woodward, R.C., Iu, H.H., and Borle, L.J., Current sensing techniques: A review, *IEEE Sensors Journal*, 9(4), 354–376, 2009.
4. Ripka, P. (ed.), *Magnetic Sensors*, Artech, Boston, MA, 2001.
5. Ripka, P., Kaspar, P., and Saneistr, J., Geometrical selectivity of current sensors, *Przeglad Elektrotechniczny*, 88(5a), 38–39, 2012.
6. Di Rienzo, L., Bazzocchi, R., and Manara, A., Circular arrays of magnetic sensors for current measurement, *IEEE Transactions on Instrumentation and Measurement*, 50(5), 1093–1096, 2001.
7. Olson, E.R. and Lorenz, R.D., Effective use of miniature multipoint field-based current sensors without magnetic cores, *IEEE Transactions on Industry Applications*, 46, 901–909, 2010.
8. Filipski, P.S., Boecker, M., and Garcocz, M., 20-A to 100-A AC-DC coaxial current shunts for 100-kHz frequency range, *IEEE Transactions on Instrumentation and Measurement*, 57, 1637–1641, 2008.
9. Lorentz, V.R.H. et al., Lossless average inductor current sensor for CMOS integrated DC–DC converters operating at high frequencies, *Analog Integrated Circuits and Signal Processing*, 62, 333–344, 2010.
10. Keithley, Inc., Low current measurements, AN 100, 2007, www.keithley.com, accessed August 3, 2013.
11. McNeill, N., Dymond, H., and Mellor, P.H., High-fidelity low-cost electronic current sensor for utility power metering, *IEEE Transactions on Power Delivery*, 26, 2309–2317, 2011.

12. Kondrath, N. and Kazimierczuk, M.K., Bandwidth of current transformers, *IEEE Transactions on Instrumentation and Measurement*, 58(6), 2008–2016, 2009.
13. Draxler, K. and Styblikova, R., Influence of instrument transformers on quality of electrical power and energy measurement, *Proceedings of the IEEE International Symposium on Industrial Electronics*, Vols. 1–8, p. 1317, 2007.
14. Mlejnek, P. and Kaspar, P., Drawback of using double core current transformers in static watt-hour meters, *Sensor Letter*, 7, 394–396, 2009.
15. Koon, W., Current sensing for energy metering, 2009, http://www.analog.com, accessed August 3, 2013.
16. Suomalainen, E. and Hallstrom, J.K., Onsite calibration of a current transformer using a Rogowski coil, *IEEE Transactions on Instrumentation and Measurement*, 58(4), 1054–1058, 2009.
17. Abdi-Jalebi, E. and McMahon, R., High-performance low-cost rogowski transducers and accompanying circuitry, *IEEE Transactions on Instrumentation and Measurement*, 56(3), 753–759, 2007.
18. Kojovic, L., High-precision Rogowski coils for improved relay protection, control and measurements, *The Line*, 7–8, 2002, http://www.cooperindustries.com/, accessed August 3, 2013.
19. Qing, C., Hong-bin, L., Ben-xiong, H., and Qiao-qi, D., Rogowski sensor for plasma current measurement in J-TEXT, *IEEE Sensors Journal*, 9(3), 293–296, 2009.
20. Rojas-Moreno, M.V. et al., Study of an inductive sensor for measuring high frequency current pulses, *IEEE Transactions on Instrumentation and Measurement*, 60, 1893–1900, 2011.
21. Xiaoguang, Y., Bo, Z., Youhua, W. et al., The optimization of dual-core closed-loop fluxgate technology in precision current sensor, *Journal of Applied Physics*, 111, 07E722, 2012.
22. Waeckerlé, T., Fraisse, H., Furnemont, Q., and Bloch, F., Upgrade Fe-50%Ni alloys for open-loop dc current sensor: Design and alloy-potential characteristics, *Journal of Magnetism and Magnetic Materials*, 304(2), p. e850, 2006.
23. Popovic, R., Drljaca, P., and Kejik, P., CMOS magnetic sensors with integrated ferromagnetic parts, *Sensors and Actuators A, Physical*, 129(1–2), 94–99, 2006.
24. Frick, V., Poure, P., Hebrard, L., Anstotz, F., and Braun, F., Electromagnetically compatible CMOS auto-balanced current sensor for highly integrated power control System-On-Chip, *International Journal of Electronics*, 94(1), 9–22, 2007.
25. Dalessandro, L., Karrer, N., and Kolar, J.W., High-performance planar isolated current sensor for power electronics applications, *IEEE Transactions on Power Electronics*, 22(5), 1682–1692, 2007.
26. Pastre, M., Kayal, M., and Blanchard, H., A Hall sensor analog front end for current measurement with continuous gain calibration, *IEEE Sensors Journal*, 7(5–6), 860–867, 2007.
27. Silva, R.M. et al., Optical current sensors for high power systems: A review, *Applied Sciences*, 2(3), 602–628, 2012; doi: 10.3390/app2030602
28. Fisher, N.E. and Jackson, D.A., Vibration immunity and Ampere's circuital law for a near perfect triangular Faraday current sensor, *Measurement Science and Technology*, 7(8), 1099–1102, 1996.
29. Perciante, C.D. and Ferrari, J.A., Magnetic crosstalk minimization in optical current sensors, *IEEE Transactions on Instrumentation and Measurement*, 57(10), 2304–2308, 2008.
30. Bohnert, K., Brandle, H., Brunzel, M.G., Gabus, P., and Guggenbach, P., Highly accurate fiber-optic dc current sensor for the electrowinning industry, *IEEE Transactions on Industry Applications*, 43(1), 180–187, 2007.
31. de Nazare, F.V.B. and Werneck, M.M., Hybrid optoelectronic sensor for current and temperature monitoring in overhead transmission lines, *IEEE Sensors Journal*, 12, 1193–1194, 2012.
32. Drung, D., Beyer, J., Peters, M., Storm, J.-H., and Schurig, T., Novel SQUID current sensors with high linearity at high frequencies, *IEEE Transactions on Applied Superconductivity*, 19(3), 772–777, 2009.

List of Manufacturers and Suppliers

Allegro http://www.allegromicro.com
Amploc http://www.ampsolution.com
ABB http://www.abb.com
Bergoz http://www.bergoz.com
CR Magnetics http://www.crmagnetics.com/
LEM http://www.lem.com
Magnelab http://www.magnelab.com/
Melexis http://www.melexis.com/
Sensitec http://www.sensitec.com
Honeywell http://www.sensing.honeywell.com

17

Nonlinear System Modeling with Invariance to Fourier Transform for Fault Diagnosis: Application to Power Transformers

Gerasimos Rigatos
Harper Adams University College

Pierluigi Siano
University of Salerno

17.1 Introduction

A neural network with basis functions that remain invariant under the Fourier transform is proposed for fault diagnosis of nonlinear systems. The considered neural network is of the feedforward type and uses Gauss–Hermite polynomial basis functions. This neural model follows the concept of wavelet networks [1–3] and employs basis functions that are localized both in space and frequency, thus allowing better approximation of the multifrequency characteristics of monitored nonlinear system [4–8]. Gauss–Hermite basis functions have some interesting properties [9,10]: (1) They remain almost unchanged by the Fourier transform and satisfy an orthogonality property, which

means that the weights of the associated neural network demonstrate the energy that is distributed to the various eigenmodes of the nonlinear system's dynamics, and (2) unlike wavelet basis functions, the Gauss–Hermite basis functions have a clear physical meaning since they represent the solutions of differential equations describing stochastic oscillators (see [11]) and each neuron can be regarded as the frequency filter of the respective eigenfrequency.

The concept of the proposed fault detection and isolation (FDI) technique is as follows: The neural network with Gauss–Hermite polynomial activation functions is used for approximating the nonlinear system's dynamics out of a set of input–output data. Thus, the output of the neural network provides a series expansion that takes the form of a weighted sum of Gauss–Hermite basis functions. Knowing that the Gauss–Hermite basis functions remain unchanged under the Fourier transform, subjected only to a change of scale, one has that the considered neural network provides the spectral analysis of the output of the monitored system. Actually, the squares of the weights of the output layer of the neural network denote the distribution of the energy of the monitored signal from the nonlinear system into the associated spectral components. Moreover, since Gauss–Hermite basis functions satisfy an orthogonality property, the sum of the square of the output layer weights of these neural networks stands for a measure of the energy contained in the output of the monitored system. By observing changes in the amplitude of the aforementioned spectral components of the nonlinear system's dynamics, one can have also an indication about malfunctioning of the monitored system and can detect the existence of failures. Moreover, since specific faults are associated with amplitude changes of specific spectral components of the system, fault isolation can be also performed.

The proposed FDI method can be applied to several electromechanical systems, for example, vehicles, electric motors, and power generators. In this chapter, as a case study, the problem of fault diagnosis of electric power transformers is considered. Significant information about the thermal condition of oil-immersed power transformer and about their ageing and failure risks can be obtained through monitoring of the transformer's hot-spot temperature (HST). The HST can be measured with the placement of sensors at a specific point of the mineral oil volume that serves as both the insulating and cooling material for the transformer's functioning. A deviation of HST from the anticipated temperature profile is probably an indication of ageing of the transformer or in some cases of prefailure situations. Analytical and numerical (neural/fuzzy) models for HST have been developed [12–15]. These models associate the HST to inputs such as the ambient temperature, the top-oil temperature (TOT) rise over the ambient temperature, and the maximum winding hot-spot rise over the TOT. By modeling the HST variations with the use of neural network that contains Gauss–Hermite polynomial basis functions, one obtains (1) a numerical model that associates the HST with other parameters of the power transformer such as ambient temperature, TOT, and load current and (2) indications about the spectral characteristics of the HST signal and the distribution of its energy content to various spectral components associated with the basis functions. By recording changes in the amplitude of these spectral components, one can detect the existence of failures in the power transformer and can identify which are the components of the transformer that are responsible for malfunctioning.

The structure of the chapter is as follows: In Section 17.2 FNNs are analyzed and their use in nonlinear systems modeling is explained. In Section 17.3 FNNs with Gauss–Hermite activation functions are introduced and their distinctive properties are explained such as orthogonality of the basis functions and invariance to Fourier transform. In Section 17.4 basic principles of signals spectral analysis are presented and the use of Fourier transform in calculating a signal's energy content and power spectral density is explained. In Section 17.5 modeling issues for condition monitoring of electric power transformers are analyzed. In Section 17.6 the use of neural networks with Gauss–Hermite basis functions in modeling the thermal dynamics of electric power transformers is explained. It is shown how these neural networks enable spectral analysis of the HST signal and how based on this signal's spectral content one can perform FDI for the power transformer. Finally, in Section 17.7 concluding remarks are stated.

17.2 Feedforward Neural Networks for Nonlinear Systems Modeling

The proposed fault diagnosis approach for nonlinear dynamical systems can be substantiated with the use of FNNs. FNNs serve as powerful computational tools in a diversity of applications including function approximation, classification, and pattern recognition. When equipped with procedures for learning from measurement data, they can generate models of unknown systems. FNNs are the most popular neural architectures due to their structural flexibility, good representational capabilities, and availability of a large number of training algorithms.

The idea of function approximation with the use of FNNs comes from generalized Fourier series [16]. It is known that any function $\psi(x)$ in an L^2 space can be expanded, using generalized Fourier series in a given orthonormal basis, that is,

$$\psi(x) = \sum_{k=1}^{\infty} c_k \psi_k(x) \quad \alpha \le x \le \beta \tag{17.1}$$

Truncation of the series yields in the sum

$$S_M(x) = \sum_{k=1}^{M} a_k \psi_k(x) \tag{17.2}$$

If the coefficients a_k are taken to be equal to the generalized Fourier coefficients, that is, when $a_k = c_k = \int_a^b \psi(x)\psi_k(x)dx$, then Equation 17.2 is a mean square optimal approximation of $\psi(x)$.

Unlike generalized Fourier series, in FNN the basis functions are not necessarily orthogonal. The hidden units in an FNN usually have the same activation functions and are often selected as sigmoidal functions or Gaussians. A typical FNN consists of n inputs x_i, $i = 1, 2, \ldots, n$, a hidden layer of m neurons with activation function $h{:}R \to R$, and a single output unit (see Figure 17.1a). The FNN's output is given by

$$\psi(x) = \sum_{j=1}^{n} c_j h\left(\sum_{i=1}^{n} w_{ji} x_i + b_j\right) \tag{17.3}$$

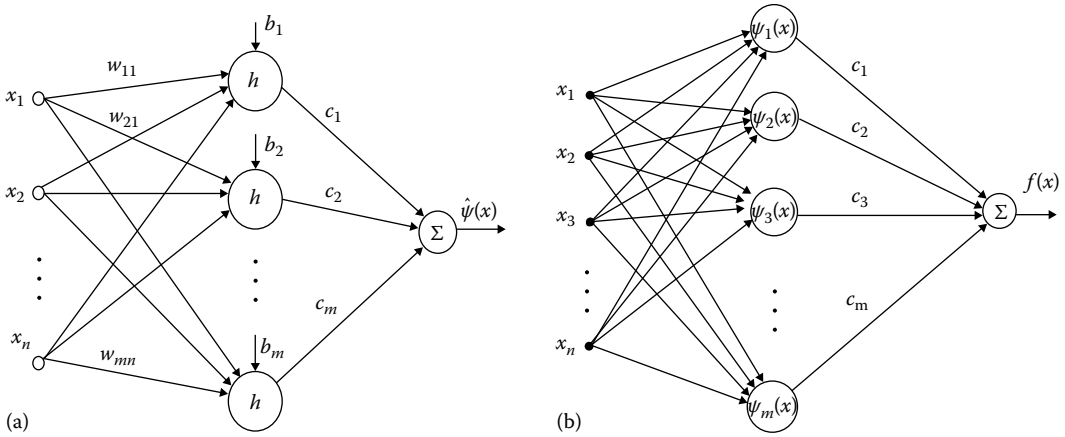


FIGURE 17.1 (a) FNN and (b) neural network with Gauss–Hermite basis functions.

The root mean square error (RMSE) in the approximation of function $\psi(x)$ by the FNN is given by

$$E_{\text{RMS}} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\psi(x^k) - \hat{\psi}(x^k))^2} \tag{17.4}$$

where $x^k = [x_1^k, x_2^k, \ldots, x_n^k]$ is the kth input vector of the neural network. The activation function is usually a sigmoidal function $h = 1/(1 + e^{-x})$, while in the case of radial basis functions networks, it is a Gaussian [17]. Several learning algorithms for neural networks have been studied. The objective of all these algorithms is to find numerical values for the network's weights so as to minimize the mean square error, E_{RMS}, of Equation 17.4. The algorithms are usually based on first- and second-order gradient techniques. These algorithms belong to (1) batch-mode learning, where, to perform parameters update, the outputs of a large training set are accumulated and the mean square error is calculated (backpropagation algorithm, Gauss–Newton method, Levenberg–Marquardt method, etc.) and, (2) pattern-mode learning, in which training examples are run in cycles and the parameters' update is carried out each time a new datum appears (extended Kalman filter algorithm) [18].

Unlike conventional FNN with sigmoidal or Gaussian basis functions, Hermite polynomial-based FNNs remain closer to Fourier series expansions by employing activation functions that satisfy the property of orthogonality [16]. Other basis functions with the property of orthogonality are Legendre, Chebyshev, and Volterra polynomials [9–12].

17.3 Neural Networks Using Gauss–Hermite Activation Functions

17.3.1 Gauss–Hermite Series Expansion

Next, as orthogonal basis functions of the FNN, Gauss–Hermite activation functions are considered. These are the spatial components $X_k(x)$ of the solution of Schrödinger's differential equation and describe a stochastic oscillation:

$$X_k(x) = H_k(x)e^{-x^2/2}, \quad k = 0, 1, 2, \ldots \tag{17.5}$$

where $H_k(x)$ are the Hermite orthogonal functions (Figure 17.2). The Hermite functions $H_k(x)$ are the eigenstates of the quantum harmonic oscillator. The general relation for the Hermite polynomials is

$$H_k(x) = (-1)^k e^{x^2} \frac{d^{(k)}}{dx^{(k)}} e^{-x^2} \tag{17.6}$$

According to Equation 17.6, the first five Hermite polynomials are

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2, \quad H_3(x) = 8x^3 - 12x, \quad H_4(x) = 16x^4 - 48x^2 + 12$$

It is known that Hermite polynomials are orthogonal, that is, it holds

$$\int_{-\infty}^{+\infty} e^{-x^2} H_m(x)H_k(x)dx = \begin{cases} 2^k k!\sqrt{\pi} & \text{if } m = k \\ 0 & \text{if } m \neq k \end{cases} \tag{17.7}$$

Using now Equation 17.7, the following basis functions can be defined [9]:

$$\psi_k(x) = [2^k \pi^{1/2} k!]^{-1/2} H_k(x)e^{-x^2/2} \tag{17.8}$$

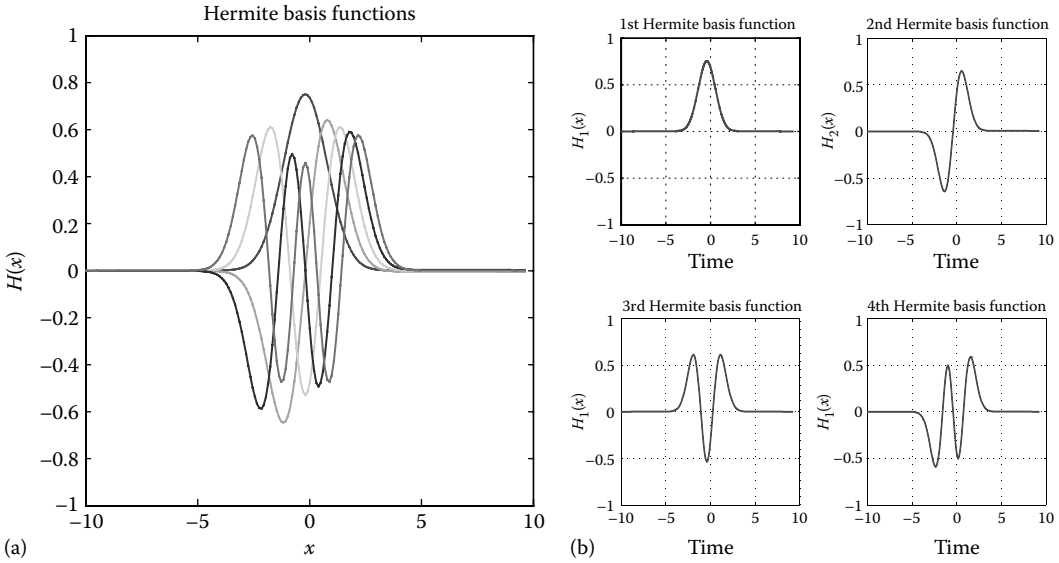


FIGURE 17.2 (a) First five 1D Hermite basis functions and (b) analytical representation of the 1D Hermite basis function.

where $H_k(x)$ is the associated Hermite polynomial. From Equation 17.7, the orthogonality of basis functions of Equation 17.8 can be concluded, which means

$$\int_{-\infty}^{+\infty} \psi_m(x)\,\psi_k(x)dx = \begin{cases} 1 & \text{if } m = k \\ 0 & \text{if } m \neq k \end{cases} \tag{17.9}$$

Moreover, to succeed multiresolution analysis, Gauss–Hermite basis functions of Equation 17.8 are multiplied with the scale coefficient α. Thus, the following basis functions are derived:

$$\beta_k(x, \alpha) = a^{-1/2}\psi_k(a^{-1}x) \tag{17.10}$$

which also satisfy orthogonality condition

$$\int_{-\infty}^{+\infty} \beta_m(x, \alpha)\beta_k(x, \alpha)dx = \begin{cases} 1 & \text{if } m \neq k \\ 0 & \text{if } m = k \end{cases} \tag{17.11}$$

Any function $f(x)$, $x \in R$, can be written as a weighted sum of the previous orthogonal basis functions, that is,

$$f(x) = \sum_{k=0}^{\infty} c_k \beta_k(x, a) \tag{17.12}$$

where coefficients c_k are calculated using the orthogonality condition

$$c_k = \int_{-\infty}^{+\infty} f(x)\beta_k(x, a)dx \tag{17.13}$$

Assuming now that instead of infinite terms in the expansion of Equation 17.12 M terms are maintained, then an approximation of $f(x)$ is succeeded. The expansion of $f(x)$ using Equation 17.12 is a Gauss–Hermite series. Equation 17.12 is a form of Fourier expansion for $f(x)$ and can be considered as the Fourier transform of $f(x)$ subject only to a scale change. Indeed, the Fourier transform of $f(x)$ is given by

$$F(s) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(x)e^{-jsx}dx \Rightarrow f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(s)e^{jsx}ds \tag{17.14}$$

The Fourier transform of the basis function $\psi_k(x)$ of Equation 17.8 satisfies [9]

$$\Psi_k(s) = j^k\psi_k(s) \tag{17.15}$$

while for the basis functions $\beta_k(x, a)$ using scale coefficient α, it holds that

$$B_k(s, a) = j^k\beta_k(s, a^{-1}) \tag{17.16}$$

Therefore, it holds

$$f(x) = \sum_{k=0}^{\infty} c_k\beta_k(x, a) \xrightarrow{F} F(s) = \sum_{k=0}^{\infty} c_k j^k\beta_k(s, a^{-1}) \tag{17.17}$$

which means that the Fourier transform of Equation 17.12 is the same as the initial function, subject only to a change of scale. The structure of an FNN with Hermite basis functions is depicted in Figure 17.1b.

17.3.2 Neural Networks Using 2D Hermite Activation Functions

2D Hermite polynomial-based neural networks can be constructed by taking products of the 1D basis functions $B_k(x, a)$ [9]. Thus, setting $x = [x_1 \quad x_2]^T$, one can define the following basis functions:

$$B_k(x, a) = \frac{1}{a}B_{k_1}(x_1, a)B_{k_2}(x_2, a) \tag{17.18}$$

These 2D basis functions are again orthonormal, that is, it holds

$$\int d^2xB_n(x, a)B_m(x, a) = \delta_{n_1m_1}\delta_{n_2m_2} \tag{17.19}$$

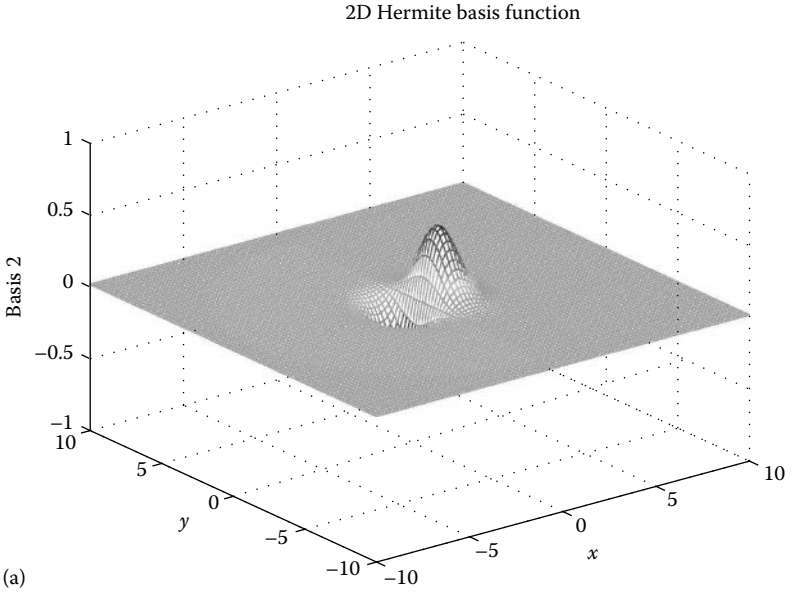
The basis functions $B_k(x)$ are the eigenstates of the 2D harmonic oscillator and form a complete basis for integrable functions of two variables. A 2D function $f(x)$ can thus be written in the series expansion

$$f(x) = \sum_{k_1,k_2}^{\infty} c_kB_k(x, a) \tag{17.20}$$

The choice of an appropriate scale coefficient a and maximum order $k_{\max}$ is of practical interest. The coefficients c_k are given by

$$c_k = \int dx^2 f(x) B_k(x, a) \tag{17.21}$$

Indicative basis functions $B_2(x, a)$, $B_6(x, a)$, $B_9(x, a)$, $B_{11}(x, a)$, $B_{13}(x, a)$, and $B_{15}(x, a)$ of a 2D FNN with Hermite basis functions are depicted in Figures 17.3 through 17.5. Following the same method, N-dimensional Hermite polynomial-based neural networks ($N > 2$) can be constructed. The associated high-dimensional Gauss–Hermite activation functions preserve the properties of orthogonality and invariance to Fourier transform.



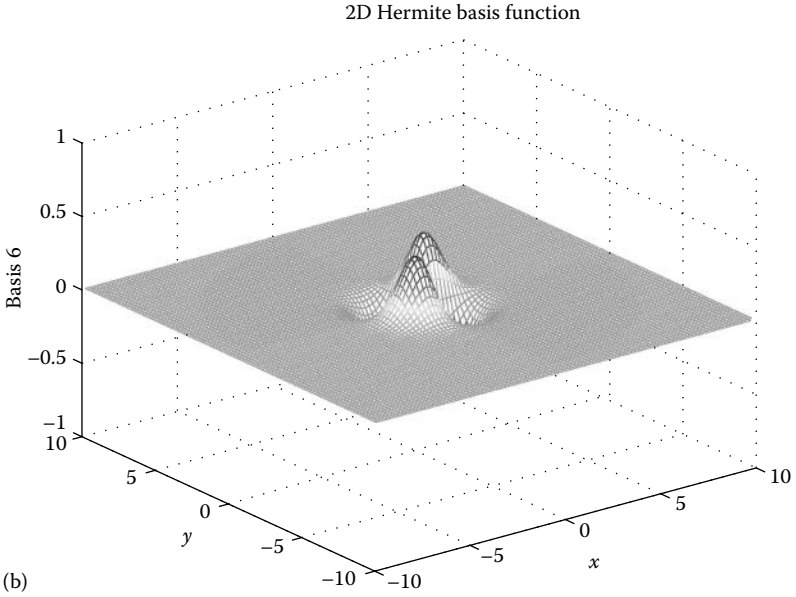


FIGURE 17.3 2D Hermite polynomial activation functions: (a) basis function $B_2(x, a)$ and (b) basis function $B_6(x, a)$.

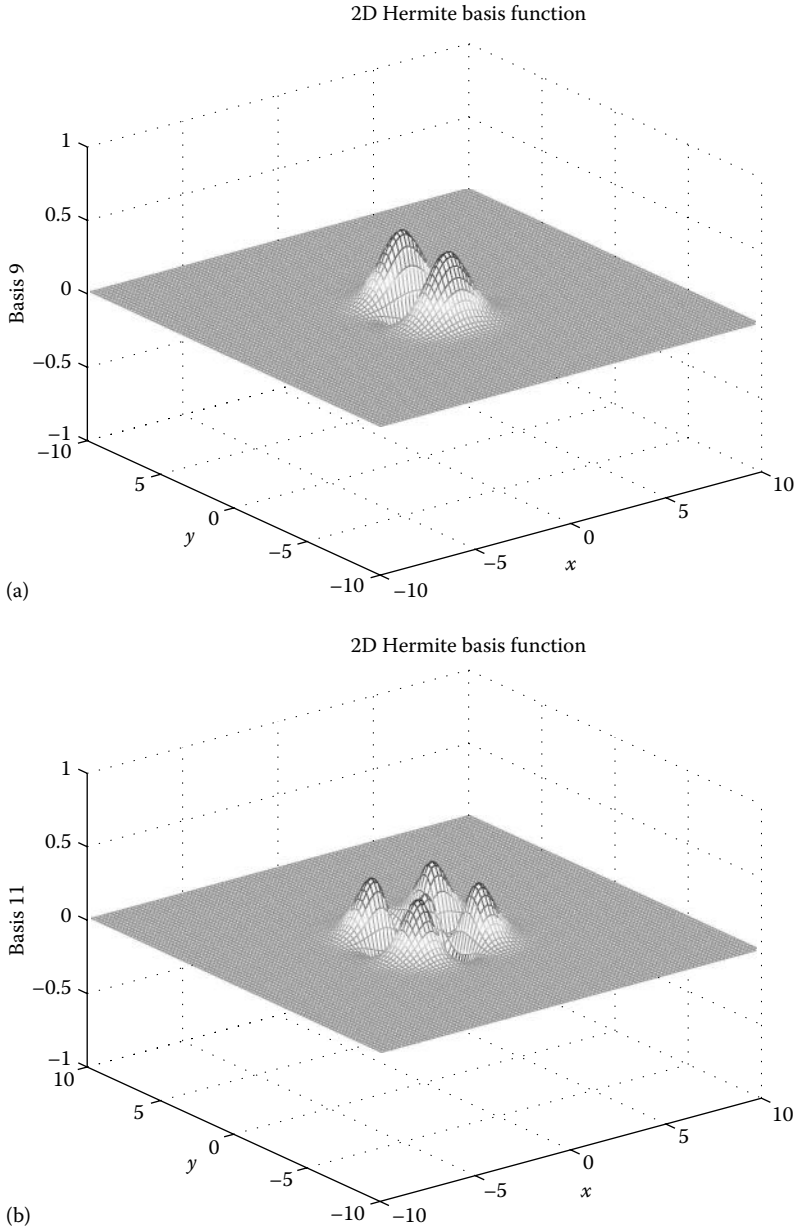


FIGURE 17.4 2D Hermite polynomial activation functions: (a) basis function $B_9(x, a)$ and (b) basis function $B_{11}(x, a)$.

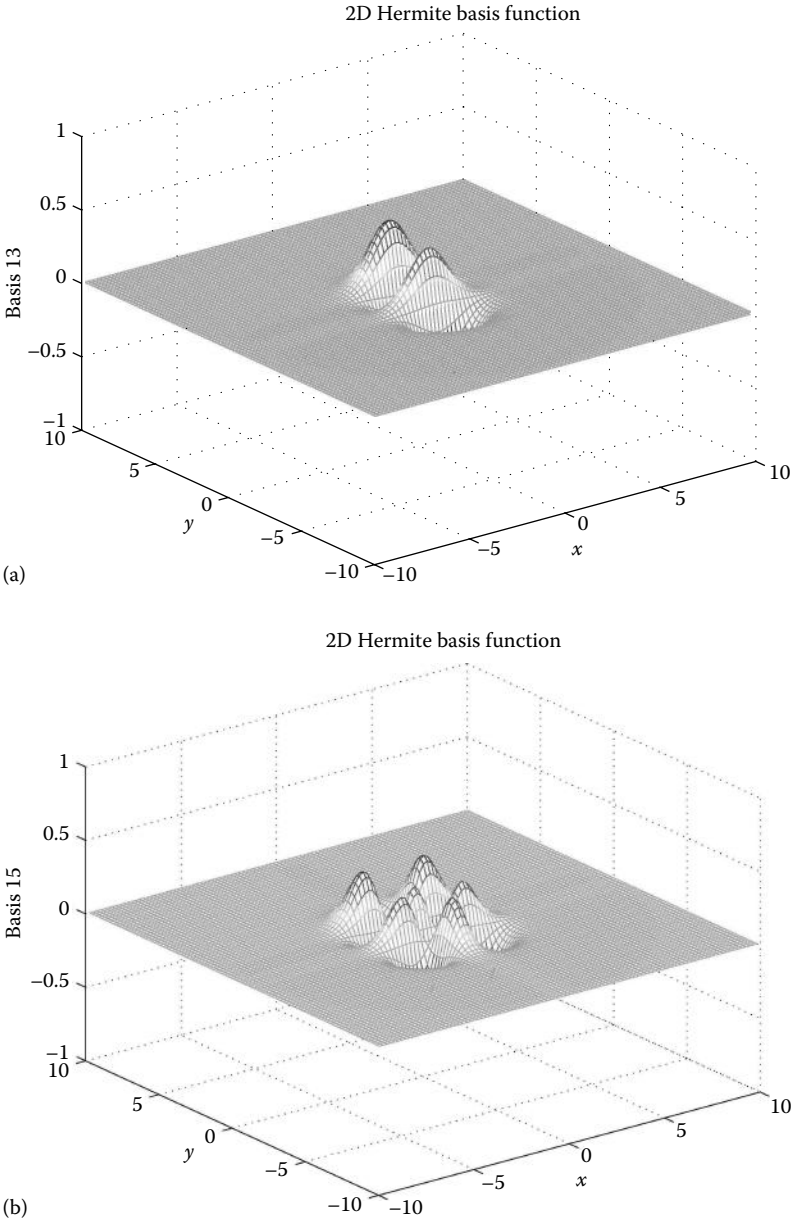


FIGURE 17.5 2D Hermite polynomial activation functions: (a) basis function $B_{13}(x, a)$ and (b) basis function $B_{15}(x, a)$.

17.4 Signal Power Spectrum and the Fourier Transform

17.4.1 Parseval's Theorem

To find the spectral density of a signal $\psi(t)$ with the use of its Fourier transform, the following definition of the signal's energy is used:

$$E_\psi = \int_{-\infty}^{+\infty} (\psi(t))^2 dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \psi(t) \left(\int_{-\infty}^{+\infty} \Psi(j\omega) e^{j\omega t} d\omega \right) dt, \quad \text{i.e.}$$
$$E_\psi = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Psi(j\omega) \left(\int_{-\infty}^{+\infty} \psi(t) e^{j\omega t} d\omega \right) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Psi(j\omega)\Psi(-j\omega) d\omega \tag{17.22}$$

Taking that $\psi(t)$ is a real signal, it holds that $\Psi(-j\omega) = \Psi^*(j\omega)$, which is the signal's complex conjugate. Using this in Equation 17.22, one obtains

$$E_\psi = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Psi(j\omega)\Psi^*(j\omega) d\omega \quad \text{or}$$
$$E_\psi = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\Psi(j\omega)|^2 \, d\omega \tag{17.23}$$

This means that the energy of the signal is equal to $1/2\pi$ times the integral over frequency of the square of the magnitude of the signal's Fourier transform. This is *Parseval's theorem*. The integrated term $|\Psi(j\omega)|^2$ is the energy density per unit of frequency and has units of magnitude square per hertz.

17.4.2 Power Spectrum of Infinite Energy Signals

The concept of Fourier transform is also applied to infinite time extent and energy signals. The ordinary Fourier transform is defined for signals that are absolutely integrable, that is, $\int_{-T}^{+T} |\psi(t)| \, dt < \infty$, so it cannot be directly used for infinite energy signals. Some of these signals can be handled by allowing the Fourier transform to contain impulses. The average power of a signal is defined as

$$P_\psi = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |\psi(t)|^2 \, dt \tag{17.24}$$

and a *power signal* is one for which $0 < P_\psi < \infty$, which means that the power is finite (but not necessarily the energy). The Fourier transform is redefined to handle also the case of signals of infinite energy. One approach for power signals is to take a time-limited sample of an infinite duration signal, where the time window is allowed to approach infinity. For example, such a time-limited sample can be of the form

$$\psi_T(t) = \psi(t) \quad |t| \le T < \infty$$
$$\psi_T(t) = 0 \quad |t| > T \tag{17.25}$$

and the corresponding Fourier transform is

$$\Psi_T(j\omega) = \int_{-\infty}^{+\infty} \psi_T(t) e^{-j\omega t} dt = \int_{-T}^{T} \psi(t) e^{-j\omega t} dt \tag{17.26}$$

Using a generalization of Parseval's theorem, the average power of a real-valued power signal $\psi_T(t)$ is computed from the Fourier transform as

$$P_\psi = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |\psi(t)|^2 dt \quad \text{or}$$

$$P_\psi = \lim_{T\to\infty} \frac{1}{4\pi T} \int_{-\infty}^{\infty} |\Psi_T(j\omega)\Psi_T(-j\omega) d\omega \quad \text{or} \tag{17.27}$$

$$P_\psi = \lim_{T\to\infty} \frac{1}{4\pi T} \int_{-\infty}^{\infty} |\Psi(j\omega)|^2 \, d\omega$$

The power spectral density of the signal is computed from the Fourier transform as

$$S_\psi(\omega) = \lim_{T\to\infty} \frac{|\Psi_T(j\omega)|^2}{2T} \tag{17.28}$$

17.4.3 Power Spectrum of Stochastic Signals

The power spectral density is a stationary stochastic signal (its statistical characteristics do not change in time) that converges and that is defined as an expected (mean) value. As with the case of power signals, one can consider first a time-limited sample of an infinite duration signal and then extend the duration to infinity. The average power of a signal over a finite time window $|t| < T$ is

$$\frac{1}{2T} \int_{-\infty}^{+\infty} |\psi_T(t)|^2 dt = \frac{1}{4\pi T} \int_{-\infty}^{+\infty} |\Psi_T(j\omega)|^2 \, d\omega \tag{17.29}$$

From Equation 17.29 and using the statistical expectation, it holds

$$E\left[\frac{1}{2T} \int_{-\infty}^{+\infty} [\psi_T(t)]^2 dt\right] = E\left[\frac{1}{4\pi T} \int_{-\infty}^{+\infty} [\Psi_T(j\omega)]^2 d\omega\right] \quad \text{i.e.} \tag{17.30}$$

and consequently one obtains

$$\lim_{T\to\infty} E\left[\frac{1}{2T} \int_{-\infty}^{+\infty} [\psi_T(t)]^2 dt\right] \quad \text{i.e.}$$

$$\lim_{T\to\infty} E\left[\frac{1}{4\pi T} \int_{-\infty}^{+\infty} [\Psi_T(j\omega)]^2 dt\right] \quad \text{i.e.} \tag{17.31}$$

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \lim_{T\to\infty} E\left[\frac{[\Psi_T(j\omega)]^2}{2T}\right] d\omega$$

The integrand of the last row of Equation 17.31 is the power spectral density of the signal:

$$S_{\psi}(\omega) = \lim_{T\to\infty} E\left[\frac{|\Psi_T(j\omega)|^2}{2T}\right] \tag{17.32}$$

17.4.4 Power Spectrum of the Signal Using the Gauss–Hermite Expansion

It has been shown that the Gauss–Hermite basis functions satisfy the orthogonality property, that is, according to Equation 17.9 for these functions, it holds

$$\int_{-\infty}^{+\infty} \psi_m(x)\psi_k(x)dx = \begin{cases} 1 & \text{if } m = k \\ 0 & \text{if } m \neq k \end{cases}$$

Therefore, using the definition of the signal's energy, one has

$$E = \int_{-\infty}^{+\infty} (\psi(t))^2 dt = \int_{-\infty}^{+\infty} \left[\sum_{k=1}^{N} c_k \psi_k(t)\right]^2 \tag{17.33}$$

and exploiting the orthogonality property one obtains

$$E = \sum_{k=1}^{N} c_k^2 \tag{17.34}$$

Therefore, the square of the coefficients c_k provides an indication of the distribution of the signal's energy to the associated basis functions. One could arrive at the same results using the Fourier-transformed description of the signal and Parseval's theorem. It has been shown that the Gauss–Hermite basis functions remain invariant under the Fourier transform subject only to a change of scale. Denoting by $\Psi(j\omega)$ the Fourier-transformed signal of and by $\Psi_k(j\omega)$ the Fourier transform of the kth Gauss–Hermite basis function, one obtains

$$\Psi(j\omega) = \sum_{k=1}^{N} c_k \Psi_k(j\omega) \tag{17.35}$$

and the energy of the signal is computed as

$$E_{\psi} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\Psi(j\omega)|^2 \, d\omega \tag{17.36}$$

Substituting Equation 17.35 into Equation 17.36, one obtains

$$E_{\psi} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left|\sum_{k=1}^{N} c_k \Psi_k(j\omega)\right|^2 d\omega \tag{17.37}$$

and using the invariance of the Gauss–Hermite basis functions under Fourier transform, as well as their orthogonality property, one gets that the signal's energy is proportional to the sum of the squares of the coefficients c_k that are associated with the Gauss–Hermite basis functions, that is, a relation of the form of Equation 17.34.

17.5 Application for Condition Monitoring of Electric Power Transformers

17.5.1 Failures in Electric Power Transformers

It will be shown that the proposed method for modeling nonlinear systems, which is characterized by invariance to Fourier transform, is suitable for performing FDI in several electromechanical systems such as vehicles, electric motors and power generators, or components of the electric power transmission and distribution system. As a case study the problem of fault diagnosis for electric power transformers will be presented.

One can distinguish between several classes of power transformers using two major classification criteria. The first criterion has to do with the insulating material used in the transformer (e.g., oil-immersed, gas-immersed, and dry-type transformers). The second criterion has to do with the incoming and outgoing voltage levels of the power transformers and their role in the electric power grid (e.g., power transformers in generation stations, power transformers in the transmission system, power transformers in the distribution system, distribution substation transformers, or distribution network transformers).

The minimization of the risk of failures in power transformers is important for improving the operation of the power grid. FDI for power transformers aims at continuously assessing the transformer's state through the monitoring of associated critical parameters and at determining if the transformer is on the verge of a failure (this can be due to an internal fault or due to ageing) [19–21]. Common failures in electric power transformers are (1) insulation breakdown in windings, (2) the on-load tap changer (OLTC) failures, and (3) bushing failures (see Figure 17.6).

Among other condition monitoring methods for power transformers, *thermal analysis* aims at monitoring of the evolution in time of the transformer's HST, which can be an indication of the ageing and degradation of the windings or of operating the transformer under overload conditions [22,23]. Most of the faults cause change in the thermal behavior of the transformer. Such abnormal conditions can be detected by analyzing the HST [24,25]. The most common abnormal condition of the transformer that can be detected with the use of thermal analysis is the overload. Transformer life is severely affected if the HST remains for long time intervals more than 110 °C. In this chapter a method for preventive maintenance of power transformers based on HST monitoring (thermal analysis) will be investigated.

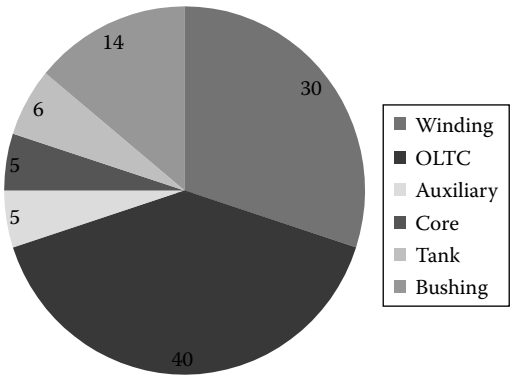


FIGURE 17.6 Frequency of faults in components of electric power transformers.

17.5.2 Analytical Thermal Model of Power Transformers

The stages for obtaining an analytical model of the power transformer's thermal behavior are as follows [14]:

- Calculate at each time step the ultimate top oil temperature (TOT) rise in the transformer from the load current at that instant, using

$$\Delta\Theta_{TO,U} = \Delta\Theta_{TO,R}\left[\frac{I_L^2 R + 1}{R + 1}\right]^q \tag{17.38}$$

  where
  $\Delta\Theta_{TO,U}$ is the ultimate TOT rise, °C
  $\Delta\Theta_{TO,R}$ is the rated TOT rise over ambient, °C
  I_L is the load current normalized to rated current, p.u.
  q is an empirically derived exponent to approximately account for effects of change of resistance with change in load
  R is the ratio of rated-load loss to no-load loss at applicable tap position

- Calculate the increment in the TOT from the ultimate top-oil rise and the ambient temperature at each time step using the following differential equation:

$$\tau_{TO}\frac{d\Theta_{TO}}{dt} = [\Delta\Theta_{TO,U} + \Theta_A] - \Theta_{TO} \tag{17.39}$$

  where
  Θ_{TO} is the TOT, °C
  τ_{TO} is the top-oil rise time constant
  Θ_A is the ambient temperature, °C

- Calculate the ultimate HST rise using

$$\Delta\Theta_{HS,U} = \Delta\Theta_{HS,R} I_L^{2\beta} \tag{17.40}$$

  where
  β is an empirically derived exponent, dependent on the cooling method
  $\Delta\Theta_{HS,U}$ is the ultimate HST rise over top oil (for a given load current), °C
  $\Delta\Theta_{HS,R}$ is the rated HST rise over top oil (for rated-load current), °C

- Calculate the increment in the HST rise, using the following differential equation:

$$\tau_{HS}\left\{\frac{d\Delta\Theta_{HS}}{dt}\right\} = \Delta\Theta_{HS,U} - \Delta\Theta_{HS} \tag{17.41}$$

  where
  Θ_{HS} is the hot-spot winding temperature, °C
  $\Delta\Theta_{HS}$ is the HST rise above top oil, °C
  τ_{HS} is the hot-spot rise time constant, h

- Finally, add the TOT to the HST rise to get the HST, using

$$\Theta_{HS} = \Theta_{TO} + \Delta\Theta_{HS} \tag{17.42}$$

The model of Equations 17.38 through 17.42 named top-oil rise model is based on some simplifying assumptions, and its accuracy can deteriorate due to parameter variations. As a result, in order to protect power transformers, conservative safety factors have been introduced that prevent the transformer's overheating. Consequently, the calculated maximum power transfer may be 20%–30% less or worse than the real transformer capability.

17.6 Gauss–Hermite Neural Modeling of Power Transformers Thermal Condition

The previously described methodology was employed to identify a neural model able to estimate with high accuracy the winding HST of a laboratory prototype mineral-oil-immersed power transformer. A measurement station has been set up consisting of thermocouples that were monitoring (a) the HST of the medium and voltage windings and (b) the TOT. The HST could have been also measured with optical fiber sensors. The manufacturer's specifications give the most probable hot-spot position. A Hall effect current transducer has been used in order to measure the load current.

Neural modeling of the power transformer has been carried out. The electric power transformer was modeled with the use of a neural network, having as output the estimated HST (k) and as inputs past values of the TOT, for example, $\Theta_{TO}(k-1)$, $\Theta_{TO}(k-2)$, and past values of the load current, for example, $I_L(k-1)$. Neural models with the same output, such as HST (k) and a larger number of inputs, that is, including more past values of the TOT and of the load current, could be also considered.

The transformer's model has been identified considering the previously analyzed neural network with Gauss–Hermite basis functions. The neural network was used to model the variations of the power transformer HST. The *inputs/outputs* configuration of the neural model of the power transformer thermal dynamics is shown in Figure 17.7. The RMSE of training the Gauss–Hermite neural model was of the order of 4×10^{-3}.

RMSE, defined as $\text{RMSE} = \sqrt{1/N \sum_{k=1}^{N} \left(y_k - y_k^d \right)^2}$, gives a measure of the performance of the neural network with Gauss–Hermite basis functions. It is noted that training affects only the output weights and can be performed with second-order gradient algorithms. However, since the speed of convergence is not the primary objective of this study, the least mean squares (LSM) algorithm is sufficient for training. Thus, the update of the output layer weights of the neural network is given by a gradient-type equation of the form

$$w^i(k+1) = w^i(k) - \eta e(k) \phi^T(k) \tag{17.43}$$

where

$e(k) = y(k) - y_d(k)$ is the output estimation error at time instant k

$\phi^T(k)$ is the regressor vector at time instant k having as elements the values $\phi(x(k))$ of the Gauss–Hermite basis functions for input vector $x(k)$

Training with the *data set* of dimension N is repeated in successive cycles (epochs) until no further drop of the approximation error (RMSE) can be observed. Diagrams depicting the approximation of the HST signal provided by the neural network with Gauss–Hermite polynomial basis functions, in comparison to the real HST signal, are given in Figures 17.8 and 17.9. Four different HST profiles have been considered, each one associated with different operating conditions for the transformer. As shown in Figures 17.8 and 17.9, thanks to the inherent multifrequency characteristics of the Hermite polynomial basis functions, such a neural model can capture with increased accuracy spikes and abrupt changes in the HST profile [1,11,18].

In Figures 17.8 and 17.9, diagrams are provided about the approximation of the HST signal by the neural network with Gauss–Hermite basis functions, and the estimated HST variations are compared

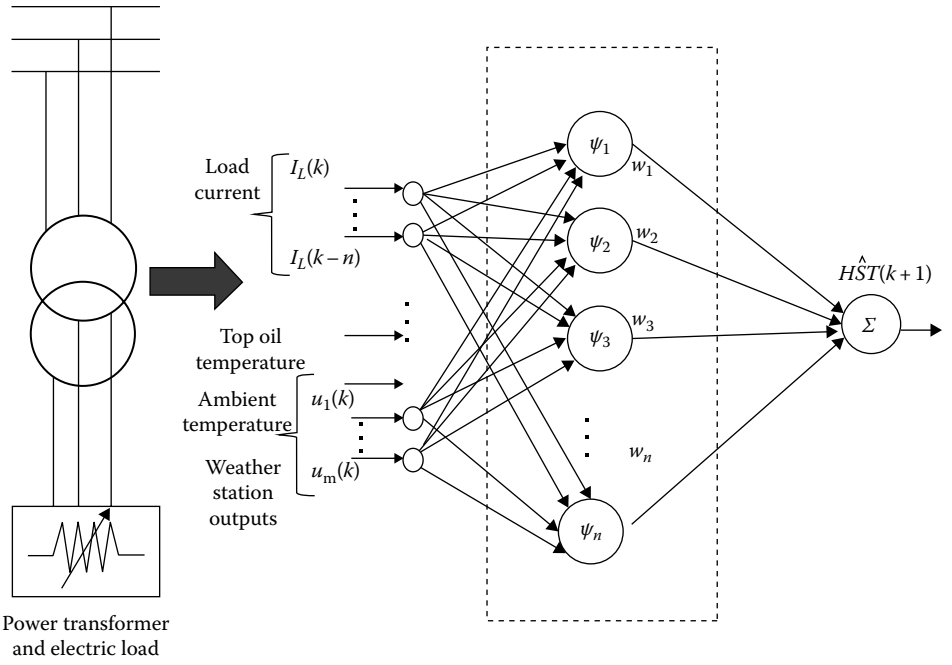


FIGURE 17.7 Inputs/outputs configuration of the neural model of the power transformer thermal dynamics.

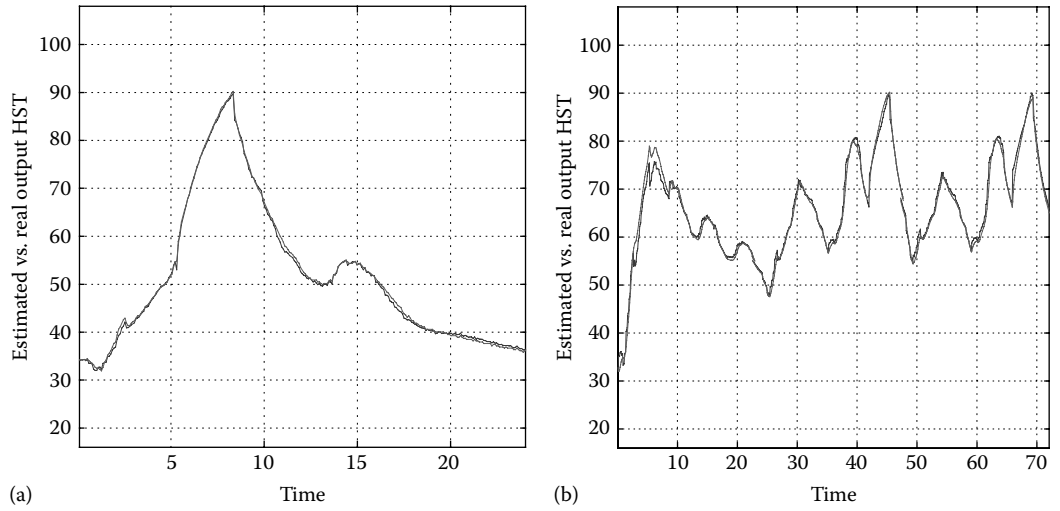


FIGURE 17.8 Approximation of the HST of the electric power transformer; by a neural network with Hermite polynomial basis functions: (a) HST time variation (profile 1) and (b) HST time variation (profile 2).

against the real HST measurements. Four different HST profiles have been considered, each one associated with different operating conditions of the power transformer.

To approximate the HST variations described in a data set consisting of 870 quadruplets of the form $[\Theta_{TO}(k-1), \Theta_{TO}(k-2), I_L(k-1)|HST(k)]$, an FNN with 3D Gauss–Hermite basis functions has been used, containing 64 nodes in its hidden layer. The spectral components of the HST signal for both the fault-free and the under-fault operation of the power transformer have been shown in Figures 17.10a through 17.13a. Note that after a fault has taken place in the power transformer, the amplitude of the

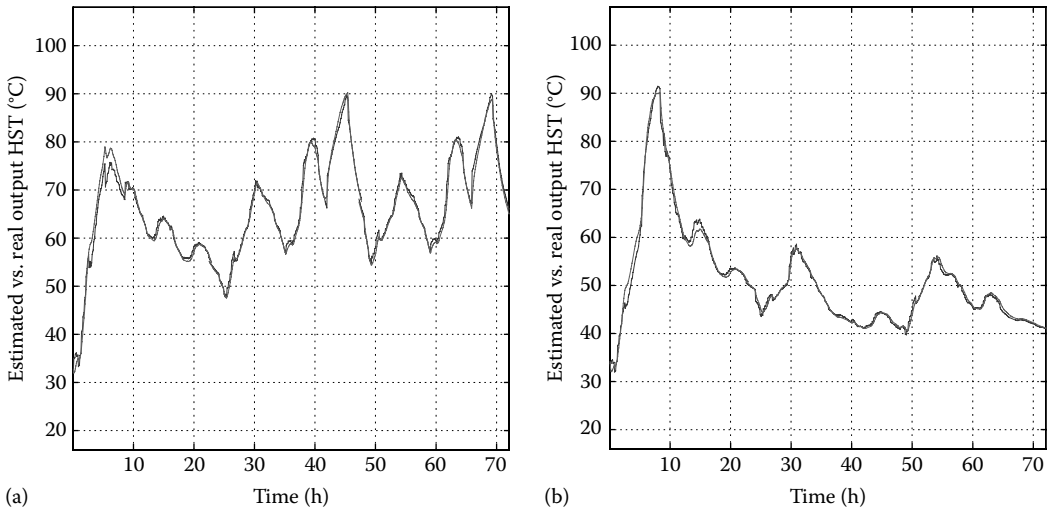


FIGURE 17.9 Approximation of the HST of the electric power transformer; by a neural network with Hermite polynomial basis functions: (a) HST time variation (profile 3) and (b) HST time variation (profile 4).

aforementioned spectral components changes, and this can be a strong indication about failure of the monitored transformer.

Obviously, the proposed spectral decomposition of the monitored signal, using the series expansion in Gauss–Hermite basis functions, can be used for fault detection tasks. As it can be seen in Figures 17.10b through 17.13b, in case of failure the spectral components of the monitored signal differ from the ones that are obtained when the system is free of fault. Moreover, the fact that certain spectral components exhibit greater sensitivity to the fault and change value in a more abrupt manner is a feature that can be exploited for fault isolation. Specific failures can be associated with variations of specific spectral components of the monitored signal. Therefore, they can provide indication about the appearance of specific types of failures and specific malfunctioning components.

It is noted that the problem of thermal condition monitoring is apparent in several systems and devices. For example, through thermal condition monitoring, one can draw conclusions about the existence of

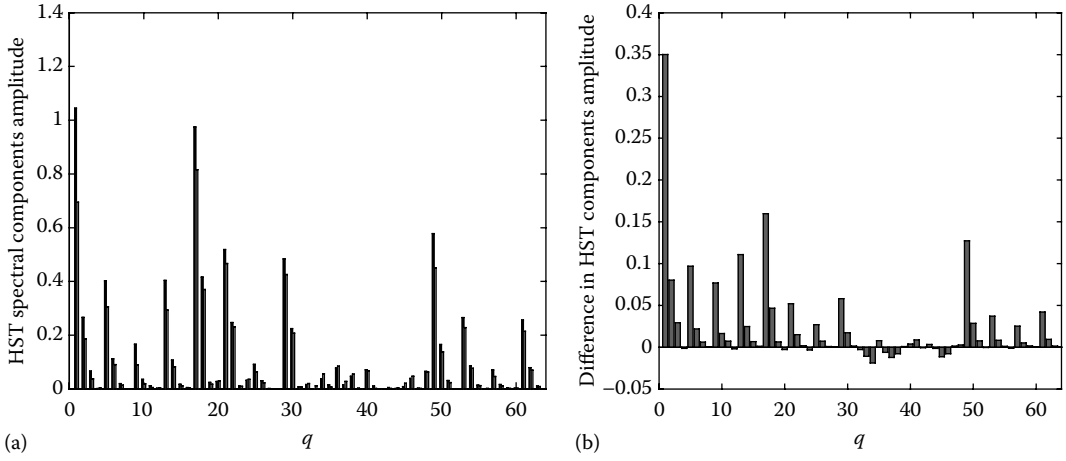


FIGURE 17.10 HST time variation (profile 1): (a) amplitude of the spectral components of the HST signal measured from the electric power transformer in the fault-free case (black line) and when a fault had taken place (gray line) and (b) differences in the amplitudes of the spectral components between the fault-free and the faulty case.

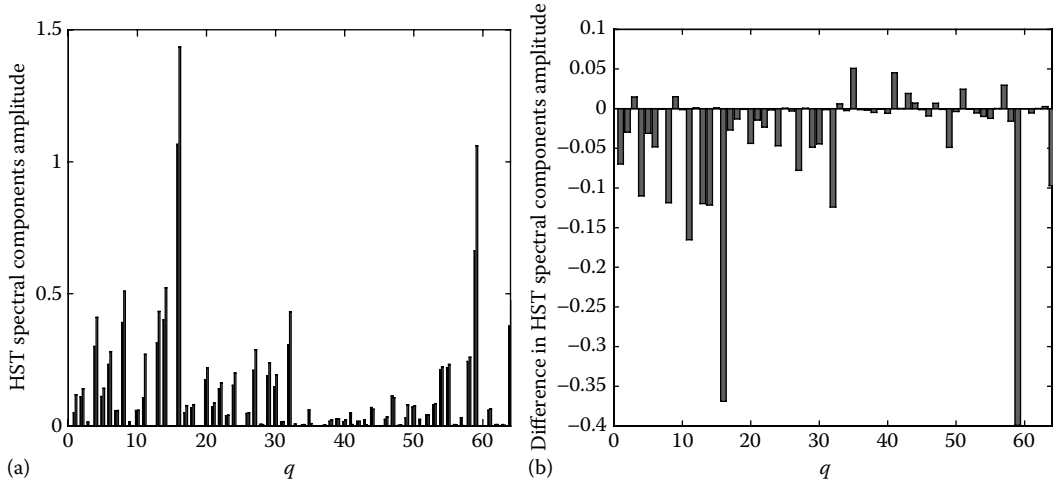


FIGURE 17.11 HST time variation (profile 2): (a) amplitude of the spectral components of the HST signal measured from the electric power transformer in the fault-free case (black line) and when a fault had taken place (gray line) and (b) differences in the amplitudes of the spectral components between the fault-free and the faulty case.

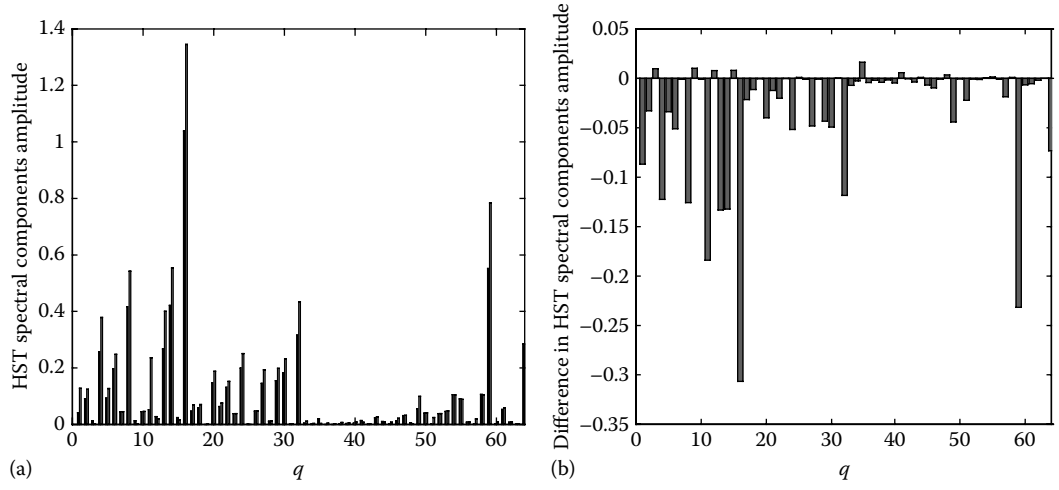


FIGURE 17.12 HST time variation (profile 3): (a) amplitude of the spectral components of the HST signal measured from the electric power transformer in the fault-free case (black line) and when a fault had taken place (gray line) and (b) differences in the amplitudes of the spectral components between the fault-free and the faulty case.

failures not only in power transformers but also in transmission lines and in electric machines such as motors and generators. Temperature histories strongly influence an electric machine's condition and future capability. High temperatures inside an electric machine, especially the winding temperatures, affect the functioning of the machine. For instance, winding insulation materials suffer irreversible damage when thermal stresses exceed their permissible limits. These temperature-dependent changes affect the machine's performance, can shorten its useful life, and may lead to failure. Therefore, the previously analyzed approach for nonlinear dynamics modeling under invariance to Fourier transform can be a useful method for performing systematic condition monitoring of dynamical systems subject to thermal stresses or other harsh operating conditions.

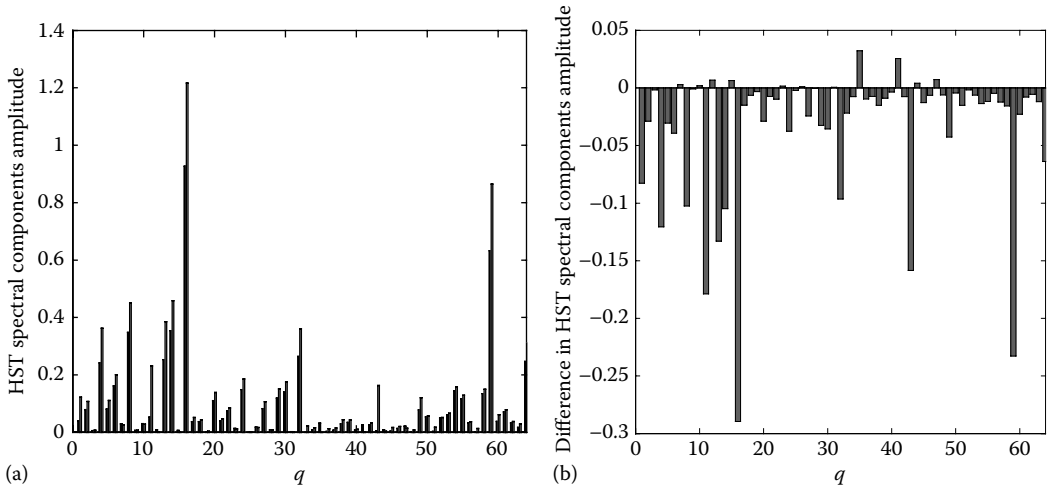


FIGURE 17.13 HST time variation (profile 4): (a) amplitude of the spectral components of the HST signal measured from the electric power transformer in the fault-free case (black line) and when a fault had taken place (gray line) and (b) differences in the amplitudes of the spectral components between the fault-free and the faulty case.

17.7 Conclusions

A new method for fault diagnosis of nonlinear systems has been proposed based on the modeling of the system's dynamics with FNNs that use orthogonal basis functions exhibiting invariance to Fourier transform. A neural network with Gauss–Hermite polynomial activation functions has been used for approximating the nonlinear system's dynamics out of a set of input–output data. The output of the neural network has taken the form of a series expansion that was written as a weighted sum of Gauss–Hermite basis functions. Using the Fourier transform property of the Gauss–Hermite basis functions, it was shown that the considered neural network could provide spectral analysis of the output of the monitored system. Moreover, using the orthogonality property of the Gauss–Hermite basis functions, it was shown that the sum of the square of the output layer weights of these neural networks stands for a measure of the energy contained in the output of the monitored system.

The monitoring of changes in the amplitude of the aforementioned spectral components provides an indication about malfunctioning of the monitored system and a tool for detecting the existence of failures. Additionally, since specific faults are associated with amplitude changes of specific spectral components of the system, fault isolation can be also performed. The proposed FDI method can be applied to several electromechanical systems, for example, vehicles, electric motors, and power generators. As a case study, the problem of fault diagnosis of electric power transformers has been examined. The considered neural network with Gauss–Hermite polynomial activation functions enabled to obtain information about the thermal condition of oil-immersed power transformers and about their ageing and failure risks through the approximation and spectral decomposition of a critical variable of the transformer known as HST. Evaluation tests have confirmed the efficiency of the proposed fault diagnosis method.

References

1. Q. Zhang and A. Benveniste, Wavelet networks, *IEEE Transactions on Neural Networks*, 4, 889–898, 1993.
2. P.S. Addison, *The Illustrated Wavelet Transform Handbook*, Institute of Physics Publishing, Bristol, PA, 2002.

3. H.R. Karimi and B. Lohmann, A computational method to robust vibration control of vehicle engine-body system using Haar wavelets, in *Proceedings of the 2006 IEEE International Conference on Control Applications*, pp. 169–174, Munich, Germany, October 2006.
4. M. Cannon and J.J.E. Slotine, Space-frequency localized basis function networks for nonlinear system estimation and control, *Neurocomputing*, 9, 293–342, 1995.
5. C.P. Bernard and J.J. Slotine, Adaptive control with multiresolution bases, in *Proceedings of the 36th IEEE International Conference on Decision and Control (IEEE CDC 97)*, San Diego, CA, December 1997.
6. A. Krzyzak and J. Sasiadek, Flexible robot identification using nonparametric techniques, in *Proceedings of the 30th IEEE Conference on Decision and Control*, Brighton, U.K., December 1991.
7. C.K. Lin, Nonsingular terminal sliding mode control of robot manipulators using fuzzy wavelet networks, *IEEE Transactions on Fuzzy Systems*, 14, 849–859, 2006.
8. N. Sureshbabu and J.A. Farell, Wavelet-based system identification for nonlinear control, *IEEE Transactions on Automatic Control*, 44, 412–417, 1999.
9. A. Refregier and I. Shapelets, A method for image analysis, *Monthly Notices of the Royal Astronomical Society*, 338, 35–47, 2003.
10. G.G. Rigatos and S.G. Tzafestas, Feed-forward neural networks using Hermite polynomial activation functions, *Lecture Notes in Artificial Intelligence*, 3925, 323–333, 2006.
11. G.G. Rigatos and S.G. Tzafestas, Neural structures using the eigenstates of the Quantum Harmonic Oscillator, *Open Systems and Information Dynamics*, 13(1), 27–41, 2006.
12. S.S. Yang and C.S. Cheng, An orthogonal neural network for function approximation, *IEEE Transactions on Systems, Man and Cybernetics—Part B: Cybernetics*, 26, 779–784, 1996.
13. V. Galdi, L. Ippolito, A. Piccolo, and A. Vaccaro, Neural diagnostic system for transformer thermal overload protection, *IEE Proceedings on Power Applications*, 147(5), 415–421, 2000.
14. L. Ippolito and P. Siano, Identification of Tagaki-Sugeno-Kang fuzzy model for power transformers predictive overload system, *IEE Proceedings Generation, Transmission and Distribution*, 151(5), 582–589, 2004.
15. G. Rigatos, P. Siano, and A. Piccolo, Incipient fault detection for electric power transformers using neural modeling and the local statistical approach to fault diagnosis, in *Proceedings of the 2012 IEEE Sensors Applications Symposium (IEEE SAS 2012)*, University of Brescia, Brescia, Italy, February 2012.
16. W. Zuo, Y. Zhu, and L. Cai, Fourier-Neural-Network-Based Learning control for a class of nonlinear systems with flexible components, *IEEE Transactions on Neural Networks*, 20, 139–151, 2009.
17. S. Haykin, *Neural Networks: A Comprehensive Foundation*, McMillan, New York, 1994.
18. G. Rigatos and Q. Zhang, Fuzzy model validation using the local statistical approach, *Fuzzy Sets and Systems*, 60(7), 882–904, 2009.
19. A. Piccolo, P. Siano, and G. Rigatos, An adaptive framework for power components dynamic loadability, in *Modelling and Automation of Intelligent Industrial Systems*, G.G. Rigatos (Ed.), IGI Publications, Hershey, PA, pp. 321–345, 2010.
20. V.M. Catterson, E.M. Davidson, and S.D.J. McArthur, Embedded intelligence for electrical network operation and control, *IEEE Intelligent Systems*, 26(2), 38–45, 2011.
21. V.M. Catterson, E.M. Davidson, and S.D.J. McArthur, Agents for active network management and condition monitoring in the smart grid, in *Proceedings of the First International Workshop on Agent Technologies for Energy Systems*, ATES, Toronto, Ontario, Canada, 2010.
22. I.A. Metwally, Failures monitoring and new trends of power transformers, *IEEE Potentials*, 30(3), 36–43, 2011.

23. M. Arshad and S.M. Islam, A novel fuzzy logic technique for power transformer asset management, *IEEE Industry Applications Conference, Conference Record of the 2006 41st IAS Annual Meeting*, pp. 276–286, October 2006, Tampa, FL, 2006.
24. A.E.B. Abu-Elanien and M.M.A. Salama, Asset management techniques for transformers, *Electric Power Systems Research*, 80, 465–464, 2010.
25. J.L. Velasquez-Contreras, M.A. Sanz-Bobi, and S.G. Arellano, General asset management model in the context of an electric utility: Application to power transformers, *Electric Power Systems Research*, 81, 2015–2037, 2011.

18

Meter Voltage Measurement

Alessandro Ferrero
Politecnico di Milano

Instruments for the measurement of electric voltage are called *voltmeters*. Correct insertion of a voltmeter requires the connection of its terminals to the points of an electric circuit across which the voltage has to be measured, as shown in Figure 18.1. To a first approximation, the electric equivalent circuit of a voltmeter can be represented by resistive impedance Z_v (or a pure resistance R_v for dc voltmeters). This means that any voltmeter, once connected to an electric circuit, draws a current I_v given by

$$I_v = \frac{V}{Z_v} \tag{18.1}$$

where V is the measured voltage. The higher the value of the internal impedance, the higher the quality of the voltmeter, since it does not significantly modify the status of the electric circuit under test.

Different operating principles are used to measure an electric voltage. The mechanical interaction between currents, between a current and a magnetic field, or between electrified conductors was widely adopted in the past to generate a mechanical torque proportional to the voltage or the squared voltage to be measured. This torque, balanced by a restraining torque, usually generated by a spring, causes the instrument pointer, which can be a mechanical or a virtual optical pointer, to be displaced by an angle proportional to the driving torque and hence to the voltage or the squared voltage to be measured. The value of the input voltage is therefore given by the reading of the pointer displacement on a graduated scale. The thermal effects of a current flowing in a conductor are also used for measuring electric voltages, although they have not been adopted as widely as the previous ones. More recently, the widespread diffusion of semiconductor devices led to the development of a completely different class of voltmeters: *electronic* voltmeters. They basically attain the required measurement by processing the input signal by means of electronic semiconductor devices. According to the method, analog or digital, the input signal is processed; the electronic voltmeters can be divided into *analog*

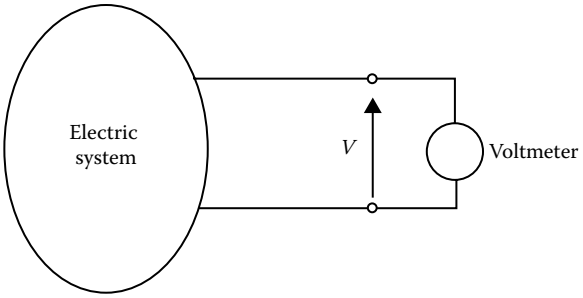


FIGURE 18.1 Voltmeter insertion.

TABLE 18.1 Classification of Voltage Meters

Class	Operating Principle	Subclass	Application Field
Electromagnetic	Interaction between currents and magnetic fields	Moving magnet	dc voltage
		Moving coil	dc voltage
		Moving iron	dc and ac voltage
Electrodynamic	Interactions between currents	—	dc and ac voltage
Electrostatic	Electrostatic interactions	—	dc and ac voltage
Thermal	Current's thermal effects	Direct action	dc and ac voltage
		Indirect action	dc and ac voltage
Induction	Magnetic induction	—	ac voltage
Electronic	Signal processing	Analog	dc and ac voltage
		Digital	dc and ac voltage

electronic voltmeters and *digital* electronic voltmeters. Table 18.1 shows a rough classification of the most commonly employed voltmeters, according to their operating principle and their typical application field.

This chapter section briefly describes the most commonly employed voltmeters, both electromechanical and electronic.

18.1 Electromechanical Voltmeters

Electromechanical voltmeters measure the applied voltage by transducing it into a mechanical torque. This can be accomplished in different ways, basically because of the interactions between currents (*electrodynamic voltmeters*), between a current and a magnetic field (*electromagnetic voltmeters*), between electrified conductors (*electrostatic voltmeters* or *electrometers*), and between currents induced in a conducting vane (*induction voltmeters*). According to the different kinds of interactions, different families of instruments can be described, with different application fields. Moving-coil electromagnetic voltmeters are restricted to the measurement of dc voltages; moving-iron electromagnetic, electrodynamic, and electrostatic voltmeters can be used to measure both dc and ac voltages, while induction voltmeters are restricted to ac voltages.

The most commonly employed electromechanical voltmeters are the electromagnetic and electrodynamic ones. Electrostatic voltmeters have been widely employed in the past (and are still employed) for the measurement of high voltages, both dc and ac, up to a frequency on the order of several megahertz. Induction voltmeters have never been widely employed, and their present use is restricted to ac voltages.

Therefore, only the electromagnetic, electrodynamic, and electrostatic voltmeters will be described in the following sections.

18.2 Electromagnetic Voltmeters

18.2.1 DC Moving-Coil Voltmeters

The structure of a dc moving-coil meter is shown in Figure 18.2. A small rectangular pivoted coil is wrapped around an iron cylinder and placed between the poles of a permanent magnet. Because of the shape of the poles of the permanent magnet, the induction magnetic field B in the air gap is radial and constant.

Suppose that a dc current I is flowing in the coil, the coil has N turns, and that the length of the sides that cut the magnetic flux (active sides) is l; the current interacts with the magnetic field B, and a force F is exerted on the conductors of the active sides. The value of this force is given by

$$F = NBlI \tag{18.2}$$

Its direction is given by the right-hand rule. Since the two forces applied to the two active sides of the coil are directed in opposite directions, a torque arises in the coil, given by

$$T_i = Fd = NBldI \tag{18.3}$$

where d is the coil width. Since N, B, l, and d are constant, Equation 18.3 leads to

$$T_i = k_i I \tag{18.4}$$

showing that the mechanical torque exerted on the coil is directly proportional to the current flowing in the coil itself.

Because of T_i, the coil rotates around its axis. Two little control springs, with k_r constant, provide a restraining torque T_r. The two torques balance when the coil is rotated by an angle δ so that

$$k_i I = k_r \delta \tag{18.5}$$

that leads to

$$\delta = \frac{k_i}{k_r} I \tag{18.6}$$

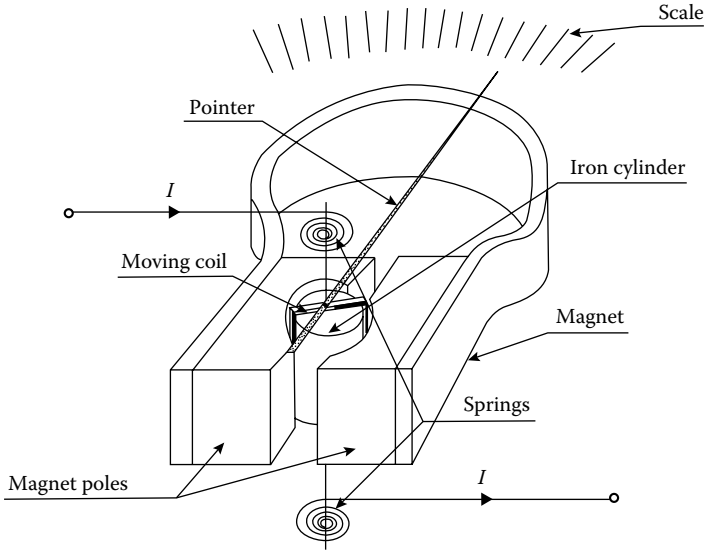


FIGURE 18.2 DC moving-coil meter.

Equation 18.6 shows that the rotation angle of the coil is directly proportional to the dc current flowing in the coil. If a pointer with length h is keyed on the coil axes, a displacement $\lambda = h\delta$ can be read on the instrument scale. Therefore, the pointer displacement is proportional to the current flowing in the coil, according to the following relationship:

$$\lambda = h\frac{k_i}{k_r}I \tag{18.7}$$

This instrument is hence intrinsically a current meter. A voltmeter can be obtained by connecting an additional resistor in series with the coil. If the coil resistance is R_c and the resistance of the additional resistor is R_a, the current flowing in the coil when the voltage V is applied is given by

$$I = \frac{V}{R_a + R_c} \tag{18.8}$$

and therefore the pointer displacement is given by

$$\lambda = h\delta = h\frac{k_i}{k_r}I = h\frac{k_i}{k_r(R_a + R_c)}V \tag{18.9}$$

and is proportional to the applied voltage. Because of this proportionality, moving-coil dc meters show a proportional-law scale, where the applied voltage causes a proportional angular deflection of the pointer.

Because of the operating principle expressed by Equation 18.3, these voltmeters can measure only dc voltages. Due to the inertia of the mechanical part, ac components typically do not cause any coil rotation, and hence these meters can be also employed to measure the dc component of a variable voltage. They have been widely employed in the past for the measurement of dc voltages up to some thousands volts with a relative measurement uncertainty as low as 0.1% of the full-scale value. At present, they are being replaced by electronic voltmeters that feature the same or better accuracy at a lower cost.

18.2.2 DC Galvanometer

18.2.2.1 General Characteristics

A galvanometer is used to measure low currents and low voltages. Because of the high sensitivity that this kind of measurement requires, galvanometers are widely employed as null indicators in all dc balance measurement methods (like the bridge and potentiometer methods) [1,2].

A dc galvanometer is, basically, a dc moving-coil meter, and the relationship between the index displacement and the current flowing in the moving coil is given by Equation 18.7. The instrument constant

$$k_a = h\frac{k_i}{k_r} \tag{18.10}$$

is usually called the galvanometer *current constant* and is expressed in mm μA^{-1}. The galvanometer *current sensitivity* is defined as $1/k_a$ and is expressed in $\mu A\ mm^{-1}$.

According to their particular application field, galvanometers must be chosen with particular care. If k_a is taken into account, note that once the full-scale current and the corresponding maximum pointer displacement are given, the value of the ratio hk_i/k_r is also known. However, the single values of h, k_i, and k_r can assume any value and are usually set in order to reduce the friction effects. In fact, if the restraining friction torque T_f is taken into account in the balance equation, Equation 18.5 becomes

$$k_i I = k_r\frac{\lambda}{h} \pm T_f \tag{18.11}$$

where the ± sign shows that the friction torque does not have its own sign but always opposes the rotation.

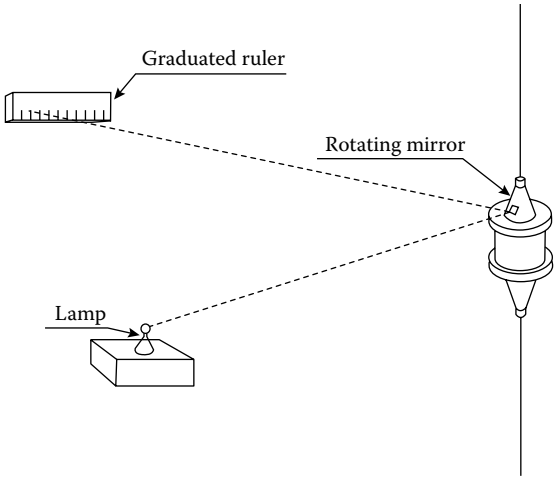


FIGURE 18.3 Virtual optical pointer structure in a dc galvanometer.

The effects of T_f can be neglected if the driving torque hk_iI and the restraining torque $k_r\lambda$ are sufficiently greater than T_f. Moreover, since the galvanometer is employed as a null indicator, a high sensitivity is needed; hence, k_a must be as high as possible. According to Equations 18.10 and 18.11, this requires high values of hk_i and low values of k_r. A high value of h means a long pointer; a high value of k_i means a high driving torque, while a low value of k_r means that the inertia of the whole moving system must be low.

The pointer length can be increased without increasing the moving system inertia by employing virtual optical pointers: a little, light concave mirror is fixed on the moving-coil axis and is lit by an external lamp. The reflected light hits a translucid, graduated ruler, so that the mirror rotation can be observed (Figure 18.3). In this way, a virtual pointer is obtained, whose length equals the distance between the mirror and the graduated ruler.

The reduction of the moving system inertia is obtained by reducing the weight and dimension of the moving coil and reducing the spring constant. This is usually done by suspending the moving coil with a thin fiber of conducting material (usually bronze). Thus, the friction torque is practically removed, and the restraining spring action is given by the fiber torsion.

According to Equations 18.3 and 18.4, the driving torque can be increased by increasing the coil flux linkage. Three parameters can be modified to attain this increase: the induction field B, the coil section ld, and the number of turns N of the coil winding.

The induction field B can be increased by employing high-quality permanent magnets, with high coercive force, and minimizing the air gap between the magnet's poles. This minimization prevents the use of moving coils with a large section. Moreover, large coil sections lead to heavier coils with greater inertia, which opposes the previous requirement of reduced inertia. For this reason, the coil section is usually rectangular (although a square section maximizes the flux linkage) and with $l > d$.

If the galvanometer is used to measure a low-voltage V, the *voltage sensitivity*, expressed in μV mm^{-1}, is the inverse of

$$k_v = \frac{\lambda}{V} \tag{18.12}$$

where k_v is called the galvanometer's *voltage constant* and is expressed in mm μV^{-1}.

18.2.2.2 Mechanical Characteristics

Due to the low inertia and low friction, the galvanometer moving system behaves as an oscillating mechanical system. The oscillations around the balance position are damped by the electromagnetic

forces that the oscillations of the coil in the magnetic field exert on the coil active sides. It can be proved [1] that the oscillation damping is a function of the coil circuit resistance, that is, the coil resistance r plus the equivalent resistance of the external circuit connected to the galvanometer. In particular, the damping effect is nil if the coil circuit is open and maximum if the coil is short-circuited.

In practical situations, a resistor is connected in series with the moving coil, whose resistance is selected in such a way to realize a critical damping of the coil movement. When this situation is obtained, the galvanometer is said to be *critically damped* and reaches its balance position in the shortest time, without oscillations around this position.

18.2.2.3 Actual Trends

Moving-coil dc galvanometers have been widely employed in the past when they represented the most important instrument for high-sensitivity measurements. In more recent years, due to the development of the electronic devices and particularly high-gain, low-noise amplifiers, the moving-coil galvanometers are being replaced by electronic galvanometers, which feature the same, or even better, performance than the electromagnetic ones.

18.3 Electrodynamic Voltmeters

18.3.1 AC Moving-Coil Voltmeters

The structure of an ac moving-coil meter is shown in Figure 18.4. It basically consists of a pivoted moving coil, two stationary field coils, control springs, a pointer, and a calibrated scale. The stationary coils are series connected, and when a current i_f is applied, a magnetic field B_f is generated

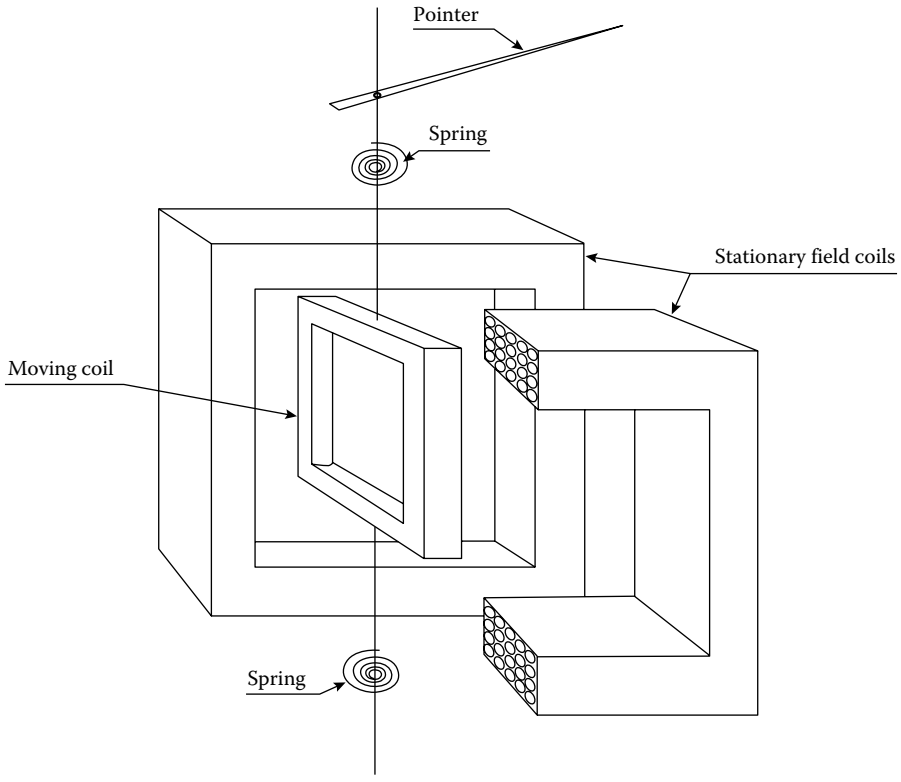


FIGURE 18.4 AC moving-coil meter.

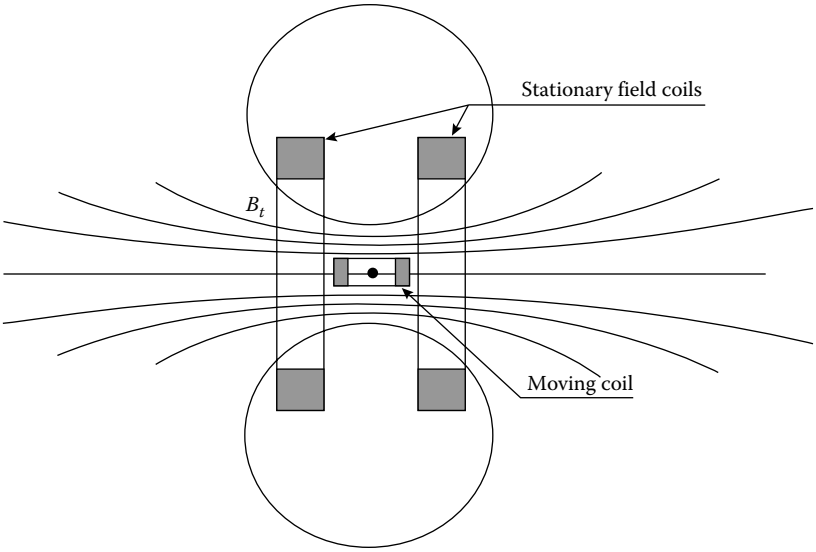


FIGURE 18.5 Magnetic field generated by the field coils in an ac moving-coil meter.

along the axis of the stationary coils, as shown in Figure 18.5. A magnetic flux is therefore generated, whose instantaneous values are given by

$$\varphi_f(t) = k' m_f i_f(t) \tag{18.13}$$

where

m_f is the number of turns of the stationary coil
k' is a proportionality factor

When a current i_m is applied to the moving coil, a torque arises, whose instantaneous values are proportional to the product of φ_f and i_m instantaneous values:

$$T_i(t) = k''\varphi_f(t) i_m(t) = k i_f(t) i_m(t) \tag{18.14}$$

The driving torque is therefore proportional to the instantaneous product of the currents flowing in the two coils. Due to this driving torque, the moving element is displaced by an angle (δt), until the spring restraining torque $T_s(t) = k_s\delta(t)$ balances the driving torque. The moving element rotation is thus given by

$$\delta(t) = \frac{k}{k_s} i_f(t) i_m(t) \tag{18.15}$$

and, if the pointer length is h, the following pointer displacement can be read on the scale:

$$\lambda(t) = h\frac{k}{k_s} i_f(t) i_m(t) \tag{18.16}$$

The proportionality factor k is generally not constant, since it depends on the mutual inductance between the two coils and thus on their number of turns, shape, and relative position. However, if the two coils are carefully designed and placed, the magnetic field can be assumed to be constant and radial in the rotation area of the moving coil. Under this condition, k is virtually constant.

Because the bandwidth of the moving element is limited to a few hertz, due to its inertia, the balance position is proportional to the average value of the driving torque when the signal bandwidth exceeds

this limit. If i_f and i_m currents are sinusoidal, with I_f and I_m rms values, respectively, and with a relative phase displacement β, the driving torque average value is given by

$$\overline{T_i} = kI_f I_m \cos\beta \tag{18.17}$$

and, thus, the pointer displacement in Equation 18.16 becomes

$$\lambda = h\frac{k}{k_s} I_f I_m \cos\beta \tag{18.18}$$

In order to realize a voltmeter, the stationary and moving coils are series connected, and a resistor, with resistance R, is also connected in series to the coils. If R is far greater than the resistance of the two coils and if it is also far greater than the coil inductance, in the frequency operating range of the voltmeter, the rms value of the coils' currents is given by

$$I_f = I_m = \frac{V}{R} \tag{18.19}$$

V being the applied voltage rms value. From Equation 18.18, the pointer displacement is therefore given by

$$\lambda = h\frac{k}{k_s}\frac{V^2}{R^2} = k_v V^2 \tag{18.20}$$

Because of Equation 18.20, the voltmeter features a square-law scale, with k_v constant, provided that the coils are carefully designed and that the coils' inductance can be neglected with respect to the resistance of the coils themselves and the series resistor. This last condition determines the upper limit of the input voltage frequency.

These voltmeters feature good accuracy (their uncertainty can be as low as 0.2% of the full-scale value), with full-scale values up to a few hundred volts, in a frequency range up to 2 kHz.

18.4 Electrostatic Voltmeters

The action of electrostatic instruments is based on the force exerted between two charged conductors. The conductors behave as a variable plate air capacitor, as shown in Figure 18.6. The moving plate, when charged, tends to move so as to increase the capacitance between the plates. The energy stored in the capacitor, when the applied voltage is V and the capacitance is C, is given by

$$W = \frac{1}{2}CV^2 \tag{18.21}$$

This relationship is valid under both dc and ac conditions, provided that the voltage rms value V is considered for ac voltage.

When the moving plate is displaced horizontally by ds, while the voltage is held constant, the capacitor energy changes in order to equal the work done in moving the plate. The resulting force is

$$F = \frac{\mathrm{d}W}{\mathrm{d}s} = \frac{V^2}{2}\frac{\mathrm{d}C}{\mathrm{d}s} \tag{18.22}$$

For a rotable system, Equation 18.21 leads similarly to a resulting torque:

$$T = \frac{\mathrm{d}W}{\mathrm{d}\vartheta} = \frac{V^2}{2}\frac{\mathrm{d}C}{\mathrm{d}\vartheta} \tag{18.23}$$

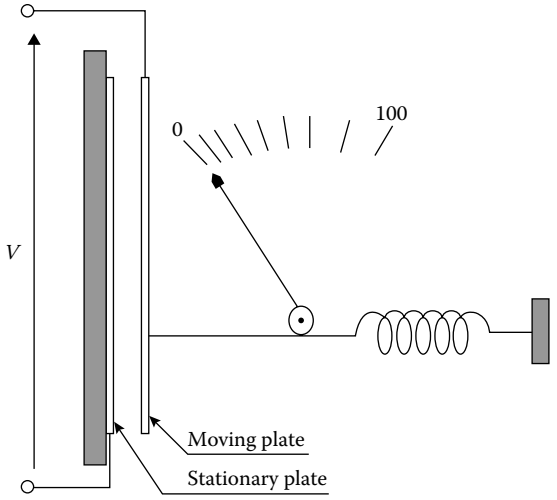


FIGURE 18.6 Basic structure of an electrostatic voltmeter.

If the action of a control spring is also considered, both Equations 18.22 and 18.23 show that the balance position of the moving plate is proportional to the square of the applied voltage, and hence electrostatic voltmeters have a square-law scale. These equations, along with Equation 18.21, show that these instruments can be used for the measurement of both dc and ac rms voltages. However, the force (or torque) supplied by the instrument schematically represented in Figure 18.6 is generally very weak [2], so that its use is very impractical.

18.4.1 Electrometer

A more useful configuration is the quadrant electrometer, shown in Figure 18.7. Four fixed plates realize four quadrants and surround a movable vane suspended by a torsion fiber at the center of the system. The opposite quadrants are electrically connected together, and the potential difference $(V_1 - V_2)$ is applied. The moving vane can be either connected to potential V_1 or V_2, or energized by an independent potential V_3.

Let the zero torque position of the suspension coincide with the symmetrical $X - X$ position of the vane. If $V_1 = V_2$, the vane does not leave this position; otherwise, the vane will rotate.

Let C_1 and C_2 be the capacitances of quadrants 1 and 2, respectively, relative to the vane. They both are functions of ϑ, and according to Equation 18.23, the torque applied to the vane is given by

$$T = \frac{(V_3 - V_1)^2}{2}\frac{\mathrm{d}C_1}{\mathrm{d}\vartheta} + \frac{(V_3 - V_2)^2}{2}\frac{\mathrm{d}C_2}{\mathrm{d}\vartheta} \tag{18.24}$$

Since the vane turns out of one pair of quadrants as much as it turns into the other, the variations of C_1 and C_2 can be related by

$$-\frac{\mathrm{d}C_1}{\mathrm{d}\vartheta} = \frac{\mathrm{d}C_2}{\mathrm{d}\vartheta} = k_1 \tag{18.25}$$

Taking into account the suspension restraining torque $T_r = k_2\vartheta$, the balance position can be obtained by Equations 18.24 and 18.25 as

$$\vartheta = \frac{k_1}{2k_2}[(V_3 - V_2)^2 - (V_3 - V_1)^2] \tag{18.26}$$

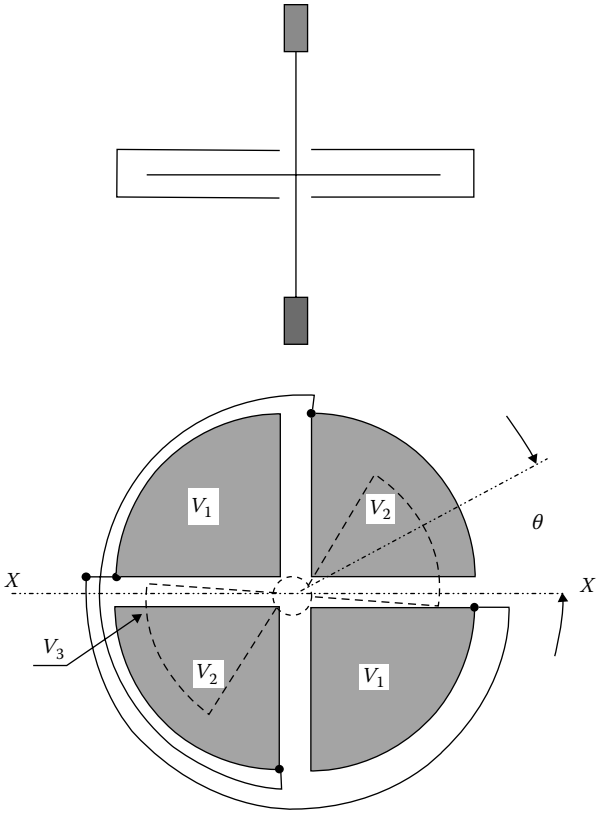


FIGURE 18.7 Quadrant electrometer structure.

If the vane potential V_3 is held constant and is large compared to the quadrant potentials V_1 and V_2, Equation 18.26 can be simplified as follows:

$$\vartheta = \frac{k_1}{k_2} V_3 (V_1 - V_2) \tag{18.27}$$

Equation 18.27 shows that the deflection of the vane is directly proportional to the voltage difference applied to the quadrants. This method of use is called the *heterostatic* method.

If the vane is connected to quadrant 1, $V_3 = V_1$ follows, and Equation 18.26 becomes

$$\vartheta = \frac{k_1}{2k_2} (V_1 - V_2)^2 \tag{18.28}$$

Equation 18.28 shows that the deflection of the vane is proportional to the square of the voltage difference applied to the quadrants, and hence this voltmeter has a square-law scale. This method of use is called the *idiostatic* method and is suitable for the direct measurement of dc and ac voltages without an auxiliary power source.

The driving torque of the electrometer is extremely weak, as in all electrostatic instruments. The major advantage of using this kind of meter is that it allows for the measurement of dc voltages without drawing current by the voltage source under test. Now, due to the availability of operational amplifiers with extremely high input impedance, they have been almost completely replaced by electronic meters with high input impedance.

18.5 Electronic Voltmeters

Electronic meters process the input signal by means of semiconductor devices in order to extract the information related to the required measurement [3,4]. An electronic meter can be basically represented as a three-port element, as shown in Figure 18.8.

The input signal port is an input port characterized by high impedance, so that the signal source has very little load. The measurement result port is an output port that provides the measurement result (in either an analog or digital form, depending on the way the input signal is processed) along with the power needed to energize the device used to display the measurement result. The power supply port is an input port that the electric power required to energize the meter internal devices and the display device flows through.

One of the main characteristics of an electronic meter is that it requires an external power supply. Although this may appear as a drawback of electronic meters, especially where portable meters are concerned, note that, this way, the energy required for the measurement is no longer drawn from the signal source.

The high-level performance of modern electronic devices yields meters that are as accurate (and sometime even more accurate) as the most accurate electromechanical meters. Because they do not require the extensive use of precision mechanics, they are presently less expensive than electromechanical meters and are slowly, but constantly, replacing them in almost all applications.

Depending on the way the input signal is processed, electronic meters are divided into *analog* and *digital* meters. Analog meters attain the required measurement by analog, continuous-time processing of the input signal. The measurement result can be displayed both in analog form using, for example, an electromechanical meter, and in digital form by converting the analog output signal into digital form. Digital meters attain the required measurement by digital processing of the input signal. The measurement result is usually displayed in digital form. Note that the distinction between analog and digital meters is not due to the way the measurement result is displayed but to the way the input signal is processed.

18.5.1 Analog Voltmeters

An electronic analog voltmeter is based on an electronic amplifier and an electromechanical meter to measure the amplifier output signal. The amplifier operates to make a dc current, proportional to the input quantity to be measured, flow into the meter. This meter is hence a dc moving-coil milliammeter.

Different full-scale values can be obtained using a selectable-ratio voltage divider if the input voltage is higher than the amplifier dynamic range or by selecting the proper amplifier gain if the input voltage stays within the amplifier dynamic range.

The main features of analog voltmeters are high input impedance, high possible gain, and wide possible bandwidth for ac measurements. The relative measurement uncertainty can be lower than 1% of full-scale value. Because of these features, electronic analog voltmeters can have better performance than the electromechanical ones.

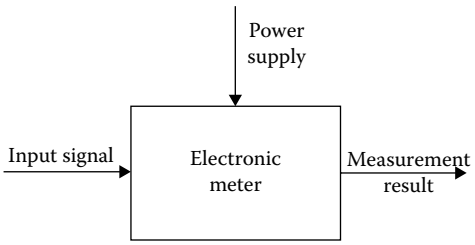


FIGURE 18.8 Electronic meter.

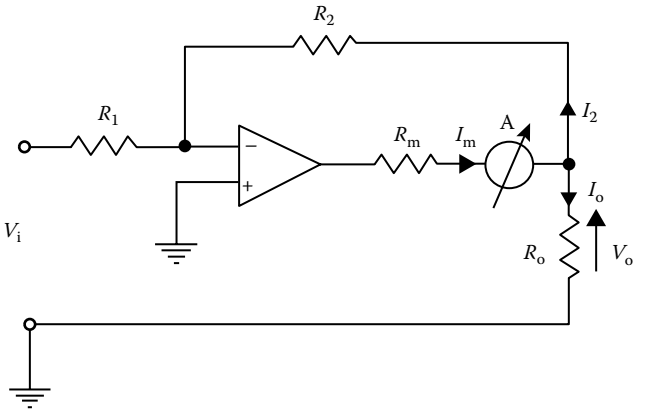


FIGURE 18.9 Electronic dc analog voltmeter schematics.

18.5.1.1 DC Analog Voltmeters

Figure 18.9 shows the circuit for an electronic dc analog voltmeter. Assuming that the operational amplifier exhibits ideal behavior, current I_m flowing in the milliammeter A is given by

$$I_m = I_o + I_2 = \frac{V_o}{R_o} + \frac{V_o}{R_2} = -V_i \frac{R_2}{R_1} \frac{R_2 + R_o}{R_2 R_o} = -\frac{V_i}{R_1}\left(1 + \frac{R_2}{R_o}\right) \tag{18.29}$$

If $R_1 = R_2$ and the same resistances are far greater than R_o, Equation 18.29 can be simplified to

$$I_m = -\frac{V_i}{R_o} \tag{18.30}$$

Equation 18.30 shows that the milliammeter reading is directly proportional to the input voltage through resistance R_o only. This means that, once the milliammeter full-scale value is set, the voltmeter full-scale value can be changed, within the dynamic range of the amplifier, by changing the R_o value. This way, the meter full-scale value can be changed without changing its input impedance.

18.5.1.2 Rectifier-Based AC Analog Voltmeters

Analog meters for ac voltages can be obtained starting from the dc analog voltmeters, with a rectifying input stage. Figure 18.10 shows how the structure in Figure 18.9 can be modified in order to realize an ac voltmeter.

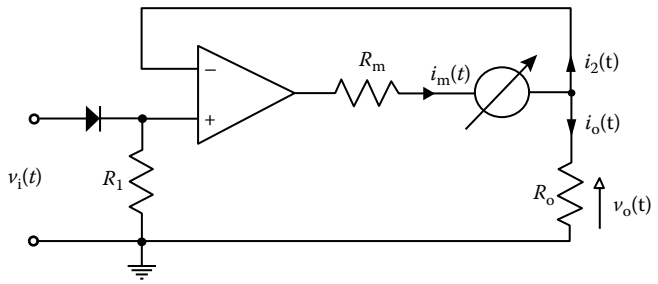


FIGURE 18.10 Electronic, rectifier-based ac analog voltmeter schematics.

Because of the high input impedance of the electronic amplifier, $i_2(t) = 0$, and the current $i_m(t)$ flowing in the milliammeter A is the same as current $i_o(t)$ flowing in the load resistance. Since the amplifier is connected in a voltage-follower configuration, the output voltage is given by

$$v_o(t) = v_i(t) \tag{18.31}$$

Due to the presence of the input diode, current $i_m(t)$ is given by

$$i_m(t) = \frac{v_i(t)}{R_o} \tag{18.32}$$

when $v_i(t) > 0$, and

$$i_m(t) = 0 \tag{18.33}$$

when $v_i(t) \leq 0$. If $v_i(t)$ is supposed to be a sine wave, the waveform of $i_m(t)$ is shown in Figure 18.11.

The dc moving-coil milliammeter measures the average value $\overline{I}_m$ of $i_m(t)$, which, under the assumption of sinusoidal signals, is related to the rms value V_i of $v_i(t)$ by

$$\overline{I}_m = \frac{2\sqrt{2}}{\pi R_o} V_i \tag{18.34}$$

The performance of the structure in Figure 18.10 can be substantially improved by considering the structure in Figure 18.12, which realizes a full-wave rectifier. Because of the presence of diodes D_1 and D_2, the output of amplifier A_1 is given by

$$v_1(t) = \begin{cases} -v_i(t) & \text{for } v_i(t) \geq 0 \\ 0 & \text{for } v_i(t) < 0 \end{cases} \tag{18.35}$$

where $v_i(t)$ is the circuit input voltage.

If capacitor C is supposed to be not connected, amplifier A_2 output voltage is

$$v_o(t) = -\left[v_i(t) + 2v_1(t)\right] \tag{18.36}$$

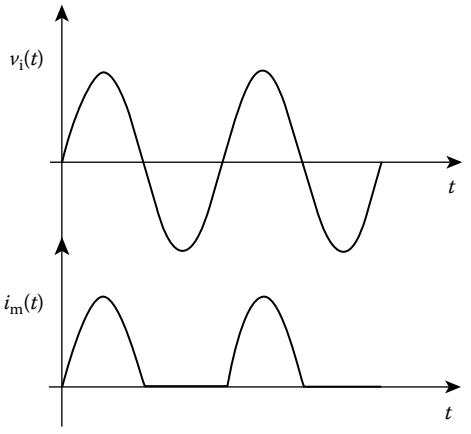


FIGURE 18.11 Signal waveforms in a rectifier-based ac analog voltmeter when the input voltage is sinusoidal.

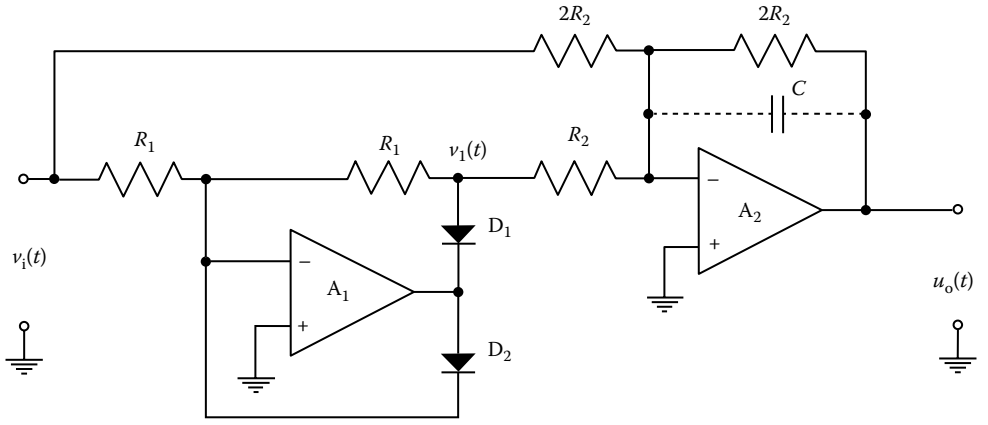


FIGURE 18.12 Electronic, full-wave rectifier-based ac analog voltmeter schematics.

that gives

$$v_o(t) = \begin{cases} v_i(t) & \text{for } v_i(t) \geq 0 \\ -v_i(t) & \text{for } v_i(t) < 0 \end{cases} \tag{18.37}$$

thus proving that the circuit in Figure 18.12 realizes a full-wave rectifier.

If $v_i(t)$ is a sine wave, the waveforms of $v_i(t)$, $v_1(t)$, and $v_o(t)$ are shown in Figure 18.13. Connecting capacitor C in the feedback loop of amplifier A_2 turns it into a first-order low-pass filter, so that the circuit output voltage equals the average value of $v_o(t)$:

$$\overline{V}_o = \overline{|v_i(t)|} \tag{18.38}$$

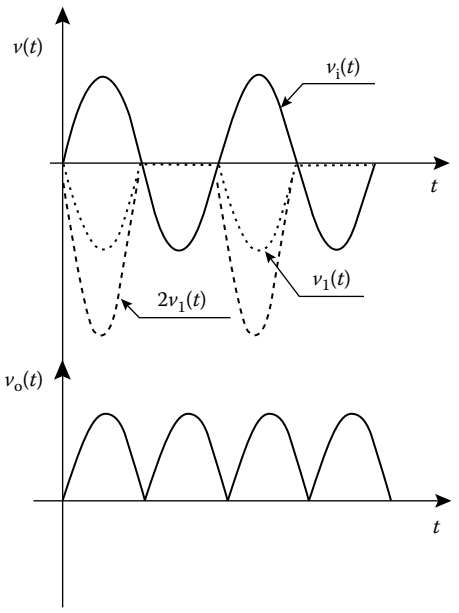


FIGURE 18.13 Signal waveforms in a full-wave rectifier-based ac analog voltmeter when the input voltage is sinusoidal.

In the case of sinusoidal input voltage with rms value V_i, the output voltage is related to this rms value by

$$\bar{V}_o = \frac{2\sqrt{2}}{\pi} V_i \tag{18.39}$$

$\bar{V}_o$ can be measured by a dc voltmeter.

Both meters in Figures 18.10 and 18.12 are actually average detectors. However, due to Equations 18.34 and 18.39, their scale can be labeled in such a way that the instrument reading gives the rms value of the input voltage, provided it is sinusoidal. When the input voltage is no longer sinusoidal, an error arises that depends on the signal form factor.

18.5.1.3 True rms Analog Voltmeters

The rms value V_i, of a periodic input voltage signal $v_i(t)$, with period T, is given by

$$V_i = \sqrt{\frac{1}{T}\int_0^T v_i^2(t)\,dt} \tag{18.40}$$

The electronic circuit shown in Figure 18.14 provides an output signal V_o proportional to the squared rms value of the input signal $v_i(t)$. The circuit section between nodes 1 and 2 is a full-wave rectifier. Hence, node 2 potential is given by

$$v_2(t) = \left|v_i(t)\right| \tag{18.41}$$

The circuit section between nodes 2 and 4 is a log multiplier. Because of the logarithmic characteristic of the feedback path due to the presence of T_1 and T_2, node 3 potential is given by

$$v_3(t) = 2k_1 \log[v_2(t)] = k_1 \log[v_2^2(t)] = k_1 \log\left[\left|v_i(t)\right|^2\right] = k_1 \log[v_i^2(t)] \tag{18.42}$$

and, due to the presence of T_3, the current flowing in node 4 is given by

$$i_4(t) = k_2 \exp\left[v_3(t)\right] = k_3 v_i^2(t) \tag{18.43}$$

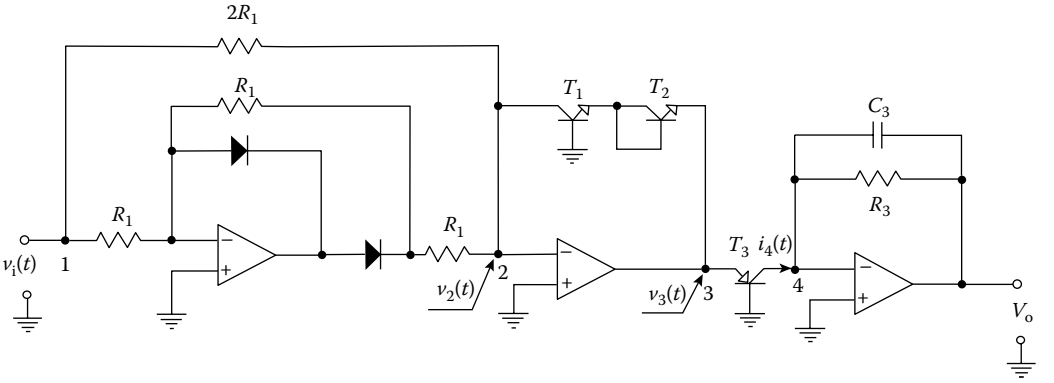


FIGURE 18.14 True rms electronic ac voltmeter schematics.

The circuit section after node 4 is a low-pass filter that extracts the dc component of the input signal. Therefore, the circuit output voltage is given by

$$V_o = \frac{k}{T}\int_0^T v_i^2(t)\mathrm{d}t = kV_i^2 \tag{18.44}$$

thus providing an output signal proportional to the squared rms value of the input signal $v_i(t)$ in accordance with Equation 18.40. Quantities k_1, k_2, and k depend on the values of the elements in the circuit in Figure 18.14. Under circuit operating conditions, their values can be considered constant, so that k_1, k_2, and k can be considered constant also.

If carefully designed, this circuit can feature an uncertainty in the range of ±1% of full scale, for signal frequencies up to 100 kHz.

18.5.2 Digital Voltmeters

A digital voltmeter (DVM) attains the required measurement by converting the analog input signal into digital and, when necessary, by discrete-time processing of the converted values. The measurement result is presented in a digital form that can take the form of a digital front-panel display or a digital output signal. The digital output signal can be coded as a decimal BCD code or a binary code.

The main factors that characterize DVMs are speed, automatic operation, and programmability. In particular, they presently offer the best combination of speed and accuracy if compared with other available voltage-measuring instruments. Moreover, the capability of automatic operations and programmability make DVMs very useful in applications where flexibility, high speed, and computer controllability are required. A typical application field is therefore that of automatically operated systems.

When a DVM is directly interfaced to a digital signal processing system and used to convert the analog input voltage into a sequence of sampled values, it is usually called an analog-to-digital converter (ADC).

DVMs basically differ in the following ways: (1) number of measurement ranges, (2) number of digits, (3) accuracy, (4) speed of reading, and (5) operating principle.

The basic measurement ranges of most DVMs are either 1 or 10 V. It is however possible, with an appropriate preamplifier stage, to obtain full-scale values as low as 0.1 V. If an appropriate voltage divider is used, it is also possible to obtain full-scale values as high as 1000 V.

If the digital presentation takes the form of a digital front-panel display, the measurement result is presented as a decimal number, with a number of digits that typically range from 3 to 6. If the digital representation takes the form of a binary-coded output signal, the number of bits of this representation typically ranges from 8 to 16, though 18 to 24 bit ADCs are now available.

The accuracy of a DVM is usually correlated to its resolution. Indeed, assigning an uncertainty lower than 0.1% of the range to a three-digit DVM makes no sense, since this is the displayed resolution of the instrument. Similarly, a poorer accuracy makes the three-digit resolution quite useless. Presently, a six-digit DVM can feature an uncertainty range, for short periods of time in controlled environments, as low as 0.0015% of reading or 0.0002% of full range.

The speed of a DVM can be as high as 1000 readings per second. When the ADC is considered, the conversion rate is taken into account instead of the speed of reading. Presently, the conversion rate for 12 bit, successive approximation ADCs can be on the order of 10 MHz. It can be in the order of 100 MHz to 1 GHz for lower resolution, flash ADCs [5].

DVMs can be divided into two main operating principle classes: the *integrating* types and the *nonintegrating* types [3]. The following sections give an example for both types.

18.5.2.1 Dual-Slope DVM

Dual-slope DVMs use a counter and an integrator to convert an unknown analog input voltage into a ratio of time periods multiplied by a reference voltage. The block diagram in Figure 18.15 shows this operating principle. The switch S1 connects the input signal to the integrator for a fixed period of time t_f. If the input voltage is positive and constant, $v_i(t) = V_i > 0$, the integrator output represents a negative-slope ramp signal (Figure 18.16). At the end of t_f, S1 switches and connects the output of the voltage reference V_R to the integrator input. The voltage reference output is negative for a positive input voltage. The integrator output starts to increase, following a positive-slope ramp (Figure 18.16). The process stops when the ramp attains the 0 V level, and the comparator allows the control logic to switch S1 again. The period of time t_v the ramp takes to increase to 0 V is variable and depends on the ramp peak value attained during period t_f.

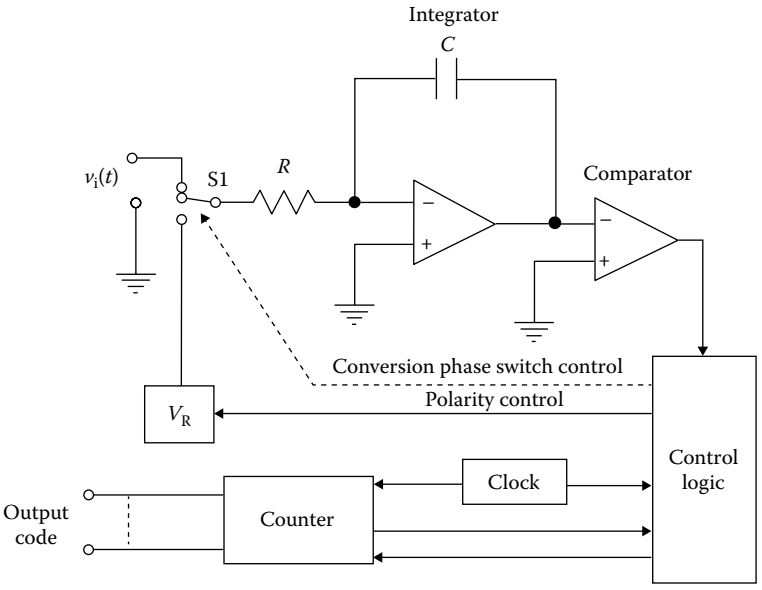


FIGURE 18.15 Dual-slope DVM schematics.

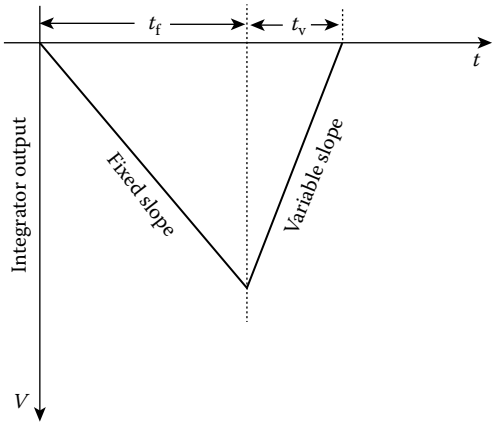


FIGURE 18.16 Integrator output signal in a dual-slope DVM.

The relationship between the input voltage V_i and the time periods t_v and t_f is given by

$$\frac{1}{RC}\int_0^{t_f} V_i \, dt = \frac{t_v}{RC} V_R \tag{18.45}$$

that, for a constant input voltage V_i, leads to

$$V_i = V_R \frac{t_v}{t_f} \tag{18.46}$$

Since the same integrating circuit is used, errors due to comparator offset, capacitor tolerances, long-term counter clock drifts, and integrator nonlinearities are eliminated. High resolutions are therefore possible, although the speed of reading is low (in the order of milliseconds).

Slowly varying voltages can be also measured by dual-slope DVMs. However, this requires that the input signal does not vary for a quantity greater than the DVM resolution during the reading time. For high-resolution DVMs, this limits the DVM bandwidth to a few hertz.

18.5.2.2 Successive Approximation ADC

The successive approximation technique represents the most popular technique for the realization of ADCs. Figure 18.17 shows the block diagram of this type of converter. The input voltage is assumed to have a constant value V_i and drives one input of the comparator. The other comparator's input is driven by the output of the digital-to-analog converter (DAC), which converts the binary code provided by the successive approximation register (SAR) into an analog voltage. Let n be the number of bits of the converter, V_R the voltage reference output, and C the code provided by the SAR. The DAC output voltage is then given by

$$V_c = \frac{C}{2^n} V_R \tag{18.47}$$

When the conversion process starts, the SAR most significant bit (MSB) is set to logic 1. The DAC output, according to Equation 18.47, is set to half the reference value and hence half the analog input full-scale range. The comparator determines whether the DAC output is above or below the input signal. The comparator output controls the SAR in such a way that, if the input signal is above the DAC output, as shown in Figure 18.18, the SAR MSB is retained and the next bit is set to logic 1.

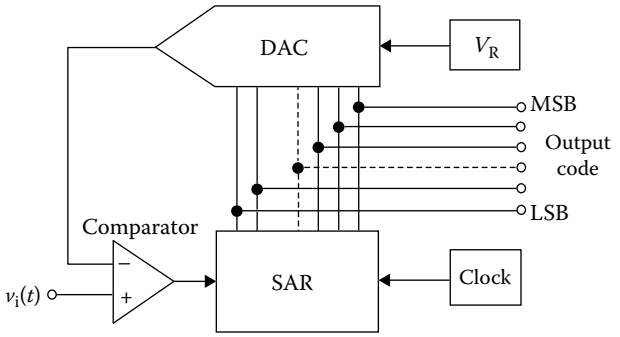


FIGURE 18.17 Successive approximation ADC schematics.

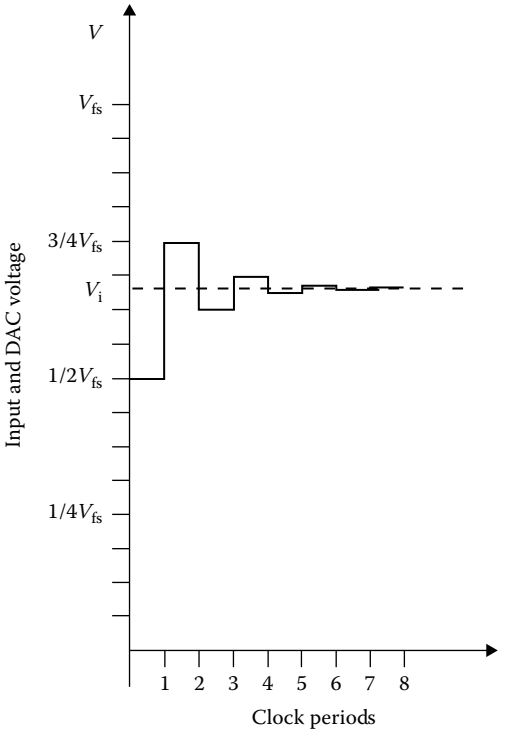


FIGURE 18.18 DAC output signal in a successive approximation ADC.

If now the input signal is below the DAC output (Figure 18.18), the last SAR bit set to logic 1 is reset to logic 0, and the next one is set to logic 1. The process goes on until the SAR least significant bit (LSB) has been set. The entire conversion process takes time $t_c = nT_c$, where T_c is the clock period. At the end of conversion, the SAR output code represents the digitally converted value of the input analog voltage V_i.

According to Equation 18.47, the ADC resolution is $V_R/2^n$, which corresponds to 1 LSB. The conversion error can be kept in the range ±½ LSB. Presently, a wide range of devices is available, with resolution from 8 to 16 bits and conversion rates from 100 μs to below 1 μs.

Varying voltages can be sampled and converted into digital by the ADC, provided the input signal does not vary by a quantity greater than $V_R/2^n$ during the conversion period t_c. The maximum frequency of an input sine wave that satisfies this condition can be readily determined starting from given values of n and t_c.

Let the input voltage of the ADC be an input sine wave with peak-to-peak voltage $V_{pp} = V_R$ and frequency f. Its maximum variation occurs at the zero-crossing time and, due to the short conversion period t_c, is given by $2\pi f t_c V_{pp}$. To avoid conversion errors, it must be

$$2\pi f t_c V_{pp} \leq \frac{V_R}{2^n} \tag{18.48}$$

Since $V_{pp} = V_R$ is assumed, this leads to

$$f \leq \frac{1}{2^n 2\pi t_c} \tag{18.49}$$

If $t_c = 1$ μs and $n = 12$, Equation 18.49 leads to $f \leq 38.86$ Hz. However, ADCs can still be employed with input signals whose frequency exceeds the value given by Equation 18.49, provided that a *sample and hold* circuit is used to keep the input voltage constant during the conversion period.

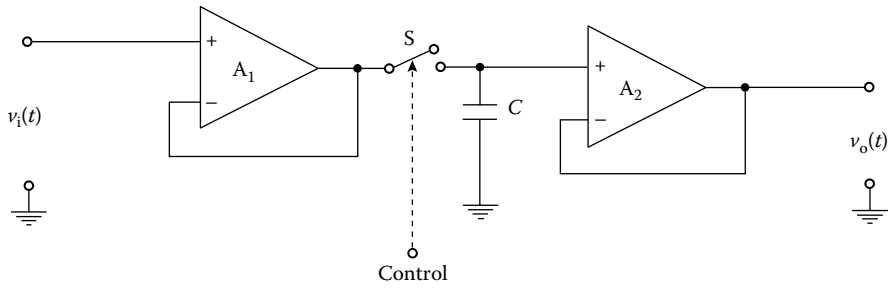


FIGURE 18.19 Sample and hold schematics.

The *sample and hold* circuit is shown in Figure 18.19. When the electronic switch S is closed, the output voltage $v_o(t)$ follows the input voltage $v_i(t)$. When switch S is open, the output voltage is the same as the voltage across capacitor *C*, which is charged at the value assumed by the input voltage at the time the switch was opened. Due to the high input impedance of the operational amplifier A_2, if a suitable value is chosen for capacitor *C*, its discharge transient is slow enough to keep the variation of the output voltage below the ADC resolution.

18.5.2.3 Delta–Sigma ADC

The modern delta–sigma (Δ–Σ) ADC combines the dual-slope technique with a DSP technique to attain the good accuracy and linearity of the dual-slope DVM with a much shorter conversion time.

The block diagram of a Δ–Σ ADC is shown in Figure 18.20. Two main blocks can be recognized: an analog modulator and a digital filter. The operations of both devices are synchronized by a high-frequency clock.

The analog modulator can be schematically represented with an integrator, a comparator, and a single-bit DAC. The comparator provides the modulator digital, single-bit output signal, called *bit stream*. The DAC provides a two-level, bipolar voltage V_{DA} that acts as the feedback signal of the modulator closed-loop structure. The absolute value of V_{DA} must be greater than the converter full-scale input voltage. It can be recognized that the modulator structure is similar, in principle, to that of the dual-slope converter of Figure 18.15.

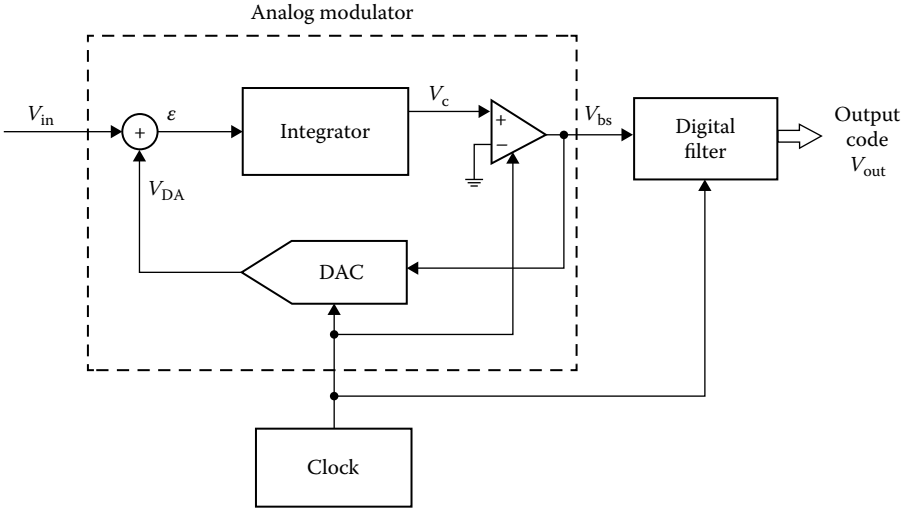


FIGURE 18.20 Δ–Σ ADC schematics.

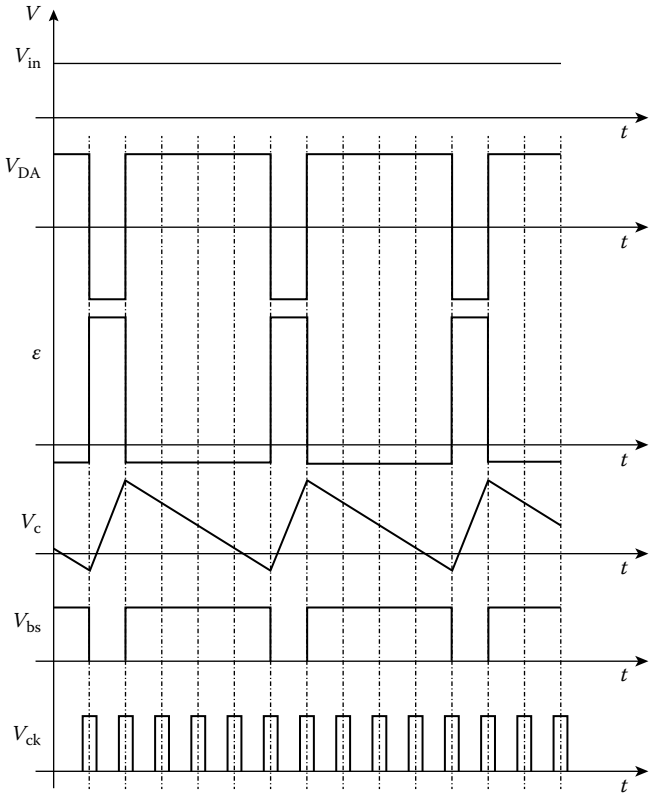


FIGURE 18.21 Operating principle of a Δ–Σ modulator.

The modulator operating principle is shown, in the time domain, in Figure 18.21. Let us suppose that the input voltage V_{in} is positive and constant and the integrator output is negative at the first clock pulse. Consequently, the comparator output V_{bs} is switched to the low level by the clock pulse. The DAC output voltage is switched to the negative $-V_{DA}$ value and subtracted to V_{in}, so that the input signal ε to the integrator is positive and high enough to force the integrator output V_c to become positive before the next clock pulse.

Consequently, the comparator output is switched to the high value at the next clock pulse, thus driving the DAC value to its positive V_{DA} value. Signal ε becomes now negative, with an absolute value that is inversely proportional to V_{in}. Therefore, the integrator output signal decreases to zero with a negative, slower slope than the positive one. Positive and negative slopes are equal only if $V_{in} = 0$. At the first clock pulse after V_c becomes negative, the whole process restarts.

For constant input signals V_{in}, the modulator output V_{bs} and comparator output V_c are periodic (Figure 18.21), with period NT_{ck}, T_{ck} being the clock period.

The relationship between V_{in} and the V_{bs} period can be obtained in a similar way as that followed to study the dual-slope DVM. It is

$$\int_0^{T_{ck}} \left(V_{in} + V_{DA}\right) dt = \int_{T_{ck}}^{(N-1)T_{ck}} \left(V_{DA} - V_{in}\right) dt \tag{18.50}$$

that leads to

$$V_{in} = \frac{N-2}{N} V_{DA} \tag{18.51}$$

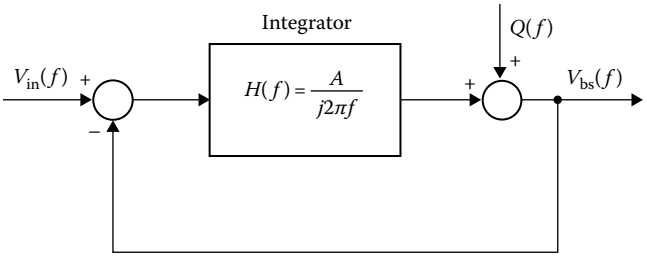


FIGURE 18.22 Linearized, frequency-domain model of a Δ–Σ modulator.

Therefore, the value of the input voltage V_{in} is related to the succession of 1 and 0 bits in the bit stream obtained as output signal V_{bs} by the comparator.

A better insight on the way the modulator works can be obtained by considering variable input signals and analyzing the modulator in the frequency domain. Figure 18.22 shows the linearized, frequency-domain model of the modulator, where the comparator effect is modeled as a quantization noise $q(t)$.

The Fourier transform $V_{bs}(f)$ of the modulator output signal is related to the Fourier transforms $V_{in}(f)$ and $Q(f)$ of the input signal and quantization noise, respectively, by the following relationship:

$$V_{bs}(f) = \frac{A}{A + j2\pi f} \cdot V_{in}(f) + \frac{j2\pi f}{A + j2\pi f} \cdot Q(f) \tag{18.52}$$

where A is the integrator constant and depends on the modulator parameters.

It can be easily recognized that the input signal and the quantization noise are processed in a quite different way by the modulator. Indeed, the first term in Equation 18.52 represents a low-pass filter, with cutoff frequency $f_c = A/2\pi$, applied to the input signal. On the other hand, the second term in Equation 18.52 represents a high-pass filter, with the same cutoff frequency f_c, applied to the quantization noise.

Therefore, the signal spectral components are separated by the quantization noise components, as shown in Figure 18.23. If the modulator parameters are such that $f_0 < f_c$, f_0 being the upper limit of the input signal band, and the clock frequency f_{ck} is significantly higher than the Nyquist frequency $2 \cdot f_0$, the input signal is not modified by the modulator, while most part of the quantization noise is moved outside the signal band, thus improving the signal-to-noise ratio inside the signal band. Because of this result, the modulator of a Δ–Σ ADC is often called *noise-shaping filter*, and its frequency response is shown in Figure 18.23.

In order to extract the input signal components from the V_{bs} bit stream modulator output, this signal has to be low-pass filtered, with cutoff frequency f_c' in the range: $f_0 < f_c' < f_c$. This is done by the digital filter in Figure 18.20. The digital filter is a finite impulse response (FIR) filter [6,7] that provides the output code as a linear combination of M previous samples of V_{bs}:

$$V_{out}(n) = \sum_{k=0}^{M} a_k \cdot V_{bs}(n-k) \tag{18.53}$$

where the filter order M and the M a_k coefficients are defined according to the desired frequency response [6,7]. Equation 18.53 shows also that, being V_{bs} coded on a single bit, the number of bits on which V_{out} is coded depends on the number of bits employed to code the a_k coefficients.

Therefore, a Δ–Σ ADC can feature a virtually unlimited number of bits. The actual limit is set by the residual noise left in the signal band by the noise-shaping filter. This residual noise can be further reduced by increasing the modulator order or, if the signal bandwidth is much lower than

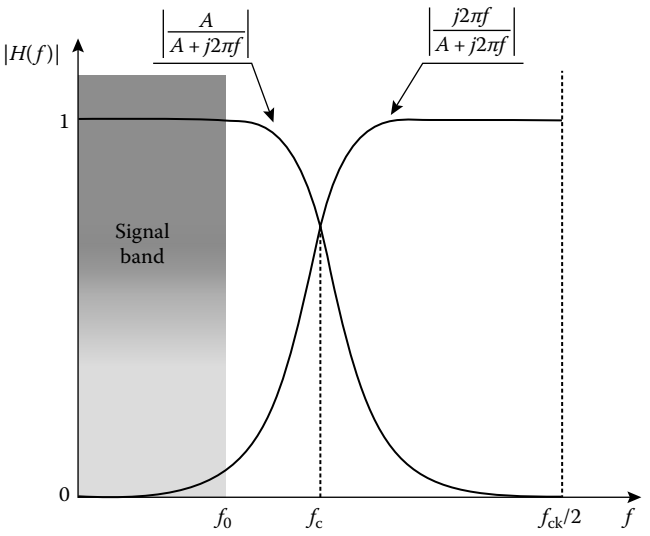


FIGURE 18.23 Frequency response of the noise-shaping filter.

half the clock frequency (which represents the actual sampling frequency), by applying decimation techniques on the V_{out} sequence.

Nowadays, 24 bit Δ–Σ ADCs are available, with sampling frequency up to 200 kHz. Higher sampling frequencies can be obtained with a lower resolution.

18.5.2.4 AC Digital Voltmeters

True rms ac voltmeters with digital reading can be obtained using an electronic circuit like the one in Figure 18.14 to convert the rms value into a dc voltage signal and measuring it by means of a DVM.

However, this structure cannot actually be called a digital structure, because the measurement is attained by means of analog processing of the input signal.

A more modern approach, totally digital, is shown in Figure 18.24. The input signal $u_i(t)$ is sampled at constant sampling rate f_s and converted into digital by the ADC. The digital samples are stored in the memory of the digital signal processor (DSP) and then processed in order to evaluate Equation 18.40 in a numerical way. Assuming that the input signal is periodic, with period T, and its frequency spectrum is upper limited by harmonic component of order N, the sampling theorem is satisfied if at least $(2N + 1)$ samples are taken over period T in such a way that $(2N + 1)T_s = T$, $T_s = 1/f_s$ being the sampling period [6,7]. If $u_i(kT_s)$ is the kth sample, the rms value of the input signal is given by, according to Equation 18.40,

$$V^2 = \sqrt{\frac{1}{2N+1}\sum_{k=0}^{2N} v_i^2(kT_s)} \tag{18.54}$$

This approach can feature a relative uncertainty as low as ±0.1% of full scale, with an ADC resolution of 12 bits. The instrument bandwidth is limited to half the sampling frequency, according to the

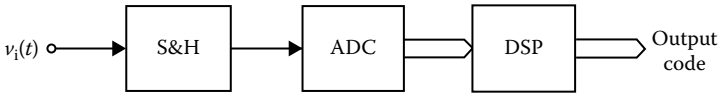


FIGURE 18.24 Block diagram of a modern digital meter.

sampling theorem. When modern ADCs and DSPs are employed, a 500 kHz to 1 MHz bandwidth can be obtained. Wider bandwidths can be obtained but with a lower ADC resolution and hence with a lower accuracy.

18.5.2.5 Frequency Response of AC Voltmeters

When the frequency response of ac voltmeters is taken into account, a distinction must be made between the analog voltmeters (both electromechanical and electronic) and DVMs, based on DSP techniques.

The frequency response of the analog meters is basically a low-pass response, well below 1 kHz for most electromechanical instruments and up to hundreds of kilohertz for electronic instruments.

When digital, DSP-based meters are concerned, the sampling theorem and aliasing effects must be considered. To a first approximation, the frequency response of a digital meter can be considered flat as long as the frequency-domain components of the input signal are limited to a frequency band narrower than half the sampling rate. If the signal components exceed this limit (the so-called Nyquist frequency), the aliasing phenomenon occurs [6]. Because of this phenomenon, the signal components at frequencies higher than half the sampling rate are folded over the lower-frequency components, changing them. Large measurement errors occur under this situation.

To prevent the aliasing, a low-pass filter must be placed at the input stage of any digital meter. The filter cutoff frequency must ensure that all frequency components above half the sampling rate are negligible. If the low-pass, antialiasing filter is used, the digital DSP-based meters feature a low-pass frequency response also.

References

1. M. B. Stout, *Basic Electrical Measurements*, Englewood Cliffs, NJ: Prentice Hall, 1960.
2. I. F. Kinnard, *Applied Electrical Measurements*, New York: John Wiley & Sons, 1956.
3. B. M. Oliver and J. M. Cage, *Electronic Measurements and Instrumentation*, London, U.K.: McGraw-Hill, Inc., 1975.
4. T. T. Lang, *Electronics of Measuring Systems*, New York: John Wiley & Sons, 1987.
5. Analog Devices, *Analog-Digital Conversion Handbook*, Englewood Cliffs, NJ: Prentice Hall, 1986.
6. A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, Englewood Cliffs, NJ: Prentice Hall, 1975.
7. G. D'Antona and A. Ferrero, *Digital Signal Processing for Measurement Systems—Theory and Applications*, New York: Springer, 2006.

19

Oscilloscope Voltage Measurement

Jerry Murphy
Hewlett-Packard

19.1 Introduction

Engineers, scientists, and other technical professionals around the world depend on oscilloscopes as one of the primary voltage measuring instruments. This is an unusual situation because the oscilloscope is not the most accurate voltage measuring instrument usually available in the lab. It is the graphical nature of the oscilloscope that makes it so valued as a measurement instrument—not its measurement accuracy.

The oscilloscope is an instrument that presents a graphical display of its input voltage as a function of time. It displays voltage waveforms that cannot easily be described by numerical methods. For example, the output of a battery can be completely described by its output voltage and current. However, the output of more complex signal source needs additional information such as frequency, duty cycle, peak-to-peak amplitude, overshoot, preshoot, rise time, fall time, and more to be completely described. The oscilloscope, with its graphical presentation of complex waveforms, is ideally suited to this task. It is often described as the "screwdriver of the electronic engineer" because the oscilloscope is the most fundamental tool that technical professionals apply to the problem of trying to understand the details of the operation of their electronic circuit or device. So, what is an oscilloscope?

The oscilloscope is an electronic instrument that presents a high-fidelity graphical display of the rapidly changing voltage at its input terminals.

The most frequently used display mode is voltage versus time. This is not the only display that could be used, nor is it the display that is best suited for all situations. For example, the oscilloscope could be called on to produce a display of two changing voltages plotted one against the other, such as a Lissajous display. To accurately display rapidly changing signals, the oscilloscope is a high-bandwidth device. This means that it must be capable of displaying the high-order harmonics of the signal being applied to its input terminals in order to correctly display that signal.

19.2 Oscilloscope Block Diagram

The oscilloscope contains four basic circuit blocks: the vertical amplifier, the time base, the trigger, and the display. This section treats each of these in a high-level overview. Many textbooks exist that cover the details of the design and construction of each of these blocks in detail [1]. This discussion will cover these blocks in enough detail so that readers can construct their own mental model of how their operation affects the application of the oscilloscope for their voltage measurement application. Most readers of this book have a mental model of the operation of the automatic transmission of an automobile that is sufficient for its successful operation but not sufficient for the overhaul or redesign of that component. It is the goal of this section to instill that level of understanding in the operation of the oscilloscope. Those readers who desire a deeper understanding will get their needs met in later sections.

Of the four basic blocks of the oscilloscope, the most visible of these blocks is the display with its *cathode-ray tube* (CRT). This is the component in the oscilloscope that produces the graphical display of the input voltage, and it is the component with which the user has the most contact. Figure 19.1 shows that the input signal is applied to the vertical axis of a CRT. This is the correct model for an analog oscilloscope, but it is overly simplified in the case of the digital oscilloscope. The important thing to learn from this diagram is that the input signal will be operated on by the oscilloscope's vertical axis circuits so that it can be displayed by the CRT. The differences between the analog and digital oscilloscope are covered in sections to follow.

The *vertical amplifier* conditions the input signal so that it can be displayed on the CRT. The vertical amplifier provides controls of volts per division, position, and coupling, allowing the user to obtain the desired display. This amplifier must have a high enough bandwidth to ensure that all of the significant frequency components of the input signal reach the CRT.

The *trigger* is responsible for starting the display at the same point on the input signal every time the display is refreshed. It is the stable display of a complex waveform that allows the user of an oscilloscope to make judgments about that waveform and its implications as to the operation of the device under test.

The final piece of the simplified block diagram is the *time base*. This circuit block is also known as the horizontal system in some literature. The time base is the part of the oscilloscope that causes the input signal

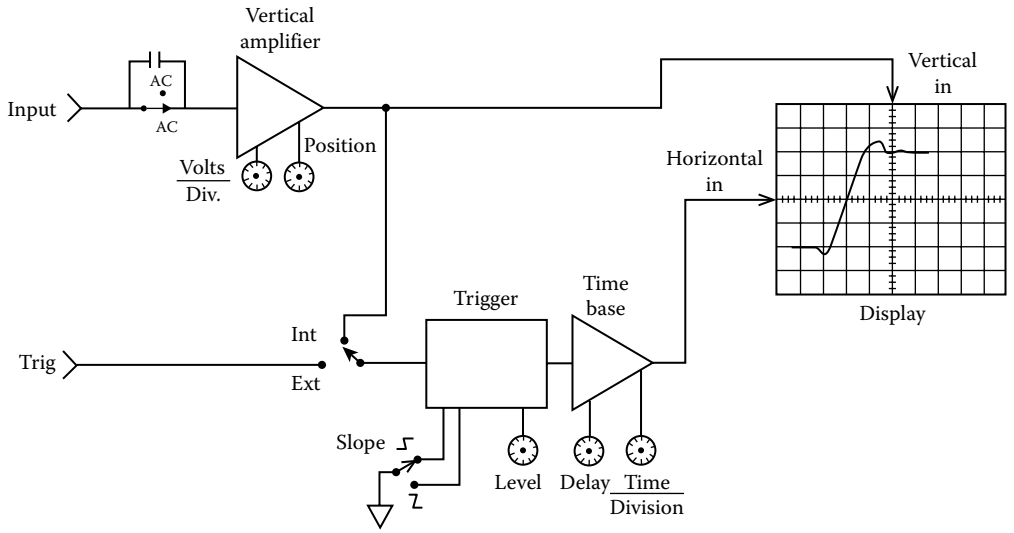


FIGURE 19.1 Simplified oscilloscope block diagram that applies to either analog or digital oscilloscopes. In the case of the digital oscilloscope, the vertical amplifier block will include the ADC and high-speed waveform memory. For the analog scope, the vertical block will include delay lines with their associated drivers and a power amplifier to drive the CRT plates.

to be displayed as a function of time. The circuitry in this block causes the CRT beam to be deflected from left to right as the input signal is being applied to the vertical deflection section of the CRT. Controls for time per division and position (or delay) allow the user of the oscilloscope to adjust the display for the most useful display of the input signal. The time-per-division controls of most oscilloscopes provide a wide range of values, ranging from a few nanoseconds (10^{-9} s) to seconds per division. To get a feeling for the magnitude of the dynamic range of the oscilloscope's time base settings, keep in mind that light travels about 1 m in 3 ns.

19.3 Oscilloscope as a Voltage Measurement Instrument

That the oscilloscope's vertical axis requires a wide-bandwidth amplifier and its time base is capable of displaying events that are as short as a few nanoseconds apart indicates that the oscilloscope can display rapidly changing voltages. Voltmeters, on the other hand, are designed to give their operator a numeric readout of steady-state or slowly changing voltages. Voltmeters are not well suited for displaying voltages that are changing levels very quickly. This can be better understood by the examination of the operation of a voltmeter as compared to that of an oscilloscope. The analog voltmeter uses the magnetic field produced by current flowing through a coil to move the pointer against the force of a spring. This nearly linear deflection of the voltmeter pointer is calibrated by applying known standard voltages to its input. Therefore, if a constant voltage is applied to the coil, the pointer will move to a point where the magnetic force being produced by the current flowing in its coil is balanced by the force of the spring. If the input voltage is slowly changing, the pointer will follow the changing voltage. This mechanical deflection system limits the ability of this measurement device to the measurement of steady-state or very low-frequency changes in the voltage at its input terminals. Higher frequency voltmeters depend on some type of conversion technique to change higher frequencies to a dc signal that can be applied to the meter's deflection coil. For example, a diode is used to rectify ac voltages to produce a dc voltage that corresponds to the average value of the ac voltage at the input terminals in average responding ac voltmeters.

The digital voltmeter is very much like the analog meter except that the mechanical displacement of the pointer is replaced with a digital readout of the input signal. In the case of the digital voltmeter, the input signal is applied to an analog-to-digital converter (ADC) where it is compared to a reference voltage and digitized. This digital value of the input signal is then displayed in a numerical display. The ADC techniques applied to voltmeters are designed to produce very accurate displays of the same signals that were previously measured with analog meters. The value of a digital voltmeter is its improved measurement accuracy as compared to that of its analog predecessors.

The oscilloscope will display a horizontal line displaced vertically from its zero-voltage level when a constant or dc voltage is applied to its input terminals. The magnitude of this deflection of the oscilloscope's beam vertically from the point where it was operating with no input being applied is how the oscilloscope indicates the magnitude of the dc level at its input terminals. Most oscilloscopes have a graticule as a part of their display, and the vertical axis of the scope is calibrated in volts per division of the graticule. As one can imagine, this is not a very informative display of a dc level, and perhaps a voltmeter with its numeric readout is better suited for such applications.

There is more to the scope–voltmeter comparison than is obvious from the previous discussion. That the oscilloscope is based on a wide-bandwidth data-acquisition system is the major difference between these two measurement instruments. The oscilloscope is designed to produce a high-fidelity display of rapidly changing signals. This puts additional constraints on the design of the oscilloscope's vertical system that are not required in the voltmeter. The most significant of these constraints is that of a constant group delay. This is a rather complex topic that is usually covered in network analysis texts. It can be easily understood if one realizes the effect of group delay on a complex input signal.

Figure 19.2 shows such a signal. The amplitude of this signal is a dc level, and the rising edge is made up of a series of high-frequency components. Each of these high-frequency components is a sine wave of specific amplitude and frequency. Another example of a complex signal is a square wave with a frequency of 10 MHz. This signal is made up of a series of odd harmonics of that fundamental frequency. These harmonics are sine

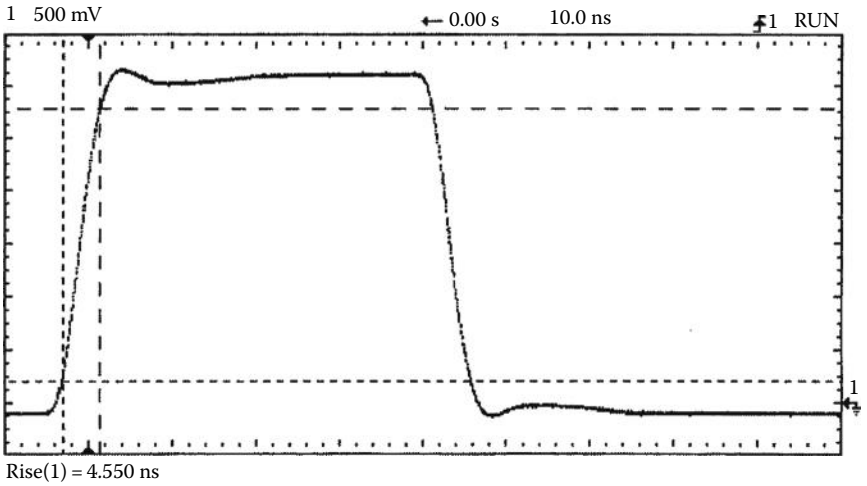


FIGURE 19.2 A typical complex waveform. This waveform is described by measurements of its amplitude, offset, rise time, fall time, overshoot, preshoot, and droop.

waves of frequencies of 10, 30, 50, 70 MHz, etc. So, the oscilloscope must pass all of these high-frequency components to the display with little or no distortion. Group delay is the measure of the propagation time of each component through the vertical system. A constant group delay means that each of these components will take the same amount of time to propagate through the vertical system to the CRT, independent of their frequencies. If the higher order harmonics take more or less time to reach the scope's deflection system than the lower harmonics, the resulting display will be a distorted representation of the input signal. Group delay (in seconds) is calculated by taking the first derivative of phase-versus-frequency response (in radians/(l/s) of an amplifier). If the amplifier has a linearly increasing phase shift with frequency, the first derivative of its phase response will be a horizontal line corresponding to the slope of the phase plot (in seconds). Amplifier systems that have a constant group delay are known as Gaussian amplifiers. They have this name because their pass band shape resembles that of the bell curve of a Gaussian distribution function (Figure 19.3). One would think that the oscilloscope's vertical amplifier should have a flat frequency response, but this is not the case because such amplifiers have nonconstant group delay [1].

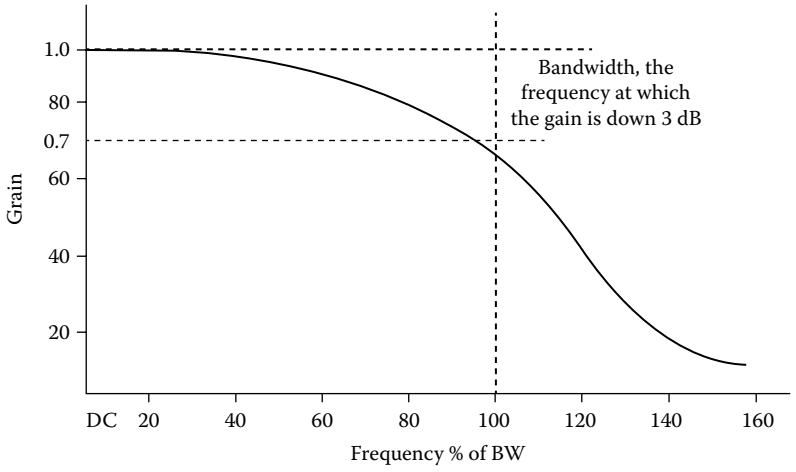


FIGURE 19.3 The Gaussian frequency response of the oscilloscope's vertical system, which is not flat in its pass band. Amplitude measurements made at frequencies greater than 20% of the scope's bandwidth will be in error.

The oscilloscope's bandwidth specification is based on the frequency where the vertical deflection will be –3 dB (0.707) of the input signal. This means that if a constant 1 V sine wave is applied to the oscilloscope's input and the signal's frequency is adjusted to higher and higher frequencies, the oscilloscope's bandwidth will be that frequency where its display of the input signal has been reduced to be 0.707 V. Noticeable errors in amplitude measurements will start at 20% of the scope's bandwidth. The oscilloscope's error-free display of complex waveforms gives it poor voltage accuracy. For the measurement of dc and single-frequency signals such as sine waves, other instruments can produce more accurate measurements.

Conclusion: The voltmeter makes the most accurate measurements of voltages that are dc, slowly changing, or can be converted to a dc analog of their ac content. The oscilloscope is not the most accurate voltage measurement instrument, but it is well suited to measurements of voltages that are changing very rapidly as a function of time. Oscilloscopes are the instrument of choice for observing and characterizing these complex voltages.

19.3.1 Analog or Digital

The world of oscilloscopes is divided into two general categories: analog and digital. The first oscilloscopes were analog. These products are based on the direct-view vector cathode-ray tube (DVVCRT or CRT for short). The analog oscilloscope applies the input signal to the vertical deflection plates of the CRT where it causes the deflection of a beam of high-energy electrons moving toward the phosphor-coated faceplate. The electron beam generates a lighted spot where it strikes the phosphor. The intensity of the light is directly related to the density of the electrons hitting a given area of the phosphor. Because this analog operation is not based on any digitizing techniques, most people have little trouble creating a very accurate and simple mental model in their minds of its operation.

The analog oscilloscope produces a display of the input signal that is bright and easy to see under most conditions. It can also contain as many as 16 shades of gray-scale information. This means that an event that occurs less frequently will appear at a lower intensity in the display than another event that occurs more frequently. This oscilloscope does not produce a continuous display of the input signal. It is blind during retrace and trigger hold-off times. Because the display depends on the production of visible light from the phosphor being excited by an electron beam, the display must be refreshed frequently. This makes the analog oscilloscope a low-dead-time display system that can follow rapidly changing signals. Also, there is little lag time in front-panel control settings.

The analog oscilloscope is not without its shortcomings. The strength of the analog oscilloscope is its CRT, but this is also the source of its weaknesses. The biggest problem with analog scopes is their dependence on a display that is constantly being refreshed. This means that these scopes do not have any waveform storage. If the input signal fails to repeat frequently, the display will simply be a flash of light when the beam sweeps by the phosphor. If the signal's repetition rate falls below 100 Hz, the display will flicker annoyingly. Figure 19.4 shows a plot of the range of an input signal's repetition frequency range from a single-shot event to the full bandwidth of a scope versus the scope's sweep speeds. The result is a map of the scope's operational area. Figure 19.4 shows that the analog oscilloscope fails to map onto the full range of possible input signals and sweep speeds.

Another problem of the analog oscilloscope is its inability to display information ahead of its trigger. This is a problem in applications where the only suitable trigger is at the end of the event of interest. Another limitation of analog scopes is their timing accuracy. The time base of the analog scope is based on the linearity of a voltage ramp. There are other sources of errors in the analog oscilloscope's horizontal axis, but the sweep nonlinearity is the major contributor. This results in these scopes having a timing accuracy of typically ±3% of their full-scale setting. Therefore, if the time base is set to 100 ns/div, in order to view a 100 ns wide pulse, the full scale will be 1000 ns or 1 μs. The accuracy of this pulse width measurement will be ±30 ns or ±30% of the pulse width!

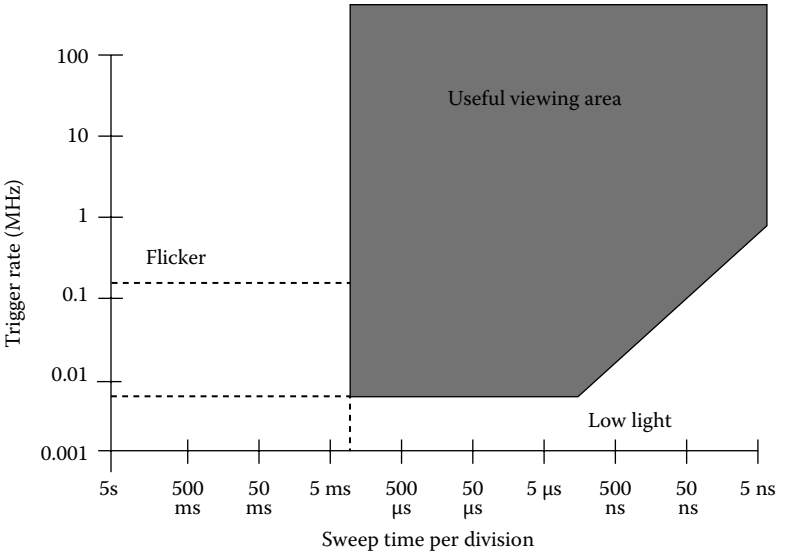


FIGURE 19.4 The operating range of the analog oscilloscope. This is a plot of input signal repetition rate from the lower limit of single shot to the full bandwidth of the scope plotted against sweep speed. The shaded area is the area where the analog oscilloscope will produce a usable display.

The digital oscilloscope or digital storage oscilloscope (DSO) differs from its analog counterpart in that the input signal is converted to digital data and therefore it can be managed by an embedded microprocessor. The waveform data can have correction factors applied to remove errors in the scope's acquisition system and can then be stored, measured, and/or displayed. That the input signal is converted from analog to digital and manipulations are performed on it by a microprocessor results in people not having a good mental model of the digital oscilloscope's operation. This would not be a problem except for the fact that the waveform digitizing process is not totally free from errors, and a lack of a correct mental model of the scope's operation on the part of its user can increase the odds of a measurement error. To make matters worse, various manufacturers of these products make conflicting claims, making it easy to propagate incorrect mental models of the digital scope's operation. It is the intention of this presentation to give the information needed to create a mental model of the operation of these devices that enable the user to perform error-free measurements with ease.

The DSO offers many advantages over its analog counterpart. The first is accuracy. The voltage measurement accuracy of the digital oscilloscope is better than that of an analog scope because the microprocessor can apply correction factors to the data to correct for errors in the calibration of the scope's vertical system. The timing accuracy of a digital oscilloscope is an order of magnitude better than that of an analog scope. The digital scope can store the waveform data for comparison to other test results or uploading to a computer for analysis or project documentation. The digital oscilloscope does not depend on the input signal being continuously updated to produce an easily viewable display. A single-shot event is displayed at the same brightness level as a signal that repeats in time periods corresponding to the full bandwidth of the scope.

The disadvantages of the digital oscilloscope are its more complex operation, aliasing, and display performance. The analog-to-digital conversion process [1] is used to convert the input signal into a series of discrete values, or samples, uniformly spaced in time, which can be stored in memory. Voltage resolution is determined by the total number of codes that can be produced. A larger number permit a smoother and more accurate reproduction of the input waveform but increase both the cost and difficulty in achieving a high sample frequency. Most digital oscilloscopes provide 8-bit resolution in their ADC. As the ADC's sampling speed is increased, the samples will be closer together, resulting in smaller gaps in the waveform record.

All digital scopes are capable of producing an aliased display. Some models are more prone to this problem than others, but even the best will alias under the right conditions. An alias is a lower-frequency false reproduction of the input signal resulting from undersampling, that is, sampling less than the Nyquist frequency. The display of the digital scope is based on computer display technology. This results in a display that is very bright and easy to see, even under conditions where an analog scope would have difficulty in producing a viewable display. The disadvantage of the digital scope's display is its lower horizontal resolution. Most of the scopes on the market have a raster scan display with a resolution of 500 lines, less than half the resolution of an analog scope's display. This is not a problem in most applications. It could become a factor where very complex waveforms, such as those found in TV systems, are being analyzed. Many digital scopes have display systems that exhibit large dead or blind times. Scopes based on a single CPU will be able to display their waveform data only after the CPU has finished all of its operations. This can result in a display that is unresponsive to front-panel control inputs as well as not being able to follow changes in the input signal.

Table 19.1 shows that both analog and digital oscilloscopes have relative advantages and disadvantages. All the major producers of oscilloscopes are pushing the development of digital scopes in an attempt to overcome their disadvantages. A few manufacturers produce scopes that are both analog and digital. These products appear to have the best of both worlds; however, they have penalties with respect to both cost and complexity of operation.

Digital systems place additional demands on the oscilloscope that exceed the capabilities of the analog scope. For example, often, in digital electronic systems, there is a need to view fast events that occur at very slow or infrequent rates. Figure 19.4 shows that these events fail to be viewable on analog scopes. Another common problem with digital systems is the location of trigger events. Often, the only usable trigger is available at the end of the event being viewed. Analog scopes can only display events that occur after a trigger event. The rapid growth of digital electronics in the late 1990s is being attributed to the lowering of the cost of single-chip microcontrollers. These devices, which contain a complete microprocessor on one integrated circuit, are responsible for the "electronics everywhere" phenomenon, where mechanical devices are becoming electronic as well as those devices that were previously electrical in nature. In 1996, Hewlett-Packard introduced a class of oscilloscope designed to meet the unique needs of the microcontroller-based applications. This class of oscilloscope is known as the mixed-signal oscilloscope or MSO [2].

TABLE 19.1 Comparison of Analog and Digital Oscilloscopes

	Analog Oscilloscope	Digital Oscilloscope
Operation	Simple	Complex
Front-panel controls	Direct access knobs	Knobs and menus
Display	Real-time vector	Digital raster scan
Gray scales	>16	>4
Horizontal resolution	>1000 lines	500 lines
Dead time	Short	Can be long
Aliasing	No	Yes
Voltage accuracy	±3% of full scale	±3% of full scale
Timing accuracy	±3% of full scale	±0.01% of full scale
Single-shot capture	None	Yes
Glitch capture	Limited	Yes
Waveform storage	None	Yes
Pretrigger viewing	None	Yes
Data out to a computer	No	Yes

19.4 Voltage Measurements

Voltage measurements are usually based on comparisons of the waveform display to the oscilloscope's graticule. Measurements are made by counting the number of graticule lines between the end points of the desired measurement and then multiplying that number by the sensitivity setting. This was the only measurement available to most analog scope users, and it is still used by those performing troubleshooting with their digital scope as a time-saving step. (Some late-model analog oscilloscopes incorporate cursors to enhance their measurement ability.) For example, a waveform that is 4.5 divisions high at a vertical sensitivity of 100 mV/div would be 450 mV high.

Switching the scope's coupling between ac and dc modes will produce a vertical shift in the waveform's position that is a measurement of its dc component. This technique can be applied to either analog or digital scopes. Simply note the magnitude of the change in waveform position and multiply by the channel's sensitivity.

Additional measurements can be performed with an analog oscilloscope, but they usually require more skill on the part of the operator. For example, if the operator can determine the location of the top and base of a complex waveform, its amplitude can be measured. Measurements based on percentages can be made using the scope's vernier to scale the waveform so that its top and bottom are five divisions apart. Then, each division represents 20% of the amplitude of the waveform being studied. The use of the vernier, which results in the channel being uncalibrated, prevents performance of voltage measurements. Many analog scopes have a red light to warn the operator that the scope is uncalibrated when in vernier mode.

The digital oscilloscope contains an embedded microprocessor that automates the measurement. This measurement automation is based on a histogramming technique, where a histogram of all the voltage levels in the waveform is taken from the oscilloscope's waveform data. The histogram is a plot of the voltage levels in the waveform plotted against the number of samples found at each voltage level. Figure 19.5 shows the histogramming technique being applied to the voltage measurements of complex waveforms.

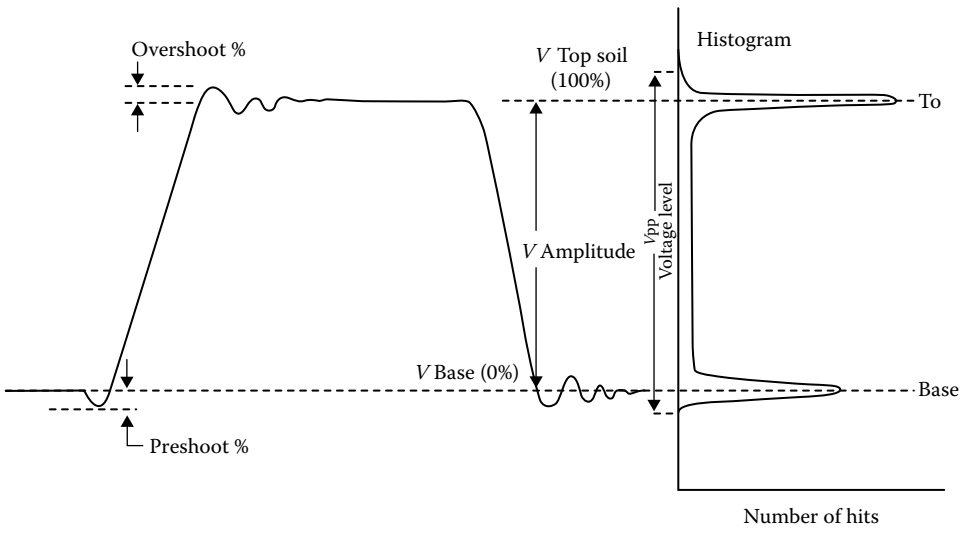


FIGURE 19.5 Voltage histograms as applied by a digital oscilloscope. The complex waveform is measured by the use of the voltage histogram. This histogram is a plot of each voltage level in the display and the number of data points at that level.

19.5 Understanding the Specifications

The oscilloscope's vertical accuracy is one place that a person's mental model of the scope's operation can lead to measurement trouble. For example, the oscilloscope's vertical axis has a frequency response that is not flat across its pass band. However, as noted earlier, the scope has a Gaussian frequency response to produce the most accurate picture of complex signals. This means that the oscilloscope's accuracy specification of ±3% is a dc-only specification. If one were to attempt to measure the amplitude of a signal whose frequency is equal to the bandwidth of the scope, one would have to add another 29.3% to the error term, for a total error of ±32.3%. This is true for both analog and digital oscilloscopes. This limitation can be overcome by carefully measuring the frequency response of the oscilloscope's vertical channels. One will need to repeat this process every time the scope is serviced or calibrated, because the various high-frequency adjustments that may need to be made in the scope's vertical axis will affect the scope's frequency response. One is probably asking, why don't the manufacturers do this for me? The answer is twofold. The first is cost, and the second is that this is not the primary application of an oscilloscope. There are other instruments that are much better suited to the measurement of high-frequency signals. The spectrum analyzer would be this author's first choice.

Additionally, the vertical accuracy is a full-scale specification. This means that at 1 V/div, the full-scale value is typically 8 V. The measurement error for a scope with a ±3% specification under these conditions will be ±0.24 V. If the signal being measured is only 1 V high, the resulting measurement will be ±24% of reading. Check the manual for the scope being used, as some manufacturers will specify full-scale as being 10 or even 10.2 divisions. This will increase the error term because the full-scale term is larger.

In digital oscilloscopes, the vertical accuracy is often expressed as a series of terms. These attempt to describe the analog and digital operations the scope performs on the input signal. Terms might include digitizing resolution, gain, and offset (sometimes called as position). They also might be called out as single- and dual-cursor accuracies. The single-cursor accuracy is a sum of all three terms. In the dual-cursor case, where the voltage measurement is made between two voltage cursors, the offset term will cancel out, leaving only the digitizing resolution and gain errors. For example, the Hewlett-Packard model 54603B has a single-cursor accuracy specification of ±1.2% of full scale, ±0.5% of position value, and a dual-cursor specification of ±0.4% of full scale.

Hint: Always try to make the voltage measurements on the largest possible vertical and widest possible display of the signal.

The horizontal accuracy specifications of analog and digital scopes are very different; however, both are based on a full-scale value. In the analog scope, many manufacturers limit accuracy specifications to only the center eight divisions of their display. This means that a measurement of a signal that starts or ends in either the first or ninth graticule will be even more error prone than stated in the scope's specifications. To the best of this author's knowledge, this limitation does not apply to digital scopes. The horizontal specifications of digital scopes are expressed as a series of terms. These might include the crystal accuracy, horizontal display resolution, and trigger placement resolution. These can be listed as cursor accuracy. For example, the Hewlett-Packard model 54603B has a horizontal cursor accuracy specification of ±0.01% ±0.2% full-scale ±200 ps. In this example, the first term is the crystal accuracy, the second is the display resolution (500 lines), and the final term is twice the trigger placement error. By comparing the analog and digital scopes' horizontal specifications, it can be seen that in either case, the measurement is more accurate if it can be made at full screen. The digital scope is more accurate than its analog counterpart.

Digital scopes also have acquisition system specifications. Here is another place where the operator's mental model of the operation of a digital scope can produce measurement errors. All manufacturers of digital scopes specify the maximum sampling speed of their scope's acquisition system

as well as its memory depth and number of bits. The scope's maximum sampling speed does not apply to all sweep speeds; only memory depth and number of bits applies to all sweep speeds. The scope's maximum sampling speed applies only to its fastest sweep speeds.

The complexity of the digital scope results from the problem of having to sample the input. There is more to be considered than Nyquist's sampling theorem in the operation of a digital scope. For example, how does the scope's maximum sampling rate relate to the smallest time interval that the scope can capture and display? A scope that samples at 100 MSa/s takes a sample every 10 ns; therefore, in principle, it cannot display any event that is less than 10 ns wide because that event will fall between the samples. In practice, however, this limit can—under certain circumstances—be extended. If the scope is operating in an "equivalent time" or "random repetitive" mode and if the signal is repetitive, even if very infrequently, the scope will be able to capture any event that is within its vertical system bandwidth. Figure 19.6 shows an infrequently occurring pulse that is 25 ns wide embedded into a data stream being captured and displayed on an oscilloscope with a maximum sampling speed of 20 MSa/s (sampling interval of 50 ns). Figure 19.6b shows this pulse at a faster sweep speed. An analog scope would produce a similar display of this event, with the infrequent

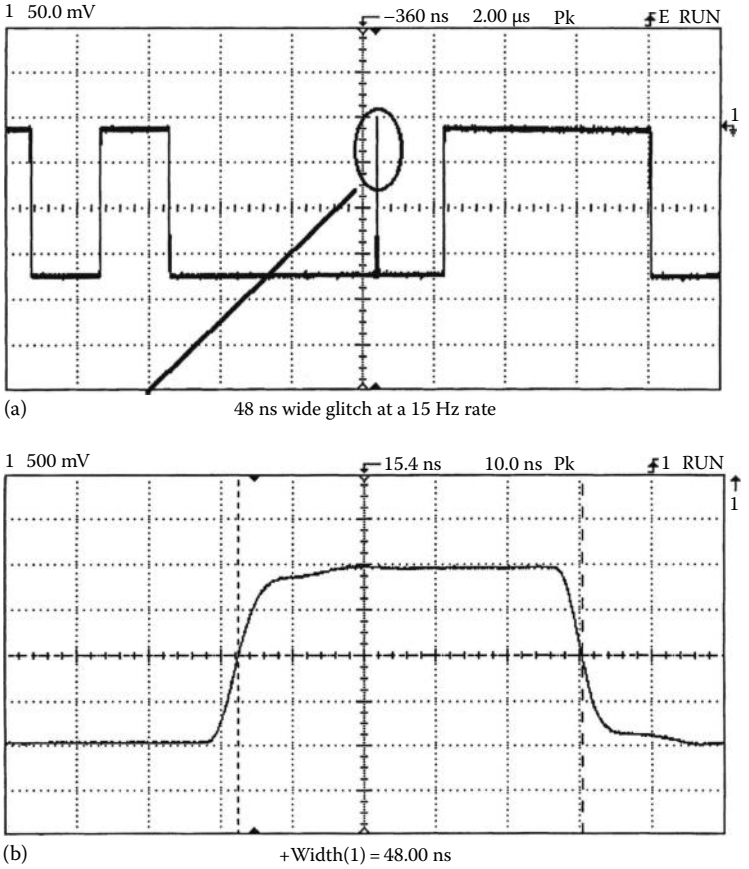


FIGURE 19.6 An infrequently occurring event as displayed on a digital oscilloscope with random repetitive sampling; (a) the event embedded in a pulse train, and (b) shows the same event at a faster sweep speed. The fact that the waveform baseline is unbroken under the narrow pulse indicates that it does not occur in every sweep. The width of this pulse is less than half the scope's sampling period in (b). Both traces are from a Hewlett-Packard model 54603B dual-channel 60 MHz scope.

event being displayed at a lower intensity than the rest of the trace. Notice that the infrequent event does not break the baseline of the trace.

The correct mental model of the digital scope's ability to capture signals needs to be based on the scope's bandwidth, operating modes, and timing resolution. It is the timing resolution that tells the operator how closely spaced the samples can be in the scope's data record.

The most common flaw in many mental models of the operation of a digital scope is related to its maximum sampling speed specification. As noted, the maximum sampling speed specification applies only to the scope's fastest sweep speeds. Some scope manufacturers will use a multiplex A/D system that operates at its maximum sampling speed only in single-channel mode. The scope's memory depth determines its sampling speed at the sweep speed being used for any specific measurement. The scope's memory depth is always equal to the scope's horizontal full-scale setting. For scopes with no offscreen memory, this in 10× the time base setting. If the scope has offscreen memory, this must be taken into account. For example, assume that one has two scopes with a maximum sampling speed of 100 MSa/s. One scope has a memory depth of 5 K points and the other only 1 K. At a sweep speed of 1 μs per division, both scopes will be able to store data into their memory at their full sampling speed, and each will be storing 100 data points per division, for a total of 1000 data points being stored. The scope with the 5 K memory will have a data point in one of every five memory locations, and the scope with the 1 K memory will have a data point in every memory location. If one reduces the sweep speed to 5 μs/div, the deeper memory scope will now fill every one of its memory locations with data points separated by 10 ns. The scope with only 1 K of memory would produce a display only two divisions wide if its sampling speed is not reduced. Scope designers believe that scope users expect to see a full-length sweep at every sweep speed. Therefore, the 1 K scope must reduce its sampling speed to one sample every 50 ns, or 20 MSamples/s, to be able to fill its memory with a full sweep width of data. This 5:1 ratio of sampling speeds between these two scopes will be maintained as their time bases are set to longer and longer sweeps. For example, at 1 s/div, the 5 K scope will be sampling at 500 samples per second, while the 1 K scope will be sampling at only 100 samples/s. One can determine a scope's sampling speed for any specific time base setting from Equation 19.1:

$$S\,(\text{samples/second}) = \frac{\text{Memory depth (samples)}}{\text{Full-scale time base (seconds)}} \tag{19.1}$$

or the scope's maximum sampling speed, whichever is less.

One must look closely at the application to determine if a specific scope is best suited to that application. As a rule, the deeper the memory, the faster the scope will be able to sample the signal at any given time base setting. Memory depth is not free. High-speed memory required to be able to store the data out of the scope's A/D is costly, and deeper memory takes longer to fill, thus reducing the scope's display update rate. Most scopes that provide memory depths of 20 K or more will also give the user a memory depth selection control so that the user can select between fast and deep. (In 1996, Hewlett-Packard Co. introduced two scopes based on an acquisition technology known as MegaZoom (TM) [2] that removes the need for a memory depth control.) A correct mental model for the sampling speed of a digital scope is based on Equation 19.1 and not just on the scope's maximum performance specifications.

Some digital oscilloscopes offer a special sampling mode known as *peak detection*. Peak detection is a special mode that has the effect of extending the scope's sampling speed to longer time records. This special mode can reduce the possibility of an aliased display. The performance of this special mode is specified as the minimum pulse width that the peak detection system can capture. There are several peak detection systems being used by the various manufacturers. Tektronix has an analog-based peak detection system in some of its models, while Hewlett-Packard has a digital system in all of its models. Both systems perform as advertised, and they should be evaluated in the lab to see

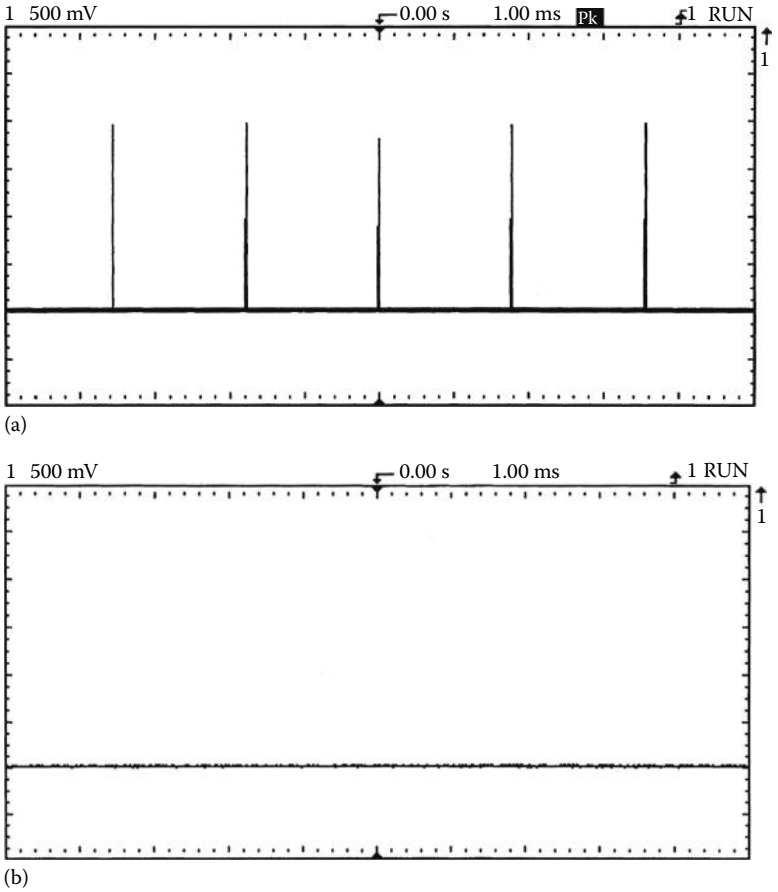


FIGURE 19.7 Peak detection. This special mode has the effect of increasing the scope's sampling speed at time base settings where it would be decimated. In operation, each memory location contains either the maximum or minimum value of the waveform at that location in time. (a) A series of 300 ns wide pulses being captured at a slow sweep speed and (b) the same setup with peak detection disabled. These narrow pulses would appear as intermittent pulses if the scope could be seen in operation with peak detection disabled.

which system best meets one's needs. There is a downside to peak detection systems and that is that they display high-frequency noise that might not be within the bandwidth of the system under test. Figure 19.7 shows a narrow pulse being captured by peak detection and being missed when the peak detection is off.

What effect does display dead time have on the oscilloscope's voltage measurement capabilities? Display dead time applies to both analog and digital oscilloscopes, and it is that time when the oscilloscope is not capturing the input signal. This is also a very important consideration in the operation of a digital scope because it determines the scope's ability to respond to front-panel control commands and to follow changing waveforms. A digital scope that produces an incorrect display of an amplitude-modulated signal is not following this rapidly changing signal because its display update rate is too low. Sampling speed is not related to display update rate or dead time. Display dead time is a function of the scope's ability to process the waveform data from its A/D and plot it on the display. Every major oscilloscope manufacturer has been working on this problem. Tektronix offers a special mode on some of its products known as InstaVu (TM) [3]. This special mode allows these scopes to process up

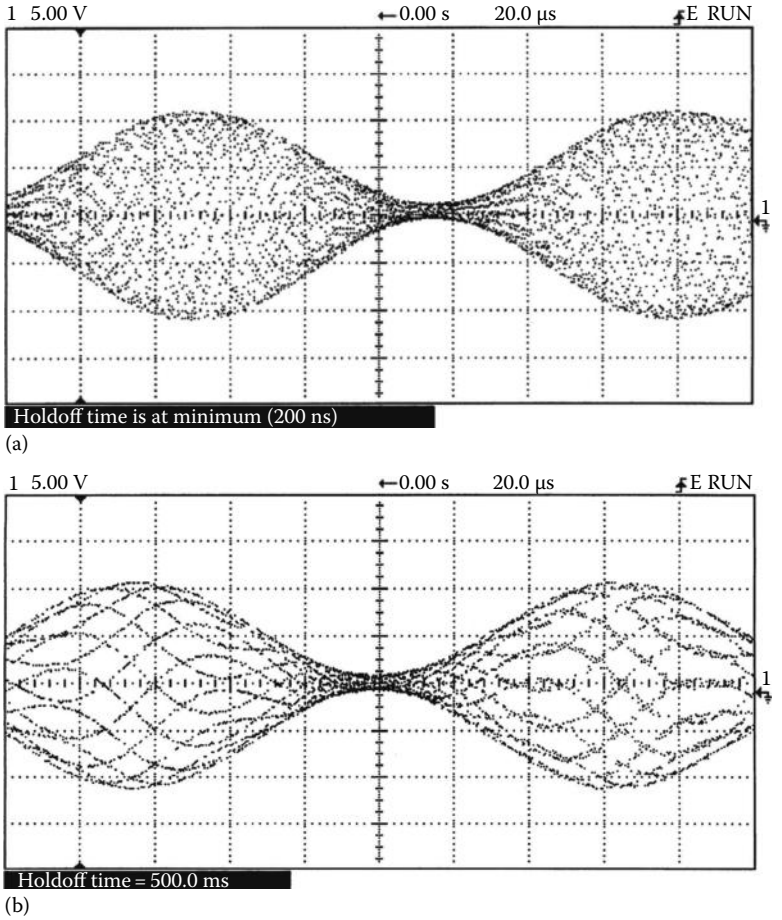


FIGURE 19.8 Display dead time. The time that an oscilloscope is blind to the input signal has an effect on the scope's ability to correctly display rapidly changing signals: (a) an amplitude-modulated signal with a high-speed display and (b) the same signal with the dead time increased by the use of hold-off.

to 400,000 waveforms per second to their display. Hewlett-Packard has developed a multiple parallel processor technology [4] in the HP 54600 series of benchtop scopes that provides a high-speed, low-dead-time display in a low-cost instrument. These instruments can plot 1,500,000 points/s to their display, and they have no dead time at their slower sweep speeds. LeCroy has been applying the PowerPC as an embedded processor for its scopes to increase display throughput. There are other special modes being produced by other vendors, so be sure to understand what these can do before selecting an oscilloscope. Figure 19.8 shows the effect of display update rate on a rapidly changing waveform. An amplitude-modulated signal is displayed with a high-speed display and with the display speed reduced by the use of hold-off.

19.5.1 Triggering

The trigger of the oscilloscope has no direct effect on the scope's ability to measure a voltage except that the trigger does enable the oscilloscope to produce a stable display of the voltage of interest. Reference [5] presents a thorough discussion of this subject.

Conclusion: The mental model that oscilloscope users have created in their minds of the oscilloscope's operation can be helpful in reducing measurement errors. If the operator's mental model is based on the following facts, measurement errors can be minimized:

- Oscilloscopes have a frequency response that affects measurement accuracy.
- Digital scopes are more accurate than analog scopes.
- Analog scopes do not have continuous displays.
- Oscilloscope accuracy specifications always contain a percent of full-scale term.
- Measurements should be made at the largest possible deflection in order to minimize errors.
- Maximum sampling speed is available only at the scope's fastest sweep speeds.
- Deeper memory depth allows faster sampling at more sweep speeds.
- All digital scopes can produce aliases, some more than others.
- Display dead time is an important characteristic of digital scopes that is often not specified.
- Display dead time affects measurement accuracy because it can cause a distorted display.
- The scope with the highest maximum sampling speed specification might not be the most accurate or have the lowest display dead time.
- The operator must have some knowledge of the signals being measured to be able to make the best possible measurements.

The person who has the mental model of the oscilloscope that takes these factors into account will be able to purchase the scope that is best suited to his or her application and not spend too much money on unnecessary additional performance. In addition, that person will be able to make measurements that are up to the full accuracy capabilities of the scope.

19.6 Selecting the Oscilloscope

There are 10 points to consider when selecting an oscilloscope. This author has published a thorough discussion of these points [6], and they are summarized as follows:

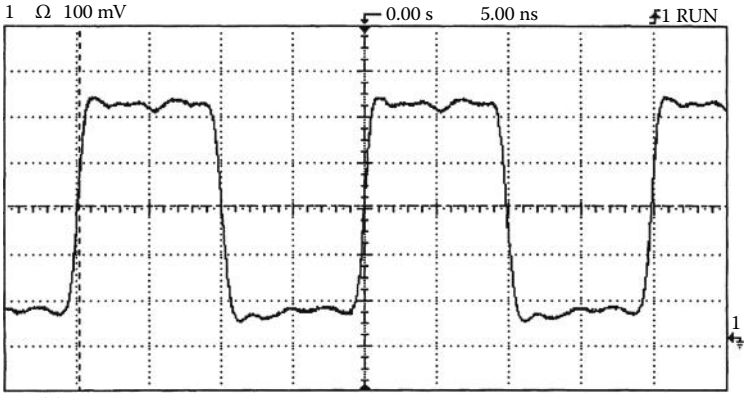
1. *Analog or digital*? There are a few places where the analog scope might be the best choice, and the reader can make an informed selection based on the information presented here.
2. *How much bandwidth*? This is a place where the person selecting an oscilloscope can save money by not purchasing more bandwidth than is needed. When analog oscilloscopes were the only choice, many people were forced to purchase more bandwidth than they needed because they needed to view infrequent or low repetition signals. High-bandwidth analog scopes had brighter CRTs so that they were able to display high-frequency signals at very fast time base settings. At a sweep speed of 5 ns/div, the phosphor is being energized by the electron beam for 50 ns, so the electron beam had to be very high energy to produce a visible trace. This situation does not apply to digital scopes. Now, one needs to be concerned only with the bandwidth required to make the measurement. Figure 19.9 shows the effect of oscilloscope bandwidth on the display of a 50 MHz square wave.

   The oscilloscope's bandwidth should be >2× the fundamental highest frequency signal to be measured.

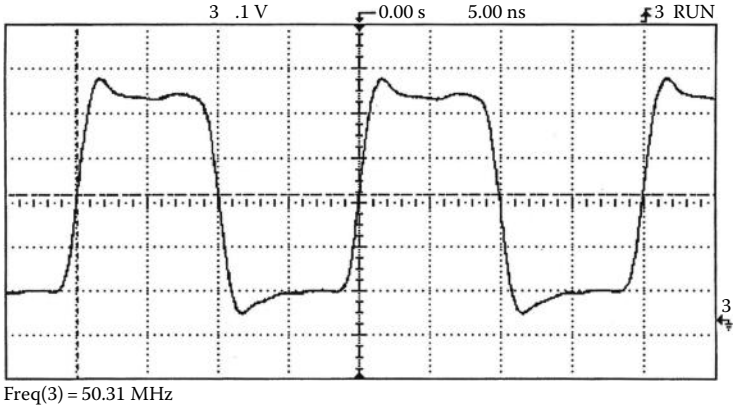
   The bandwidth of the scope's vertical system can affect the scope's ability to correctly display narrow pulses and to make time interval measurements. Because of the scope's Gaussian frequency response, one can determine its ability to correctly display a transient event in terms of rise time with Equation 18.2:

$$t_r = \frac{0.35}{\text{BW}} \tag{19.2}$$

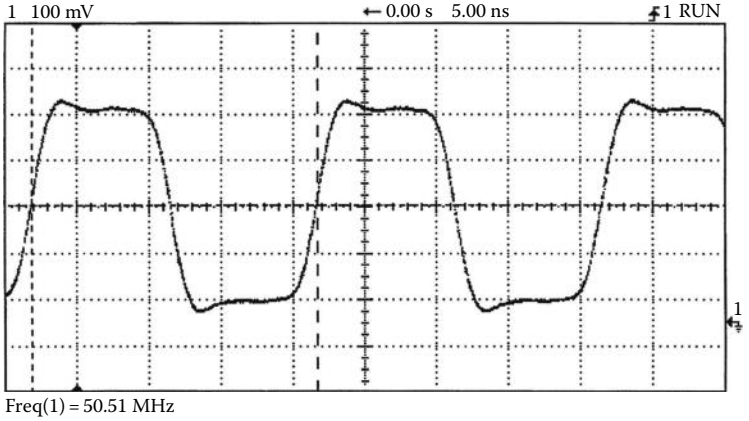
   Therefore, a 100 MHz scope will have a rise time of 3.5 ns. This means that if the scope were to have a signal at its input with zero rise time edges, it would be displayed with 3.5 ns edges. This will affect



Freq(1) = 50.44 MHz
(a) BW = 500 MHz



Freq(3) = 50.31 MHz
(b) BW = 250 MHz



Freq(1) = 50.51 MHz
(c) BW = 100 MHz

FIGURE 19.9 The effect of the scope's bandwidth is shown in this set of waveforms. The same 50 MHz square wave is shown as it was displayed on different bandwidth scopes: (a) 500 MHz bandwidth, (b) 250 MHz bandwidth, (c) 100 MHz bandwidth,

(*continued*)

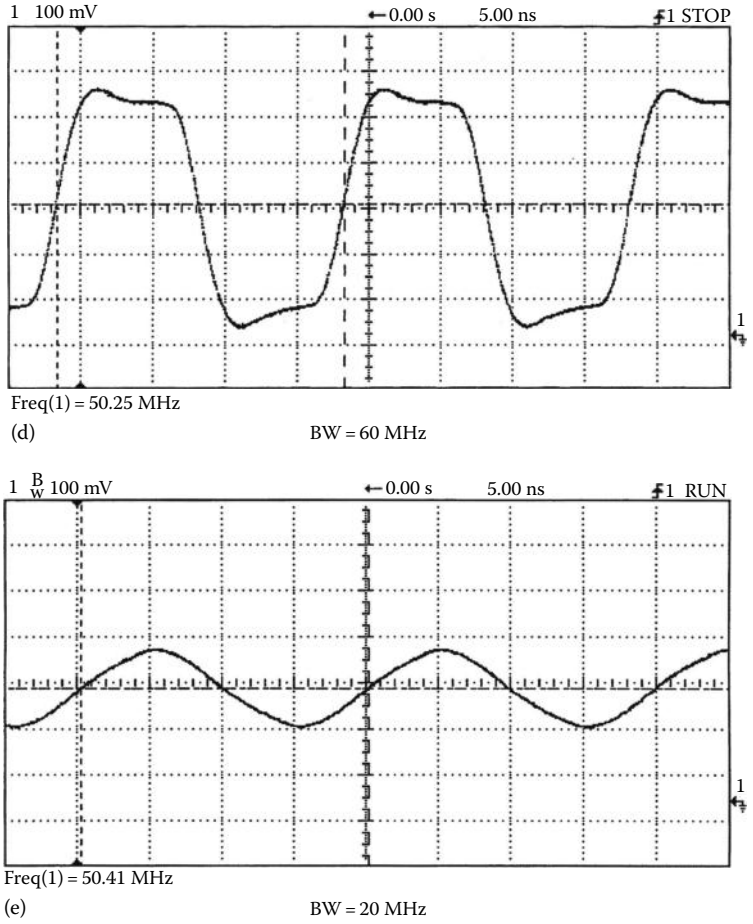


FIGURE 19.9 (continued) The effect of the scope's bandwidth is shown in this set of waveforms. The same 50 MHz square wave is shown as it was displayed on different bandwidth scopes: (d) 60 MHz bandwidth, and (e) 20 MHz bandwidth. Notice that the 100 MHz scope produced a usable display although it was missing the high-frequency details of the 500 MHz display. The reason that the 100 MHz scope looks so good is the fact that its bandwidth is slightly greater than 100 MHz. This performance, which is not specified on any data sheet, is something to look for in any evaluation.

the scope's measurements in two ways. First are narrow pulses. Figure 19.10 shows the same 5 ns wide pulse being displayed on oscilloscopes of 500 and 60 MHz bandwidths, and the effect of the lower bandwidth on this event that is closest to the rise time of the slower scope is apparent.

The second is fast time interval measurements. A measurement of signal rise time is an example. The observed rise time on the scope's display is according to Equation 19.3:

$$t_{\text{observed}} = (t_{\text{signal}}^2 + t_{\text{scope}}^2)^{1/2} \tag{19.3}$$

If a 10 ns rise time were to be measured with a 100 MHz scope, one would obtain a measurement of 10.6 ns based on Equation 19.3. The scope would have made this measurement with a 6% reading error before any other factors, such as time base accuracy, are considered.

The scope's rise time should be at least no more than 1/5 of the shortest time interval to be measured. For time interval measurements, this should be >1/10.

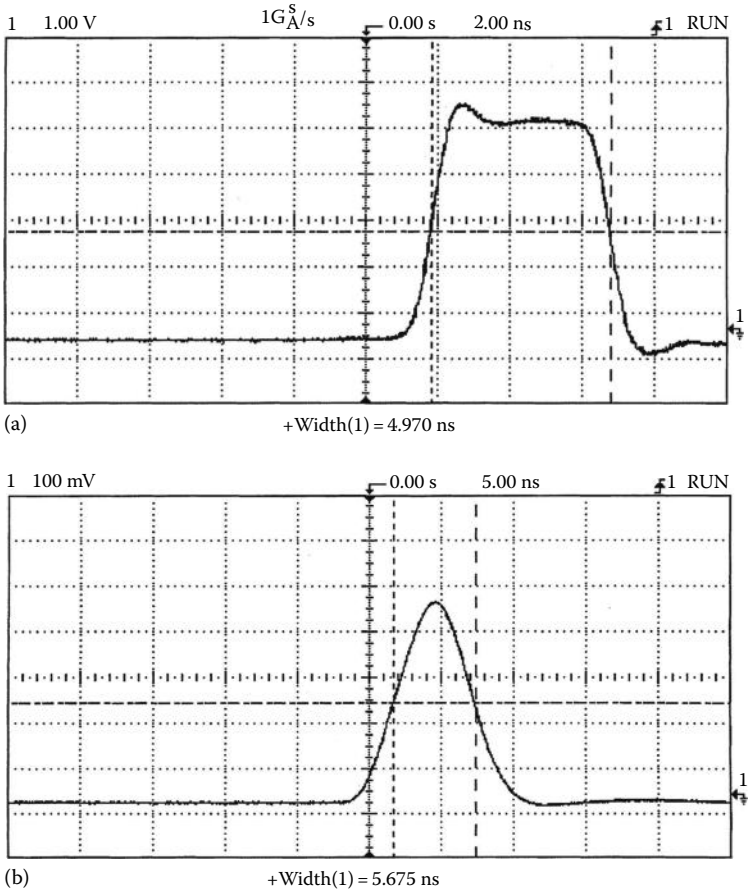


FIGURE 19.10 Bandwidth and narrow events. (a) A 5 ns wide pulse as displayed on a 500 MHz scope and (b) the same pulse displayed on a 60 MHz scope. The 60 MHz scope has a rise time of 5.8 ns, which is longer than the pulse width. This results in the pulse shape being incorrectly displayed and its amplitude being in error.

3. *How many channels*? Most oscilloscopes in use today are dual-channel models. In addition, there are models described as being 2 + 2 and four channels. This is one time where 2 + 2 is not equal to 4. The 2 + 2 models have limited features on two of their channels and cost less than 4-channel models. Most oscilloscope suppliers will hold the 4-channel description only for models with four full-featured channels, but user should be sure to check that the model under consideration so as to be sure if it is a 4- or 2 + 2 model. Either of the four channel classes is useful for applications involving the testing and development of digital-based systems where the relationship of several signals must be observed.

   Hewlett-Packard introduced a new class of oscilloscopes that is tailored for the applications involving both analog and digital technologies and mixed-signal systems. The MSO [3] provides two scope channels and 16 logic channels so that it can display both the analog and digital operations of a mixed-signal system on its display.
4. *What sampling speed*? Do not simply pick the scope with the highest banner specification. One needs to ask, what is the sampling speed at the sweep speeds that my application is most likely to require? As observed in Equation 19.1, the scope's sampling speed is a function of memory depth and full-scale time base setting. If waveforms are mostly repetitive, one can save money by selecting an oscilloscope that provides equivalent time or random repetitive sampling.

5. *How much memory*? As previously discussed, memory depth and sampling speed are related. The memory depth required depends on the time span needed to measure and the time resolution required. The longer the time span to be captured and the finer the resolution required, the more memory one will need. High-speed waveform memory is expensive. It takes time to process a longer memory, so the display will have more dead time in a long memory scope than a shallow memory model. All the suppliers of deep memory scopes provide a memory depth control. They provide this control so that the user can choose between a high-speed display and deep memory for the application at hand. Hewlett-Packard introduced MegaZoom (TM) technology [7] in 1996; it produces a high-speed low-dead-time display with deep memory all the time.
6. *Triggering*? All scope manufacturers are adding new triggering features to their products. These features are important because they allow for triggering on very specific events. This can be a valuable troubleshooting tool because it will let the user prove whether a suspected condition exists or not. Extra triggering features add complexity to the scope's user interface; so be sure to try them out to make sure that they can be applied.
7. *Trustworthy display*? Three factors critically affect a scope's ability to display the unknown and complex signals that are encountered in oscilloscope applications. If the user loses confidence in the scope's ability to correctly display what is going on at its probe tip, productivity will take a real hit. These are display update rate, dead time, and aliasing.

   Because all digital scopes operate on sampled data, they are subject to aliasing. An alias is a false reconstruction of the signal caused by undersampling the original. An alias will always be displayed as a lower frequency than the actual signal. Some vendors employ proprietary techniques to minimize the likelihood of this problem occurring. Be sure to test any scope being considered for purchase on your worse-case signal to see if it produces a correct or aliased display. Do not simply test it with a single-shot signal that will be captured at the scope's fastest sweep speed because this will fail to test the scope's ability to correctly display signals that require slower sweep speeds.
8. *Analysis functions*? Digital oscilloscopes with their embedded microprocessors have the ability to perform mathematical operations that can give additional insight into waveforms. These operations often include addition, subtraction, multiplication, integration, and differentiation. An FFT can be a powerful tool, but do not be misled into thinking that it is a replacement for a spectrum analyzer. Be sure to check the implementation of these features in any scope being considered. For example, does the Fast Fourier Transform (FFT) provide a selection of window functions? Are these analysis functions implemented with a control system that only their designer could apply?
9. *Computer I/O*? Most of the digital scopes on the market today can be interfaced to a PC. Most of the scope manufacturers also provide some software that simplifies the task of making the scope and PC work together. Trace images can be incorporated into documents such as PC Paintbrush (.pcx), tagged image file format (.tiff), Joint Photographic Expert Group (.jpg) files. Waveform data can be transferred to spreadsheet applications for additional analysis. Some scope models are supplied with a disk drive that can store either waveform data or trace images.
10. *Try it out*? Now, one has the information to narrow oscilloscope selection to a few models based on bandwidth, sampling speed, memory depth, and budget requirements. Contact the scope vendors (Table 19.2) and ask for an evaluation unit. While the evaluation unit is in the lab, look for the following characteristics:
    a. *Control panel responsiveness*: Does the scope respond quickly to inputs or does it have to think about it for a while?
    b. *Control panel layout*: Are the various functions clearly labeled? Does the user have to refer to the manual even for simple things?
    c. *Display speed*: Turn on a couple of automatic measurements and check that the display speed remains fast enough to follow changing signals.
    d. *Aliasing*: Does the scope produce an alias when the time base is reduced from fast to slow sweep speeds? How does the display look for the toughest signal?

TABLE 19.2 Major Suppliers of Oscilloscopes and Their Web Addresses

Vendor	Description	Web Address
B&K Precision, 6460 W. Cortland St., Chicago, IL 60635	Analog and digital scopes and Metrix scopes in France	http://bkprecision.com
Boonton Electronics Corp., 25 Estmans Road, P.O. Box 465, Parsippany, NJ 07054-0465	US importer for Metrix analog, mixed analog, and digital scopes from France	http://www.boonton.com
Fluke, P.O. Box 9090, Everett, WA 98206-9090	Handheld, battery-powered scopes (ScopeMeter), analog scopes, and CombiScopes(R)	http://www.fluke.com
Gould, Roebuck Road, Hainault, Ilford, Essex IG6 3UE, England	200 MHz DSO products	http://www.gould.co.uk
Hewlett-Packard Co., Test & Measurement, Mail Stop 51LSJ, P.O. Box 58199, Santa Clara, CA 95052-9952	A broad line of oscilloscopes and the MSO for technical professionals	http://www.tmo.hp.com/tmo/pia search on "oscilloscopes"
LeCroy Corp., 700 Chestnut Ridge Road, Chestnut Ridge, NY 10977	Deep memory oscilloscopes for the lab	http://www.lecroy.com
Tektronix, Inc., Corporate Offices, 26600 SW Parkway, P.O. Box 1000, Watsonville, OR 97070-1000	The broad line oscilloscope supplier with products ranging from handheld to high-performance lab scopes	http://www.tek.com/measurement search on "oscilloscopes"
Yokogawa Corp. of America, Corporate offices, Newnan, GA, 1-800-258-2552	Digital oscilloscopes for the lab	http://www.yca.com

The oscilloscope is undergoing a period of rapid change. The major manufacturers of oscilloscopes are no longer producing analog models, and the digital models are evolving rapidly. There is confusion in the oscilloscope marketplace because of the rapid pace of this change. Hopefully, this discussion will prove valuable to the user in selecting and applying oscilloscopes in the lab in the years to come.

References

1. A. DeVibiss, Oscilloscopes, Chapter 14, p.14.1–60 in C. F. Coombs, Jr. (ed.), *Electronic Instrument Handbook*, 2nd edn., McGraw-Hill, New York, 1995.
2. R. A. Witte, A family of instruments for testing mixed-signal circuits and systems, *Hewlett-Packard J.*, April 1996, Hewlett-Packard Co., Palo Alto, CA.
3. InstaVu acquisition mode, *Tektronix Measurement Products Catalog*, Tektronix, Inc., Beaverton, OR, 1996, p. 69.
4. M. S. Holcomb and D. P. Timm, A high-throughput acquisition architecture for a 100 MHz digitizing oscilloscope, *Hewlett-Packard J.*, February 1992, Hewlett-Packard Co., Palo Alto, CA.
5. R. A. Witte, *Electronic Test Instruments, Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1993.
6. J. Murphy, Ten points to ponder in picking an oscilloscope, *IEEE Spectrum*, 33(7), 69–77, 1996.
7. M. S. Holcomb, S. O. Hall, W. S. Tustin, P. J. Burkart, and S. D. Roach, Design of a mixed signal oscilloscope, *Hewlett-Packard J.*, April 1996, Hewlett-Packard Co., Palo Alto, CA.

20

Inductive and Capacitive Voltage Measurement

Cipriano Bartoletti
Sapienza University of Rome

Luca Podestà
Sapienza University of Rome

Giancarlo Sacerdoti
Sapienza University of Rome

20.1 Introduction

Capacitive and inductive voltage sensors are mainly utilized in low-frequency electric measurements. In this chapter, the capacitive and inductive measurement with large voltage ranges, from 10^{-10} to 10^{7} V, has been discussed. The associated voltage waveforms can be periodic or impulsive. In periodic waveforms, it is often important to measure the spectrum and the subharmonic components, which can vary depending on the position of the measuring devices within the same network. In impulsive measurements, obtaining the maximum voltage values, the pulse lengths, etc., is preferred.

20.2 Capacitive Sensors

The voltage to be measured can be reduced by means of capacitive dividers (Figure 20.1). Capacitive dividers are affected by temperature and frequency and therefore are not important, at least in Europe. Capacitive sensors detect voltage by different methods:

1. Electrostatic force (or torque)
2. Kerr or Pockels effect
3. Josephson effect
4. Transparency through a liquid-crystal device
5. Change in refractive index of the optic fiber or in light pipe

1. The relations that rule the listed capacitive voltage sensors are reported in the succeeding text. The force between two electrodes is (Figure 20.2)

$$F = \varepsilon_0 \frac{S}{d}(V_1 - V_2)^2 \tag{20.1}$$

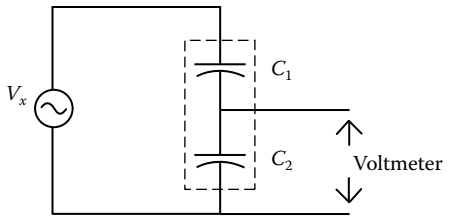


FIGURE 20.1 Schematic arrangement of a capacitive divider.

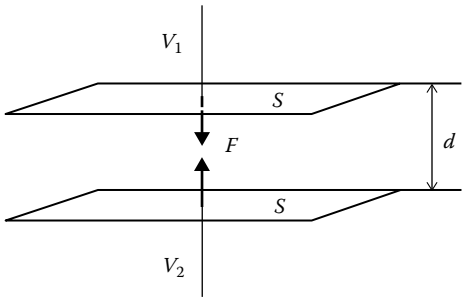


FIGURE 20.2 Force between two electrodes with an applied voltage.

where
ε_0 is the dielectric constant
S is the area of the electrode
d is the distance
V_1, V_2 are the potentials of the electrodes

The *torque* between electrostatic voltmeter quadrants (Figure 20.3) is given by

$$T = \frac{1}{2}\frac{\partial C}{\partial \theta}(V_1 - V_2)^2 \tag{20.2}$$

where
C is the capacitance
θ is the angle between electrodes

To get the torque from the rate of change (derivative) of electrostatic energy versus the angle is easy. Obtaining the torque by mapping the electric field is difficult and requires long and complex field computing.

2. The rotation of the polarization plane of a light beam passing through a KDP crystal under the influence of an electric field (*Pockels effect*) is expressed by (Figure 20.4)

$$\theta = k_p l(V_1 - V_2) \tag{20.3}$$

where
k_p is the electro-optic constant
l is the length of crystal

One obtains a rotation of $\pi/2$ by applying a voltage of the order of 1 kV to a KDP crystal of a few centimeters in length.

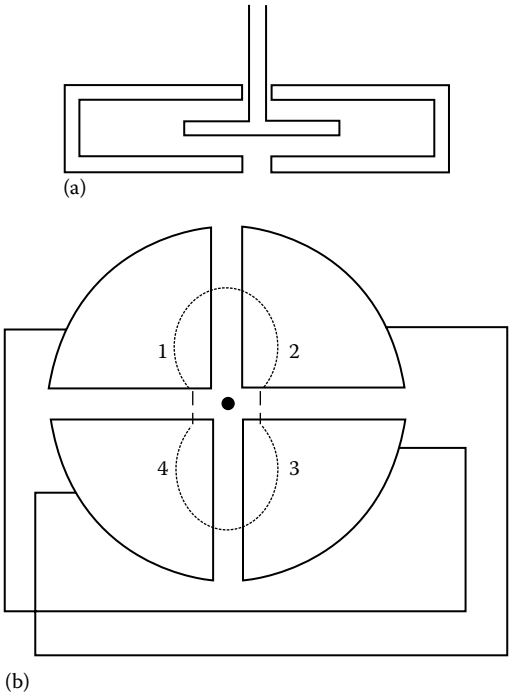


FIGURE 20.3 Scheme of an electrostatic voltmeter. (a) Lateral view and (b) top view: (1), (2), (3), and (4) are the static electrodes; the moving vane is shown in transparency.

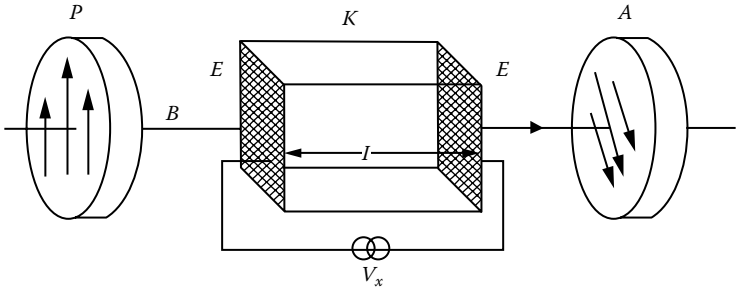


FIGURE 20.4 Scheme of an electro-optic KDP device. *B*, a light beam; *P*, a polarizer; *A*, an analyzer; *K*, a KDP crystal, with the voltage to be measured V_x applied to its transparent electrodes; *E*, transparent electrodes.

If a light beam passes through a light pipe that performs the *Kerr effect*, one observes a quadratic dependence of the rotation versus *V*:

$$\theta \equiv kE^2 \equiv k'V^2 \tag{20.4}$$

3. The *Josephson effect* consists of translation of a voltage into a periodical signal of a certain frequency, carried out by a special capacitive sensor. There is an array of *N* layers of Josephson superconducting junctions; the frequency of emitted signal, when a voltage *V* is applied, is given by

$$\nu = \frac{2eV}{Nh} \tag{20.5}$$

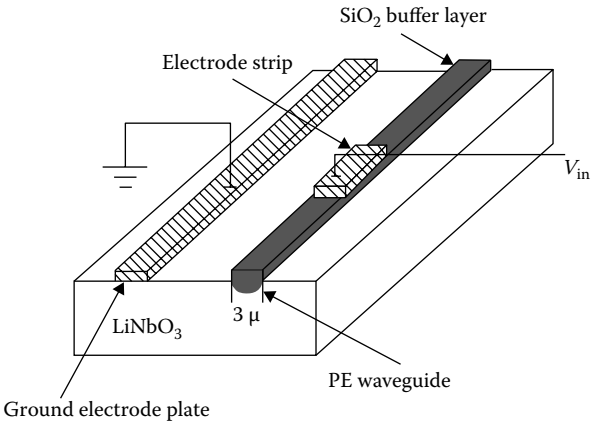


FIGURE 20.5 Li-Nb optical wave guide device.

4. The *transparency* of a liquid-crystal device depends on the difference of potential applied. There are liquid-crystal devices working in transmission or in reflection. A change in transparency is obtained when a difference of potential of a few volts is applied.
5. The *change in refractive index* due to the presence of an electric field can be detected by
   a. Interferometric methods (where the velocity of light is equal to c/n)
   b. Change in light intensity in a beam passing through an optical wave guide device like Li-Nb (Figure 20.5)

By means of method 1, many kinds of instruments (voltmeters) can be realized. Methods 2 through 5 are used in research laboratories but are not yet used in industrial measurements.

20.3 Inductive Sensors

20.3.1 Voltage Transformers

Voltage transformers (VTs) have two different tasks:

- Reduction in voltage values for meeting the range of normal measuring instruments or protection relays
- Insulation of the measuring circuit from power circuits (necessary when voltage values are over 600 V)

VTs are composed of two windings—one primary and one secondary winding. The primary winding must be connected to power circuits; the secondary to measuring or protection circuits. Electrically, these two windings are insulated but are connected magnetically by the core.

One can define the following

$$\text{Nominal ratio} = K_n = \frac{V_{1n}}{V_{2n}} \tag{20.6}$$

as the ratio between the magnitude of primary and secondary rated voltages.

$$\text{Actual ratio} = K = \frac{V_1}{V_2} \tag{20.7}$$

as the ratio between the magnitudes of primary and secondary actual voltages.

Burden is the value of the apparent power (normally at cos $\varphi = 0.8$) that can be provided on the secondary circuit (instruments plus connecting cables).

Burden limits the maximum value of secondary current and then the minimum value of impedance of the secondary circuit is

$$Z_{\min} = \frac{V_{2n}^2}{A_n} \tag{20.8}$$

where A_n is the VT burden.

For example, if $A_n = 25$ VA and $V_{2n} = 100$ V, one obtains

$$Z_{\min} = \frac{100}{0.25} = 400 \text{ W} \tag{20.9}$$

There are two kinds of errors:

1. Ratio error $= h_{\%} = \dfrac{K_n - K}{K}$ (20.10)
2. Angle error = the phase displacement between the primary voltage and the secondary voltage (positive if the primary voltage lags the secondary one).

VTs are subdivided into accuracy classes related to the limits in ratio and angle error (according to CEI and IEC normative classes 0.1, 0.2, 0.5, 1, 3; see Table 20.1). To choose the VT needed, the following technical data must be followed:

- Primary and secondary voltage (rated transformation ratio). Normally, the secondary value is 100 V.
- Accuracy class and rated burden in VA: for example, cl. 0.5 and $A_n = 10$ VA.
- Rated working voltage and frequency.
- Insulation voltage.
- Voltage factor: the ratio between maximum operating voltage permitted and the rated voltage. The standard voltage factor is 1.2 V_n (i.e., the actual primary voltage) for an unlimited period of time (with VT connected with phases) and is 1.9 V_n for a period of 8 h for VT connected between phase and neutral.
- Thermal power is the maximum burden withstood by VT (errors excluded).

TABLE 20.1 Angle and Ratio Error Limit Table Accepted by CEI and IEC Standards

Class	Percentage Voltage (Ratio) Error (±)	Phase Displacement Minutes (±)	Centiradians (±)
0.1	0.1	5	0.15
0.2	0.2	10	0.3
0.5	0.5	20	0.6
1	1	40	1.2
3	3	—	—
3P	3	120	3.5
6P	6	240	7

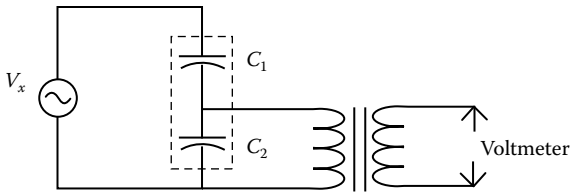


FIGURE 20.6 Capacitive divider and VT device for extremely high voltage.

For extremely high voltage values, both capacitive dividers and VTs are normally used, as shown in Figure 20.6. The capacitive impedance must compensate for the effect of the transformer's internal inductive impedance at the working frequency.

20.3.2 Other Methods

The ac voltage inductive sensors act by interaction between a magnetic field (by an electromagnet excited by voltage to be measured) and the eddy current induced in an electroconductive disk, producing a force or a torque. This can be achieved by the scheme shown in Figure 20.7. The weight of many parts of the indicator can be some tens of grams. The power absorbed is on the order of a few watts. The precision is not high, but it is possible to get these sensors or instruments as they are similar to the widely produced induction energy meters. They are quite robust and are priced between \$50 and \$100, but they are not widely used. The relation between torque and voltage is quadratic:

$$T = k_i V^2 \tag{20.11}$$

The proportionality factor k_i depends on magnet characteristics and disk geometry.

GEC, Landis & Gyr, ABB, Schlumberger, etc., are the major companies that furnish components and instruments measuring voltage by inductive and capacitive sensors.

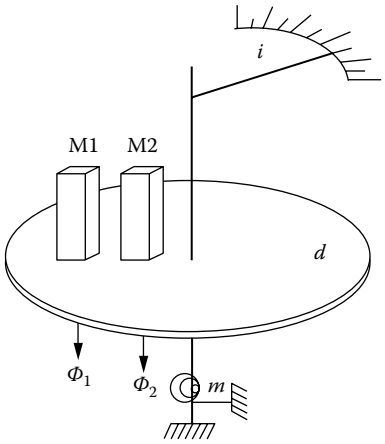


FIGURE 20.7 Schematic inductive voltmeter. *i*, index; *d*, metallic disk; M1 and M2, electromagnets; *m*, spring; Φ_1 and Φ_2, generated fluxes.

Defining Terms

CEI: Comitato Elettrotecnico Italiano

IEC: International Electric Committee

KDP: potassium dihydrogen phosphate

Li-Nb: ($LiNbO_3$) lithium niobate

Further Information

Moeller, J. and G. Rosenberger, Instrument transformers for HV systems, Siemens Power Engineering III (1981) Special Issue, *High-Voltage Technology*.

Sacerdoti, G., O. Jappolo, and R. Paggi, *Misure Elettriche, Vol. I Strumenti*, Zanichelli, Bologna, Italy, 1994.

Zingales, G., *Metodi e Strumenti per le Misure Elettriche*, UTET, Bologna, Italy, 1976.

21

Current Measurement

Douglas P. McNutt
The MacNauchtan Laboratory

Current-measuring devices are selected for the following:

- Precision
- Stability
- Frequency response, including direct current (dc)
- Galvanic isolation
- Presentation of the data
- Effect on measured circuit

In this chapter, the most common current sensors will be introduced in detail. Some less common sensors will follow after definitions and a bit of magnetic theory, which can be skipped. Any magnetic field sensor can be used as a current sensor, and there are some exotic examples such as quantum effects in low-temperature superconductors used to measure currents in neurons within the brain. This discussion will be limited to measurement of currents in wires with commercially practical devices.

An isolated current sensor is free of any metallic connection to the circuit being measured. It is also essentially free of capacitive coupling so that it is safe to use with grounded amplifiers and other equipment. The quality of the isolation is measured in volts and is usually the breakdown potential of an insulator. 5 kV is typical for personal safety around light industrial power.

By far, the simplest current-to-voltage converter is the resistor. In current-measuring service, it is called a shunt even though it is typically placed in series with the load. That is because shunts are sometimes used to increase the range of another current-measuring device using a connection that bypasses part of the current around a meter. Frequency response of a shunt is good and includes dc. Shunts produce a voltage output that can be presented by a variety of secondary meters including analog meters, digital meters, oscilloscopes, and 4–20 mA converters. Shunts provide no isolation and have a potentially unacceptable effect on the circuit being measured. Shunts used for dc are as accurate as the resistance and the associated voltmeter.

The common moving-coil meter, the D'Arsonval movement [1,2], probably with an internal shunt and/or rectifier, is an easily used device, which is still available. Its isolation is by means of the human eye since that is the only way to read out the result. It is useful for power panels where an operator needs quick data. Accuracy is no better than 2%.

For power frequency, 25–400 Hz service, the current transformer, called a "donut" transformer or CT in the trade, is commonly employed. The current-carrying conductor is passed through the hole in a toroid of magnetic material. A shorted secondary winding of n turns carries current that is $1/n$ times the measured current and is typically passed to another ammeter or used as the current input to a power-measuring device. Isolation is as good as the insulation on the primary conductor, frequency response is fair but does not include dc, there is a minimal effect on the measured circuit, and cost is low. Operational safety is an issue; see the note in the succeeding text.

A variety of noncontact sensors are available for dc sensing. Most depend on the Hall effect, and all require a source of operating power. Frequency response from dc to 200 kHz is advertised. Because operating power is available, output to later processing can be voltage, current, or digital. Accuracy is whatever you want to pay for. Long-term stability depends on dc-coupled operational amplifiers and may exhibit zero drift. Externally, these look much like CTs.

CTs, Hall devices, and other similar noncontact schemes are available in wraparound form so they can be installed without disconnecting power. The wrapping process always involves breaking a magnetic path, and consistency of reassembly becomes a limit on precision. Everything from current sensing probes for oscilloscopes to CTs for 10,000 A circuits can be found as wraparounds.

21.1 Definition of the Ampere

There is a perpetual argument about the number of "fundamental" quantities required to describe our environment. Is there a fundamental unit of electricity or are the electric units derived from more basic things such as mass, length, and time? The U.S. National Institute of Standards and Technology (NIST) currently defines the ampere as "that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross section, and placed 1 m apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} Newton per meter of length" [3]. Work in progress [4] aims to change that definition to "that current corresponding to the flow of $6.241509468 \times 10^{18}$ elementary charges per second." The change is in line with a recent redefinition of the meter in terms of the velocity of light with the idea of eliminating physical standards, like the standard kilogram used in the current definition of the ampere.

Changes like that are always given long, but exact, scale values so that engineering practices are not affected. For the rest of this discussion, I shall make the assumption that instrumentation is "traceable to the NIST," meaning that one way or another, an electric unit as measured is compared to a standard ampere maintained by NIST. Absolute calibrations using the quantum properties of the Hall effect and the Josephson junction are not practical in the field.

For practical reasons, it is often easier to distribute a voltage standard than a current standard. Chemical cells and, more recently, semiconductor voltage references are quite stable and do not depend on the length of the wires used to connect them together during calibration and use. As a result, current measuring is usually a matter of conversion of current to an equivalent voltage.

An exception is current measurement by comparison to magnetic forces provided by a permanent magnet, and that may be why the older unrationalized centimeter–gram–second, the cgs units with no separate electric unit, are still found in specifications of magnetic quantities. The gauss and the oersted are particular examples of these now deprecated units, but they are still found in the datasheets for magnetic parts. It is this confusion of units that deters many who would measure current away from even attempting calculations involving magnetic devices.

21.2 Magnetics

Magnetic current sensors have advantages over shunts. To understand them, we need to delve into the interaction between currents and magnetic fields. Following Maxwell by way of Sommerfeld [5], it is convenient to describe magnetic effects in terms of two vector fields, **B** and **H**. **H** is the field that is created by an electric current, and **B** is the field that acts on a moving charge or a current-carrying wire. **B** and **H** are related by characteristics of the material in which they coexist. Strictly speaking, **B** is the flux density and **H** is the field, but they are both called the field in less than precise usage.

The SI units of **B** and **H** are the tesla (T) and the ampere per meter (A m^{-1}). They are called rationalized because a ubiquitous 4π has been suppressed in the underlying equations.

For practical engineering, it is still necessary to understand the unrationalized equivalents, the gauss and the oersted, because they are still used in specifications for magnetic materials. To convert from gauss to tesla, divide by 10^4. To convert from oersted to amperes per meter, multiply by $1000/(4\pi)$, a number that is commonly approximated as simply 80, but strictly speaking, the units of **H** in the two systems are dimensionally different and cannot be converted.

The relationship between **B** and **H** is most generally a tensor that reflects spatial anisotropy in the material, but for common magnetic materials used in current sensing, a scalar value μ applies. In SI units, **B** and **H** have different physical dimensions and are not equal in vacuum. The "permeability of free space" μ_0 is defined so that the μ for a material is a dimensionless constant. We say

$$\mathbf{B} = \mu\mu_0\mathbf{H} \tag{21.1}$$

where μ_0 is exactly $4\pi \times 10^{-7}$ H m^{-1}. (H here is the Henry, the SI unit of inductance, base units m^2 kg s^{-2} A^{-2}). For many problems, μ will be a constant of the magnetic material, but when magnetic saturation needs to be considered, it will be a variable. For engineering calculations, μ is often treated as though it depends on frequency. Values of μ range from less than 100 for a high-frequency ferrite to 5000 for transformer iron and to 10^5 for carefully annealed magnetic alloys:

$$\mathbf{H} = \mathbf{I}/(2\pi r) \tag{21.2}$$

The field due to a long straight wire is shown in Figure 21.1 and Equation 21.2. The **H** vector obeys the right-hand rule and is everywhere perpendicular to the wire. The amplitude of **H** is proportional to the current and falls off linearly with the radial distance from the wire. This is the field that makes noncontact sensing of current possible.

The field at the center of a circular loop of wire is shown in Figure 21.2 and Equation 21.3. It obeys the right-hand rule, is proportional to the current, and is inversely proportional to the radius of the loop:

$$\mathbf{H} = \mathbf{I}/r \tag{21.3}$$

A magnetic field can often be calculated by direct application of Maxwell's relations in integral form [6]. The line integral over any closed path of $\mathbf{H} \cdot d\mathbf{l}$ is equal to the current passing through a surface delimited by the closed path.

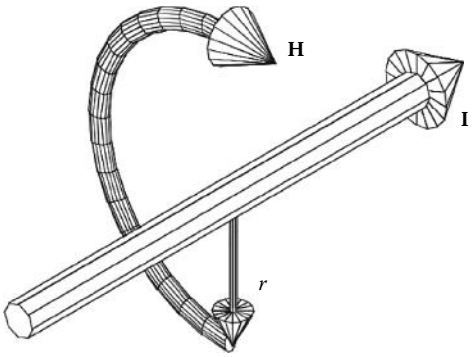


FIGURE 21.1 The magnetic field associated with a long wire carrying current **I**. It is this field that makes contactless sensing of current possible.

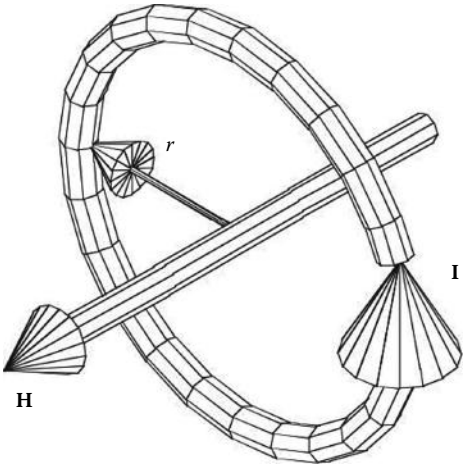


FIGURE 21.2 The magnetic field at the center of a loop of wire carrying current **I**. This is the starting point for many calculations involving inductors and transformers.

Magnetic flux passing through a surface, usually denoted by Φ, is the integral of $\mathbf{B} \cdot \mathbf{n}\, dA$ over the surface with normal vector **n**. The SI unit of Φ is the weber. The unrationalized unit, the line of force, is best relegated to history except that the lines, which form continuous loops, dramatize the fact that it doesn't matter how a surface is drawn to close a bounding line. The flux is the same. It is convenient to think of magnetic lines, forming closed loops, which can be bent but not interrupted or removed.

A carrier of charge, an electron in a wire, a particle in vacuum, or a hole in a semiconductor, moving in a magnetic field, is acted on by a force that is perpendicular to the field and the velocity. For positive charge, the force obeys the right-hand rule. The magnitude of the force is proportional to the charge q, the magnitude of **B**, the velocity **v**, and the sine of the angle between them. It is a vector cross product:

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} \tag{21.4}$$

A carrier, moving or not, is affected in a similar way by a changing magnetic field. The result is an electromotive force (EMF) in a loop of wire through which a changing magnetic flux passes. The EMF is equal to the rate of change of the flux enclosed by the loop with a change of sign. That is Faraday's law of induction:

$$\text{EMF} = -d\Phi / dt \tag{21.5}$$

Most magnetic materials exhibit hysteresis. That is, the relationship between **B** and **H** depends on the history of the applied **H**. A plot of **B** versus **H** is called a hysteresis loop, and a sample is shown as Figure 21.3.

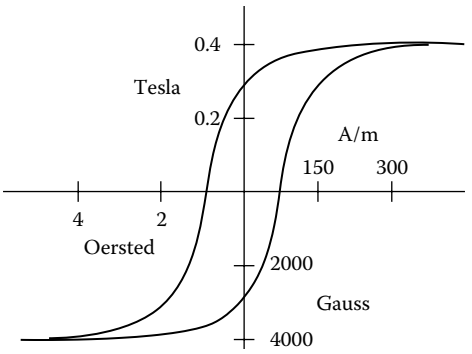


FIGURE 21.3 A hysteresis curve for some rather poor transformer iron. Unrationalized units are shown below and left of the origin because that remains standard practice in the industry.

The area of the loop represents energy. If a magnetic core is repeatedly driven around its hysteresis loop, there is a power loss that is proportional to the area of the loop and the frequency. A material with very large hysteresis is a permanent magnet.

Magnetic materials that are also electric conductors have free carriers, which are affected by alternating magnetic fields. As they move back and forth, they encounter electric resistance and dissipate energy in the form of heat. These eddy currents can be minimized by the use of high-resistance ferrites and powdered iron, by laminating cores, or by keeping the flux small.

21.3 Shunts

Shunts were introduced earlier. They dissipate power as heat, and the resistance changes in response to the rising temperature. The dissipation is proportional to the voltage across the shunt, and a design compromise must be made because low voltage implies less accuracy in the voltmeter. A standard of 50 mV has evolved. Shunts do not provide galvanic isolation between the measured circuit and the measurement device.

Measurement of alternating current (ac) using shunts is also affected by skin effect and the inductance of the shunt. Skin effect can be minimized in the design of the shunt by the use of several parallel sheets of thin metal (Figure 21.4), a feature that also improves heat dissipation. There isn't much to be done about inductance except to minimize size.

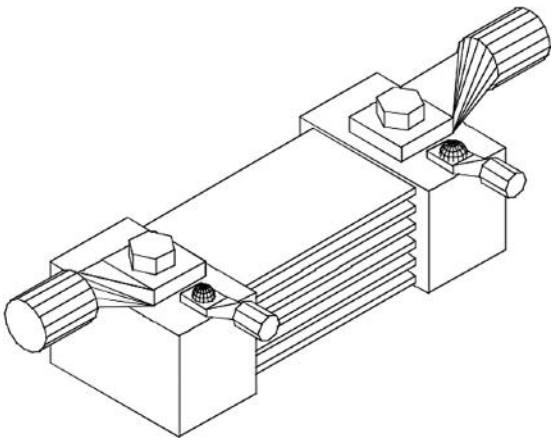


FIGURE 21.4 Multiple sheets of conductor are provided in this shunt to reduce skin effect and allow air cooling. Kelvin connections are provided for the voltmeter so that voltage drops in the high-current connectors are not inadvertently included in the measurement.

Safety is enhanced if shunts are placed in the ground leg of a circuit for that way the output leads, usually only at 50 mV, are near ground. But that introduces a resistance in the ground path and can interfere with common mode requirements of interdevice signal connections. If the shunt is placed in the high side, care must be taken to protect the wiring and the meter to which it is attached from accidental grounds.

Sometimes sufficient accuracy can be obtained by measuring the voltage drop along a length of conductor, which is otherwise required in the installation. In a vehicular service, it is common to sense the voltage drop in a battery cable using a millivoltmeter. Such installations are always high side and should be protected with fuses or other current limiting devices installed near the points of measurement. It is also wise to provide a connection means that is independent of contact resistance where the current-carrying conductor is installed—a Kelvin connection. Including one lead of an in-line fuse holder in the terminals before they are crimped is one such technique.

21.4 Moving Magnet Meter

The simplest current indicator balances the force on a permanent magnet created by current in a wire against a spring. A magnetic material is usually placed around the conductor to concentrate the field and reduce the effect of the magnetic field of the earth. The use is limited to low-precision indicators such as a battery charging meter for a vehicle. It is a dc device.

21.5 D'Arsonval Meter

This indicator balances the force on a current-carrying wire due to a permanent magnet against a spring. The measured current flows in a coil of wire supported in bearings. It is described in elementary texts [1,2], and I will not use space on its details. It is generally a dc instrument, but chart recorders have been built with frequency response in the kHz range using mirrors on the moving coil. They are then called galvanometers. For ac service, these meters are often equipped with internal copper oxide rectifiers and a nonlinear scale to correct for the diode drop. For current ranges above a few mA, they will have an internal shunt.

The moving magnet meter and the D'Arsonval movement are the only current sensors that do not convert current to voltage and then depend on other devices to read out the voltage.

21.6 Electrodynamometer

A variation of the D'Arsonval meter for ac service can be built by replacing the permanent magnet with an electromagnet. The force on the moving coil becomes proportional to both the current being measured and the voltage applied to the coil of the electromagnet. It is sensitive to the relative phase of the voltage and current in just the right way to be useful for measurement of power in a circuit with correction for power factor. An electrodynamometer in a power panel is often the load for a CT.

21.7 RF Ammeter and True rms

Current to a radio transmitting antenna is commonly passed through a small resistor, a shunt, which is thermally connected to a thermocouple or other thermometers and mounted in a thermally insulating blanket. The rise in temperature is a measure of the current and is often sensed with a thermocouple.

This is an example of true root-mean-square (rms) indication. The rms current is the square root of the integral of the square of the instantaneous current over an unspecified time divided by that time. It is intended to represent a stationary ac waveform by a single value, which is equal to the dc, which would dissipate the same power in a resistive load. The radio frequency (RF) ammeter does that precisely. The indication is not particularly linear, but it can easily be calibrated by applying dc to the input.

Other schemes for measuring rms current depend on analog multipliers and subsequent integration. They are limited by crest factor, the ratio of highest instantaneous peak to the rms over a measuring period. Inexpensive meters simply measure the peak, assume a sinusoidal waveform, and scale to rms.

21.8 Current Transformer

Galvanic isolation of current-measuring schemes becomes more and more important when there is no human observer looking at a meter that might be running at a high voltage. The CT remains a popular way to accomplish isolation for ac systems. Other techniques to follow are chosen mostly for their ability to provide isolation.

Consider the magnetics of a toroidal core of high-μ material through which a current-carrying conductor passes. Include a secondary winding of n turns as shown in Figure 21.5. The secondary is connected to a low-resistance load.

In this current transformer, universally referred to as a CT, ac in the single-turn primary attempts to magnetize the core but in so doing creates an EMF and current in the secondary, which tends to cancel the field. If the secondary truly has zero resistance, the current in it exactly cancels the field due to the primary. The result is a secondary current equal to the primary current divided by the number of secondary turns. The secondary current is in phase with the primary current. Because of the tightly closed magnetic loop, there is little effect from nearby conductors or position of the primary wire in the hole.

The secondary circuit can now be connected to a low-resistance current or power sensing device with assurance of calibration. But the secondary resistance is never really zero, and the magnetic coupling is never perfect so there are other considerations.

First, the concept of secondary burden is introduced. It is called that to avoid calling it a "load" because it behaves differently; the best burden is a short circuit. Burden is sometimes expressed as a secondary resistance in ohms but more often as an equivalent power in kVA for a defined current without consideration of phase. When the burden is not a perfect short circuit, energy is dissipated and the magnetic field present in the core is no longer zero. The secondary current leads the primary current with a phase, which depends on frequency.

Manufacturers of CTs have techniques to optimize accuracy of the CT when specified for a particular burden. The finished units may not have the number of secondary turns you would expect but will nonetheless provide results accurate to a percent or so. They have laminations selected to minimize heating of

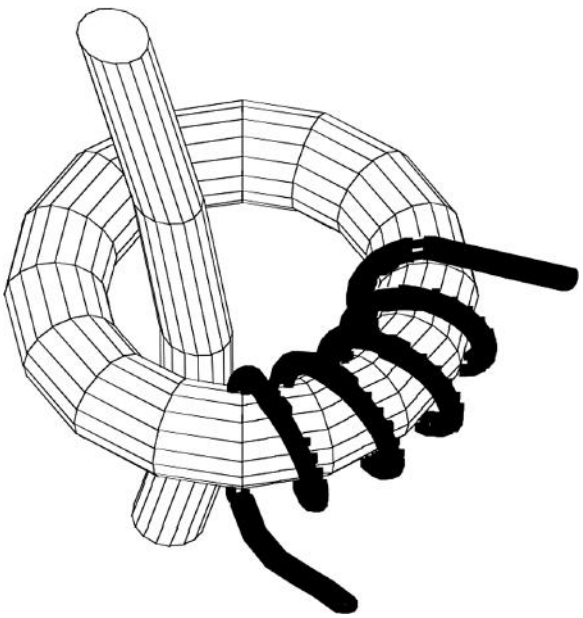


FIGURE 21.5 The ideal CT is tightly coupled with no magnetic gap. Current in the secondary exactly balances current in the single-turn primary so that the magnetic flux in the core is zero.

the core. You will see ratings like 100:5, meaning 100 A in the primary will produce 5 A in the secondary rather than "20 turns." They should be installed in a circuit that provides the burden for which they were calibrated. The voltage across a burden resistor is commonly amplified and passed to a data collection device. CTs for large currents need to be large to avoid magnetic saturation when burdened. Cores are prepared with laminates of silicon iron in the form of disks, concentric rings, or tape, which is wound on a bobbin. Even with the best of materials, eddy current and hysteresis loss are present. When power dissipation is unacceptable, another choice of sensor may be preferable.

Most CTs are used for measurement of low-frequency power and energy. They are found at the front end of kilowatt hour meters used by power providers. RF current in transmitting antennas can be measured with suitable core material. Ferrite cores with appropriate compensation are used for sensing pulse-width-modulated current in switching power supplies. Very large cores are used to sense pulsing beams of high-energy particles. Some oscilloscope probes are highly compensated CTs with a core that can be opened to allow a current-carrying wire to be introduced. With modern winding equipment for toroids, it is possible to put 2000 or more turns on a secondary.* The CT then begins to look more like a current-to-voltage converter in its own right without the need for very small values of the burden resistor and subsequent voltage amplification.

Safety Note: The secondary of a CT should always remain shorted, while there is any possibility of current flow in the primary. With an open secondary, the core can be driven back and forth between saturation in opposite directions resulting in high-speed changes in the internal **B** field. The result is very high, dangerous to life, voltages on the open secondary. Insulation in the secondary circuit can be damaged by arcing. Many CTs are made with a provision for shorting the secondary if the circuit must be opened. Use it.

21.9 Gapped Inductive Sensors

It is common practice in the design of transformers to introduce a small gap in the magnetic path. For even very small gaps, the magnetic properties of the magnetic loop become almost completely determined by the length of the gap, the rest of the material serving only to contain the lines of flux. The analysis of such a device begins with the understanding that the **B** field is continuous in the core and through the gap. The **H** field is not, but it still satisfies the relation that the line integral of $\mathbf{H} \cdot d\mathbf{l}$ around the core is equal to the linked current. For a magnetic path of length s in material of permeability μ with a gap g, the ratio **B**/**H**, the effective permeability, is given by

$$\mu_{\text{eff}} = s/(g + s/\mu) \tag{21.6}$$

that applies for g much smaller than s. Note that when μ is sufficiently large, the effective permeability becomes independent of μ.

Introducing a gap into what would otherwise be a CT and drastically increasing the secondary turns count to 10,000 or more results in a current sensor, which is intrinsically safe because the core cannot saturate (Figure 21.6).† Because the **B** field is always small, the heating effect of eddy currents is less important than in the CT. When loaded with an appropriate resistor, the high inductance of the secondary causes the sensor to act like a current source, which generates a voltage across the load proportional to the primary current with better than 1% linearity. Useful bandwidth of a sensor can be over three decades. The output impedance is high and requires the use of electronic voltmeters. Power dissipation is low even for very high-current models. Output voltage is high enough that simple diode rectifiers can be used to provide for dc output to further processing. In many cases, such a sensor can be used without any special electronics other than a voltmeter.

* C.R. Magnetics is one source of high turns current transformers. Smith Research & Technology, SRT is no longer in active production.

† Smith Research and Technology. Inc., SRT of Colorado Springs once sold sensors like this. They are not currently available.

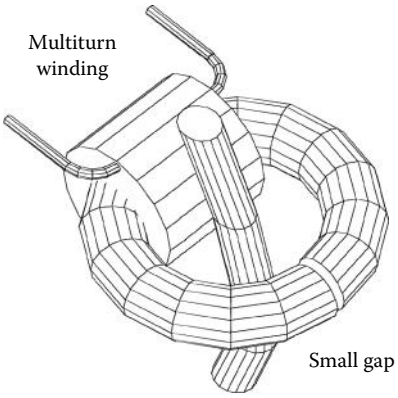


FIGURE 21.6 Placing a gap in the core and dramatically increasing the secondary turn count result in a current-to-voltage converter with high output voltage. The gap can be distributed using a powdered-iron core.

21.10 Hall Effect Sensor

The Hall effect as a sensor for magnetic fields is described in another chapter of this volume and in [7]. It depends on a semiconductor crystal selected for its high carrier mobility and placed in a magnetic field. A current is passed through the crystal along an axis perpendicular to the field. The carriers assume a mean velocity, which causes them to be acted upon by the field, and they move toward the other axis perpendicular to the field. The result is an EMF at the faces of the crystal that can be measured. The EMF is proportional to the field, the bias current, and the mobility.

In principle, such a field sensor could be placed near a current-carrying wire and oriented to sense the field created by the current (Figure 21.1), but the sensitivity is insufficient, and there would always be interfering fields from currents in other nearby wires. A flux concentrator that looks like a CT with a large gap is always used (see Figure 21.7).

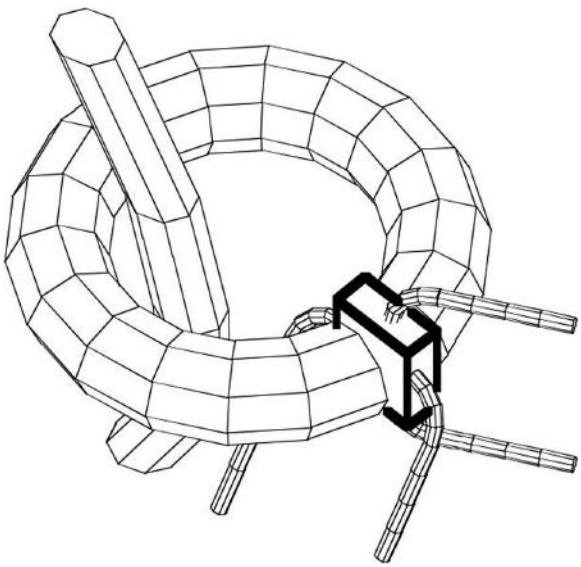


FIGURE 21.7 A semiconducting crystal is placed in the gap of a flux concentrating magnetic core. Bias current on one axis of the crystal produces a Hall voltage on the other.

The device is sensitive to dc and the polarity is preserved. The Hall voltage is a few millivolts, and amplification is always required. Zero drift in the amplifiers must be properly compensated though this is not so important for ac service. The bias current must be carefully controlled, and it can be used to provide intrinsic analog multiplication and metering of power if it is made proportional to the circuit voltage.

Sensitivity is best with the smallest gap, but there must be room for the crystal so gaps are larger than in the gapped inductive sensors. Fringing of the field in the larger gap reduces the natural shielding of the toroid from unwanted magnetic fields.

The accuracy and linearity of the Hall effect sensor can be improved in closed-loop mode. A feedback winding is added to the core and driven by a servo amplifier. The EMF from the Hall device is used to drive the servo amplifier until the field is zero. The output is then the feedback current, which is less than the sensed current by the number of turns in the feedback winding. The frequency response of the closed-loop system is surprisingly good, hundreds of kHz.*

21.11 Clamp-On Sensors

It is often desirable to measure current in an existing system without removing power in order to install a device so most of the magnetic devices are available in a clamp-on configuration. The variety ranges from tiny oscilloscope probes to clamps for 3000 A power buses.

Accuracy is always reduced in a clamp-on mode because the clamp itself constitutes a gap that is uncontrollable and subject to wear. Some manufacturers provide highly polished surfaces that slide together. Others have iron fingers that interlace as the clamp is closed. Still, others don't worry about it because the instrument is intended for field use where accuracy is not so critical.

Some units have handles for one-hand operation while others require a wrench or other tools. An interesting variation is the flexible core by Flexcorp.† Installation is by bending the core and allowing it to spring back to shape.

21.12 Magnetoresistive Sensors

Most of the features of a Hall effect sensor are available if the Hall crystal is replaced by a device whose resistance changes with magnetic field. The discovery of giant magnetoresistive devices has recently made this idea attractive.‡

Such devices still exhibit rather small resistance change and are sensitive to other effects such as temperature so it is imperative that they be used in self-compensating bridge circuits in the manner of a strain gauge. They are also insensitive to polarity of the field.

Zetex has delivered magnetoresistive current sensors using thin-film permalloy in a variety of printed circuit packages. They are constructed in the form of a bridge and require bias and a differential amplifier. Measured current up to 20 A passes through the chip via its solder pins.

21.13 High-Side Sensing

In fairly low-voltage, especially battery-powered, equipment, it has become necessary to monitor and control current used by various parts of an electronic device in the interest of extending battery life and presenting "fuel gauge" information on the state of charge to a user. The systems typically require a reliable common ground that would be unavailable if shunt resistors were to be installed on various parts for monitoring of current on the ground side. A simplified circuit that has been around for a while

* F. W. Bell offers considerable knowledge on its website. There is also a Wikipedia entry at Hall sensor which offers other references.

† Flex-Core offers Split-Core current transformers that can be wrapped around a large conductor.

‡ NVE corporation provides a discussion of the giant magnetoresistive effect at GMR.

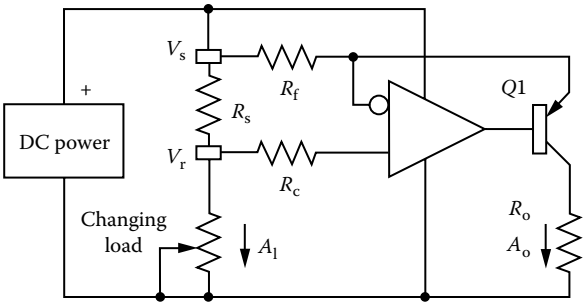


FIGURE 21.8 Schematic for an operational amplifier that converts the voltage across a high-side shunt into a ground-referenced current.

has found a new life in the form of dedicated integrated circuit chips that produce a ground-referenced measure of the voltage across a shunt in the positive power rail.

Referring to Figure 21.8, the current from a power source drives a changing load via a shunt resistor R_s, which is chosen to drop 10–100 mV or so at the maximum current to be measured. The load current creates voltages V_r and V_s with the difference between them equal to the load current A_l times the resistance R_s. The operational amplifier requires that the voltage at its inverting input match the reference voltage V_r, which it accomplishes by turning on PNP transistor $Q1$ just enough that the current in R_f produces the same drop as that across the shunt. The current, A_o, in output resistor R_o, providing that the current gain of $Q1$ is high, is the same as the current in R_f. The overall result is that $R_fA_o = R_sA_l$, and A_o is the same as the load current scaled by R_s/R_f. The current A_o can be used directly as the input of a current switching analog-to-digital converter, or it can be converted to a voltage with resistor R_o or with two op-amps wired as an inverting current-to-voltage converter followed by an inverter.

There are some serious considerations. The op-amp must operate from the possibly high power voltage. Resistor R_c is not critical but is often set to equal R_f to compensate for the less than infinite input resistance of the op-amp. The circuit as shown monitors the current in the load while ignoring the current used by the op-amp and the output. The details of the op-amp, including internal biasing that allows for common mode input at the supply voltage, are nontrivial, but they have been approached by designers of integrated circuits made for the purpose. Those chips may also include R_f, R_c, $Q1$, and an output amplifier leaving only the shunt to be selected. National Semiconductor and Zetex (see the table of suppliers) among others have offerings.

21.14 Fluxgates

In its simplest form, a fluxgate magnetometer uses a driving coil to drive a high-permeability rod into saturation first in one direction and then in the other. A second winding observes the rate of change of the **B** field inductively. In the absence of an external magnetic field, the observed signal is symmetric, but when an external field shifts, the hysteresis loop to the right or left the symmetry is lost. An amplifier tuned and phase locked to the second harmonic of the driving frequency can be used to determine the amplitude of the external field.

In principle, such a fluxgate could be used to sense the field in the gap of a flux concentrator as with the Hall sensor, but it's too big and the linearity would be unacceptable. It's better to think of a way to drive the whole flux concentrator with a pumping frequency similar to that used in a magnetic amplifier.

Driving a toroid this way has the undesirable effect of coupling the pump energy into the circuit being measured, but one can pump two toroids in opposite directions. Now, a current to be sensed passes through both cores and biases one in one direction and one in the other relative to the pump. A pickup winding senses second harmonic, which is demodulated with a phase-locked sensor to preserve the direction of the sensed current.

Linearity is improved by adding one more winding through both toroids, which is driven by a servo amplifier that is satisfied only when the amplitude of the second harmonic is zero; the flux at the measured frequency is zero. This is now the ideal CT. At the frequency of the sensed current, which might be dc, there is zero average flux in the toroids, and the output of the servo amplifier is a true representation of the current being measured reduced by the turn count of the servo winding.

Unfortunately, complicated electronics and considerable power are required to accomplish all of this, but the resulting accuracy is significantly better than anything other than precision shunts. Suppliers* provide equipment, which uses these principles. Many of the details are either patented or proprietary.

21.15 Optical Sensors

The Faraday effect is a rotation of the plane of polarization of light as it passes in a transparent medium parallel to a magnetic field. Optical fibers that do not randomize polarization are available with a small Faraday coefficient. Winding such a fiber on a form to be installed around a wire so that the light propagates parallel to the field produced by current in the wire (Figure 21.1) results in a sensor. A measurable rotation proportional to the current comes about because of the long optical path. When the myriad problems of preparation are solved, the result is a sensor that looks externally like a CT but has no wires. Using a reflecting ring made of yttrium–iron–garnet with a large Faraday coefficient, NIST reports sensitivity of 220 nA. Matsushita Electric Industrial Company makes sensors using thin garnet films. More information can be found in Refs. [8–11].

The analysis of the polarization requires a polarized light source and a polarization analyzer at the receiving end. An advantage of such a sensor is that the fiber leads to and from can be long enough to provide isolation for very high voltages.

Winding and annealing of fiber sensors without destroying the polarization-preserving properties of the fiber and temperature sensitivity of the Faraday coefficient must be addressed. Further information on these sensors can be found in [6].† ABB, Sweden, reports a maximum detectable current of >23 kA, a sensitivity of about 2 A, and a relative error of ±0.15% [6].

21.16 Fault Indicators

Latching indicators that save an indication of high pulse current, which was present sometime in the past, are needed for power distribution systems subject to lightning strikes. Such indicators are often a bar magnet hanging on a transmission wire, which moves into a latched position when a current pulse occurs. Readout is visual and they can be reset and installed on live wires using high-voltage tools.

21.17 Other Schemes

The double balanced mixer, familiar to RF engineers, is also a sensor for low-frequency current. The IF port is usually dc coupled through diodes in a ring configuration. The dc applied there will modulate an RF pump applied to one of the RF ports and recovered on the other. There are no known products that use this principle.

Magnetic amplifiers take advantage of nonlinearity of the **B**–**H** curve in a transformer core. The efficiency of a transformer is adjusted by a dc or low-frequency current, which moves the operating point on the hysteresis curve. Excitation, a pump, is required at a higher frequency, which is passed through the transformer and then synchronously rectified and filtered into a higher-power representation of the

* Fluxgates GMW, Holec, Others Wiki is a start. A Google search for "+Fluxgate +magnetometer +'current sensor' -eddy -patent" generates 310 good places to look for more.

† C.R. Magnetics is one source of high turns current transformers. Smith Research & Technology, SRT is no longer in active production; Smith Research and Technology. Inc., SRT of Colorado Springs once sold sensors like this. They are not currently available.

low-frequency signal. The magnetic windings are designed so that the pump power cancels in the signal circuit. As a current-measuring device, the magnetic amplifier leaves much to be desired in linearity and frequency response, but it does provide isolation and is limited in sensitivity only by the number of turns placed in the sensed winding. The 20 mA dc off-hook current in a telephone is one such application that is probably not used at all anymore.

21.18 Some Generalities and Warnings

Except for very simple loads, the waveform of the current drawn by a load does not resemble the voltage waveform. Besides the well-known phase shift and power factor, current flows in harmonics and subharmonics of the power frequency. Bridge rectifiers with capacitor input filters draw pulses of current near the peaks of the voltage waveform. Triacs cause a phase shift of the fundamental and introduce odd harmonics. Triacs in a full-cycle mode with zero current switching may draw full current for a few cycles followed by zero current for a few more introducing frequency components below the power frequency. Frequency changers are likely to draw current in pulse-width-modulated bursts.

A classic error is to measure true rms voltage and true rms current and multiply the two to get power. Even after having corrected for the phases, this can still be wrong. In short, accurate measurement demands some knowledge of the characteristics of the load. Beware of sensors labeled "true rms" for they may be anything but.

21.19 Current Actuated Switches and Indicators

Another form taken by current sensors is the current actuated switch. There was a time when contacts were placed on the pointer of a D'Arsonval movement to establish upper and lower limits for a process variable. The modern way is to configure a sensor so that it operates a solid-state switch.

When operating power is available, any of the current sensors can be configured as a switch, but when the switch must be controlled solely by the power that can be extracted from the current being measured, the gapped toroid is superior. The high-voltage available can be rectified to control MOSFETs or the base of an open collector Darlington transistor.*

One company [CR Magnetics] markets a light-emitting diode (LED), indicator that shows the presence of ac without any connection or external source of power.

Circuit breakers can be implemented with a magnetic coil or with a bimetallic thermal element. For completeness, there is also the fuse.

Ground fault breakers use a toroidal CT through which the line current is passed twice, once on the way to the load and again on the way back to the neutral wire in such a way as to cancel. Any fault current to ground associated with the load causes the return current to be less than the source and induces voltage on the secondary of the toroid. This is amplified and used to trip a switch. This application demands absolute insensitivity to the position of a current-carrying wire in the hole of the toroid. Gapped sensors typically are not suitable because of magnetic leakage.

21.20 Where to Get Current Sensors

CTs, shunts, and some other sensors are commodity items available from a variety of manufacturers and distributors. Well-known manufacturers of switchgear, General Electric and Westinghouse, have their own CTs for use with their power meters. Test equipment manufacturers, Simpson, Tektronix, Fluke, Extech, and the like, offer handheld current meters. Table 21.1 shows distributors and manufacturers of current sensing equipment. There is no way that a list such as this can be complete. It has selected itself from those companies willing to contribute to the author's library.

* Unpowered current actuated switches are possible using high turn count gapped CTs and power MOSFETs. They are no longer available from SRT.

TABLE 21.1 Selected Manufacturers and Distributors of Current Sensors

ABB, Inc.

Home: http://www.abb.com/

A big company that offers a family of high-accuracy fiber-optic current sensors.

AEMC through Optimum Energy Products Ltd.

#333, 11979 40 St SE, Calgary, AB, Canada T2Z 4M3

e-mail: info@optimumenergy.com

Home: http://www.myaemcstore.com/

AEMC Instruments is a manufacturer of meters, testers, and analyzers, such as current probes and dataloggers.

AEMC clamp-on CTs and other sensors

http://www.myaemcstore.com/c1084/clamp_on_meters.php

http://www.myaemcstore.com/c1104/flexible_current_probes.php

AEMC manufactures clamp-on current transformers and other sensors.

AEMC Instruments

200 Foxborough Blvd., Foxborough, MA 02035

Home: http://www.aemc.com/

Makes CT-based current sensors.

American Aerospace Controls

570 Smith St., Farmingdale, NY 11735

Home: http://www.a-a-c.com/

Has a line of ac and dc current sensors for mounting in remote locations with wiring to collect data.

Amprobe

e-mail: info@amprobe.com

Home: http://amprobe.com

Offers handheld, mostly clamp-on, meters for a wide range of ac currents.

Coilcraft

1102 Silver Lake Road, Cary, IL 60013

e-mail: cust_service@coilcraft.com

Home: http://www.Coilcraft.com

Manufactures smaller CTs mostly for mounting on printed circuit boards.

CR Magnetics, Inc.

3500 Scarlet Oak Blvd., St. Louis, MO 63122

Home: http://www.crmagnetics.com

Provides a full line of CTs and current actuated switches. Their technical references are worth a look.

Dranetz Technologies, Inc.

1000 New Durham Rd., Edison, NJ 08818 4019

Home: http://www.dranetz-bmi.com

Offers a line of equipment for analysis of ac power and is included here although it does not sell instruments dedicated only to current sensing.

Extech Instruments Corporation

285 Bear Hill Road, Waltham, MA 02451 1064

Home: http://www.extech.com/instruments/

Sells a line of handheld meters suitable for current measurement to 2000 A.

TABLE 21.1 (continued) Selected Manufacturers and Distributors of Current Sensors

F. W. Bell

6120 Hanging Moss Rd., Orlando, FL 32807

Home: http://www.fwbell.com

http://www.oeco.com/HallEffectSensors.html

http://www.oeco.com/Current_Sensors.html

F. W. Bell is now part of Pacific Scientific and Danaher. Information on Hall effect and magnetoresistive current sensors can be found on the Internet.

Flex-Core division of Morlan and Associates

Office: 6625 McVey Blvd., Columbus, OH 43235

Factory: 4970 Scioto-Darby Rd., Hilliard, OH 43026

e-mail: flexcore@msn.com

Home: http://www.flex-core.com/

Makes a line of sensors that cover the range of CTs. Of particular interest is their flexible iron core, which can be installed over an existing conductor by bending it.

Fluke

Home: http://www.fluke.com/

http://www.fluke.com/Fluke/usen/products/Clamp-Meters.htm

Fluke, with numerous sales offices and distributors, has a line of sensors for use with their handheld instruments.

GMW Danfysik

955 Industrial Road, San Carlos, CA 94070

Home: http://www.gmw.com/index.html

Deals in laboratory grade equipment for precision noncontact measurement. Precision in the ppm range and prices to match. Up to 10000 ADC and more than 200 kHz.

Holec division of Eaton Electric B.V.

Based in the Netherlands

e-mail: hppbv@pi.net

Home: http://www.eaton.com/EatonCom/ProductsServices/Holec/

Offers zero-flux CTs using servo amplifiers to drive the secondary actively. Their document 5.20.2 helps to understand the principles involved that operate from dc to 500 kHz.

LEM U.S.

Home: http://www.lem.com/

http://www.lem.com/hq/en/content/view/161/153/Open loop

http://www.lem.com/hq/en/content/view/168/155/Closed loop

Offers Hall effect sensors in both open- and closed-loop form for ac and dc currents.

MICRO SWITCH division of Honeywell, Inc.

Home: http://sensing.honeywell.com/

Inductive with digital output, Hall effect devices. Search for "current sensor" on their home page.

Neilsen-Kuljian

Supported by Clearwater Technologies, Inc.

1025 Exchange Street, Boise, ID 83716

e-mail: info@clrwtr.com

Home: http://www.clrwtr.com/

http://www.clrwtr.com/NK-Technologies-Product-Selection-Guide.htm

CTs, current switches, current transducers, ground fault sensors.

(*continued*)

TABLE 21.1 (continued) Selected Manufacturers and Distributors of Current Sensors

Nonvolatile Electronics (NVE)
11409 Valley View Road, Eden Prairie, MN 55344
e-mail: info@nve.com
Home: http://www.nve.com
Giant magnetoresistive sensor and others. Interesting catalog at http://www.nve.com/Downloads/catalog.pdf

Ohio Semitronics
4242 Reynolds Drive, Hilliard, OH 43026
Home: http://www.ohiosemi.com
Offers a wide range of current, voltage, and power transducers for ac and dc. Network-capable digital output devices.

Omega Engineering
Home: http://www.grainger.com
Offers a variety of current sensors and CTs mostly for industrial use.

Pearson Electronics, Inc.
4009 Transport Street, Palo Alto, CA 94303
Home: http://www.pearsonelectronics.com/
Provides accurate high-speed compensated CTs for pulse work with oscilloscopes.

Smith Research & Technology, Inc. (SRT)
Home: http://www.srt-inc.com
Once offered gapped inductive sensors with various output features. Current actuated switches. Website maintained but new products are not available.

SSAC
Now owned by SymCom
2880 North Plaza Drive, Rapid City, SD 57702
e-mail: CustomerService@ssac.com
Home: http://www.ssac.com/
Has a line of CTs and switching devices.

Zetex
Home: http://www.digikey.com/
Now owned by Diodes Inc. and well supported by DigiKey
Their ZXCT1020 high-side current monitoring IC is an example of the technology.

References

1. I. Genzer and P. Youngner, *Physics*, Chapter 10.6, Morristown, NJ: General Learning Corporation, 1973, A high school physics text.
2. H. E. White, *Modern College Physics*, Chapter 51, New York: D. Van Nostrand, 1952, Wiki/Ammeter has a nice sketch of a D'Arsonval ammeter.
3. B. N. Taylor, *Guide for the Use of the International System of Units (SI)*, NIST Special Publication 811, Washington, DC: U.S. Government Printing Office, 1995, Appendix A.5.
4. P. Mohr, Proposal to change the definition of the ampere. A PDF of a presentation, 2006.
5. A. Sommerfeld, Electrodynamics, *Lectures on Theoretical Physics*, Vol. III, New York: Academic Press, 1952, Web based sources of the same information are at magnetic field and magnetism.
6. R. C. Dorf (ed.), *The Electrical Engineering Handbook*, Chapter 49, The Hall effect, Boca Raton, FL: CRC Press, 1993.

7. Honeywell, Hall Effect Sensors, Microswitch Sensing and Control, http://www.honeywell-sensor.com.cn/prodinfo/magnetic_position/technical/chapter2.pdf, accessed August 15, 2013.
8. K. B. Rochford, A. H. Rose, M. N. Deeter, and G. W. Day, Faraday effect optical current sensor with improved sensitivity-bandwidth product, *Optics Letters* 19, 1903–1905, 1994.
9. S. R. Forrest (ed.), *JTEC Panel Report on Optoelectronics in Japan and the United States*, Baltimore, MD: International Technology Research Institute, Loyola College, February 1996, NTIS PB96-152202. http://itri.loyola.edu/opto/c6_s3.htm
10. M. N. Deeter, Domain effects in Faraday effect sensors based on iron garnets, *Applied Optics*, 34, 655, 1995.
11. M. N. Deeter, A. H. Rose, and G. W. Day, Faraday effect magnetic field sensors based on substituted iron garnets, in *Fiber Optic and Laser Sensors VIII*, R. P. DePaula and E. Udd (eds.), *Proc. Soc. Photo-Opt Instrumentation. Eng*. 1367, 243–248, 1990.

22

Power Measurement

Pasquale Arpaia
Università del Sannio

Francesco Avallone
Università di Napoli Federico II

Aldo Baccigalupi
Università di Napoli Federico II

Carmine Landi
Università de L'Aquila

In this chapter, the concept of electric power is first introduced, and then the most popular power measurement methods and instruments in *dc*, *ac*, and *pulse* waveform circuits are illustrated.

Power is defined as the *work performed per unit time*. So, dimensionally, it is expressed as joules per second, J s^{-1}. According to this general definition, electric power is the electric work or energy dissipated per unit time, and dimensionally, it yields

$$\mathrm{J\,s^{-1}} = \mathrm{J\,C^{-1}} \times \mathrm{C\,s^{-1}} = \mathrm{V} \times \mathrm{A} \tag{22.1}$$

where
J = Joules
s = seconds
C = Coulombs
V = volts
A = amperes

The product *voltage times current* gives an electric quantity equivalent to *power*.

22.1 Power Measurements in DC Circuits

Electric power (P) dissipated by a load (L) fed by a dc power supply (E) is the product of the voltage across the load (V_L) and the current flowing in it (I_L):

$$P = V_L \times I_L \tag{22.2}$$

Therefore, a power measurement in a dc circuit can be generally carried out using a voltmeter (V) and an ammeter (A) according to one of the arrangements shown in Figure 22.1. In the arrangement

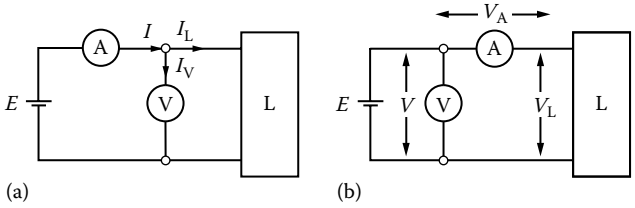


FIGURE 22.1 Two arrangements (a and b) for dc power measurement circuits.

of Figure 22.1a, the ammeter measures the current flowing into the voltmeter, as well as that into the load, whereas in the arrangement of Figure 22.1b, this error is avoided, but the voltmeter measures the voltage drop across the ammeter in addition to that dropping across the load. Thus, both arrangements give a surplus of power measurement absorbed by the instruments. The corresponding measurement errors are generally referred to as insertion errors.
According to the notation,

- I, current measured by the ammeter (Figure 22.1a)
- V, voltage measured by the voltmeter (Figure 22.1b)
- R_V, R_A, internal resistance of the voltmeter and the ammeter, respectively
- R_L, load resistance
- I_V, current flowing into the voltmeter (Figure 22.1a)
- V_A, voltage drop across the ammeter (Figure 22.1b)

The following expressions between the measurand electric power P and the measured power $V \times I$ are derived by analyzing the circuits of Figure 22.1a and b, respectively:

$$P = V_L \times I_L = V \times I \times \left(\frac{R_V - R_L}{R_V} \right) \tag{22.3}$$

$$P = V_L \times I_L = V \times I \times \left(\frac{R_L - R_A}{R_L} \right) \tag{22.4}$$

If

- I_V, compared with I
- V_A, compared with V

are neglected for the arrangements of Figure 22.1a and b, respectively, it approximately yields

$$\frac{I_V}{I} = \frac{R_L}{R_V + R_L} \cong \frac{R_L}{R_V} \cong 0; \quad \frac{V_A}{V} \cong \frac{R_A}{R_A + R_L} \cong \frac{R_A}{R_L} \cong 0 \tag{22.5}$$

consequently, measured and measurand power will be coincident.

On this basis, from Equations 22.3 through 22.5, analytical corrections of the insertion errors can be easily derived for the arrangement of Figure 22.1a and b, respectively.

The instrument most commonly used for power measurement is the *dynamometer*. It is built by (1) two fixed coils, connected in series and positioned coaxially with space between them, and (2) a moving coil, placed between the fixed coils and equipped with a pointer (Figure 22.2a).

The torque produced in the dynamometer is proportional to the product of the current flowing into the fixed coils times that in the moving coil. The fixed coils, generally referred to as *current coils*,

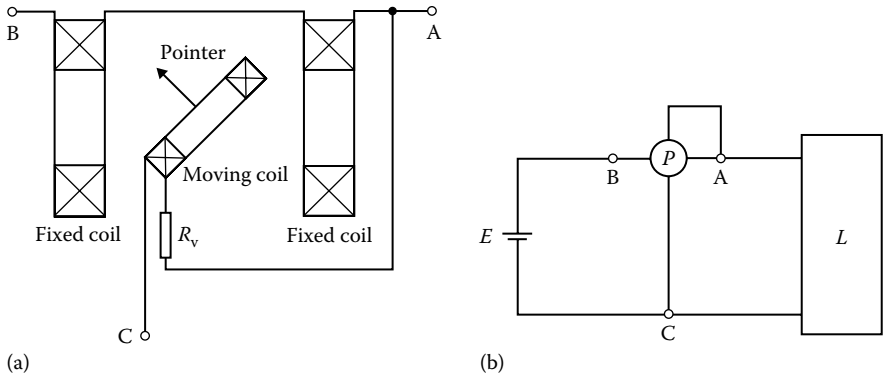


FIGURE 22.2 Power measurement with a dynamometer: (a) working principle and (b) measurement circuit.

carry the load current, while the moving coil, generally referred to as *voltage coil*, carries a current that is proportional, via the multiplier resistor R_V, to the voltage across the load resistor R_L. As a consequence, the deflection of the moving coil is proportional to the power dissipated into the load.

As for the case of Figure 22.1, insertion errors are also present in the dynamometer power measurement. In particular, by connecting the voltage coil between A and C (Figure 22.2b), the current coils carry the surplus current flowing into the voltage coil. Consequently, the power P_L dissipated in the load can be obtained by the dynamometer reading P as

$$P_L = P - \frac{V^2}{R'_v} \tag{22.6}$$

where R'_v is the resistance of the voltage circuit ($R'_v = R_v + R_{vc}$, where R_{vc} is the resistance of the voltage coil). By connecting the moving coil between B and C, this current error can be avoided, but now the voltage coil measures the surplus voltage drop across the current coils. In this case, the corrected value is

$$P_L = P - I^2 R_C \tag{22.7}$$

where R_C is the resistance of the current coil.

22.2 Power Measurements in AC Circuits

22.2.1 Definitions

All the earlier considerations relate to *dc* power supplies. Now look at power dissipation in *ac*-fed circuits. In this case, electric power, defined as voltage drop across the load times the current flowing through it, is the function

$$p(t) = v(t) \times i(t) \tag{22.8}$$

referred to as the *instantaneous power*. In ac, one is mainly interested in the mean value of instantaneous power for a defined time interval. In circuits fed by periodic ac voltages, it is relevant to define the mean power dissipated in one period T (*active power P*):

$$p = \frac{1}{T}\int_0^T p(t)dt \tag{22.9}$$

The simplest case is a sinusoidal power supply feeding a purely resistive load. In this case, $v(t)$ and $i(t)$ are in phase and $p(t)$ is given by

$$p(t) = VI[1 - \cos(2\omega t)] \tag{22.10}$$

where

V and I are the rms values of $v(t)$ and $i(t)$, respectively
ω is the power supply angular frequency

Therefore, the instantaneous power is given by a constant value VI plus the *ac* quantity oscillating with twice the angular frequency of the power supply; thus, the active power is simply the product VI. In this case, all the earlier considerations referring to active power for *dc* circuits are still correct, but voltages and currents must be replaced by the corresponding rms values.

The case of purely reactive loads is the opposite; the voltage drop across the load and current flowing through it are out of phase by 90°. Instantaneous power $p(t)$ is given by

$$p(t) = VI\cos(2\omega t) \tag{22.11}$$

Thus, the active power dissipated by a reactive load is zero, owing to the phase introduced by the load itself between voltage and current.

The simplest cases of sinusoidal power sources supplying purely resistive and purely reactive loads have been discussed. In these cases, the load is expressed by a real or a pure imaginary number. In general, the load is represented by a complex quantity (the impedance value). In this case, load impedance can be represented by its equivalent circuit (e.g., a pure resistance and a pure reactance in series). With this representation in mind, the electric power dissipated in the load Z_L (Figure 22.3) can be expressed by the sum of power components separately dissipated by resistance R_{EQ} and reactance X_{EQ} of the equivalent circuit Z_L.

Considering that no active power is dissipated in the reactance X_{EQ}, it yields

$$P = V_{REQ}I_L = V_L I_L \cos\varphi \tag{22.12}$$

The term $\cos\varphi$ appearing in Equation 22.12 is referred to as the *power factor*. It considers that only a fraction of voltage V_L contributes to the power; in fact, its component V_{XEQ} (the drop across the reactance) does not produce any active power, as it is orthogonal to the current I_L flowing into the load.

Figure 22.4 plots the waveforms of instantaneous power $p(t)$, voltage $v(t)$, and current $i(t)$. The effect of the power factor is demonstrated by a dc component of $p(t)$ that varies from a null value (i.e., $v(t)$ and $i(t)$ displaced by 90°) toward the value VI (i.e., $v(t)$ and $i(t)$ in phase).

The term

$$P_A = V_L I_L \tag{22.13}$$

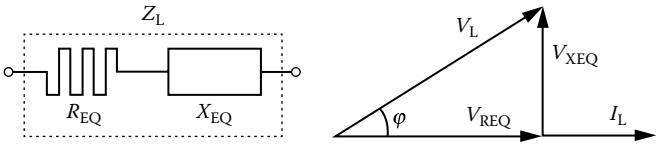


FIGURE 22.3 Voltage drop on the load and on its equivalent components.

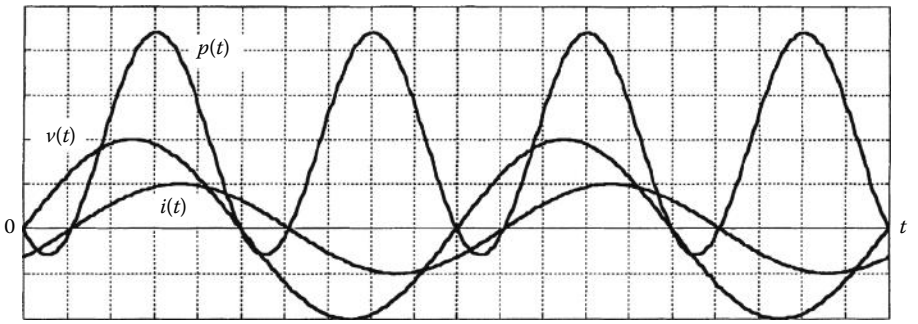


FIGURE 22.4 Waveforms of instantaneous power (p), voltage (v), and current (i).

is called the *apparent power*, while the term

$$Q = V_{XEQ} I_L = V_L I_L \sin\varphi \tag{22.14}$$

is called the *reactive power* because it represents a quantity that is dimensionally equivalent to power. This is introduced as a consequence of the voltage drop across a pure reactance and, therefore, does not give any contribution to the active power. From Figure 22.3, the relationship existing between *apparent power*, *active power*, and *reactive power* is given by

$$P_A = \sqrt{P^2 + Q^2} \tag{22.15}$$

Dynamometers working in ac are designed to integrate instantaneous power according to Equation 22.9. Insertion errors can be derived by simple considerations analogous to the dc case. However, in ac, a phase uncertainty due to the not purely resistive characteristic of voltage circuit arises. In sinusoidal conditions, if ε_w (in radians) is the phase of the real coil impedance and $\cos\varphi$ is the load power factor, the relative uncertainty in active power measurements can be shown to be equal to $-\varepsilon_w T_g \varphi$ where T_g is period. The phase uncertainty depends on the frequency. By using more complex circuits, the frequency range of the dynamometer can be extended to a few tens of kilohertz.

The aforementioned has presented the power definitions applied to *ac* with the restrictions of sinusoidal quantities. In the most general case of distorted quantities, obviously symbolic representation can no longer be applied. In any case, active power is always defined as the mean power dissipated in one period.

As far as methods and instruments for *ac* power measurements are concerned, some circuit classification is required. In fact, the problems are different, arising in circuits as the frequency of power supply increases. Therefore, in the following, *ac* will be classified into (1) line-frequency circuits, (2) low- and medium-frequency circuits (up to a few megahertz), and (3) high-frequency circuits (up to a few gigahertz). Line-frequency circuits will be discussed separately from low-frequency circuits, principally because of the existence of problems related specifically to the three-phase power supply of the main.

22.2.2 Low- and Medium-Frequency Power Measurements

In the following, the main methods and instruments for power measurements at low and medium frequencies are considered.

22.2.2.1 Three-Voltmeter Method

The power dissipation in the load L can be measured using a noninductive resistor R and measuring the three voltages shown in Figure 22.5 [1]. Although one of the voltages might appear redundant on a first analysis of the circuit, in actual fact, three independent data are needed in order to derive power from

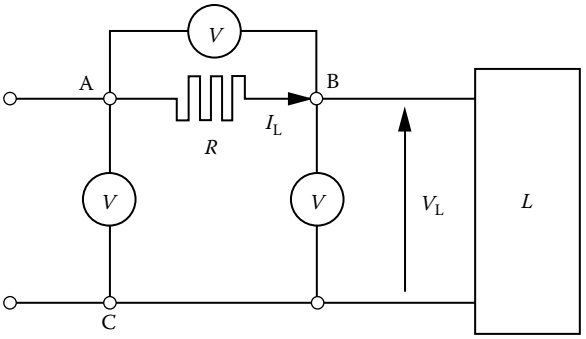


FIGURE 22.5 Three-voltmeter method.

Equation 22.12. In particular, from voltage drops v_{AB} and v_{BC}, the load current and load voltage can be directly derived; instead, v_{AC} is used to retrieve information about their relative phase.

If currents derived by voltmeters are neglected and the current i_L flowing into the load L is assumed to be equal to that flowing into the resistor R, the statement can be demonstrated as follows:

$$\begin{aligned} v_{AC} &= v_L + Ri_L \\ v_{AC}^2 &= R^2 i_L^2 + v_L^2 + 2Rv_L i_L \end{aligned} \tag{22.16}$$

where the small characters indicate instantaneous values. By computing rms values (indicated as capital characters), one obtains the power P_L:

$$\begin{aligned} &\frac{1}{T}\int_0^T v_{AC}^2 dt = \frac{1}{T}\int_0^T R^2 i_L^2 dt = \frac{1}{T}\int_0^T v_L^2 dt + \frac{1}{T}\int_0^T 2Rv_L i_L dt \\ &V_{AC}^2 = RI_L^2 + V_L^2 + 2RP_L \\ &P_L = \frac{V_{AC}^2 - R^2 I_L^2 - V_L^2}{2R} = \frac{V_{AC}^2 - V_{AB}^2 - V_{BC}^2}{2R} \end{aligned} \tag{22.17}$$

Equation 22.17 is also the same in dc by replacing rms values with dc values. Since the result is obtained as a difference, problems arise from the relative uncertainty when the three terms have a sum near zero.

Such a method is still used for high-accuracy applications.

22.2.2.2 Thermal Wattmeters

The working principle of thermal wattmeters is based on a couple of twin thermocouples whose output voltage is proportional to the square of the rms value of the currents flowing into the thermocouple heaters [2].

The principal circuit of a thermal wattmeter is shown in Figure 22.6a. Without the load, with the hypothesis $S \ll r_1$ and $S \ll r_2$, the two heaters are connected in parallel, and, if they have equal resistance r ($r_1 = r_2 = r$), they are polarized by the same current i_p:

$$i_1 = i_2 = \frac{i_p}{2} = \frac{v}{2R + r} \tag{22.18}$$

In this case, the output voltages of the two thermocouples turn out to be equal ($e_1 = e_2$); thus, the voltage Δe measured by the voltmeter is null. In Figure 22.6b, this situation is highlighted by the working point T equal for both thermocouples. By applying a load L with a corresponding current i_L, a voltage

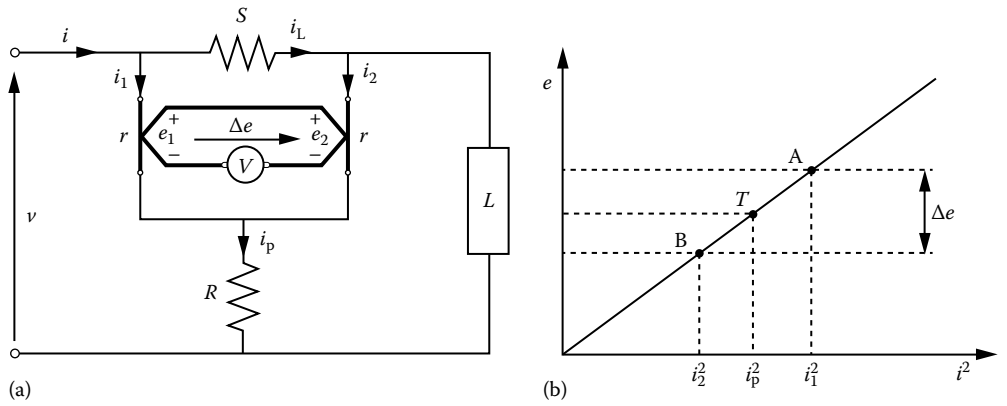


FIGURE 22.6 Thermal wattmeter based on twin thermocouples (a) and working characteristic in ideal conditions (b).

drop across S arises, causing an imbalance between currents i_1 and i_2. With the hypothesis that $r \ll R$, the two heaters are in series; thus, the current imbalance through them is

$$i_1 - i_2 = \frac{Si_L}{2R} \tag{22.19}$$

This imbalance increases the current i_1 and decreases i_2. Therefore, the working points of the two thermocouples change: the thermocouple polarized by the current i_1 operates at A, and the other thermocouple operates at B (Figure 22.6b). In this situation, with the earlier hypotheses, the voltmeter measures the voltage imbalance Δe proportional to the active power absorbed by the load (except for the surplus given by the powers dissipated in R, S, r_1, and r_2):

$$\begin{aligned}\Delta e &= k\left(\left\langle i_1^2\right\rangle - \left\langle i_2^2\right\rangle\right) = k\left(\left\langle (i_p + i_L)^2\right\rangle - \left\langle (i_P - i_L)^2\right\rangle\right) \\ &= k\left\langle 4i_p i_L\right\rangle = k_1\left\langle v(t)i(t)\right\rangle = k_1 P\end{aligned} \tag{22.20}$$

where the notation $\langle i \rangle$ indicates the time average of the quantity i.

If the two thermocouples cannot be considered as twins and linear, the power measurement accuracy will be obviously compromised. This situation is shown in Figure 22.7 where the two

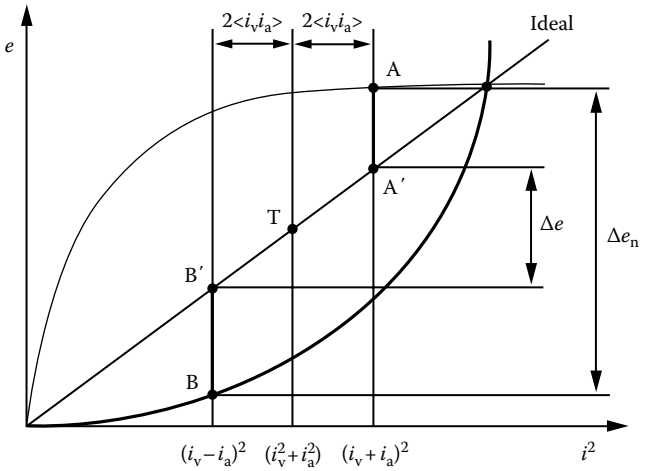


FIGURE 22.7 Ideal and actual characteristics of thermal wattmeter thermocouples.

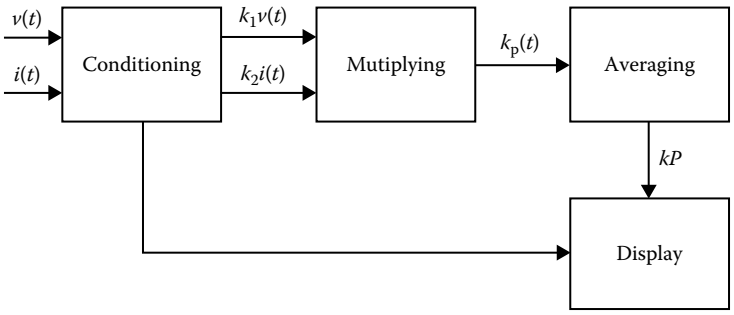


FIGURE 22.8 Block diagram of a multiplier-based wattmeter.

thermocouples are supposed to have two quite different nonlinear characteristics. In this case, the voltage measured by voltmeter will be Δe_n instead of Δe.

Wattmeters based on thermal principle allow high accuracy to be achieved in critical cases of highly distorted wide-band spectrum signals.

22.2.2.3 Wattmeters Based on Multipliers

The multiplication and averaging processes (Figure 22.8) involved in power measurements can be undertaken by electronic means.

Electronic wattmeters fall into two categories, depending on whether multiplication and averaging operations are performed in a continuous or discrete way. In continuous methods, multiplications are mainly carried out by means of analog electronic multipliers. In discrete methods, sampling wattmeters take simultaneous samples of voltage and current waveforms, digitize these values, and provide multiplication and averaging using digital techniques.

Analogous to the case of dynamometers, the resistances of the voltage and current circuits have to be taken into account (see Equations 22.6 and 22.7). Also, phase errors of both current ε_{wc} and voltage ε_{wv} circuits increase the relative uncertainty of power measurement, for example, in case of sinusoidal conditions increased at $(\varepsilon_{wc} - \varepsilon_{wv})T_g\varphi$.

22.2.2.4 Wattmeters Based on Analog Multipliers

The main analog multipliers are based on a transistor-based popular circuit such as a four-quadrant multiplier [3], which processes voltage and current to give the instantaneous power and an integrator to provide the mean power (Figure 22.9). More effective solutions are based on (1) time division multipliers (TDMs) and (2) Hall effect-based multipliers.

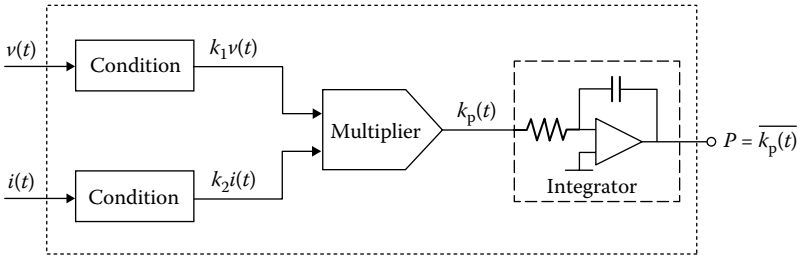


FIGURE 22.9 Block diagram of a four-quadrant, multiplier-based wattmeter.

22.2.2.4.1 TDM-Based Wattmeters

The block diagram of a wattmeter based on a TDM is shown in Figure 22.10 [4]. A square wave v_m (Figure 22.11a) with constant period T_g, and duty cycle and amplitude determined by $i(t)$ and $v(t)$, respectively, is generated. If T_g is much smaller than the period of measurands $v_x(t)$ and $v_\gamma(t)$, these voltages can be considered as constant during this time interval.

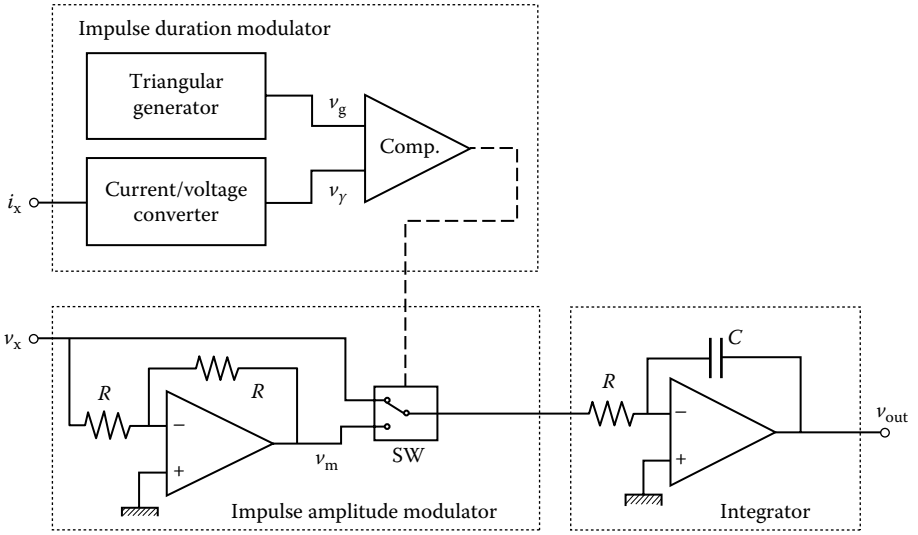


FIGURE 22.10 Block diagram of a TDM-based wattmeter.

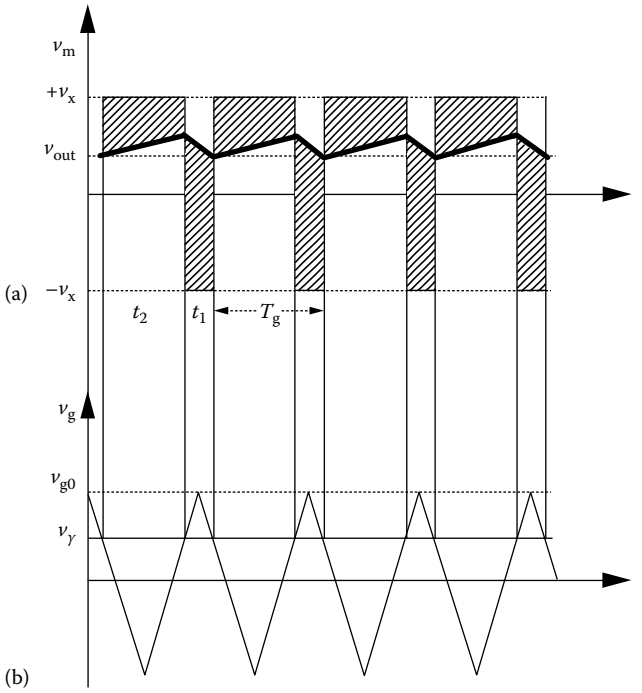


FIGURE 22.11 Waveform of the TDM-based power measurement: (a) impulse amplitude modulator output and (b) ramp generator output.

The duty cycle of v_m is set by an impulse duration modulator circuit (Figure 22.10). The ramp voltage $v_g(t)$ (Figure 22.11b) is compared to the voltage $v_\gamma(t)$ proportional to $i(t)$, and a time interval t_2, whose duration is proportional to $v_\gamma(t)$, is determined. If

$$v_g(t) = \frac{4V_{g0}}{T_g}t \quad \text{when } 0 \le t \le \frac{T_g}{4} \tag{22.21}$$

then from simple geometrical considerations, one obtains

$$t_2 = 2\left(\frac{T_g}{4} - \frac{v_y T_g}{4V_{g0}}\right) \tag{22.22}$$

and

$$t_1 - t_2 = \frac{T_g}{V_{g0}}v_y \tag{22.23}$$

The amplitude of $v_m(t)$ is set by an impulse amplitude modulator circuit. The output square wave of the impulse duration modulator drives the output $v_m(t)$ of the switch SW to be equal to $+v_x$ during the time interval t_1 and to $-v_x$ during the time interval t_2 (Figure 22.11a).

Then, after an initial transient, the output voltage $v_{out}(t)$ of the low-pass filter (integrator) is the mean value of $v_m(t)$:

$$v_{out} = \frac{1}{RC}\int_0^t v_m(t)dt = K'\left(\int_0^{t_1} v_x(t)dt - \int_{t1}^{t_1+t_2} v_x(t)dt\right) = K'v_x(t_1 - t_2) = Kv_x v_y \tag{22.24}$$

The high-frequency limit of this wattmeter is determined by the low-pass filter, and it must be smaller than half of the frequency of the signal $v_g(t)$. The frequency limit is generally between 200 Hz and 20 kHz and can reach 100 kHz. Uncertainties are typically 0.01%–0.02% [5].

22.2.2.4.2 Hall Effect-Based Wattmeters

As is well known, in a Hall effect transducer, the voltage $v_H(t)$ is proportional to the product of two time-dependent quantities [6]:

$$v_H(t) = R_H i(t)B(t) \tag{22.25}$$

where

R_H is the hall constant
$i(t)$ is the current through the transducer
$B(t)$ is the magnetic induction

In the circuit of Figure 22.12a, the power P is determined by measuring $v_H(t)$ through a high-input impedance averaging voltmeter and by considering that $v_x(t) = ai(t)$ and $i_x(t) = bB(t)$, where a and b are proportionality factors:

$$P = \frac{1}{T}\int_0^T v_x(t)\cdot i_x(t)dt = ab\frac{1}{T}\int_0^T i(t)\cdot B(t)dt = abR_H V_H \tag{22.26}$$

where

T is the measurand period
V_H is the mean value of $v_H(t)$

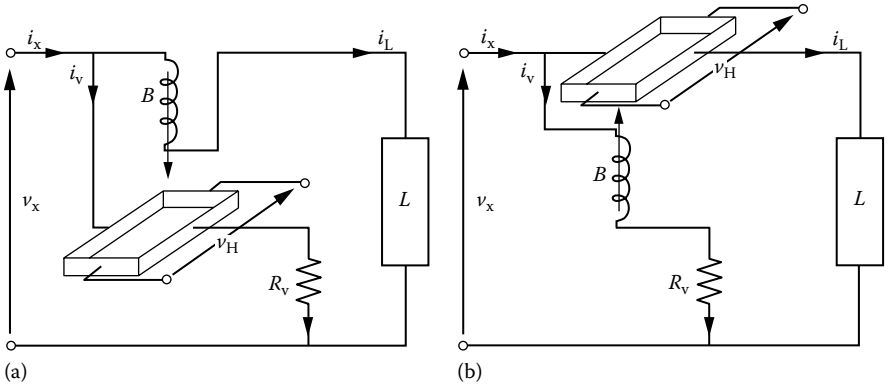


FIGURE 22.12 Configurations of the Hall effect-based wattmeter: (a) polarizing current set by resistor and (b) polarizing current set by load.

In the usual realization of the Hall multiplier (0.1% up to a few megahertz), shown in Figure 22.12a, the magnetic induction is proportional to the load current, and the optimal polarizing current i_v is set by the resistor R_v.

For the frequency range up to megahertz, an alternative arrangement is shown in Figure 22.12b, in which the load current I_L flows directly into the Hall device, acting as a polarizing current, and the magnetic field is generated by the voltage v. In this same way, the temperature influence is reduced for line-frequency applications with constant-amplitude voltages and variable load currents.

In the megahertz to gigahertz range, standard wattmeters use probes in waveguides with rectifiers.

22.2.2.5 Wattmeters Based on Digital Multipliers

22.2.2.5.1 Sampling Wattmeters

The most important wattmeter operating on discrete samples is the sampling wattmeter (Figure 22.13). It is essentially composed of two analog-to-digital input channels, each constituted by (1) a conditioner (C), (2) a sample/hold (S/H), (3) an analog-to-digital converter (ADC), (4) a digital multiplier (MUL), and (5) summing (SUM), dividing (DIV), and displaying units (DISP). The architecture is handled by a processing unit not shown in Figure 22.13.

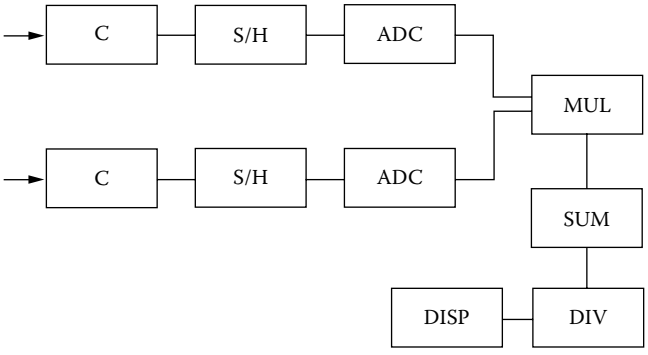


FIGURE 22.13 Block diagram of the sampling wattmeter.

If samples are equally spaced, the active power is evaluated as the mean of the sequence of instantaneous power samples $p(k)$:

$$\overline{P} = \frac{1}{N}\sum_{k=0}^{N-1} p(k) = \frac{1}{N}\sum_{k=0}^{N-1} v(k)i(k) \tag{22.27}$$

where

N represents the number of samples in one period of the input signal
$v(k)$ and $i(k)$ are the kth samples of voltage and current, respectively

A previous estimation of the measurand fundamental period is made to adjust the summation interval of Equation 22.27 and/or the sampling period in order to carry out a synchronous sampling [7]. The sampling period can be adjusted by using a frequency multiplier with phase-locked loop (PLL) circuit driven by the input signal [8]. Alternatively, the contribution of the sampling error is reduced by carrying out the mean on a high number of periods of the input signal.

In the time domain, the period estimation of highly distorted signals, such as pulse-width modulation (PWM), is made difficult by the numerous zero crossings present in the waveform. Some types of digital filters can be used for this purpose. An efficient digital way to estimate the period is the discrete integration of the PWM signal. In this way, the period of the fundamental harmonic is estimated by detecting the sign changes of the cumulative sum function [9]:

$$S(k) = \sum_{i=1}^{k} p_i \quad k = 1,2,\ldots,N \tag{22.28}$$

If the summation interval is extended to an integer number of periods of the $S(k)$ function, a "quasi-synchronous" sampling [10] is achieved through a few simple operations (cumulative summation and sign detection), and the maximum synchronization error is limited to a sampling period. Through relatively small increases in computational complexity and memory size, the residual error can be further reduced through a suitable data processing algorithm, that is, the multiple convolution in the time domain of triangular windows [9]. Hence, the power measurement can be obtained as

$$P_{(B)} = \frac{1}{\sum_{k=0}^{2B(N^{\star}-1)} w(k)} \sum_{k=0}^{2B(N^{\star}-1)} w(k)p(k) \tag{22.29}$$

where

$p(k)$ is the kth sample of the instantaneous power
$w(k)$ is the kth weight corresponding to the window obtained as the convolution of B triangular windows [10]

Another way to obtain the mean power is through the consideration of the harmonic components of voltages and currents in the frequency domain using the discrete Fourier transform [11]. In particular, a fast Fourier transform algorithm is used in order to improve efficiency. Successively, a two-step research of the harmonic peaks is carried out: (1) the indexes of the frequency samples corresponding to the greatest spectral peaks provide a rough estimate of the unknown frequencies when the wide-band noise superimposed onto the signal is below threshold; (2) a more accurate estimate of harmonic frequencies is carried out to determine the fractional bin frequency (i.e., the harmonic determination under the frequency resolution); to this aim, several approaches such as zero padding, interpolation techniques, and flattop window-based technique can be applied [12].

22.2.3 Line-Frequency Power Measurements

For line applications where the power is directly derived by the source network, the assumption of infinite power source can be reliably made, and at least one of the two quantities voltage or current can be considered as sinusoidal. In this case, the definition of the power as the product of voltage and current means that only the power at the fundamental frequency can be examined [13].

22.2.3.1 Single-Phase Measurements

Single-phase power measurements at line frequency are carried out by following the criteria previously mentioned. In practical applications, the case of a voltage greater than 1000 V is relevant; measurements must be carried out using voltage and current transformers inserted as in the example of Figure 22.14. The relative uncertainty is equal to

$$\frac{\Delta P}{P} = \left(\eta_w + \eta_a + \eta_v\right) + \left(\varepsilon_w + \varepsilon_a + \varepsilon_v\right) tg\varphi_c \tag{22.30}$$

where

η_w and ε_w are the instrumental and phase uncertainty of the wattmeter
η_a and η_v are the ratio uncertainties of current (CT) and voltage (VT) transformers
ε_a and ε_v are their phase uncertainties, respectively

If the load current exceeds the current range of the wattmeter, a current transformer must be used, even in the case of low voltages.

22.2.3.2 Polyphase Power Measurements

Three-phase systems are the polyphase systems most commonly used in practical industrial applications. In the following, power measurements on three-phase systems will be derived as a particular case of polyphase systems (systems with several wires) and analyzed for increasing costs: (1) balanced and symmetrical systems, (2) three-wire systems, (3) two-wattmeter-based measurements, (4) unbalanced systems, (5) three-wattmeter-based measurements, and (6) medium-voltage systems.

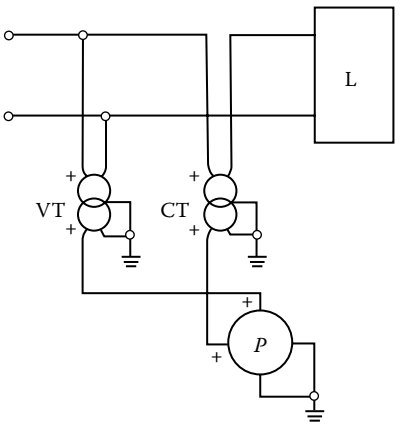


FIGURE 22.14 Single-phase power measurement with voltage (VT) and current (CT) transformers.

22.2.3.3 Measurements on Systems with Several Wires

Consider a network with sinusoidal voltages and currents composed by n wires. For the currents flowing in such wires, the following relation is established:

$$\sum_{1}^{n} \dot{I}_{\mathrm{i}} = 0 \tag{22.31}$$

The network can be thought as composed of $n - 1$ single-phase independent systems, with the common return on any one of the wires (e.g., the sth wire). Then, the absorbed power can be measured as the sum of the readings of $n - 1$ wattmeters, each one inserted with the current circuit on a different wire and the voltmeter circuit between such a wire and the sth one (Figure 22.15):

$$P = \sum_{1}^{n-1} \left(\dot{V}_{\mathrm{is}} \times \dot{I}_{\mathrm{i}} \right) \tag{22.32}$$

The absorbed power can be also measured by referring to a generic point O external to the network. In this case, the absorbed power will be the sum of the readings of n wattmeters, each inserted with the ammeter circuit on a different wire and the voltmeter circuit connected between such a wire and the point O:

$$P = \sum_{1}^{n} \left(\dot{V}_{\mathrm{io}} \times \dot{I}_{\mathrm{i}} \right) \tag{22.33}$$

22.2.3.4 Power Measurements on Three-Wire Systems

Active power in a three-phase power system can generally be evaluated by three wattmeters connected as shown in Figure 22.16.

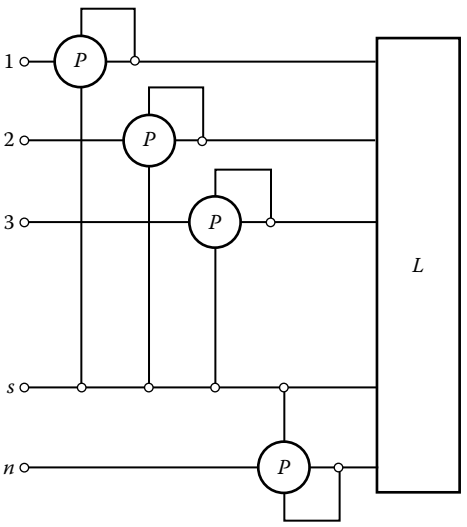


FIGURE 22.15 Power measurement on systems with several wires.

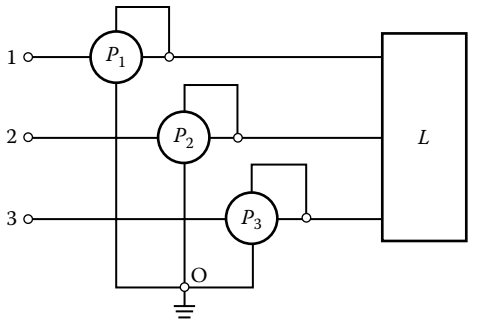


FIGURE 22.16 Power measurement on three-wire systems.

For each power meter, the current lead is connected on a phase wire, and the voltmeter lead is connected between the same wire and an artificial neutral point O, whose position is fixed by the voltmeter impedance of power meters or by suitable external impedances.

Under these conditions, absorbed power will be the sum of the three-wattmeter indications:

$$P = \sum_{1}^{3} \left(\dot{V}_{io} \times \dot{I}_i \right) \tag{22.34}$$

If the three-phase system is provided by four wires (three phases with a neutral wire), the neutral wire is utilized as a common wire.

22.2.3.5 Symmetrical and Balanced Systems

The supply system is symmetrical, and the three-phase load is balanced; that is,

$$\begin{cases} V_1 = V_2 = V_3 \\ I_1 = I_2 = I_3 \end{cases} \tag{22.35}$$

In Figure 22.17, the three possible kinds of insertion of an instrument S (an active power or a reactive power meter) are illustrated. The first (Figure 22.17) was described in the last section; if S is a wattmeter, the overall active power is given by three times its indication and similarly for the reactive power if S is a reactive power meter. Notice that a couple of twin resistors with the same resistance *R* of the voltage circuit of S are placed on the other phases to balance the load.

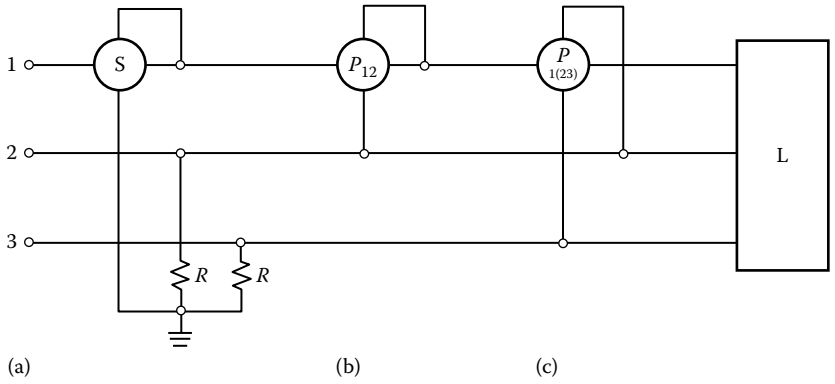


FIGURE 22.17 The three kinds of insertion of a power meter.

The other two insertions are indicated by the following convention: S_{ijk} indicates a reading performed with the current leads connected to the line "i" and the voltmeter leads connected between the phases "j" and "k." If "i" is equal to "j," one is omitted (e.g., the notation P_{12} (Figure 22.17b). The active power absorbed by a single phase is usually referred to as P_1.

The wattmeter reading corresponding to the (c) case in Figure 22.17 is equal to the reactive power Q_1 involved in phase 1, save for the factor $\sqrt{3}$. Hence, in the case of symmetrical and balanced systems, the overall reactive power is given by

$$Q = 3Q_1 = \frac{3P_{1(23)}}{\sqrt{3}} = \sqrt{3}P_{1(23)} \tag{22.36}$$

In fact, one has

$$P_{1(23)} = \dot{I}_1 \times \dot{V}_{23} \tag{22.37}$$

but

$$\begin{aligned} \dot{V}_{12} + \dot{V}_{23} + \dot{V}_{31} = 0 \quad &\Rightarrow \quad P_{1(23)} = \dot{I}_1 \times (-\dot{V}_{12} - \dot{V}_{31}) \\ \dot{V}_{13} = -\dot{V}_{31} \quad &\Rightarrow \quad P_{1(23)} = -\dot{I}_1 \times \dot{V}_{12} + \dot{I}_1 \times \dot{V}_{13} \\ \begin{cases} \dot{I}_1 \times \dot{V}_{12} = P_{12} \\ \dot{I}_1 \times \dot{V}_{13} = P_{13} \end{cases} \quad &\Rightarrow \quad P_{1(23)} = P_{13} - P_{12} \end{aligned}$$

In the same manner, the following relationships, which are valid for any kind of supply and load, can be all proved:

$$\begin{aligned} P_{1(23)} &= P_{13} - P_{12} \\ P_{2(31)} &= P_{21} - P_{23} \\ P_{3(12)} &= P_{32} - P_{31} \end{aligned} \tag{22.38}$$

If the supply system is symmetrical, $P_{1(23)} = \sqrt{3}Q_1$.

In fact, moving from the relationship (Figure 22.18),

$$P_{1(23)} = \dot{I}_1 \times \dot{V}_{23} = I_1 V_{23} \cos\beta \tag{22.39}$$

where $\beta = 90° - \varphi_1$, one obtains $P_{1(23)} = \sqrt{3}E_1 I_1 \sin\varphi_1 = \sqrt{3}Q_1$.

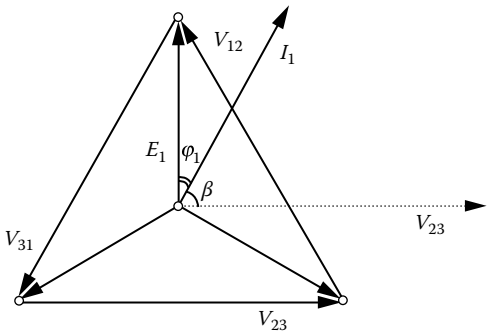


FIGURE 22.18 Phasor diagram for a three-phase symmetrical and balanced system.

In the same manner, the other two corresponding relationships for $P_{2(31)}$ and $P_{3(12)}$ are derived. Hence,

$$\begin{aligned} P_{1(23)} &= \sqrt{3}Q_1 = P_{13} - P_{12} \\ P_{2(31)} &= \sqrt{3}Q_2 = P_{21} - P_{23} \\ P_{3(12)} &= \sqrt{3}Q_3 = P_{32} - P_{31} \end{aligned} \tag{22.40}$$

22.2.3.6 Power Measurements Using Two Wattmeters

The overall active power absorbed by a three-wire system can be measured using only two wattmeters. In fact, Aron's theorem states the following relationships:

$$\begin{aligned} P &= P_{12} + P_{32} \\ P &= P_{23} + P_{13} \\ P &= P_{31} + P_{21} \end{aligned} \tag{22.41}$$

Analogously, the overall reactive power can be measured by using only two reactive power meters:

$$\begin{aligned} Q &= Q_{12} + Q_{32} \\ Q &= Q_{23} + Q_{13} \\ Q &= Q_{31} + Q_{21} \end{aligned} \tag{22.42}$$

Here, one of the previous statements, that is,

$$P = P_{12} + P_{32}$$

is proved. The two wattmeters connected as shown in Figure 22.19 furnish P_{12} and P_{32}.

Hence, the sum of the two readings gives

$$\begin{aligned} P_{12} + P_{32} &= \dot{I}_1 \times \dot{V}_{12} + \dot{I}_3 \times \dot{V}_{32} = \dot{I}_1 \times \left(\dot{E}_1 - \dot{E}_2\right) + \dot{I}_3 \times \left(\dot{E}_3 - \dot{E}_2\right) \\ &= \dot{I}_1 \times \dot{E}_1 - \dot{I}_1 \times \dot{E}_2 + \dot{I}_3 \times \dot{E}_3 - \dot{I}_3 \times \dot{E}_2 = \dot{I}_1 \times \dot{E}_1 + \dot{I}_3 \times \dot{E}_3 - \left(\dot{I}_1 + \dot{I}_3\right) \times \dot{E}_2 \\ &= \dot{I}_1 \times \dot{E}_1 + \dot{I}_3 \times \dot{E}_3 + \dot{I}_2 \times \dot{E}_2 = P_1 + P_2 + P_3 = P \end{aligned} \tag{22.43}$$

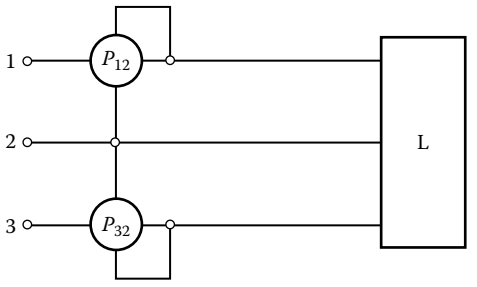


FIGURE 22.19 Power measurements using two wattmeters.

Provided that the system has only three wires, Aron's theorem applies to any kind of supply and load. In the case of symmetrical and balanced systems, it also allows the reactive power to be evaluated:

$$Q = \sqrt{3} \cdot (P_{32} - P_{12}) \tag{22.44}$$

Using Equations 22.41 and 22.44, the power factor is

$$\cos\varphi = \frac{P_{12} + P_{32}}{\sqrt{(P_{12} + P_{32})^2 + 3(P_{32} - P_{12})^2}} = \frac{P_{12} + P_{32}}{\sqrt{4P_{12}^2 + 4P_{32}^2 - 4P_{12}P_{32}}} = \frac{1 + (P_{12}/P_{32})}{2\sqrt{(P_{12}/P_{32})^2 - (P_{12}/P_{32}) + 1}} \tag{22.45}$$

Aron's insertion cannot be utilized when the power factor is low. In fact, if the functions

$$\begin{aligned} \cos(\varphi + 30) &= \frac{P_{12}}{VI} \\ \cos(\varphi - 30) &= \frac{P_{32}}{VI} \end{aligned} \tag{22.46}$$

are considered (Figure 22.20), it can be argued that (1) for $\varphi \leq 60°$, P_{12} and P_{32} are both greater than zero; (2) for $\varphi > 60°$, $\cos(\varphi - 30)$ is still greater than zero; and $\cos(\varphi + 30)$ is lower than zero.

The absolute error in the active power is

$$\Delta P = \frac{\partial\left(P_{12} + P_{32}\right)}{\partial P_{12}}\Delta P_{12} + \frac{\partial\left(P_{12} + P_{32}\right)}{\partial P_{32}}\Delta P_{32} = \Delta P_{12} + \Delta P_{32} \tag{22.47}$$

This corresponding relative error is greater as P_{12} and P_{32} have values closer to each other and are opposite in polarity; in particular, for $\cos\varphi = 0$ ($\varphi = 90°$), the error is infinite.

If η_w and ε_w are the wattmeter amplitude and phase errors, respectively, then the error in the active power is

$$\frac{\Delta P}{P} = \frac{\left(\eta_w + \varepsilon_w tg\varphi_{12}\right)P_{12} + \left(\eta_w + \varepsilon_w tg\varphi_{32}\right)P_{32}}{P_{12} + P_{32}} = \eta_w + \varepsilon_w \frac{Q}{P} \tag{22.48}$$

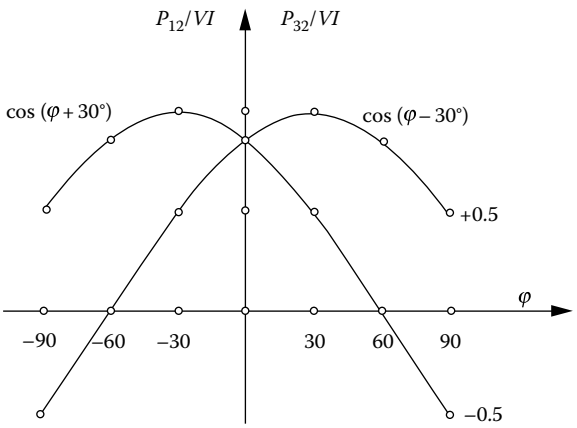


FIGURE 22.20 Sign of powers in Aron's insertion.

Let two wattmeters with nominal values V_0, I_0, $\cos\varphi_0$, and class c be considered; the maximum absolute error in each reading is

$$\Delta P = \frac{cV_0 I_0 \cos\varphi_0}{100} \tag{22.49}$$

Therefore, the percentage error related to the sum of the two indications is

$$\frac{\Delta P}{P} = \frac{cV_0 I_0 \cos\varphi_0}{\sqrt{3}VI\cos\varphi} = 1.11\frac{cV_0 I_0 \cos\varphi_0}{VI\cos\varphi} \tag{22.50}$$

equal to approximately the error of only one wattmeter inserted in a single-phase circuit with the same values of I, V, and $\cos\varphi$. Consequently, under the same conditions, the use of two wattmeters involves a measurement uncertainty much lower than that using three wattmeters.

If Aron's insertion is performed via current and voltage transformers, characterized by ratio errors η_a and η_v and phase errors ε_a and ε_v, respectively, the active power error is

$$\frac{\Delta P}{P} = \frac{(\eta_{TOT} + \varepsilon_{TOT} tg\varphi_{12})P_{12} + (\eta_{TOT} + \varepsilon_{TOT} tg\varphi_{32})P_{32}}{P_{12} + P_{32}} = \eta_{TOT} + \varepsilon_{TOT}\frac{Q}{P} = \eta_{TOT} + \varepsilon_{TOT} tg\Phi_c \tag{22.51}$$

where $\cos\Phi_c$ is the conventional power factor

$$\left.\begin{aligned}\eta_{TOT} &= \eta_w + \eta_a + \eta_v \\ \varepsilon_{TOT} &= \varepsilon_w + \varepsilon_a + \varepsilon_v\end{aligned}\right\}$$ the error sums with η_w and ε_w being the wattmeter errors

22.2.3.7 Symmetrical Power Systems Supplying Unbalanced Loads

If the load is unbalanced, the current amplitudes are different from each other, and their relative phase is not equal to 120°. Two wattmeters and one voltmeter have to be connected as proposed by Barbagelata [13] (Figure 22.21). The first wattmeter can provide P_{12} and P_{13}, and the second one gives P_{31} and P_{32}.

From Aron's theorem, the active power is

$$P = P_{12} + P_{32} \tag{22.52}$$

and then the reactive power Q is

$$Q = Q_1 + Q_2 + Q_3 = \frac{1}{\sqrt{3}}[P_{13} - P_{12} + \underline{P_{21}} - \underline{P_{23}} + P_{32} - P_{31}] \tag{22.53}$$

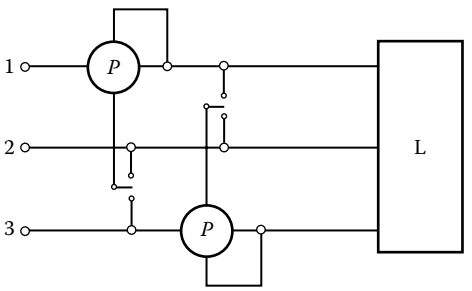


FIGURE 22.21 Barbagelata's insertion for symmetrical and unbalanced systems.

For the underlined terms, from Aron's theorem, it follows that

$$P = P_{13} + P_{23} = P_{12} + P_{32} = P_{21} + P_{31} \tag{22.54}$$

then

$$P_{13} + P_{23} = P_{21} + P_{31} \quad \Rightarrow \quad P_{21} - P_{23} = P_{13} - P_{31}$$

Thus, one obtains

$$Q = \frac{1}{\sqrt{3}}[2(P_{13} - P_{31}) + P_{32} - P_{12}] \tag{22.55}$$

Therefore, using only four power measurements, the overall active and reactive powers can be obtained.

The main disadvantage of this method is that the four measurements are not simultaneous; therefore, any load variations during the measurement would cause a loss in accuracy. In this case, a variation proposed by Righi [13] can be used. In this variation, three wattmeters are connected as shown in Figure 22.22 and give simultaneously P_{12}, P_{32}, and $P_{2(31)}$. Reactive power is

$$Q = \frac{1}{\sqrt{3}}\left[\underline{P_{13}} - P_{12} + P_{21} - P_{23} + P_{32} - \underline{P_{31}}\right] \tag{22.56}$$

Analogously as aforementioned, from Aron's theorem, it follows that

$$P_{21} - P_{23} = P_{13} - P_{31} \quad \Rightarrow \quad P_{2(31)} = P_{21} - P_{23} = P_{13} - P_{31} \tag{22.57}$$

then

$$Q = \frac{1}{\sqrt{3}}[P_{32} - P_{12} + 2P_{2(31)}] \tag{22.58}$$

For symmetrical and unbalanced systems, another two-wattmeter insertion can be carried out (Figure 22.23). The wattmeters give

$$\begin{aligned} P_{1(30)} &= \dot{E}_3 \times \dot{I}_1 = j\frac{\dot{V}_{12}}{\sqrt{3}} \times \dot{I}_1 = -\frac{Q_{12}}{\sqrt{3}} \\ P_{3(10)} &= \dot{E}_1 \times \dot{I}_3 = j\frac{\dot{V}_{23}}{\sqrt{3}} \times \dot{I}_3 = -\frac{Q_{32}}{\sqrt{3}} \end{aligned} \tag{22.59}$$

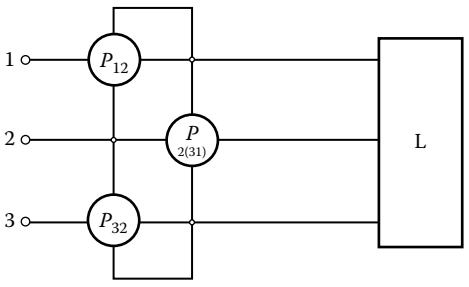


FIGURE 22.22 Righi's insertion for symmetrical and unbalanced systems.

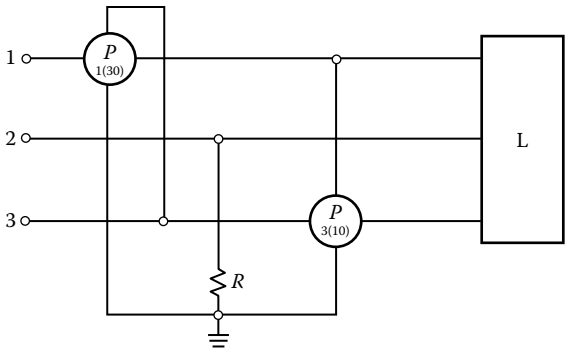


FIGURE 22.23 Two-wattmeter-based insertion for symmetrical and unbalanced systems.

Hence, the overall reactive power is

$$Q = Q_{12} + Q_{32} = \sqrt{3}\left[-P_{1(30)} + P_{3(10)}\right] \tag{22.60}$$

22.2.3.8 Three-Wattmeter Insertion

A three-wire, three-phase system can be measured by three wattmeters connected as in Figure 22.24. The artificial neutral point position does not affect the measurement; it is usually imposed by the impedance of the voltmeter leads of the wattmeters.

22.2.3.9 Medium-Voltage, Three-Wattmeter Insertion

Analogously to the single-phase case, for medium-voltage circuits, the three-wattmeter insertion is modified as in Figure 22.25.

22.2.3.10 Method Selection Guide

For three-wire systems, the flow chart of Figure 22.26 leads to selecting the most suitable method according to system characteristics.

22.2.4 High-Frequency Power Measurements

Meters used for power measurements at radio or microwave frequencies are generally classified as *absorption type* (containing inside their own load, generally 50 Ω. for radio frequency [RF] work) and *transmitted* or *through-line type* (where the load is remote from the meter). Apart from the type, power meters are mainly based on thermistors, thermocouples, diodes, or radiation sensors. Therefore, to work properly,

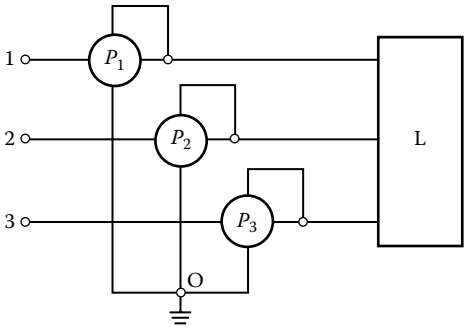


FIGURE 22.24 Three-wattmeter-based insertion for three-wire, three-phase systems.

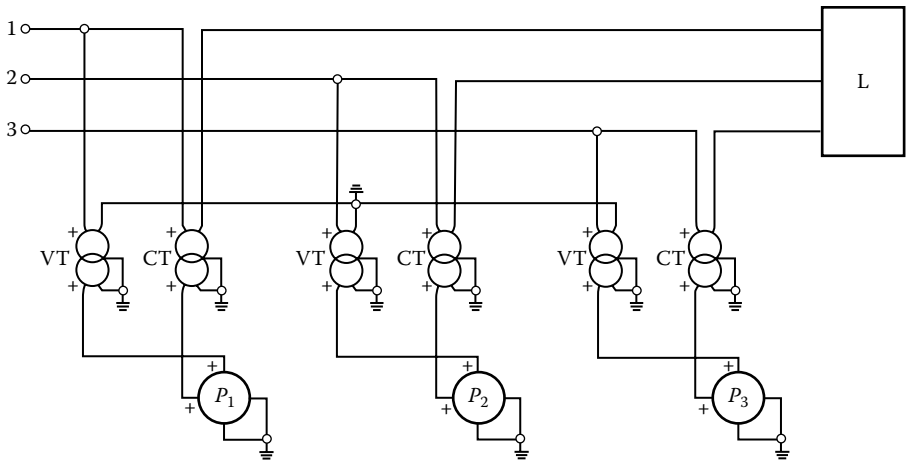


FIGURE 22.25 Medium-voltage, three-wattmeter insertion.

the sensor should sense all the RF power (P_{LOAD}) incoming into the sensor itself. Nevertheless, line-to-sensor impedance mismatches cause partial reflections of the incoming power ($P_{INCIDENT}$) so that a meter connected to a sensor does not account for the total amount of reflected power ($P_{REFLECTED}$). The relationship existing among power dissipated on the load, power incident, and power reflected is obviously

$$P_{LOAD} = P_{INCIDENT} - P_{REFLECTED} \tag{22.61}$$

Directional couplers are instruments generally used for separating incident and reflected signals so that power meters can measure each of them separately. In Figure 22.27, the longitudinal section of a directional coupler for waveguides is sketched. It is made up by two waveguides properly coupled through two holes. The upper guide is the *primary waveguide* and connects the power source and load; the lower guide is the *secondary waveguide* and is connected to the power meter. To explain the workings of directional couplers, incident and reflected waves have been sketched separately in Figure 22.27a and b. In particular, section (a) depicts a directional coupler working as incident wave separator, whereas section (b) shows the separation of the reflected wave. The correct working is based on the assumption that the distance between the holes matches exactly *one quarter of the wave length* (λ). In fact, in the secondary waveguide, each hole will give rise to two waves going in opposite directions (one outside and the other inside the waveguide); consequently, in front of each hole, two waves are summed with their own phases. The assumption made on the distance between the holes guarantees that, in front of one hole, (1) the two waves propagating outside the waveguide will be in phase, causing an enforcing effect in that direction, (2) while, in front of the other hole, the two waves (always propagating outside) will be in opposition, causing a canceling effect in that direction. The enforcing and canceling effects for incident and reflected waves are opposite. In particular, according to the directions chosen in Figure 22.27, incident power propagates on the right side and is canceled on the left side (Figure 22.27a), while reflected power propagates on the left side and is canceled on the right side (Figure 22.27b). Therefore, directional couplers allow separate measurement of incident and reflected power by means of power meters applied, respectively, on the right and on the left sides of the secondary waveguide.

In any case, the secondary waveguide must be correctly matched from the impedance point of view at both sides (by adaptive loads and/or a proper choice of the power meter internal resistance) in order to avoid unwanted reflections inside the secondary waveguide.

Directional couplers are also used to determine the *reflection coefficient* ρ of the sensor, which takes into account mismatch losses and is defined as

$$P_{REFLECTED} = \rho^2 \times P_{INCIDENT} \tag{22.62}$$

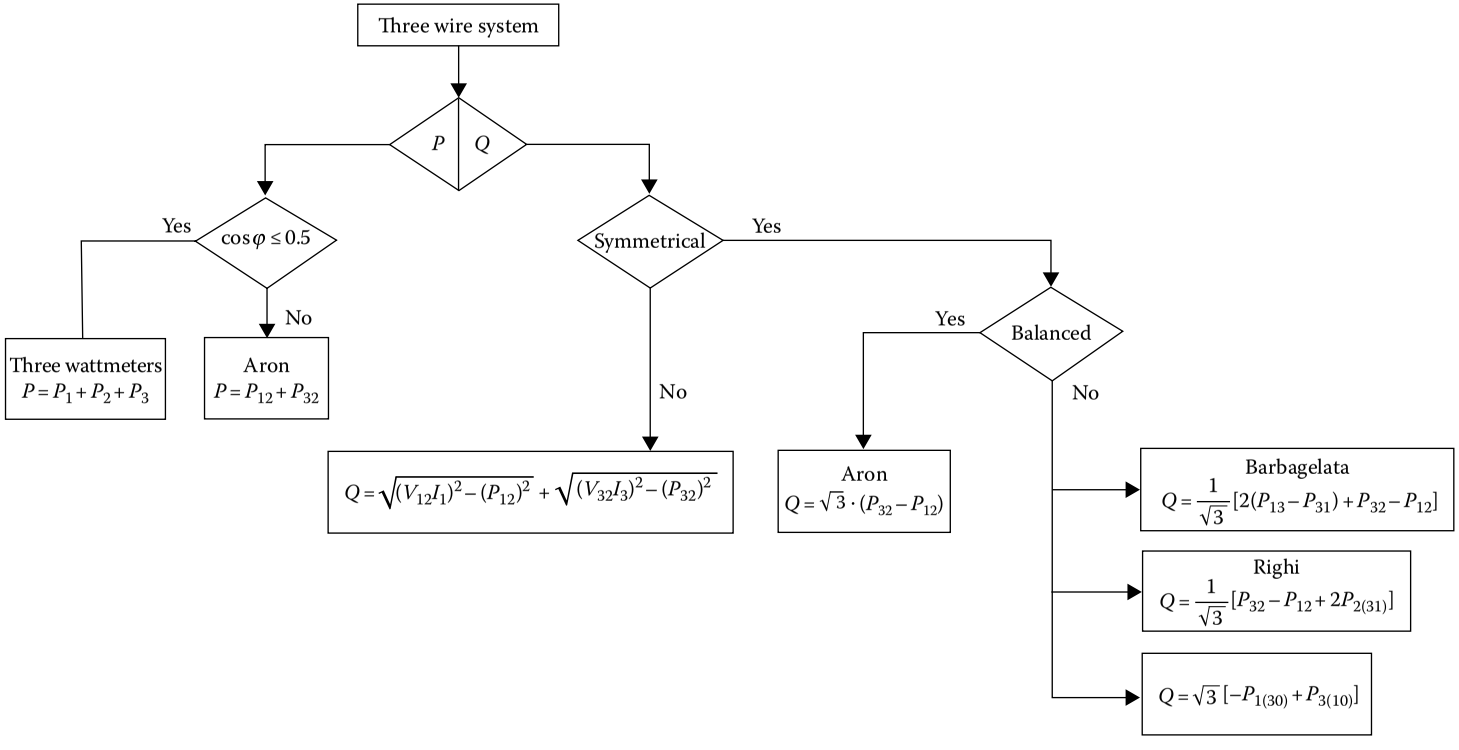


FIGURE 22.26 Method selection guide for power measurements on three-wire systems.

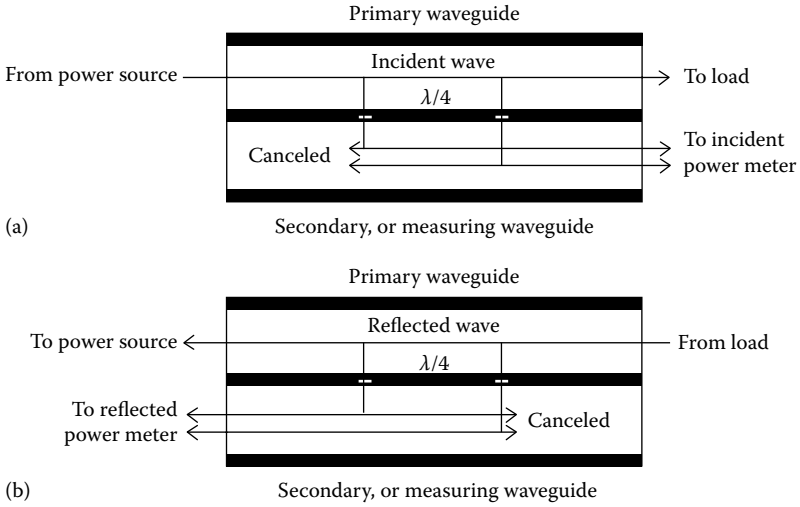


FIGURE 22.27 Directional couplers for separating incident (a) from reflected (b) power.

In order to take into account also the absorptive losses due to dissipation in the conducting walls of the sensor, leakage into instrumentation, power radiated into space, etc., besides the reflection coefficient, the *effective efficiency* η_C of the sensor should also be considered. Generally, the reflection coefficient and effective efficiency are included into the *calibration factor K*, defined as

$$K = \eta_C(1-\rho^2)\times 100 \tag{22.63}$$

For example, a calibration factor of 90% means that the meter will read 10% below the incident power. Generally, calibration factors are specified by sensor manufacturers at different values of frequency.

22.2.4.1 Thermal Methods

In this section, the main methods based on power dissipation will be examined, namely, (1) thermistor based, (2) thermocouple based, and (3) calorimetric.

22.2.4.1.1 Thermistor-Based Power Meters

A thermistor is a resistor made up of a compound of highly temperature-sensitive metallic oxides [14]. If it is used as a sensor in a power meter, its resistance becomes a function of the temperature rise produced by the applied power. In Figure 22.28, typical power-resistance characteristics are reported for several values of the operating temperature.

The working principle of the thermistor power meter is illustrated in Figure 22.29 [15]: two thermistors (R_{T1} and R_{T2}) are connected (1) in parallel, for measurand signals appearing at the RF input (P_{RF}), and (2) in series, for the following measuring circuit (e.g., a bridge). The capacitance C_1 prevents the power dc component from flowing to the thermistors; the C_2 stops the RF power toward the bridge.

A bridge with a thermistor or a barretter in one arm is called a *bolometer*. Bolometer-based measurements can be performed with (1) a manual bolometer with variation of the bias current, (2) a manual bolometer with substitution method, or (3) a self-balancing bolometer.

The *manual bolometer with a variation of the bias current* is illustrated in Figure 22.30. Its working principle consists of two steps. In the first, no RF power is applied to the sensor; the equilibrium is obtained by varying the dc power supply E until the sensor resistance R_B, related to the dc power flowing in it, is equal to R. In this condition, let the current I flowing into the sensor be equal to I_1.

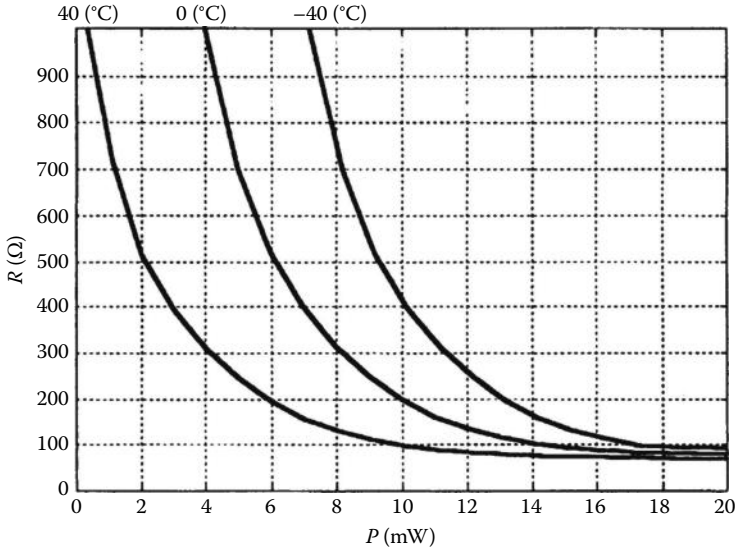


FIGURE 22.28 Typical power-resistance characteristics of commercial thermistors.

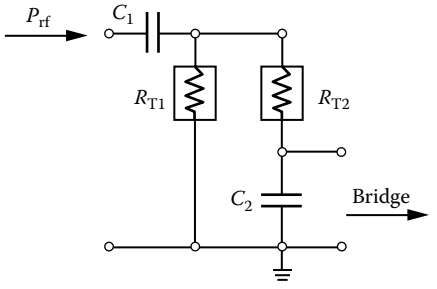


FIGURE 22.29 Working principle of the thermistor-based power meter.

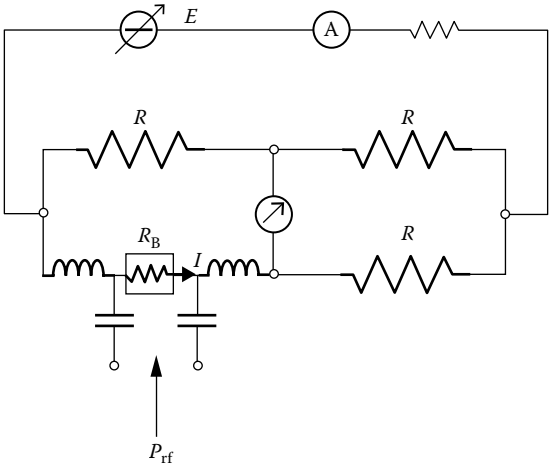


FIGURE 22.30 The manual bolometer.

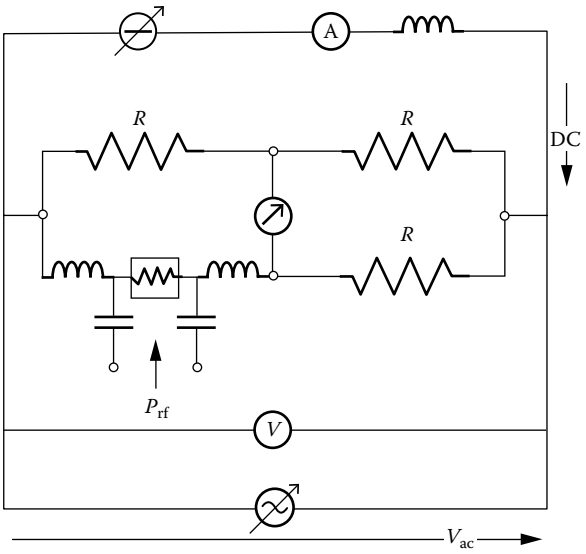


FIGURE 22.31 Manual bolometer with substitution method.

In the second step, an RF power P_{RF} is fed to the sensor; the power increase must be compensated by a dc power decrease, which is performed by lowering the bridge dc supply voltage E; in this case, let I be equal to I_2.

Since the power dissipated in the sensor has been maintained constant in both steps, the power P_{RF} can be evaluated as

$$P_{RF} = \frac{R}{4}\left(I_1^2 - I_2^2\right) \tag{22.64}$$

The *manual bolometer with substitution method* (Figure 22.31) consists of two sequential steps; in the first, both RF power (P_{RF}) and dc power (P_{dc}) are present, and the power (P_d) necessary to lead the bridge to the equilibrium is

$$P_d = P_{dc} + P_{RF} \tag{22.65}$$

During the second step, P_{RF} is set to zero, and an alternative voltage V_{ac} is introduced in parallel to the dc power supply. In this case, the power P_d necessary to balance the bridge

$$P_d = P_{dc} + P_{ac} \tag{22.66}$$

is obtained by varying v_{ac}.

Since P_d is the same in both cases, the power supplied by the alternative generator is equal to P_{RF}:

$$P_{RF} = P_{ac} = \frac{V_{ac}^2}{4R} \tag{22.67}$$

Equation 22.66 implies that the RF power can be obtained by a voltage measurement. The *self-balancing bolometer* (Figure 22.32) automatically supplies a dc voltage V to balance the voltage variations due to

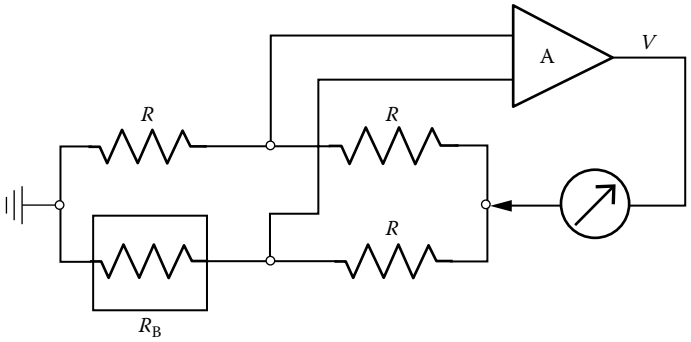


FIGURE 22.32 Self-balancing bolometer.

sensor resistance R_B changes for an incident power P_{RF}. At equilibrium, R_B is equal to R, and the RF power will then be

$$P_{RF} = \frac{V^2}{4R} \tag{22.68}$$

As mentioned earlier, the thermistor resistance depends on the surrounding temperature. This effect is compensated in an instrument based on two self-balancing bridges [15]. The RF power is input only to one of these, as shown in Figure 22.33.

The equilibrium voltages V_c and V_{RF} feed a chopping and summing circuit, whose output $V_c + V_{RF}$ goes to a voltage-to-time converter. This produces a pulse train V_1, whose width is proportional to $V_c + V_{RF}$. The chopping section also generates a signal with amplitude proportional to $V_c - V_{RF}$ and a frequency of a few kilohertz, which is further amplified. The signals V_1, and V_2 enter an electronic switch whose output is measured by a medium value meter M. This measure is proportional to the RF power because

$$P_{RF} = \frac{(V_c + V_{RF})(V_c - V_{RF})}{4R} \quad \Rightarrow \quad P_{RF} = \frac{V_c^2 - V_{RF}^2}{4R} \tag{22.69}$$

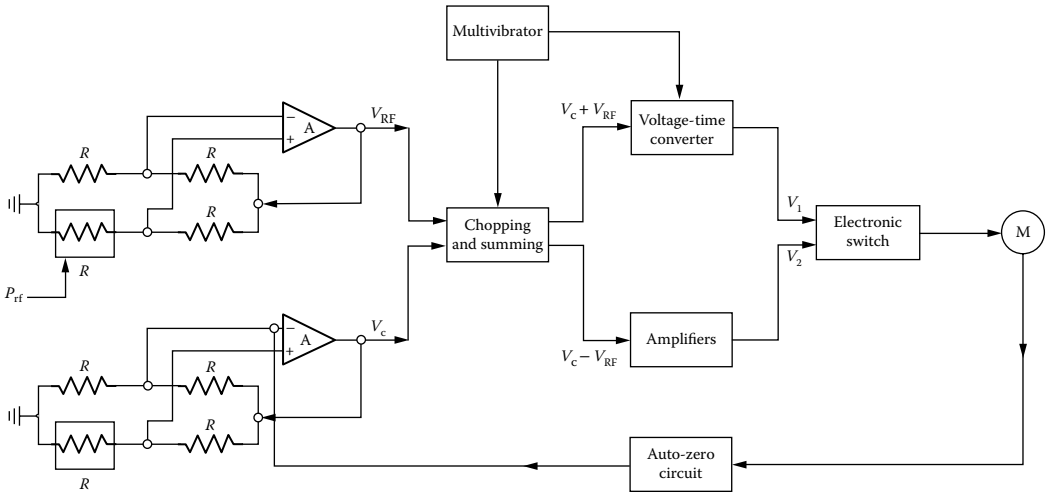


FIGURE 22.33 Power meter based on two self-balancing bridges.

Owing to the differential structure of the two bolometers, this device is capable of performing RF power measurements independent of the surrounding temperature. In addition, an offset calibration can be carried out when P_{RF} is null and V_c is equal to V_{RF}.

These instruments can range from 10 mW to 1 μW and utilize sensors with frequency bandwidths ranging from 10 kHz to 100 GHz.

22.2.4.1.2 Thermocouple-Based Power Meters

Thermocouples [14] can be also used as RF power meters up to frequencies greater than 40 GHz. In this case, the resistor is generally a thin-film type. The sensitivity of a thermocouple can be expressed as the ratio between the dc output amplitude and the input RF power. Typical values are 160 μV mW^{-1} for minimum power of about 1 μW.

The measure of voltages of about some tens of millivolts requires strong amplification, in that the amplifier does not have to introduce any offset. With this aim, a chopper microvoltmeter is utilized [16], as shown in Figure 22.34.

The thermocouple output voltage V_{dc} is chopped at a frequency of about 100 Hz; the resulting square wave is filtered of its mean value by the capacitor C and then input to an ac amplifier to further reduce offset problems. A detector, synchronized to the chopper, and a low-pass filter transform the amplified square voltage in a dc voltage finally measured by a voltmeter.

22.2.4.1.3 Calorimetric Method

For high frequencies, a substitution technique based on a calorimetric method is utilized (Figure 22.35) [17]. First, the unknown RF power P_{RF} is sent to the measuring device t, which measures the equilibrium temperature. Then, once the calorimetric fluid has been cooled to its initial temperature, a dc power P_{dc} is applied to the device and regulated until the same temperature increase occurs in the same time interval. In this way, thermal energy equivalence is established between the known P_{dc} and the measur- and P_{RF}.

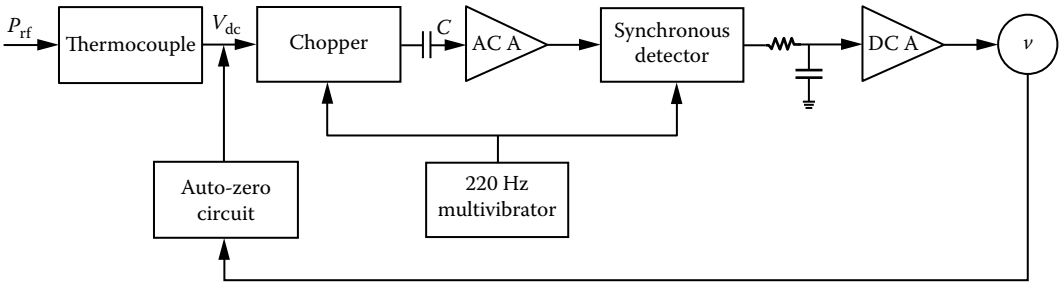


FIGURE 22.34 Power meter with thermocouple-based sensor.

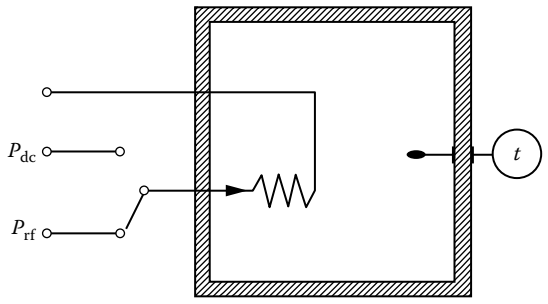


FIGURE 22.35 Calorimetric method based on a substitution technique.

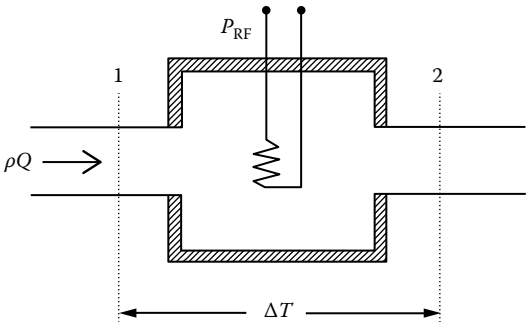


FIGURE 22.36 Calorimetric method based on a comparison technique.

A comparison version of the calorimetric method is also used for lower-frequency power measurements (Figure 22.36). The temperature difference ΔT of a cooling fluid is determined between the input (1) and the output (2) sections of a cooling element where the dissipated power P is measured. In this case, the power loss will correspond to P:

$$P = C_p \rho Q \Delta T \tag{22.70}$$

where

C_p is the specific heat
ρ is the density
Q is the volume flow, respectively, of the refreshing fluid

22.2.5 Diode Sensor-Based Power Measurements

Very sensitive (up to 0.10 nW, –70 dB m), high-frequency (10 MHz–20 GHz) power measurements are carried out through a diode sensor by means of the circuit in Figure 22.37 [18]. In particular, according to a suitable selection of the components in this circuit, (1) true-average power measurements or (2) peak power measurements can be performed.

The basic concept underlying *true-average power measurements* exploits the nonlinear squared region of the characteristic of a low-barrier Schottky diode (nondashed area in Figure 22.38). In this region, the current flowing through the diode is proportional to the square of the applied voltage; thus, the diode acts as a squared-characteristic sensor.

In the circuit of diode sensor-based wattmeters shown in Figure 22.37, the measurand v_x, terminated on the matching resistor R_m, is applied to the diode sensor D_s working in its squared region in order to produce a corresponding output current i_c in the bypass capacitor C_b. If C_b, has been suitably selected, the voltage V_c between its terminals, measured by the voltmeter amplifier V_a, is proportional to the average of i_c, that is, to the average of the squares of instantaneous values of the input signal v_x and, hence, to the true-average power.

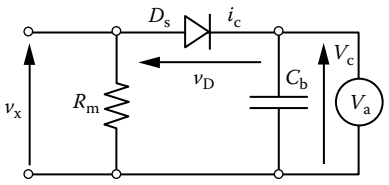


FIGURE 22.37 Circuit of the diode sensor-based power measurement.

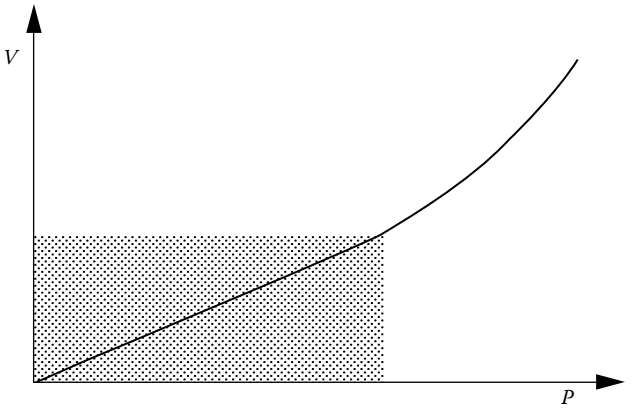


FIGURE 22.38 Characteristic of a low-barrier Schottky diode.

In the true-average power measurement of nonsinusoidal waveforms having the biggest components at low frequency, such as RF amplitude modulated (AM), the value of C_b, must also satisfy another condition. The voltage v_d on the diode must be capable of holding the diode switched on into conduction even for the smallest values of the signal. Otherwise, in the valleys of the modulation cycle, the high-frequency modulating source is disconnected by the back-biased diode, and the measurement is therefore misleading.

On the other hand, for the same signal but for a different selection of the C_b value, the circuit can act as a peak detector for *peak power measurements*. As a matter of fact, the voltage v_c on the bypass capacitor C_b during the peak of the modulation cycle is so large that in the valleys, the high-frequency peaks are not capable of switching on the diode into conduction; thus, these peaks do not contribute to the measured power level.

If higher power levels have to be measured (10–100 mW), the sensing diode is led to work out of the squared region into its linear region (dashed area in Figure 22.38). In this case, the advantage of true-average power measurements for distorted waveforms is lost, and for peak power measurements, since the diode input–output characteristic is nonlinear and the output is squared, spectral components different from the fundamental introduce significant measuring errors.

22.2.6 Radiation Sensor-Based Power Measurements

Very high-frequency power measurements are usually carried out by measuring a radiant flux of an electromagnetic radiation through a suitable sensor. In particular, semiconductor-based radiation microsensors have gained wider and wider diffusion [19], in that size reduction involves several well-known advantages such as greater portability, fewer materials, and a wider range of applications. One of the most familiar applications of radiation sensor-based power measurements is the detection of object displacement. Furthermore, they are also commonly used for low-frequency power noninvasive measurements.

Radiation sensors can be classified according to the measurand class to which they are sensitive: nuclear particles or electromagnetic radiations. In any case, particular sensors capable of detecting both nuclear particles and electromagnetic radiations, such as gamma and X-rays, exist and are referred to as nucleonic detectors.

In Table 22.1, the different types of radiation sensors utilized according to the decrease of the measurand wavelength from microwaves up to nuclear (X, gamma, and cosmic) rays are indicated.

In particular, *microwave* power radiation sensors are mainly used as noncontacting detectors relying on ranging techniques using microwaves [20]. Shorter and longer (radar) wavelength microwave devices are employed to detect metric and over-kilometer displacements, respectively.

Beyond applications analogous to microwave, power radiation *infrared* sensors also find use as contact detectors. In particular, there are two types of infrared detectors: thermal and quantum. The thermal type includes contacting temperature sensors such as thermocouples and thermopiles, as well as noncontacting pyroelectric detectors. On the other hand, the quantum type,

TABLE 22.1 Operating Field of Main Radiation Power Sensors

Operating Field (Wavelength)	Microwave (1, 10^{-3} m)	Infrared (10^{-3}, 10^{-6} m)	Visible and Ultraviolet (10^{-6}, 10^{-9} m)	Nuclear Rays (10^{-8}, 10^{-15} m)
Sensors	Noncontacting displacement sensors	Pyroelectric, photoconductive, photovoltaic	Photoconductive, photovoltaic	Scintillation counters, plastic films, solid-state, thermolum

although characterized by a strong wavelength dependence, has a faster response and includes photoconductive (spectral range: 1–12 μm) and photovoltaic (0.5–5.5 μm) devices.

The main power radiation *visible and ultraviolet* sensors are photoconductive cells, photodiodes, and phototransistors. Photodiodes are widely used to detect the presence, the intensity, and the wavelength of light or ultraviolet radiations. Compared to photoconductive cells, they are more sensitive, smaller, more stable, and linear and have lower response times. On the other hand, phototransistors are more sensitive to light.

At very low light levels, rather than silicon-based microsensors, *nuclear radiation* power microsensors are needed. In this case, the most widespread devices are scintillation counters, solid-state detectors, plastic films, and thermoluminescent devices. The scintillation counter consists of an active material that converts the incident radiation to pulses of light and a light–electric pulse converter. The active material can be a crystal, plastic fluorine, or a liquid. The scintillator size varies greatly according to the radiation energy, from thin solid films to large tanks of liquid to detect cosmic rays. A thin (5 μm) plastic polycarbonate film or a thermoluminescent material (e.g., LiF) can measure the radiation power falling on a surface. The film is mechanically damaged by the propagation of highly α-ionizing particles. Consequent etching of the film reveals tracks that can be observed and counted.

22.3 Pulse Power Measurements

Pulse waveforms are becoming more and more diffused in several fields such as telecommunications and power source applications. The *pulse power* P_p is defined as the average power P_m in the pulse width:

$$P_p = \frac{P_m}{\tau_d} \tag{22.71}$$

where τ_d is the duty cycle of the pulse waveform (i.e., the pulse width divided by the waveform period). If the pulse width cannot be accurately defined (e.g., nonrectangular pulses in the presence of noise), the pulse power P_p becomes unmeaningful. In this case, the *peak envelope power* is introduced as the maximum of the instantaneous power detected on a time interval, including several periods of the pulse waveform (but negligible with respect to the modulation period, in the case of PWM waveforms).

Several techniques are used to measure pulse power [21]. In particular, they can be classified according to the *pulse frequency* and the necessity for *constraining real-time applications* (i.e., measuring times, including a few of the modulation periods). For real-time, low-frequency applications (up to 100 kHz), the algorithms mentioned in the earlier sections on wattmeters based on digital multipliers can be applied [9].

If constraining limits of real-time do not have to be satisfied, either digital or analog techniques can be utilized. As far as the digital techniques are concerned, for high-frequency applications, if the measurand pulse waveform is stationary over several modulation periods, digital wattmeters based on equivalent sampling can be applied, with accuracies increasing according to measuring times. As far as the analog techniques are concerned, three traditional methods are still valid: (1) average power per duty cycle, (2) integration–differentiation, and (3) dc/pulse power comparison.

A block diagram of an instrument measuring *average power per duty cycle* is illustrated in Figure 22.39. At first, the mean power of the measurand pulse signal, terminated on a suitable load, is measured by

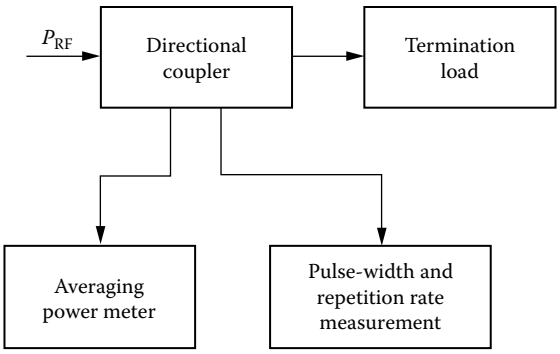


FIGURE 22.39 Block diagram of an instrument measuring average power per duty cycle.

means of an average power meter; then, the pulse width and the pulse waveform period are measured by a digital counter. Finally, the power is obtained by means of Equation 22.71.

The *integration–differentiation* technique is based on a barretter sensor capable of integrating the measurand and on a conditioning and differentiating circuit to obtain a voltage signal proportional to the measurand power. The signal is input to the barretter sensor, having a thermal constant such that the barretter resistance will correspond to the integral of the input. The barretter is mounted as an arm of a conditioning Wheatstone bridge; in this way, the barretter resistance variations are transformed into voltage variations, and an integrated voltage signal is obtained as an output of the bridge detecting arm. This signal, suitably integrated to reach a voltage signal proportional to the output, is detected by a peak voltmeter calibrated in power. Analogously to the selection of the time constant of an *RC* integrating circuit, attention must be paid to the thermal constant selection of the barretter in order to attain the desired accuracy in the integration. With respect to the measurand pulse period, a very long thermal constant will give rise to insufficient sensitivity. On the other hand, a very short constant approaching the pulse duration will give rise to insufficient accuracy.

The *dc/pulse power comparison* technique is based on the concept of first revealing the peak envelope power through a diode sensor and then comparing the peak to a known dc source with a dual-trace scope. A block diagram of an instrument based on this concept is illustrated in Figure 22.40. The peak

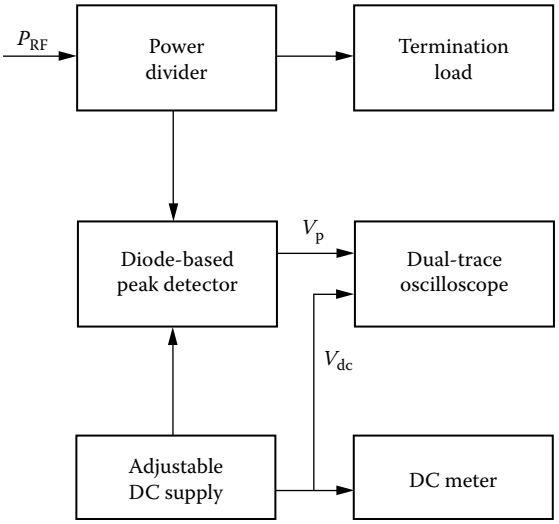


FIGURE 22.40 Block diagram of an instrument based on dc/pulse power comparison technique.

of the measurand pulse signal terminated on a suitable load is sensed by a peak detector by obtaining a proportional signal V_p. This signal is input to a channel of a dual-trace oscilloscope, and the envelope peak is displayed. An adjustable dc source is input to the other channel of the scope to obtain a signal V_{dc} to be compared to the envelope peak signal. When the two signals are made equal, a dc voltage meter directly calibrated in peak power measures the output power.

References

1. G. Zingales, *Methods and Instruments for Electrical Measurements* (in Italian), Torino, Italy: UTET, 1980, Power measurements on single-phase ac circuits (Chapter VI, 6.2).
2. G. Korànyi, Measurement of power and energy, in L. Schnell (ed.), *Technology of Electrical Measurements*, Chichester, U.K.: John Wiley & Sons, 1993.
3. J. Milmann and C.C. Halkias, *Integrated Electronics: Analog and Digital Circuits and Systems*. New York: McGraw-Hill, 1972.
4. F.F. Mazda, *Electronic Instruments and Measurement Techniques*, Cambridge, U.K.: Cambridge University Press, 1987, AC analogue instruments (Chapter VI, 6.3).
5. P.S. Filipski, A TDM wattmeter with 0.5 MHz carrier frequency, *IEEE Transactions on Instrumentation and Measurement*, IM39, 15–18, 1990.
6. J.R. Carstens, *Electrical Sensors and Transducers*. Englewood Cliffs, NJ: Prentice Hall, 1992.
7. L. Zu-Liang, An error estimate for quasi-integer-period sampling and an approach for improving its accuracy, *IEEE Transactions* on *Instrumentation and Measurement*, IM-23, 337–341, 1984.
8. V. Haasz, The error analysis of digital measurements of electrical power, *Measurement*, 6, 179–183, 1986.
9. P. Arpaia, F. Avallone, A. Baccigalupi, and C. De Capua, Real-time algorithms for active power measurements on PWM-based electric drives, *IEEE Transactions on Instrumentation and Measurement*, IM-45, 462–466, 1996.
10. X. Dai and R. Gretsch, Quasi-synchronous sampling algorithm and its applications, *IEEE Transactions on Instrumentation and Measurement*, IM-43, 204–209, 1994.
11. M. Bellanger, *Digital Processing of Signals: Theory and Practice*. Chichester, U.K.: John Wiley & Sons, 1984.
12. M. Bertocco, C. Offelli, and D. Petri, Numerical algorithms for power measurements, *European Transactions on Electrical Power*, ETEP 3, 91–101, 1993.
13. G. Zingales, *Methods and Instruments for Electrical Measurements*, (In Italian), Torino, Italy: UTET, 1980, Measurements on steady-state circuits (Chapter VI).
14. H.N. Norton, *Handbook of Transducers*, Englewood Cliffs, NJ: Prentice Hall, 1989, Thermometers (Chapter 19).
15. Anonymous, *Fundamental of RF and Microwave Power Measurements*, Application Note 64-1, Hewlett-Packard, 1978, Thermistor mounts and instrumentation (Chapter II).
16. R.E. Pratt, Power measurements (15.1–15.16), in C.F. Coombs, (ed.), in *Handbook of Electronic Instruments*, New York: McGraw-Hill, 1995.
17. F.F. Mazda, *Electronic Instruments and Measurement Techniques*, Cambridge, U.K.: Cambridge University Press, 1987, High-frequency power measurements (Chapter VIII, 8.4).
18. Anonymous, *Fundamental of RF and Microwave Power Measurements*, Application Note 64-1, Hewlett-Packard, 1978, Diode detector power sensors and instrumentation (Chapter IV).
19. J.W. Gardner, *Microsensors: Principles and Applications*. Chichester, U.K.: John Wiley & Sons, 1994.
20. H.N. Norton, *Handbook of Transducers*, Englewood Cliffs, NJ: Prentice Hall, 1989, Radiation pyrometers (Chapter 20).
21. F.F. Mazda, *Electronic Instruments and Measurement Techniques*, Cambridge, U.K.: Cambridge University Press, 1987, Pulse power measurement (Chapter VIII, 8.5).

Further Readings

Anonymous, *Fundamental of RF and Microwave Power Measurements*, Application Note 64-1, Hewlett-Packard, 1978; though not very recent, is a valid and comprehensive reference for main principles of high-frequency power measurements.

Arpaia P., G. Betta, A. Langella, and M. Vanacore, An expert system for the optimum design of measurement systems, *IEE Proceedings (Part A)*, 142, 330–336, 1995; reports about an Artificial Intelligence tool for the automatic design of power measuring systems.

Avallone F., C. De Capua, and C. Landi, Measurement station performance optimization for testing on high efficiency variable speed drives, *Proceedings of IEEE IMTC/96* (Brussels, Belgium), pp. 1098–1103, 1996; proposes an analytical model of uncertainty arising from power measurement systems working under highly distorted conditions.

Avallone F., C. De Capua, and C. Landi, Metrological performance improvement for power measurements on variable speed drives, *Measurement*, 21, 17–24, 1997; shows how compute uncertainty of measuring chain components for power measurements under highly distorted conditions.

Avallone F., C. De Capua, and C. Landi, Measurand reconstruction techniques applied to power measurements on high efficiency variable speed drives, *Proceedings of XIV IMEKO World Congress* (Tampere, Finland), 1997; proposes a technique to improve accuracy of power measurements under highly distorted conditions.

Avallone F., C. De Capua, and C. Landi, A digital technique based on real-time error compensation for high accuracy power measurement on variable speed drives, *Proceedings of IEEE IMTC/97* (Ottawa, Ontario, Canada), 1997; reports about a real-time technique for error compensation of transducers working under highly distorted conditions.

Bucci G., P. D'Innocenzo, and C. Landi, A modular high-speed dsp-based data acquisition apparatus for on-line quality analysis on power systems under non-sinusoidal conditions, *Proceedings of 9th IMEKO TC4 International Symposium* (Budapest, Hungary), pp. 286–289, 1996; shows the strategy of measurement system design for power measurements under non-sinusoidal conditions.

Clarke J.J. and J.R. Stockton, Principles and theory of wattmeters operating on the basis of regularly spaced sample pairs, *Journal of Physics E: Scientific Instruments*, 15, 645–652, 1982; gives basics of synchronous sampling for digital wattmeters.

Clarke K.K. and D.T. Hess, A 1000 A/20 kV/25 kHz-500 kHz volt-ampere-wattmeter for loads with power factors from 0.001 to 1.00, *IEEE Transactions* on *Instrumentation and Measurement*, IM-45, 142–145, 1996; provides information on the implementation of a instrument to perform an accurate measurement of currents (1–1000 A), voltages (100 V–20 kV), and powers (100 W–20 MW) over the frequency range from 25 kHz to 500 kHz.

Garverick S.L., K. Fujino, D.T. McGrath, and R.D. Baertsch, A programmable mixed-signal ASIC for power metering, *IEEE Journal of Solid State Circuits*, 26, 2008–2015, 1991; reports about a programmable mixed analog-digital integrated circuit based on six first-order sigma-delta ADCs, a bit serial DSP, and a byte-wide static RAM for power metering.

Harris F.K., The measurement of power (Chapter XI), in *Electrical Measurements*, Huntington, NY: R.E. Krieger Publishing, 1975; a clear reference for line-frequency power measurements.

In any case, *IEEE Transactions on Instrumentation and Measurement* and *Measurement* journals provide current research on power measurements.

Kolanko J.K., Accurate measurement of power, energy, and true RMS voltage using synchronous counting, *IEEE Transactions on Instrumentation and Measurement*, IM-42, 752–754, 1993; provides information on the implementation of a synchronous dual-slope wattmeter.

Montano J.C., A. Lopez, M. Castilla, and J. Gutierrrez, DSP-based algorithm for electric power measurement, *IEE Proceedings Part A*, 140, 485–490, 1993; describes a Goertzel FFT-based algorithm to compute power under nonsinusoidal conditions.

Rathore T.S., Theorems on power, mean and RMS values of uniformly sampled periodic signals, *IEE Proceedings Point A*, 131, 598–600, 1984; provides fundamental theorems for effective synchronous sampling wattmeters.

Stenbakken G.N., A wideband sampling wattmeter, *IEEE Transactions on Power Apparatus and Systems*, PAS-103, 2919–2925, 1984; gives basics of asynchronous sampling-based wattmeters and criteria for computing uncertainty in time domain.

23

Power Factor Measurement

Michael Z. Lowenstein
Harmonics Limited

Electricity can be thought of as a means of delivering power from one place to another to do work. The laws and relationships for delivering power were originally developed for direct current (dc). Power delivered, expressed in *watts*, was calculated by multiplying the voltage and current as shown in Equation 23.1:

$$P_{dc} = EI \tag{23.1}$$

The situation becomes more complex when alternating current (ac) is used to deliver power. Figure 23.1 shows a sine wave representing either ac or voltage. Since the instantaneous value of the wave is continually changing, a numerical quantity is defined to represent an average property of this wave. This quantity, the root-mean-square, or *rms* value, calculated by squaring the instantaneous value, integrating it during one cycle, dividing by the period, and taking the square root of the result, is equal to the peak value of the ac wave divided by the square root of 2, or, for ac current, $I_{rms} = i_{peak}/\sqrt{2}$. In the physical world, a sine wave ac having an rms value of 1 A (A = ampere), passed through a resistive load, produces the same heating effect as 1 A of dc. Thus, one might expect delivered ac power to be easily calculated in watts using Equation 23.1 and inserting rms values for current and voltage. While this simple relationship holds true for the instantaneous voltage and current values as shown in Equation 23.1a in general,

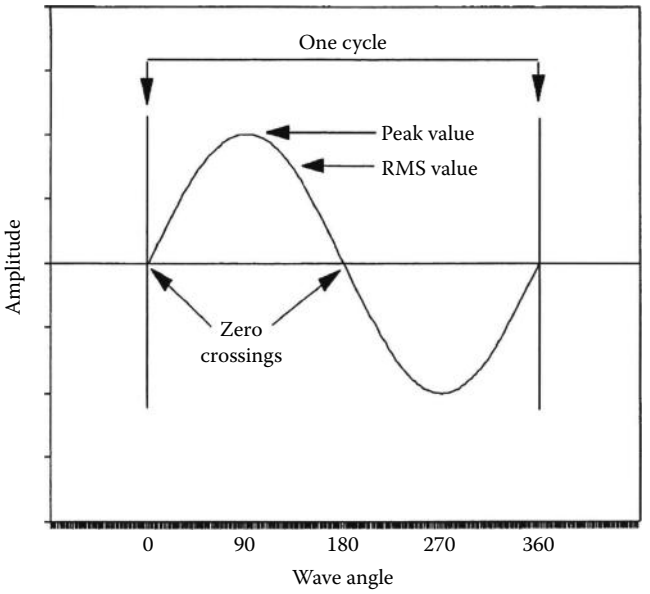


FIGURE 23.1 Sine wave characteristics. One cycle of a continuous wave is shown. The wave begins at a zero-crossing, reaches a positive peak, and continues through zero to a negative peak and back to zero. The wave repeats every 360°. The wave angle can also be expressed as a function of frequency f and time t. For a given frequency, the wave angle is related to the expression, $2\pi ft$.

it is not true for the rms quantities except for the special case when the ac current and voltage are restricted to perfect sine waves *and* the load is a pure resistance:

$$p_{\text{inst}} = ei \tag{23.1a}$$

In real-world situations where current and/or voltage waveforms are not perfectly sinusoidal and/or the loads are other than resistive, the relationships are no longer simple, and the power delivered, or *active power*, is usually less than the product of rms voltage and current, as shown in Equation 23.2:

$$p_{\text{ac}} \leq E_{\text{rms}} I_{\text{rms}} \tag{23.2}$$

The product of rms voltage and rms current does, however, define a quantity termed *apparent power*, U, as shown in Equation 23.3:

$$U = E_{\text{rms}} I_{\text{rms}} \tag{23.3}$$

A derived term, the *power factor*, F_{p}, used to express the relationship between delivered or active power, P, and apparent power, U, is defined by Equation 23.4:

$$F_{\text{p}} = \frac{P}{U} \tag{23.4}$$

From Equations 23.2 through 23.4, it is clear that the value of F_{p} must lie in a range between zero and one. This chapter focuses on (1) ac power relationships and the calculation of power factor, (2) the physical meaning of these relationships, and (3) measurement techniques and instrumentation for determining these relationships and calculating power factor.

23.1 Reasons for Interest in Power Factor

Power factor is a single number that relates the active power, *P*, to the apparent power, *U*. Electric components of a utility distribution system are designed on a kVA basis; that is, they are designed to operate at a given voltage and carry a rated current without undue temperature rise. Transformer and conductor sizes are chosen on this basis. While active power does useful work, reactive and harmonic powers do no useful work, absorb system capacity, and increase system losses, but reactive and harmonic powers are needed to provide magnetic fields or nonlinear currents. *The capacity of electric systems is limited by apparent power, not active power. Power factor expresses, with a single value, the extent to which an electric distribution system is efficiently and effectively utilized.* A low value for the power factor means that much of the system capacity is not available for useful work. From a utility viewpoint, this means reduced ability to deliver revenue-producing active power; from a user viewpoint, a low-power factor reduces the available active power or requires increased system size to deliver needed power.

23.2 AC Electric Loads

23.2.1 Linear Loads

Electric loads in ac power systems with sinusoidal voltages are categorized by the way they draw current from the system. Loads that draw sinusoidal currents, that is, the current waveshape is the same as the voltage waveshapes, are defined as *linear loads*. Historically, a high percentage of electric loads have been linear. Linear loads include (1) induction motors, (2) incandescent lighting, and (3) heaters and heating elements. Linear loads use ac electric power directly to accomplish their functions.

23.2.2 Nonlinear Loads

Electric loads that draw nonsinusoidal currents, that is, the current waveshape differs from the voltage waveshape, are defined as *nonlinear loads*. As energy savings and efficient use of electricity are emphasized, an increased percentage of nonlinear electric devices, both new and replacement, are being installed. Nonlinear loads include (1) adjustable-speed motor drives, (2) fluorescent and arc-discharge lighting, (3) computers and computerized controls, and (4) temperature-controlled furnaces and heating elements. Nonlinear loads, rather than using ac electric power directly, often convert ac power into dc before it is used to accomplish their functions. A common element in nonlinear loads is some kind of rectifier to accomplish this ac to dc conversion. Rectifiers do not draw sinusoidal currents.

23.3 AC Power Relationships

23.3.1 Sinusoidal Voltage and Current

Power calculations for sinusoidal ac electric systems require knowledge of the rms voltage, the rms current, and the phase relationships between the two. Figure 23.2 illustrates possible phase relationships between voltage and current. If the positive-going zero-crossing of the voltage is considered the reference point, then the nearest positive-going zero-crossing of the current can occur at a wave angle either less than or greater than this reference. If the current zero-crossing occurs before the reference, the current is said to *lead the* voltage. If the current zero-crossing occurs after the reference, the current is *lagging*. If the zero-crossing for the current coincides with the reference, the two waves are said to be *in phase*. The wave angle, θ, by which the current leads or lags the voltage is called the *phase angle*, in this case, 30°.

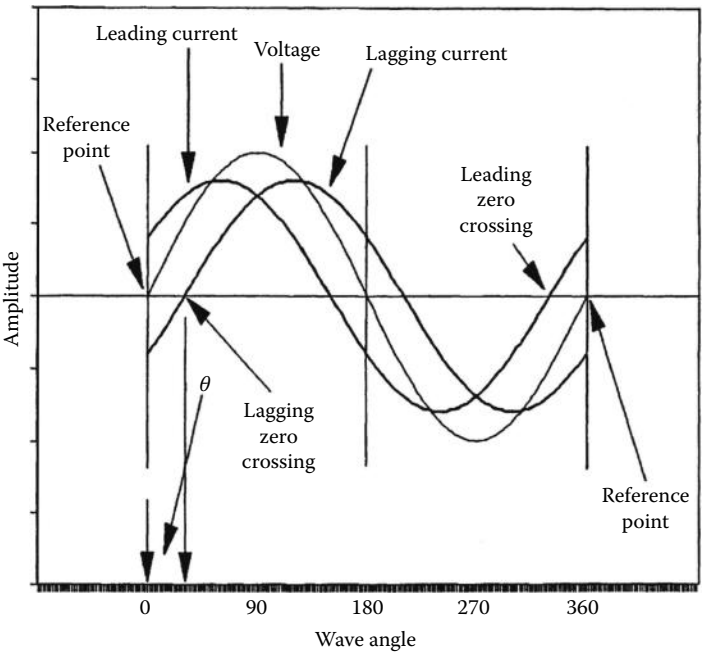


FIGURE 23.2 Sine wave phase angle. Two waves with the same zero-crossing are *in phase*. A sine wave that crosses zero before the reference wave is *leading*, and one that crosses zero after the reference wave is *lagging*. The *phase angle* θ illustrated is 30° lagging.

23.3.2 Single-Phase Circuits

23.3.2.1 Power Calculations

The power delivered to do work is easily calculated [1]. Given a sinusoidal voltage of rms magnitude E and sinusoidal current of rms magnitude I, displaced by angle θ, at time t,

$$\text{Instantaneous voltage} = e\sqrt{2}E\sin(2\pi ft)$$
$$\text{Instantaneous current} = i = \sqrt{2}I\sin(2\pi ft - \theta)(\text{note that } \theta \text{ can have a positive or negative value}) \tag{23.5a}$$

$$\text{Instantaneous power } = p = ei$$
$$p = 2E\sin(2\pi ft)\sin(2\pi ft - \theta)$$
$$p = EI\cos(\theta) - EI\cos(4\pi ft - \theta) \tag{23.5b}$$
$$\text{Average power over an integral number of cycles } p = EI\cos(\theta)$$

$$\text{Power factor } F_p = \frac{P}{U} = \frac{EI\cos(\theta)}{EI} = \cos(\theta) \tag{23.5c}$$

Equation 23.5 is the fundamental equation that defines power for systems in which the current and voltage are sinusoidal. The application of this equation is illustrated for three cases: (1) the current and voltage are in phase, (2) the current and voltage are out of phase by an angle less than 90°, and (3) the current and voltage are out of phase by exactly 90°.

23.3.2.2 AC Power Examples

Figure 23.3 shows voltage current and power when the voltage and current are in phase and the current displacement angle is zero (0). (An example would be a resistance space heater.) The power curve is obtained by multiplying together the instantaneous values of voltage and current as the wave angle is varied from 0° to 360°. Instantaneous power, the product of two sine waves, is also a sine wave. There are two zero-crossings per cycle, dividing the cycle into two regions. In region (1), both the voltage and current are positive, and the resultant product, the power, is positive. In region (2), both the voltage and current are negative, and the power is again positive. The average power in watts, given by Equation 23.5, $EI\cos(0°) = EI$, is the maximum power that can be delivered to do work. *When sinusoidal voltage and current are in phase, the delivered power in watts is the same as for dc and is the maximum that can be delivered*. The power factor $F_p = \cos(0°) = 1$ or *unity*.

Figure 23.4 shows voltage, current, and power when the current lags the voltage by 60°. (An example might be a lightly loaded induction motor.) The power sine wave again is obtained by multiplying together the instantaneous values of voltage and current. There are now four zero-crossings per cycle, dividing the cycle into four regions. In regions (2) and (4), voltage and current have the same sign, and the power is positive. In regions (1) and (3), voltage and current have opposite signs, resulting in a negative value for the power. The average power in watts, given by Equation 23.5, $EI\cos(60°) = EI(0.5)$, is less than the maximum that could be delivered for the particular values of voltage and current. *When voltage and current are out of phase, the delivered power in watts is always less than the maximum*. In this example, $F_p = \cos(60°) = 0.5$.

Figure 23.5 shows voltage, current, and power when the current lags the voltage by 90°. (This situation is not attainable in the real world.) The power sine wave again is obtained by multiplying together the instantaneous values of voltage and current. Again, four zero-crossings divide the cycle into four regions. In regions (2) and (4), the power is positive, while in regions (1) and (3), the power is negative.

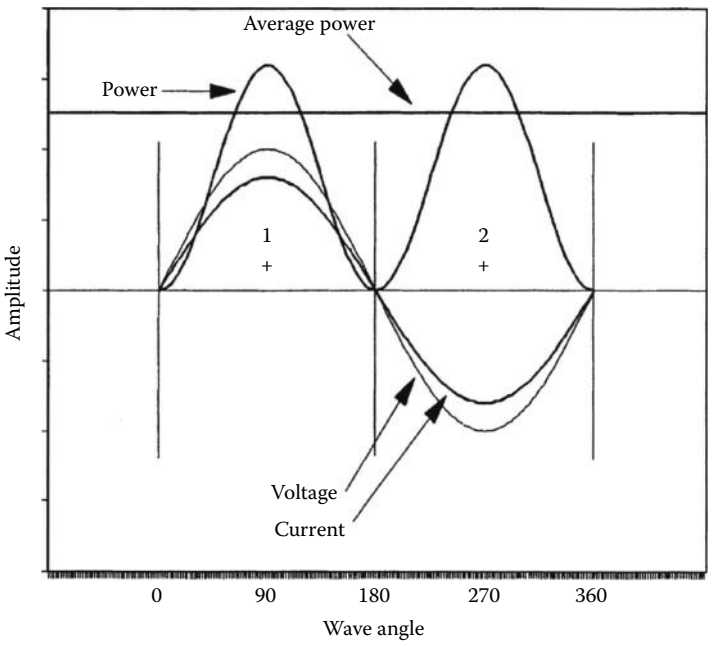


FIGURE 23.3 Voltage, current, and power for sine waves in phase. The vertical scales for voltage and current are equal. The scale for power is relative and is selected to permit the display of all three waves on a single graph. The voltage and current are in phase and both are sinusoidal. The power is positive everywhere, and average power delivered to do work is the maximum power.

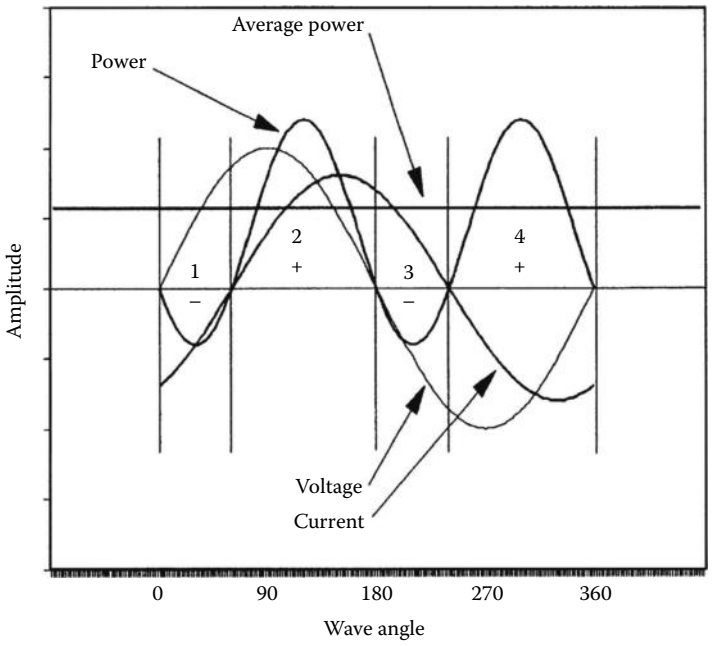


FIGURE 23.4 Voltage, current, and power for sine waves 60° out of phase. The vertical scales for voltage and current amplitudes are the same as those for Figure 23.3. Current lags voltage by 60° and both are sinusoidal. The power is positive in regions (2) and (4) and negative in regions (1) and (3). Average power delivered to do work is less than the maximum power.

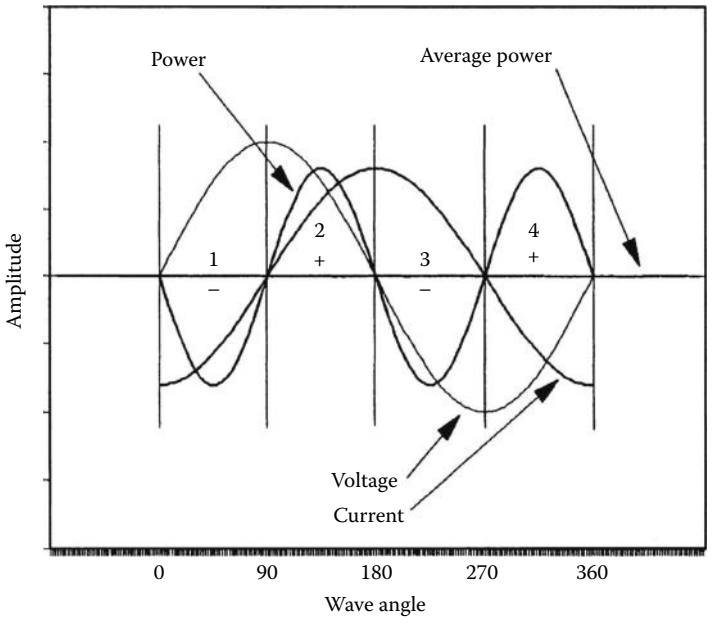


FIGURE 23.5 Voltage, current, and power for sine waves 90° out of phase. The vertical scales for voltage and current amplitudes are the same as those for Figure 23.3. Current lags voltage by 90° and both are sinusoidal. The power is positive in regions (2) and (4) and negative in regions (1) and (3) and is of equal absolute magnitude in all four regions. Average power delivered to do work is zero.

The average power in watts is given by Equation 23.5, $EI\cos(90°) = 0$. No matter what the values of voltage and current, *when voltage and current are exactly 90° out of phase, the delivered power in watts is always zero*. The power factor $F_p = \cos(90°) = 0$.

23.3.2.3 Power Factor

Resolving the current into orthogonal components on a phasor diagram illustrates how delivered power can vary from a maximum to zero, depending on the phase angle between the voltage and the current sine waves. Figure 23.6 shows the voltage vector along with the current resolved into orthogonal components [1]. The current I at an angle θ relative to the voltage can be resolved into two vectors: $I\cos(\theta)$ and $I\sin(\theta)$. The in-phase component $I\cos(\theta)$ multiplied by the voltage gives average power in watts. The current component that is 90° out of phase with the voltage, $I\sin(\theta)$, is not associated with delivered power and does not contribute to work. For want of a better name, this was often termed the *wattless component* of the current. Since this wattless current could be associated with magnetic fields, it was sometimes termed *magnetizing current* because, while doing no work, this current interacts through the inductive reactance of an ac motor winding to provide the magnetic field required for such a motor to operate.

Three types of power have been defined for systems in which both the voltage and current are sinusoidal. Throughout the years, a number of different names have been given to the three power types. The names in present usage will be emphasized.

Active power is given the symbol P and is defined by Equation 23.5:

$$P = EI\cos(\theta) \tag{23.5d}$$

Other names for active power include (1) real power and (2) delivered power. Active power is the power that does the work. Note that while all power quantities are volt–ampere products, only *active power is expressed in watts*.

Reactive power is given the symbol Q and is defined by the equation:

$$Q = EI\sin(\theta) \tag{23.6}$$

Other names for reactive power include (1) imaginary power, (2) wattless power, and (3) magnetizing power. Reactive power is expressed in $voltamperes_{reactive}$ or *vars*. If the load is predominantly inductive, current lags the voltage, and the reactive power is given a positive sign. If the load is predominantly capacitive, current leads the voltage, and the reactive power is given a negative sign.

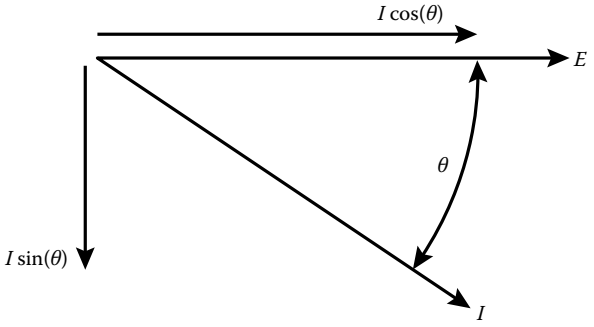


FIGURE 23.6 Phasor diagram for current and voltage. Voltage is the reference phasor. The current I has been resolved into orthogonal components $I\cos(\theta)$ and $I\sin(\theta)$. (From Bullock, D.F., Methods of measuring apparent power, GE Meter, *Missouri Valley Electrical Association Annual Engineering Conference*, Kansas City, MO, 1996.)

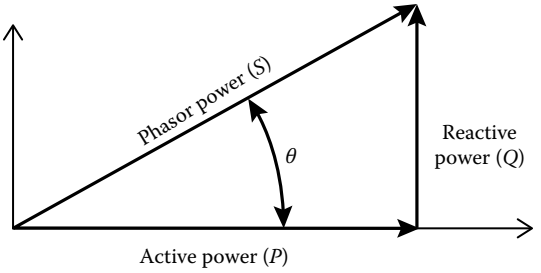


FIGURE 23.7 Power triangle showing the geometric relationships between active, reactive, and phasor power. Power factor is defined geometrically as cos θ.

Phasor power is given the symbol *S* and is defined by the equation:

$$S = \sqrt{p^2 + Q^2} \tag{23.7}$$

Phasor power was called apparent power for many years, and it will be seen in a later section that phasor power *S, for sinusoidal voltages and currents*, is identical to what is now called apparent power *U*. Phasor power is expressed in *voltamperes* or *VA*.

Figure 23.7 is a phasor diagram, often called a *power triangle*, which illustrates the relationships among the three types of power defined earlier. Reactive power is orthogonal to active power and is shown as positive for lagging current. It is clear that the definition of phasor power, Equation 23.7, is geometrically derived from active and reactive power.

Power factor is given the symbol F_p and *for sinusoidal quantities* is defined by the equation:

$$F_p = \frac{\text{Active power}}{\text{Phasor power}} = \frac{P}{S} = \frac{\text{Watts}}{\text{Voltamps}} = \cos\theta \tag{23.8}$$

Since the power factor can be expressed in reference to the displacement angle between voltage and current, power factor so defined should be termed *displacement power factor*, and the symbol is often written as $F_{p\ \text{displacement}}$. Values for displacement power factor range from one (unity) to zero as the current displacement angle varies from 0° (current and voltage in phase) to 90°. Since the cosine function is positive in both the first and fourth quadrants, the power factor is positive for both leading and lagging currents. To completely specify the voltage–current phase relationship, the words *leading* or *lagging* must be used in conjunction with power factor. Power factor can be expressed as a decimal fraction or as percent. For example, the power factor of the case shown in Figure 23.4 is expressed as either 0.5% lagging or 50% lagging.

23.3.3 Polyphase Circuits

23.3.3.1 Power Calculations

The power concepts developed for single-phase circuits with sinusoidal voltages and currents can be extended to polyphase circuits. Such circuits can be considered to be divided into a group of two-wire sets, with the neutral conductor (or a resistively derived neutral for the case of a delta-connected, three-wire circuit) paired with each other conductor. Equations 23.3 and 23.5 can be rewritten to define power terms equivalent to the single-phase terms. In these equations,

k represents a phase number, m is the total number of phases, and α and β are, respectively, the voltage and current phase angles with respect to a common reference frame:

$$p = \sum_{k=1}^{m} E_k I_k \cos(\alpha - \beta) \tag{23.9}$$

$$Q = \sum_{k=1}^{m} E_k I_k \sin(\alpha - \beta) \tag{23.10}$$

and, restating Equation 23.7,

$$S = \sqrt{P^2 + Q^2}$$

For example, a three-phase sinusoidal power distribution service, with phases a, b, and c

$$p = E_a I_a \cos(\alpha_a - \beta_a) + E_b I_b \cos(\alpha_b - \beta_b) + E_c I_c \cos(\alpha_c - \beta_c)$$

$$Q = E_a I_a \sin(\alpha_a - \beta_a) + E_b I_b \sin(\alpha_b - \beta_b) + E_c I_c \sin(\alpha_c - \beta_c)$$

$$S = \sqrt{P^2 + Q^2}$$

23.3.3.2 Power Factor

Power factor is defined by Equation 23.11. Note that it is no longer always true to say that power factor is equal to the cosine of the phase angle. In many three-phase balanced systems, the phase angles of all three phases are equal and the cosine relationship holds. In unbalanced systems, such as that represented by the phasor diagram Figure 23.8, each phase has a different phase angle, the phase voltages and currents are not equal, and the cosine relationship fails [3]:

$$F_p = \frac{\text{Total active power}}{\text{Phasor power}} = \frac{P_{\text{Equation 23.9}}}{S} = \frac{\text{Watts}}{\text{Voltamps}} \textit{often} \neq \cos\theta \tag{23.11}$$

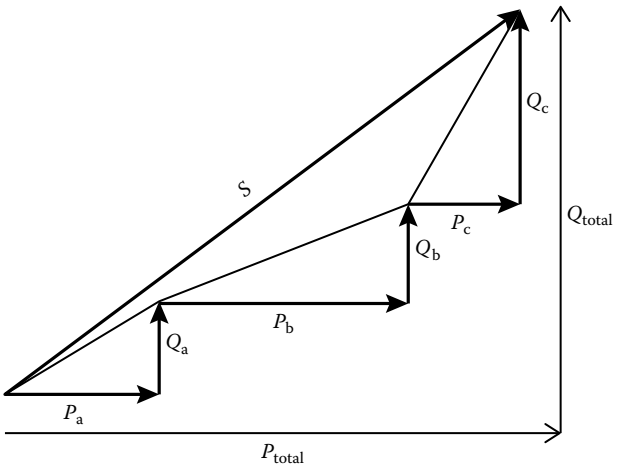


FIGURE 23.8 Phasor diagram for a sample polyphase sinusoidal service in which each phase has a different phase angle. Power factor cannot be defined as the cosine of the phase angle in this case. (From Bullock, D.F., *Phase Angle Relationships, Part 1: Theory*, GE Meter, Arkansas Electric Meter School, Fayetteville, AR, 1995.)

23.3.4 Nonsinusoidal Voltage and Current

23.3.4.1 Fourier Analysis

Figure 23.9 shows voltage, current, and power for a typical single-phase nonlinear load, a computer switch-mode power supply. Due to the nature of the bridge rectifier circuit in this power supply, current is drawn from the line in sharp spikes. The current peak is only slightly displaced from the voltage peak, and the power is everywhere positive. However, power is delivered to the load during only part of the cycle, and the average power is much lower than if the current had been sinusoidal. The current waveshape required by the load presents a problem to the ac power system, which is designed to deliver only sine wave current. The solution to this problem is based on mathematical concepts developed in 1807 for describing heat flow by Jean Baptiste Joseph Fourier, a French mathematician [4]. Fourier's theorem states that any periodic function, however complex, can be broken up into a series of simple sinusoids, the sum of which will be the original complex periodic variation. Applied to the present electric problem, Fourier's theorem can be stated: *any periodic nonsinusoidal electric waveform can be broken up into a series of sinusoidal waveforms, each a harmonic of the fundamental, the sum of which will be the original nonsinusoidal waveform*.

23.3.4.2 Harmonics

Harmonics are defined as continuous integral multiples of the fundamental waveform. Figure 23.10 shows a fundamental sine wave and two harmonic waves—the third and fifth harmonics. The harmonic numbers 3 and 5 express the number of complete cycles for each harmonic wave per cycle of the fundamental (or first harmonic). Each harmonic wave is defined by its harmonic number, its amplitude, and its phase relationship to the fundamental. Note that the fundamental frequency can have any value without changing the harmonic relationships, as shown in Table 23.1.

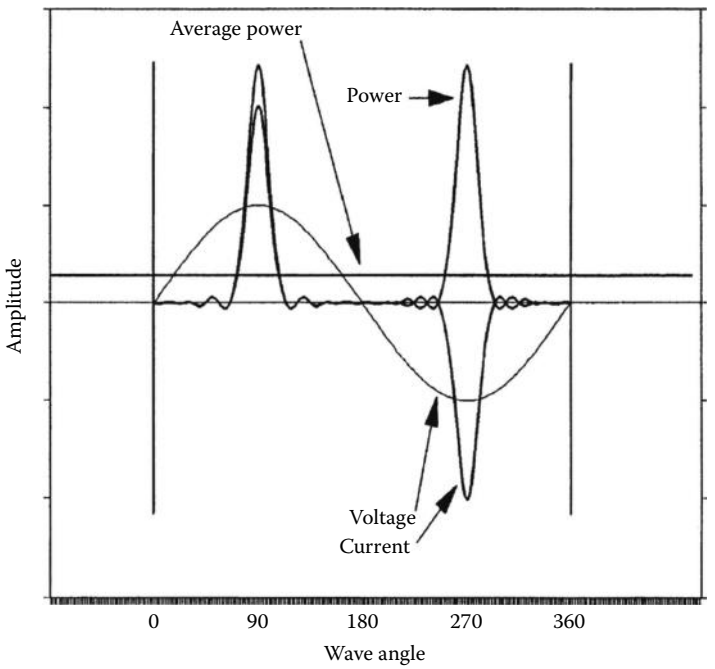


FIGURE 23.9 Voltage, current, and power for a single-phase, switch-mode computer power supply, a typical nonlinear load. The current is no longer sinusoidal.

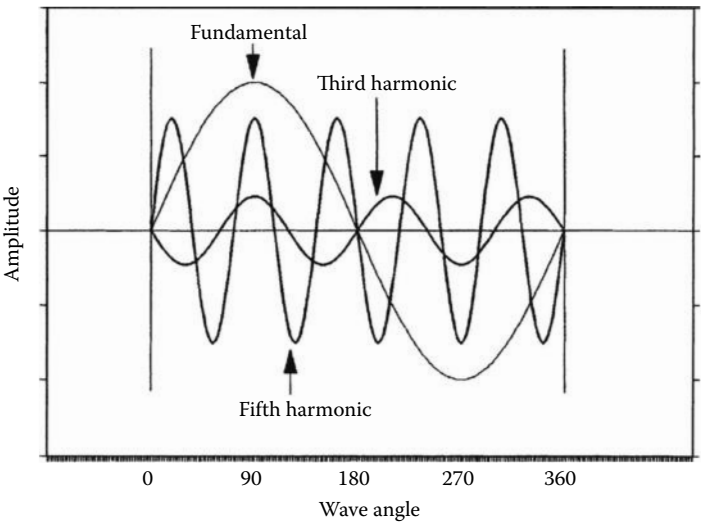


FIGURE 23.10 Harmonics are continuous integral multiples of the fundamental frequency. The fifth harmonic is shown, in phase with the fundamental, while the third harmonic is 180° out of phase.

TABLE 23.1 Harmonics and Their Relationship to the Fundamental Frequency

Harmonic Number	Frequency (Hz)	Frequency (Hz)	Frequency (Hz)
1	60	50	400
2	120	100	800
3	180	150	1200
4	240	200	1600
5	300	250	2000

23.3.4.3 Power Calculations

Calculating power delivered to do work for a nonlinear load is somewhat more complicated than if the current were sinusoidal. If the fundamental component of the voltage at frequency f is taken as a reference (the a-phase fundamental for a polyphase system), the subscript "1" means the fundamental, and E denotes the peak value of the voltage; then, the voltage can be expressed as

$$\varepsilon_{a1(t)} = E_{a1} \sin(2\pi ft + 0°)$$

The voltage fundamental will then have an amplitude E_{a1} and pass through zero in the positive-going direction at time $t = 0$. If h = harmonic number, and E_h and I_h are peak amplitudes of the harmonic voltage and current, respectively, then general expressions for any harmonic will be

$$e_{h(t)} = E_h \sin\left(2\pi fht + \alpha_h^\circ\right)$$

$$i_{h(t)} = I_h \sin\left(2\pi fht + \beta_h^\circ\right)$$

To compute the power associated with a voltage and current waveform, take advantage of the fact that products of harmonic voltages and currents of different frequency have a time average of zero.

Only products of voltages and currents of the same frequency are of interest, giving a general expression for harmonic power as

$$p_{h(t)} = E_h I_h \sin\left(2\pi fht + \alpha_h^\circ\right)\sin\left(2\pi fht + \beta_h^\circ\right)$$

Simplifying with trigonometric identities, evaluating over an integral number of cycles, and replacing peak voltage and current with rms values, the average power becomes

$$p_{h(t)} = E_h I_h \cos\left(\alpha_h^\circ - \beta_h^\circ\right)$$

For a *single-phase system* where h is the harmonic number and H is the highest harmonic, the total average power or active power is given by

$$p = \sum_{h=1}^{H} E_h I_h \cos\left(\alpha_h - \beta_h\right) \tag{23.12}$$

Total average reactive power is given by

$$Q = \sum_{h=1}^{H} E_h I_h \sin\left(\alpha_h - \beta_h\right) \tag{23.13}$$

It should be noted that in the real world, the actual contribution of harmonic frequencies to active and reactive power is small (usually less than 3% of the total active or reactive power). The major contribution of harmonic frequencies to the power mix comes as distortion power, which will be defined later.

For a *polyphase system* wherein r is the phase identification and N is the number of conductors in the system, including the neutral conductor, total average power for a polyphase system is given by

$$p = \sum_{r=1}^{N-1}\sum_{h=1}^{H} E_{rh} I_{rh} \cos(\alpha_{rh} - \beta_{rh}) \tag{23.14}$$

Total average reactive power is given by

$$Q = \sum_{r=1}^{N=1}\sum_{h=1}^{H} E_{rh} I_{rh} \sin(\alpha_{rh} - \beta_{rh}) \tag{23.15}$$

23.3.5 Power Factor

23.3.5.1 Single-Phase Systems

For a single-phase system, phasor power is again given by Equation 23.7 and illustrated by Figure 23.7, where P is the algebraic sum of the active powers for the fundamental and all the harmonics (Equation 23.12) and Q is the algebraic sum of the reactive powers for the fundamental and all the harmonics (Equation 23.13). Therefore, phasor power is based on the fundamental and harmonic active and reactive powers. It is found, however, that phasor power S is no longer equal to apparent power U and a new power phasor must be defined to recognize the effects of waveform distortion. A phasor representing the distortion, termed *distortion power* and given the symbol D, is defined by

$$D = \pm\sqrt{U^2 - S^2} \tag{23.16}$$

Without further definite information as to the sign of distortion power, its sign is selected the same as the sign of the total active power. The relationships among the various power terms are displayed in Figure 23.11, a 3D phasor diagram. Power factor, in direct parallel with sinusoidal waveforms, is defined by the equation

$$F_p = \frac{\text{Total active power}}{\text{Apparent power}} = \frac{P}{U} = \frac{\text{Watts}}{\text{Voltamps}} \tag{23.17}$$

From Equations 23.7 and 23.16, we obtain

$$S = \sqrt{p^2 + Q^2} \tag{23.7}$$

$$U = \sqrt{S^2 + D^2} = \sqrt{P^2 + Q^2 + D^2} \tag{23.18}$$

It is clear that when waveforms are sinusoidal, that is, linear loads are drawing current, there is no distortion power and Equation 23.18 reduces to Equation 23.7. Likewise as shown in Figure 23.13, as the distortion power vector goes to zero, the figure becomes 2D and reduces to Figure 23.7, and U becomes equal to 5. When, however, nonlinear loads are drawing harmonic currents from the system, U will be greater than S. As already noted, the contribution of the harmonics to the total power quantities is small, and one is frequently interested mainly in the fundamental quantities.

The power factor associated only with the fundamental voltage and current components was termed the displacement power factor $F_{\text{p displacement}}$ where Equations 23.7 and 23.8 are written as [5]

$$S_{60} = \sqrt{P_{60}^2 + Q_{60}^2}$$

and

$$F_{\text{p displacement}} = \frac{P_{60}}{S_{60}}$$

When harmonics are present, F_p is always smaller than $F_{\text{p displacement}}$.

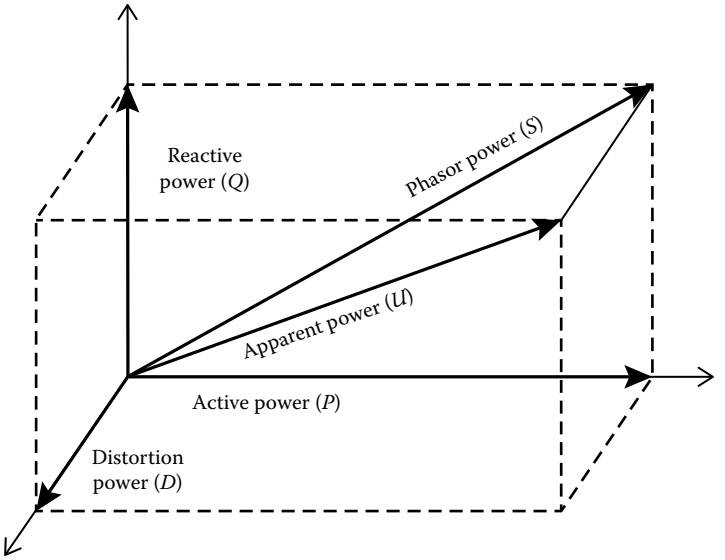


FIGURE 23.11 Phasor diagram for a single-phase, nonsinusoidal service in which the voltage and current contain harmonics. Geometric relationships are shown between active, reactive, phasor, distortion, and apparent powers.

23.3.5.2 Polyphase Systems

For a polyphase system, phasor power, *S*, is again given by Equation 23.7, but one must now use the total values for *P* and *Q* calculated using Equations 23.14 and 23.15. One can then define the apparent power *U* in one of two ways:

- *Arithmetic apparent power.* The *arithmetic apparent power* is given the symbol U_a and is defined by Equation 23.19, where E_r and I_r are the rms values for the respective phases and *M* equals the number of phases. U_a is a scalar quantity:

$$U_a = \sum_{r=1}^{M-1} E_r I_r \tag{23.19}$$

- *Apparent power. Apparent power is* given the symbol *U* and is defined by Equation 23.18 using total values for *P* and *Q* as defined by Equations 23.14 and 23.15 and a total value for *D* determined using Equation 23.16 for each phase. Figure 23.12 illustrates the two alternative concepts for polyphase apparent power [6]. Note that U_a uses arithmetic addition of vector magnitudes and is equal to apparent power *U* only if the polyphase voltages and currents have equal magnitudes and equal angular spacings, a situation that often exists in balanced three-phase systems. The two alternative definitions of apparent power, *U* and U_a, give rise to two possible values for power factor: (1) $F_p = P/U$, and (2) $F_{pa} = P/U_a$. Apparent power *U* and power factor F_p are the preferred definitions since using U_a can give unexpected results with some nonsymmetric service arrangements such as four-wire delta, and, with extremely unbalanced resistive loads, F_{pa} can exceed 1.0. Despite these shortcomings, arithmetic apparent power has become quite widely used due to the comparative simplicity of its measurement and calculation. With the advent of sophisticated digital meters, there is no longer any advantage to using arithmetic apparent power, and its use will surely decrease.

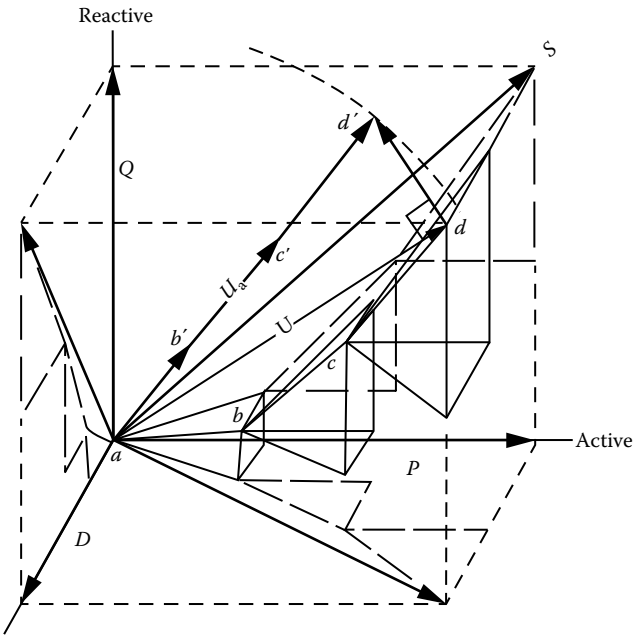


FIGURE 23.12 Phasor diagram for a three-phase, nonsinusoidal service in which the voltage and current contain harmonics. Arithmetic apparent power U_a is the length of the segmented line abed and is a scaler quantity U_a that can be represented by the line ab′c′d′. The diagonal ad, a vector quantity, is the apparent power *U*. (From Bullock, D.F. and Elmore, D.D., *MIND UR PS & QS, GE Meter EEI-AEIC Meter and Service Committee Meeting*, Dallas, TX, 1994.)

23.4 Power Factor "Measurement"

There are no instruments that measure power factor directly. (Power stations and large substations often use phase angle meters with a power factor scale representing $\cos(\theta)$ to display power factor. Such meters are accurate only for sinusoidal balanced polyphase systems.) One must remember that, of all the ac power quantities discussed, the only ones that can be directly measured are voltages, currents, and their time relationships (phase angles). All other ac power quantities are derived mathematically from these measured quantities. The only one of these derived values that has physical reality is the active power P (the quantity that does work); the others are mathematical constructs. Therefore, correct determination of power factor requires accurate measurement of voltage and current and proficient mathematics.

23.4.1 Metering for Linear Loads

By the early 1920s, the concepts of active, reactive, and apparent (since renamed phasor) power and power factor were known, and metering capabilities had been developed to enable their determination. Energy meters utilizing voltage and current coils driving an induction disk inherently measured active power P ($EI\cos\theta$), which was displayed on a mechanical register. Using the trigonometric identity $EI\sin\theta = EI\cos(90° + \theta)$, with voltage delayed 90°, a similar energy meter displayed reactive power Q, and, with these two values, displacement power factor was calculated. Voltage delay (phase shifting) was accomplished using specially wound transformers.

Through the years, the method of obtaining the 90° phase shift has been updated. Analog electronic meters are available that provide the 90° phase shift electronically within the meter. More recently, digital meters have been developed that sample voltages and currents at regular intervals and digitize the results. Voltages and currents are multiplied as they are captured to compute active power. Past voltage samples delayed by a time equal to a quarter cycle (90°) are multiplied by present current values to obtain reactive power. Active–reactive metering of this type is a widely utilized method for determining displacement power factor for utility billing. These meters do not accurately measure the effect of harmonic currents because the delay of the voltage samples is based on the fundamental frequency and is incorrect for the harmonics. (The important fifth harmonic, which is frequently the predominant harmonic component, is accurately measured because it is delayed 450° (5 × 90°), which is equivalent to the correct 90° delay).

23.4.2 Metering for Nonlinear Loads

With the application of high-speed digital computing techniques to measurement of ac currents and voltages, together with digital filtering, the quantities necessary for accurate and correct calculation of power factor are susceptible to direct computation. In practice, the ac waveforms are sampled at a frequency greater than twice the highest frequency to be measured, in compliance with well-known sampling theories. Data obtained can be treated using Fourier's equations to calculate rms values for voltage, current, and phase angle for the fundamental and each harmonic frequency. Power quantities can be obtained with digital filtering in strict accordance with their *ANSI/IEEE STD* 100 definitions. Power quantities can be displayed for the fundamental only (displacement power factor) or for fundamental plus all harmonics (power factor for nonsinusoidal waveforms).

23.4.3 Metering Applications

Instruments with the capabilities described earlier are often called *harmonic analyzers* and are available in both single-phase and polyphase versions. They can be portable, in which case they are often used for power surveys or panel mounted for utility and industrial revenue metering. Polyphase analyzers can be connected as shown in Figures 23.13 and 23.14. Care must be taken to connect the instrument properly.

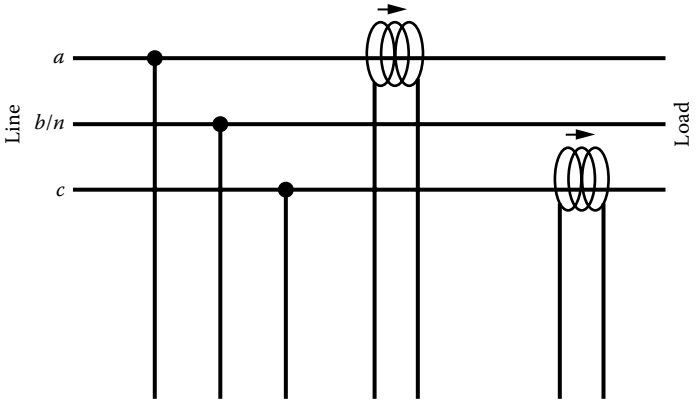


FIGURE 23.13 Meter connections for a one-, two-, or three-phase, three-wire service. Voltage leads are connected to phases *a*, *b*, and *c or* to a, c, and neutral as the system dictates. Care must be taken to connect the voltage lead for each phase to the input corresponding to the input for the current transformer reading that phase, and directional characteristics of the transformers must be observed. Only two current connections are required.

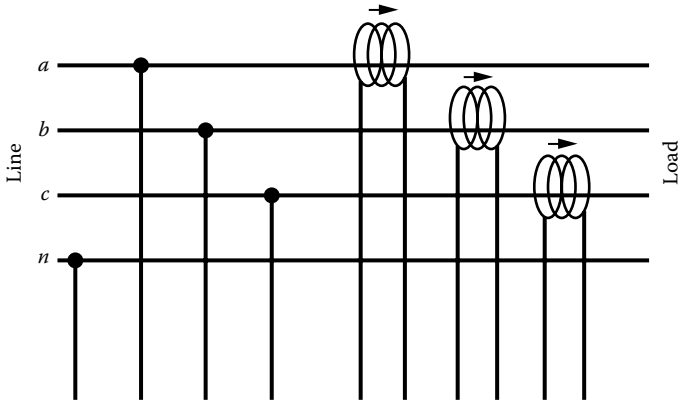


FIGURE 23.14 Meter connections for a three-phase, four-wire wye- or delta-connected service. Voltage leads are connected to phases *a*, *b*, and *c* and to the neutral. Care must be taken to connect the voltage lead for each phase to the input corresponding to the input for the current transformer reading that phase, and directional characteristics of the transformers must be observed. In situations where measurement of neutral currents is desired, a fourth current transformer can be used to read the neutral. A fourth voltage connection might be used to read neutral-ground voltages.

Both voltage leads and current transformers must be connected to the proper phases, and the current transformers must also be placed with the correct polarity. Most instruments use color-coded voltage connectors. Correct polarity is indicated on the current transformers by arrows or symbols, and complete hookup and operating instructions are included. When single-phase instruments are used, the same precautions must be followed for making connections.

23.5 Instrumentation

Table 23.2 lists a sampling of harmonic and power factor measuring instruments available from major manufacturers. All instruments listed use some type of Fourier calculations and/or digital filtering to determine power values in accordance with accepted definitions. Many can be configured to measure nonharmonic-related power quality concerns. Unless otherwise noted, all instruments require the

TABLE 23.2 Selected Instrumentation for Harmonic and Power Factor Measurement

Manufacturer	Model	*V/I* Inputs	Display Type	Communication	Special Features	List Price (U.S.$)
Handheld units						
Amprobe	HA2000 1/1	Visual	RS232	Handheld, probe included, 21 nonvolatile memories	990	—
BMI	155	1/1	Visual	Optional printer, RS232	Handheld, 1145 + 550 (printer) + 380 (probe)	—
BMI	355	4/3	Visual	Optional printer, RS232	Handheld	1995 + 550 (printer) + 380/probe
Dranetz 4300	4/4	Panel for data and graphs RS232	Handheld, battery or ac power, optional PCMCIA memory card	5000 + 450/probe	—	—
Fluke	39	1/1	Visual	None	Probe included	895
Fluke	41b	1/1	Visual	RS232	Probe included, eight memories, logging with computer and supplied software	1795
Portable units						
BMI	3030A	4/4	Built-in printer for data and graphs, optional internal modem	Portable, optional built-in disk drive for storage, long-term monitoring, optional PQ configurations	6800 + 600 (modem) + 1895 (disk storage) + 590/probe	—
Dranetz	PP1-R	4/4	Panel for data and graphs	PCMCIA card slot	Long-term monitoring, optional PQ configurations, remote control by software	12,000 + 450/probe
PC-based units						
BMI	7100	4/4	PC-based (not included)	PC connection cord	Portable, uses computer for storage, optional software, long-term monitoring, optional PQ configurations	5294 + PC + 395 (software) + 415/probe
Cooper	V-Monitor II	4/4	PC-based (not included)	Serial port	Portable, software and signal interface and data acquisition board, long-term monitoring	12,000 + 500/probe
RPM	1650	4/5	PC-based, not included	Ethernet long-term monitoring, optional PQ configurations, remote control by software	6250+ 750 software + 490/probe	

(*continued*)

TABLE 23.2 (continued) Selected Instrumentation for Harmonic and Power Factor Measurement

Manufacturer	Model	*V/I* Inputs	Display Type	Communication	Special Features	List Price (U.S.$)
Panel mounted						
Cutler-Hammer/ Westinghouse	4/4	Panel	Optional IMPACC	Panel mount to replace meters, monitoring	3690 + input devices	—
General Electric	kV	Vector electricity meter	Socket	Programmable multifunction LCD display pulse output for measured power quantities replaces industrial revenue meters and calculates and accumulates all power and revenue data	595 + 495 (software)	
Square D	PowerLogic PM620	3/3	LCD panel	RS485	Panel meter replacement, calculates and accumulates power data	1583 + probes
Square D	PowerLogic CM2350	3/3	Panel	RS485	Connect to network, remote controller, monitoring	4290 + probes

TABLE 23.3 Manufacturers of Power Factor Measuring Harmonic Analyzers

Amprobe Instruments
630 Merrick Road Lynbrook
New York 11563
Tel: (516) 593-5600
Web: http://www.amprobe.com/amprobe/usen/home/

Cooper Power Systems Division
11131 Adams Road
P.O. Box 100
Franksville, WI 53126-0100
Tel: (414) 835-2921
Web: http://www.cooperindustries.com/content/public/en.html

Dranetz Technologies, Inc.
1000 New Durham Road
Edison, NJ 08818-4019
Tel: (800) DRANTEC
Web: http://dranetz.com/

Fluke Corporation
P.O. Box 9090
Everett, WA 98206
Tel: (800) 44FLUKE
Web: http://www.fluke.com/Fluke/usen/About/Contact

GE Meter
130 Main Street
Somersworth, NH 03878
Tel: (603) 749-8477

Reliable Power Meters, Inc.
400 Blossom Hill Road
Los Gatos, CA 95032
Tel: (408) 358-5100

Square D Power Logic
330 Weakley Road
Smyrna, TN 37167-9969
Tel: (615) 459-8552

Yokogawa Corp. Meter and Instrument Div.,
Newnan, GA 30655
Tel: (800) 888-6400 (toll free)
Email: http://www.yokogawa-usa.com/

purchase of one current probe per input. Probes are available for measuring currents from 5 A to several thousand amperes. For comparison purposes, priced probes will be those with a 600 A range. Voltage leads are usually supplied as standard equipment. Table 23.3 contains addresses of these manufacturers.

Defining Terms

Active power: A term used to express the real power delivered to do work in an ac distribution system.
Apparent power: A term used to express the product of volts and amperes in an ac distribution system in which voltage and/or current is nonsinusoidal.
Harmonic power: A term used to express the power due to harmonic frequencies in an ac distribution system in which voltage and/or current is nonsinusoidal.
Phasor power: A term used to express the product of volts and amperes in an ac distribution system in which voltage and current are sinusoidal.
Power factor: A single number, calculated by dividing active power by either the phasor power or the apparent power, which describes the effective utilization of ac distribution system capacity.
Reactive power: A term used to express the imaginary power that does no work but provides magnetization to enable work in an ac distribution system.

References

1. D. F. Bullock, Methods of measuring apparent power, GE Meter, *Missouri Valley Electrical Association Annual Engineering Conference*, Kansas City, MO, 1996.
2. H. W. Beaty and D. G. Fink, *Standard Handbook for Electrical Engineers, Chapter 1: Units, Symbols, Constants, Definitions and Conversion Factors*, 16th edn., McGraw-Hill Professional, New York, 2013.
3. D. F. Bullock, *Phase Angle Relationships, Part 1: Theory*, GE Meter Arkansas Electric Meter School, Fayetteville, AR, 1995.
4. I. Assimov, *Assimov's Biographical Encyclopedia of Science and Technology*, New Rev. edn., Doubleday & Co., New York, 1972.
5. D. F. Bullock, private communication.
6. D. F. Bullock and D. D. Elmore, *MIND UR PS & QS, GE Meter EEI-AEIC Meter and Service Committee Meeting*, Dallas, TX, 1994.

24

Phase Measurement

Peter O'Shea
Royal Melbourne Institute of Technology

The notion of "phase" is usually associated with *periodic* or repeating signals. With these signals, the waveshape perfectly repeats itself every time the *period* of repetition elapses. For periodic signals one can think of the phase at a given time as the fractional portion of the period that has been completed. This is commonly expressed in degrees or radians, with full cycle completion corresponding to 360° or 2π radians. Thus, when the cycle is just beginning, the phase is zero. When the cycle is half completed, the phase is half of 360°, or 180° (see Figure 24.1). It is important to note that if phase is defined as the portion of a cycle that is completed, the phase depends on where the beginning of the cycle is taken to be. There is no universal agreement on how to specify this beginning. For a sinusoidal signal, probably the two most common assumptions are that the start of the cycle is (1) the point at which the maximum value is achieved and (2) the point at which the negative to positive zero-crossing occurs. Assumption (1) is common in many theoretical treatments of phase and for that reason is adopted in this chapter. It should be noted, however, that assumption (2) has some benefits from a measurement perspective, because the zero-crossing position is easier to measure than the maximum.

The measurement of phase is important in almost all applications where sinusoids proliferate. Many means have therefore been devised for this measurement. One of the most obvious measurement techniques is to directly measure the fractional part of the period that has been completed on a cathode-ray oscilloscope (CRO). Another approach, which is particularly useful when a significant amount of noise is present, is to take the Fourier transform of the signal. According to the Fourier theory, for a sinusoidal signal, the energy in the Fourier transform is concentrated at the frequency of the signal; the initial phase of the signal (i.e., the phase at time, $t = 0$) is the phase of the Fourier transform at the point of this energy concentration. The measurements of initial phase and frequency obtained from the Fourier transform can then be used to deduce the phase of the signal for any value of time.

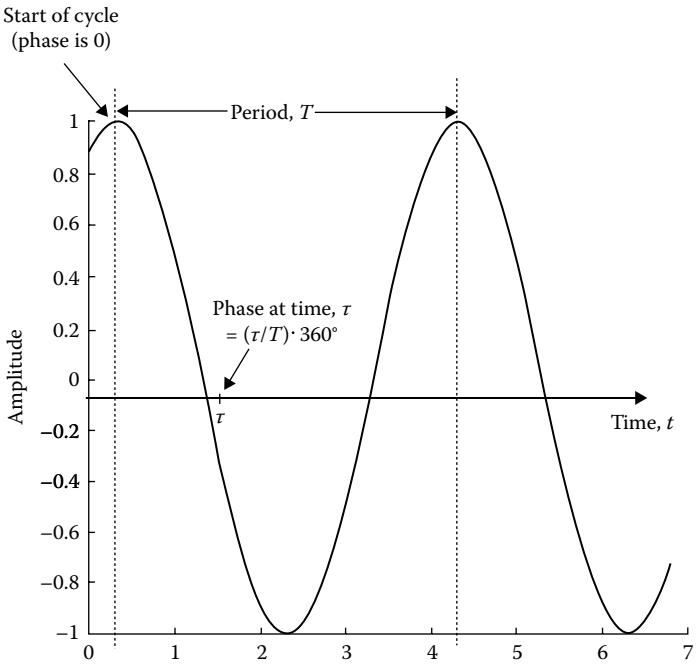


FIGURE 24.1 The phase of a periodic sinusoidal signal. The time scale is arbitrary.

Frequently what is needed in practice is a measurement of the *phase difference* between two signals of the same frequency; that is, it is necessary to measure the *relative phase* between two signals rather than the *absolute phase* of either one (see Figure 24.2). Often, in the measurement of the relative phase between two signals, both signals are derived from the same source. These signals might, for example, be the current and voltage of a power system; the relative phase, ϕ, between the current and voltage would then be useful for monitoring power usage, since the latter is proportional to the cosine of ϕ.

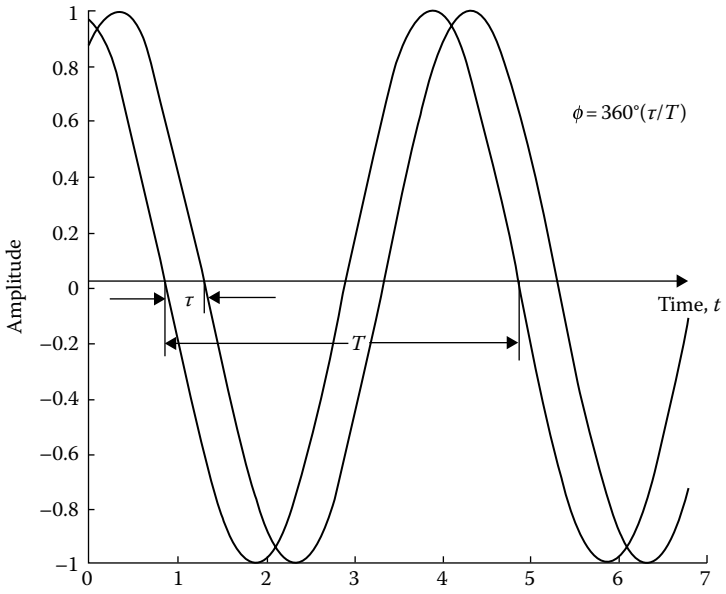


FIGURE 24.2 Two signals with a relative phase difference of ϕ between them. The time scale is arbitrary.

Several techniques are available for the measurement of "relative phase." One crude method involves forming "Lissajous figures" on an oscilloscope. In this method, the first of the two signals of interest is fed into the vertical input of a CRO, and the other is fed into the horizontal input. The result on the oscilloscope screen is an ellipse, the intercept and maximum height of which can be used to determine the relative phase. Other methods for determining relative phase include the crossed-coil meter (based on electromagnetic induction principles), the zero-crossing phase meter (based on switching circuitry for determining the fractional portion of the period completed), the three-voltmeter method (based on the use of three signals and trigonometric relationships), and digital methods (based on analog-to-digital [A/D] conversion and digital processing).

24.1 Amplitude, Frequency, and Phase of a Sinusoidal Signal

An arbitrary sinusoidal signal can be written in the following form:

$$s(t) = A\cos(2\pi ft + \phi_0) = A\cos(\omega t + \phi_0) \tag{24.1}$$

where

A is the peak amplitude
f is the frequency
ω is the angular frequency
ϕ_0 is the phase at time $t = 0$

This signal can be thought of as being the real part of a complex phasor that has amplitude, A, and which rotates at a constant angular velocity $\omega = 2\pi f$ the complex plane (see Figure 24.3).

Mathematically, then, $s(t)$ can be written as follows:

$$s(t) = \Re\mathrm{e}\left\{Ae^{j2\pi ft+\phi_0}\right\} = \Re\mathrm{e}\{z(t)\} \tag{24.2}$$

where

$z(t)$ is the complex phasor associated with $s(t)$
$\Re\mathrm{e}\{.\}$ denotes the real part

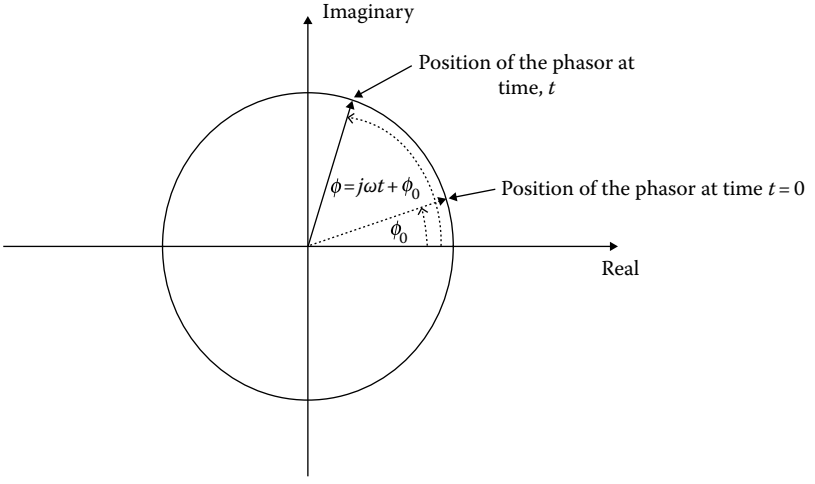


FIGURE 24.3 A complex rotating phasor, $A\exp(j\omega t + \phi_0)$. The length of the phasor is A and its angular velocity is ω. The real part of the phasor is $A\cos(\omega t + \phi_0)$.

The "phase" of a signal at any point in time corresponds to the angle that the rotating phasor makes with the real axis. The initial phase (i.e., the phase at time $t = 0$) is ϕ_0. The "frequency" f of the signal is $1/2\pi$ times the phasor's angular velocity.

There are a number of ways to define the phase of a real sinusoid with unknown amplitude, frequency, and initial phase. One way, as already discussed, is to define it as the fractional part of the period that has been completed. This is a valid and intuitively pleasing definition and one that can be readily generalized to periodic signals that contain not only a sinusoid but also a number of harmonics. It cannot, however, be elegantly generalized to allow for slow variations in the frequency of the signal, a scenario that occurs in communications with phase and frequency modulation. Gabor put forward a definition in 1946 that can be used for signals with slowly varying frequency. He proposed a mathematical definition for generating the complex phasor, $z(t)$, associated with the real signal, $s(t)$. The so-called *analytic signal* $z(t)$ is defined according to the following definition [1]:

$$z(t) = s(t) + jH\{s(t)\} \tag{24.3}$$

where $H\{\ \}$ denotes the *Hilbert transform* and is given by

$$H\{s(t)\} = p.v.\left[\int_{-\infty}^{+\infty} \frac{s(t-\tau)}{\pi\tau} d\tau\right] \tag{24.4}$$

with *p.v.* signifying the Cauchy principal value of the integral [2].

The imaginary part of the analytic signal can be generated practically by passing the original signal through a "Hilbert transform" filter. From Equations 24.3 and 24.4, it follows that this filter has impulse response given by $1/\pi t$. The filter can be implemented, for example, with one of the HSP43xxx series of integrated circuits (ICs) from Harris Semiconductor. Details of how to determine the filter coefficients can be found in [2].

Having formally defined the analytic signal, it is possible to provide definitions for phase, frequency, and amplitude as functions of time. They are given as follows:

$$\text{Phase: } \phi(t) = \arg\{z(t)\} \tag{24.5}$$

$$\text{Frequency: } f(t) = \frac{1}{2\pi}\frac{d[\arg\{z(t)\}]}{dt} \tag{24.6}$$

$$\text{Amplitude: } A(t) = \text{abs}[z(t)] \tag{24.7}$$

The definitions for phase, frequency, and amplitude can be used for signals whose frequency and/or amplitude varies slowly with time. If the frequency and amplitude do vary with time, it is common to talk about the "instantaneous frequency" or "instantaneous amplitude" rather than simply the frequency or amplitude.

Note that in the analytic signal, the imaginary part lags the real part by 90°. This property actually holds not only for sinusoids but for the real and imaginary parts of all frequency components in "multicomponent" analytic signals as well. The real and imaginary parts of the analytic signal then correspond to the "inphase (I)" and "quadrature (Q)" components used in communications systems.

In a balanced three-phase electric power distribution system, the analytic signal can be generated by appropriately combining the different outputs of the electric power signal; that is, it can be formed according to

$$z(t) = v_a(t) + \frac{j}{\sqrt{3}}[v_c(t) - v_b(t)] \tag{24.8}$$

where

$v_a(t)$ is the reference phase
$v_b(t)$ is the phase that leads the reference by 120°
$v_c(t)$ is the phase that lags the reference by 120°

24.2 Phase of a Periodic Nonsinusoidal Signal

It is possible to define "phase" for signals other than sinusoidal signals. If the signal has harmonic distortion components present in addition to the fundamental, the signal will still be periodic, but it will no longer be sinusoidal. The phase can still be considered to be the fraction of the period completed. The "start" of the period is commonly taken to be the point at which the initial phase of the fundamental component is 0, or at a zero-crossing. This approach is equivalent to just considering the phase of the fundamental and ignoring the other components. The Fourier method provides a very convenient method for determining this phase—the energy of the harmonics in the Fourier transform can be ignored.

24.3 Phase Measurement Techniques

24.3.1 Direct Oscilloscope Methods

CROs provide a simple means for measuring the phase difference between two sinusoidal signals. The most straightforward approach to use is direct measurement; that is, the signal of interest is applied to the vertical input of the CRO, and an automatic time sweep is applied to the horizontal trace. The phase difference is the time delay between the two waveforms measured as a fraction of the period. The result is expressed as a fraction of 360° or of 2π rad; that is, if the time delay is 1/4 of the period, then the phase difference is 1/4 of 360° = 90° (see Figure 24.2). If the waveforms are not sinusoidal but are periodic, the same procedure can still be applied. The phase difference is just expressed as a fraction of the period or as a fractional part of 360°.

Care must be taken with direct oscilloscope methods if noise is present. In particular, the noise can cause triggering difficulties that would make it difficult to accurately determine the period and/or the time delay between two different waveforms. The "HF reject" option, if available, will alleviate the triggering problems.

24.3.2 Lissajous Figures

Lissajous figures are sometimes used for the measurement of phase. They are produced in an oscilloscope by connecting one signal to the vertical trace and the other to the horizontal trace. If the ratio of the first frequency to the second is a rational number (i.e., it is equal to one small integer divided by another), then a closed curve will be observed on the CRO (see Figures 24.4 and 24.5). If the two frequencies are unrelated, then there will be only a patch of light observed because of the persistence of the oscilloscope screen.

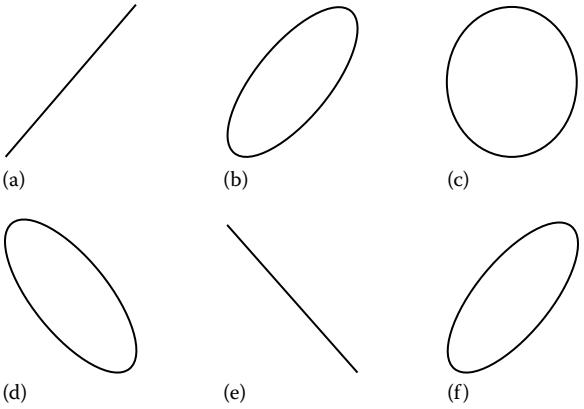


FIGURE 24.4 Lissajous figures for two equal-amplitude, frequency-synchronized signals with a relative phase difference of (a) 0, (b) $\pi/4$, (c) $\pi/2$, (d) $3\pi/4$, (e) π, and (f) $-\pi/4$.

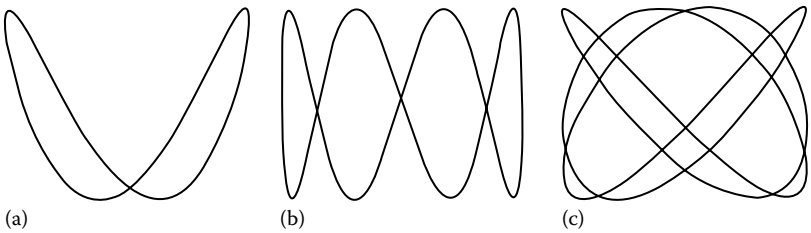


FIGURE 24.5 Lissajous figures for two signals with vertical frequency: horizontal frequency ratios of (a) 2:1, (b) 4:1, and (c) 4:3.

If the two signals have the same frequency, then the Lissajous figure will assume the shape of an ellipse. The ellipse's shape will vary according to the phase difference between the two signals and according to the ratio of the amplitudes of the two signals. Figure 24.6 shows some figures for two signals with synchronized frequency and equal amplitudes, but different phase relationships. The formula used for determining the phase is

$$\sin(\phi) = \pm\frac{Y}{H} \tag{24.9}$$

where

H is half the maximum vertical height of the ellipse
Y is the intercept on the *y-axis*

Figure 24.7 shows some figures for two signals that are identical in frequency and have a phase difference of 45°, but with different amplitude ratios. Note that it is necessary to know the direction that the Lissajous trace is moving in order to determine the sign of the phase difference. In practice, if this is not known a priori, then it can be determined by testing with a variable frequency signal generator. In this case, one of the signals under consideration is replaced with the variable frequency signal. The signal generator is adjusted until its frequency and phase equal that of the other signal input to the CRO. When this happens, a straight line will exist. The signal generator frequency is then increased a little, with the relative phase thus being effectively changed in a known direction. This can be used to determine the correct sign in Equation 24.9.

Lissajous figure methods are a little more robust to noise than direct oscilloscope methods. This is because there are no triggering problems due to random noise fluctuations. Direct methods are,

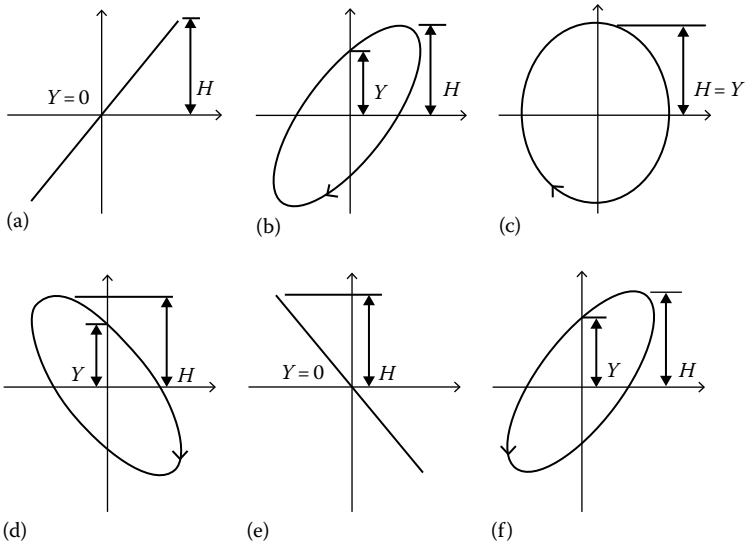


FIGURE 24.6 Lissajous figures for two signals with synchronized frequency and various phase differences: (a) phase difference = 0°, (b) phase difference = 45°, (c) phase difference = 90°, (d) phase difference = 135°, (e) phase difference = 180°, and (f) phase difference = 315°.

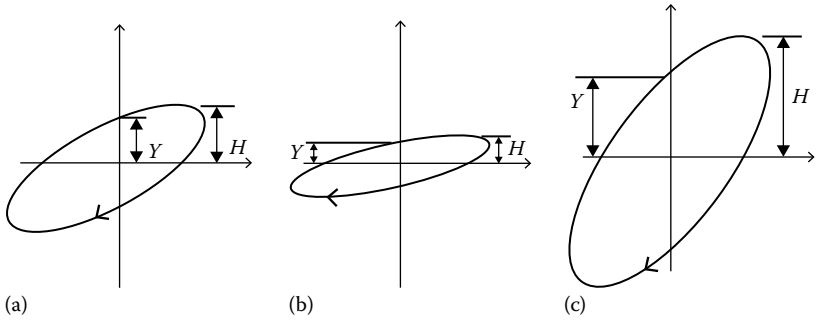


FIGURE 24.7 Lissajous figures for two signals with synchronized frequency, a phase difference of 45°, and various amplitude ratios: (a) amplitude ratio of 1, (b) amplitude ratio of 0.5, and (c) amplitude ratio of 2.

however, much easier to interpret when harmonics are present. The accuracy of oscilloscope methods is comparatively poor. The uncertainty of the measurement is typically in excess of 1°.

24.3.3 Zero-Crossing Methods

This method is currently one of the most popular methods for determining phase difference, largely because of the high accuracy achievable (typically 0.02°). The process is illustrated in Figure 24.8 for two signals, denoted *A* and *B*, that have the same frequency but different amplitudes. Each negative to positive zero-crossing of signal *A* triggers the start of a rectangular pulse, while each negative to positive zero-crossing of signal *B* triggers the end of the rectangular pulse. The result is a pulse train with a pulse width proportional to the phase angle between the two signals. The pulse train is passed through an averaging filter to yield a measure of the phase difference. It is also worth noting that if the positive to negative zero-crossings are also used in the same fashion and the two results are averaged, the effects of dc and harmonics can be significantly reduced.

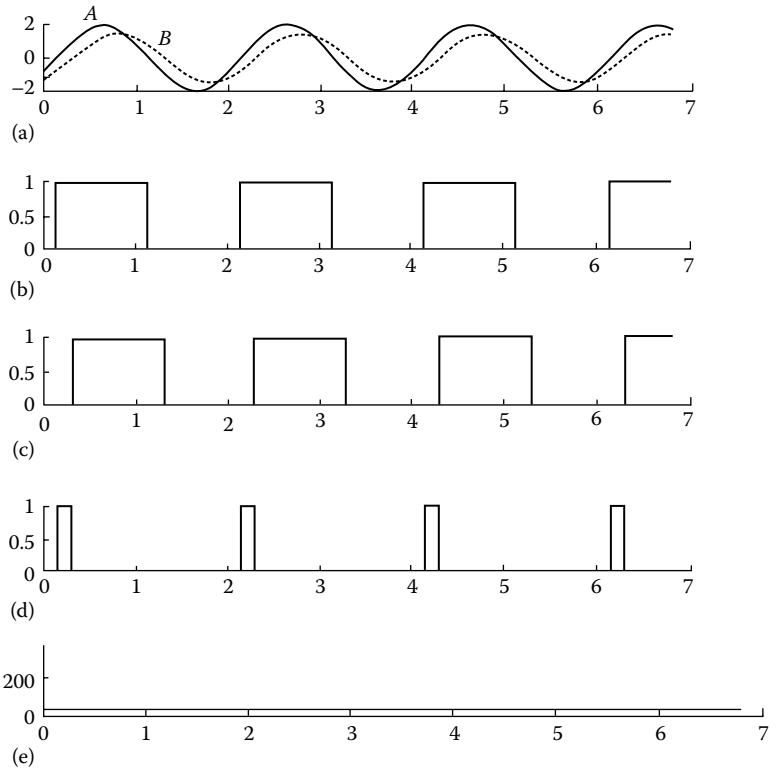


FIGURE 24.8 Input, output, and intermediate signals obtained with the zero-crossing method for phase measurement. Note that the technique is not sensitive to signal amplitude: (a) The two input signals, *A* and *B*; (b) zero-crossing detector output for *A*; (c) zero-crossing detector output for *B*; (d) bistable flip-flop output (the input signals are shown in (b) and (c)); (e) phase difference (degrees) obtained by averaging bistable output.

To implement the method practically, the analog input signals must first be converted to digital signals that are "high" if the analog signal is positive and "low" if the analog signal is negative. This can be done, for example, with a Schmitt trigger, along with an resistive-capacitive (*RC*) stabilizing network at the output. In practice, high-accuracy phase estimates necessitate that the switching of the output between high and low be very sharp. One way to obtain these sharp transitions is to have several stages of "amplify and clip" preceding the Schmitt trigger.

The digital portion of the zero-crossing device can be implemented with an edge-triggered RS flip-flop and some ancillary circuitry, while the low-pass filter on the output stage can be implemented with an RC network.

A simpler method for measuring phase based on zero-crossings involves the use of an exclusive or (XOR) gate. Again, the analog input signals must first be converted to digital pulse trains. The two inputs are then fed into an XOR gate and finally into a low-pass averaging filter. A disadvantage with this method is that it is only effective if the duty cycle is 50% and if the phase shifts between the two signals are between 0 and π rad. It is therefore not widely used.

24.3.4 Three-Voltmeter Method

The measurement of a phase difference between two voltage signals, v_{ac} and v_{bc}, can be expedited if there is a common voltage point, c. The voltage between points b and a (v_{ba}), the voltage between points b and c (v_{bc}), and the voltage between points c and a (v_{ca}) are measured with three different voltmeters. A vector diagram is constructed with the three measured voltages as shown in Figure 24.9. The phase difference

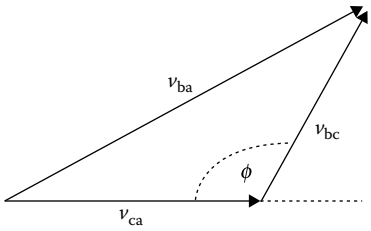


FIGURE 24.9 A vector diagram for determining the phase angle, ϕ, between two ac voltages, v_{ac} and v_{bc}, with the three-voltmeter method.

between the two vectors, v_{ac} and v_{bc}, is determined using a vector diagram (Figure 24.9) and the cos rule. The formula for the phase difference, ϕ, in radians is given by

$$\pi - \phi = \cos^{-1}\left(\frac{v_{ca}^2 + v_{bc}^2 - v_{ba}^2}{2v_{ca}v_{bc}}\right) \tag{24.10}$$

24.3.5 Crossed-Coil Method

The crossed-coil phase meter is at the heart of many analog power factor meters. It has two crossed coils, denoted A and B, positioned on a common shaft but aligned at different angles (see Figure 24.10). The two coils move as one, and the angle between them, β, never changes. There is another independent nonrotating coil, C, consisting of two separate parts, "enclosing" the rotating part (see Figure 24.10). The separation of coil C into two separate parts (forming a Helmholtz pair) allows the magnetic field of coil C to be almost constant in the region where the rotating A and B coils are positioned.

Typically the system current, I, is fed into coil C, while the system voltage, V, is applied to coil A via a resistive circuit. The current in coil A is therefore in phase with the system voltage, while the current in coil C is in phase with the system current. Coil B is driven by V via an inductive circuit, giving rise to a current that lags V (and therefore the current in coil A) by 90°. In practice, the angle between the currents in coils A and B is not quite 90° because of the problems associated with achieving purely resistive and purely inductive circuits. Assume, then, that this angle is β. If the angle between the currents in coil B and in coil C is ϕ, then the angle between the currents in coils A and C is $\beta + \phi$. The average torque induced in coil A is proportional to the product of the average currents in coils A and C and to

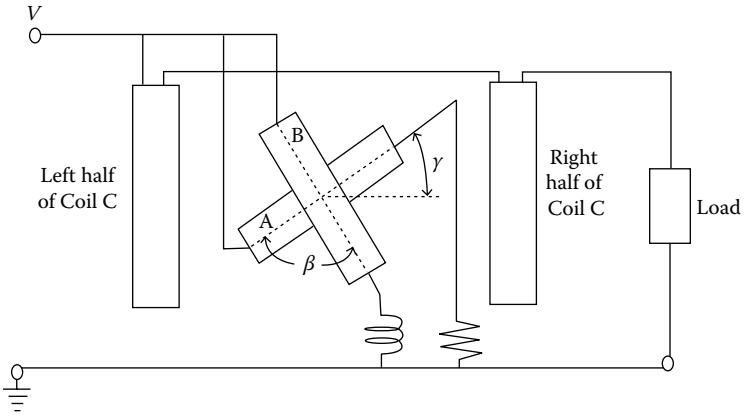


FIGURE 24.10 Diagram of a crossed-coil device for measuring phase. Coils A and B are the rotating coils. Coil C (left and right parts) is the stationary coil.

the cosine of the angle between coil A and the perpendicular to coil C. The average torque induced in coil A is therefore governed by the following equation:

$$\overline{T}_A \propto I_A I_C \cos(\phi + \beta)\cos(\gamma) = k_A \cos(\phi + \beta)\cos(\gamma) \tag{24.11}$$

where

I_A and I_C are constants
$\phi + \beta$ is the relative phase between the currents in coils A and C
γ is the angle between coil A and the perpendicular to coil C
$\propto$ signifies "is proportional to"

Assuming that the current in coil B lags the current in coil A by β, then the average torque in coil B will be described by

$$\overline{T}_B \propto I_B I_C \cos(\phi)\cos(\gamma + \beta) = k_B \cos(\phi)\cos(\gamma + \beta) \tag{24.12}$$

where

I_B is a constant
ϕ is the relative phase between the currents in coils B and C
the other quantities are as in Equation 24.11

Now, the torques due to the currents in coils A and B are designed to be in opposite directions. The shaft will therefore rotate until two torques are equal; that is, until

$$k_A \cos(\phi + \beta)\cos(\gamma) = k_B \cos(\phi)\cos(\gamma + \beta) \tag{24.13}$$

If $k_A = k_B$, then Equation 24.13 will be satisfied when $\phi = \gamma$. Thus, the A coil will be aligned in the direction of the phase shift between the load current and load voltage (apart from errors due to the circuits of the crossed coils not being perfectly resistive/inductive). Thus, a meter pointer attached to the A plane will indicate the angle between load current and voltage. In practice, the meter is usually calibrated to read the cosine of the phase angle rather than the phase angle and also to allow for the errors that arise from circuit component imperfections.

The accuracy of this method is limited, due to the heavy use of moving parts and analog circuits. Typically, the measurement can only be made accurate to within about 1° angle.

24.3.6 Synchroscopes and Phasing Lamps

The crossed-coil meter previously described is used as the basis for *synchroscopes*. These devices are often used in power generation systems to determine whether two generators are phase and frequency synchronized before connecting them together. In synchroscopes, the current from one generator is fed into the fixed coil and the current from the other generator is fed into the movable crossed coils. If the two generators are synchronized in frequency, then the meter needle will move to the position corresponding to the phase angle between the two generators. If the generators are not frequency synchronized, the meter needle will rotate at a rate equal to the difference between the two generator frequencies. The direction of rotation will indicate which generator is rotating faster.

When frequency synchronization occurs (i.e., the meter needle rotation ceases) and the phase difference is zero, the generators can be connected together. Often in practice, the generators are connected before synchronization occurs; the generator coming online is deliberately made a little higher in frequency so that it can provide extra power rather than be an extra drain on the system. The connection is still made, however, when the instantaneous phase difference is zero.

Phasing lamps are sometimes used as a simpler alternative to synchroscopes. A lamp is connected between the two generators, and any lack of frequency synchronization manifests as a flickering of the lamp. A zero-phase difference between the two generators corresponds to maximum brightness in the lamp.

24.3.7 Vector Voltmeters and Vector Impedance Methods

Alternating voltages (and currents) are often characterized as vectors consisting of a magnitude and a phase, with the phase being measured relative to some desired reference. Many instruments exist that can display the voltage amplitude and phase of a signal across a wide range of frequencies. These instruments are known as *vector voltmeters* or *network analyzers*. The phase and amplitude as a function of frequency can be obtained very simply in principle by taking the Fourier transform of the signal and simply reading the amplitude and phase across the continuum of frequencies. To achieve good accuracy, this is typically done with down-conversion and digital processing in the baseband region. The down-conversion can be analog, or it can be digital. The procedure is described more fully in the succeeding paragraphs.

To determine the real part of the voltage vector at a given frequency f, the signal is first down-converted by mixing with a local oscillator signal, $\cos(2\pi ft)$. This mixing of the signal recenters the frequency component of interest at 0 Hz. The resultant signal is low-pass filtered, digitally sampled (if not in the digital domain already), and averaged. The digital sampling and averaging enables the amplitude of the newly formed 0 Hz component to be evaluated. The imaginary part is obtained in similar fashion by mixing the signal with $\sin(2\pi ft)$, low-pass filtering, digitally sampling, and again averaging the samples. The amplitude and phase of the voltage vector, V, are obtained from the real and imaginary parts using the standard trigonometric relationships:

$$\text{Magnitude} = \text{Abs}(V) = \sqrt{[\Re\text{e}\{V\}]^2 + [\text{Im}\{V\}]^2} \tag{24.14}$$

$$\text{Phase} = \arg(V) = \arctan\left(\frac{\text{Im}\{V\}}{\Re\text{e}\{V\}}\right) \tag{24.15}$$

where $\Re\text{e}\{.\}$ and $\text{Im}\{.\}$ denote the real and imaginary parts, respectively.

The procedure for forming the vector voltage is summarized in the block diagram in Figure 24.11. In practice, the down-conversion can be carried out in more than one step. For high-frequency signals, for example, the first stage might shift a large band of frequencies to the audio region, where further down-conversion is carried out. Alternatively, the first stage might shift a band of frequencies to the intermediate frequency (IF) band and the second stage to the audio band.

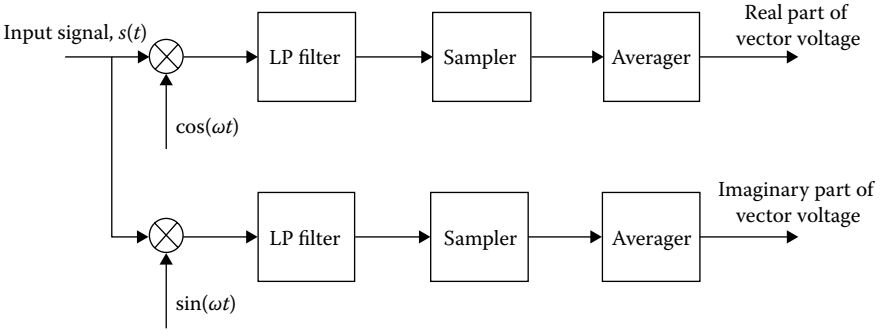


FIGURE 24.11 Vector voltmeter block diagram. The vector voltmeter determines the voltage (amplitude and phase or real and imaginary parts) of the component of the input signal at frequency f.

Just as it is possible to analyze a voltage signal and produce a magnitude and phase across any given frequency band, so it is possible to obtain a frequency profile of the magnitude and phase of a *current* signal. If current vectors and voltage vectors can be obtained for given impedance, then it is possible to obtain "vector impedance." This impedance is defined simply as the result of the complex division of voltage by current:

$$Z = \frac{V}{I} \tag{24.16}$$

The calculations of vector impedances are useful for such applications as impedance matching, power factor correction, and equalization.

Typically, much of the current processing for vector voltmeters and vector impedance meters is done digitally. One of the great advantages of this type of processing is the high accuracy achievable. Accuracies of 0.02° are common, but this figure is improving with developing technology. The high sampling rates that can be employed (typically beyond 1 GHz) cause the errors in the A/D conversion to be spread out over very large bandwidths. Since the ultimate measurement of a vector voltage or impedance is usually made over a very narrow bandwidth, the errors are substantially eliminated. The developments in technologies that enable greater accuracy are (1) sampling rate increases, (2) word-length increases, and (3) increased A/D converter fidelity.

24.3.8 Phase Standard Instruments

For high-precision phase measurements and calibration, "phase standard" instruments can be used. These instruments provide two sinusoidal signal outputs, whose phase difference can be controlled with great accuracy. They typically use crystal-controlled timing to digitally synthesize two independent sinusoids with a variable user-defined phase difference. The Clarke-Hess 5500 Digital Phase Standard is one such instrument. For this standard, the sinusoids can have frequencies ranging from 1 Hz to 100 kHz, while amplitudes can vary between 50 mV and 120 V rms. The phase can be set with a resolution of 0.001°, with a typical accuracy of about 0.003°.

24.3.9 Fast Fourier Transform Method

This method is one in which virtually all the processing is done in the digital domain. It operates on the pulse code modulated (PCM) digital samples of a signal. This and other similar methods are very promising means for measuring phase. This is because of the digital revolution that has resulted in cheap, fast, accurate, and highly versatile digital signal processors (DSPs). The latter are small computer chips capable of performing fast additions and multiplications that can be programmed to emulate conventional electronic functions such as filtering, coding, and modulation. They can also be programmed to perform a wide variety of functions not possible with analog circuitry. Up until the end of the 1980s, digital measurement was limited by the relatively inaccurate A/D conversion process required before digital processing could be performed. Developments in the early 1990s, however, saw the introduction of oversampling analog-to-digital converters (ADCs), which can achieve accuracies of about 1 part in 100,000 [3], ADC speeds as well as DSP chips are now running reliably at very high speeds.

In the fast Fourier transform (FFT) method, the digital signal samples are Fourier transformed with an FFT [2]. If the signal is sinusoidal, the initial phase is estimated as that value of the phase where the Fourier transform is maximized [4]. The frequency of the signal is estimated as that value of frequency where the Fourier transform is maximized. Once measurements of the frequency f and initial phase ϕ_0 have been obtained, the phase at any point in time can be calculated according to

$$\phi = 2\pi ft + \phi_0 \tag{24.17}$$

One important practical issue in the measurement of the frequency and initial phase with an FFT arises because the FFT yields only *samples* of the Fourier transform; that is, it does not yield a continuous Fourier transform. It is quite possible that the true maximum of the Fourier transform will fall *between* samples of the FFT. For accurate measurement of frequency and initial phase, then, it is necessary to *interpolate* between the FFT samples. An efficient algorithm to do this is described in [5].

The FFT method is particularly appealing where there is significant noise present, as it is effective down to quite low signal-to-noise ratios (SNRs). It provides an optimal estimate of the frequency and initial phase, provided that the background noise is white and Gaussian and that no harmonic components are present [4]. If harmonics are present, the estimate of the phase is commonly taken as the phase at the FFT peak corresponding to the fundamental; this is not an optimal estimate, but serves well in many applications. An optimal estimate in the case when harmonics are present can be obtained, if necessary, with the algorithms in [6–8]. DSP chips such as the Texas Instruments TMS320C3x family, the Analog Devices ADSP21020 chip, or the Motorola MC5630x series can be used to implement the real-time FFTs.

If long word-lengths are used (say 32 bit) to perform the arithmetic for the FFTs, then determination of the phase from the samples of the FFT is virtually error-free. The only significant inaccuracy incurred in determining the phase is due to the ADC errors. Moreover, the error due to the digitization will typically be spread out over a large bandwidth, only a small amount of which will be "seen" in the phase measurement. With a high-quality ADC, accuracies of less than 0.001° are possible.

24.3.10 Phase-Locked Loops

If the frequency of a signal changes significantly over the period of the measurement, the FFT method previously described will provide inaccurate results. If the signal's frequency does change substantially during measurement, one means to estimate the phase of the signal is to use a phase-locked loop (PLL). In this case, the signal, $s(t) = A\sin(\omega t + \phi(t))$, can be thought of as a constant frequency component, $A\sin(\omega t)$, which is phase modulated by a time-varying phase component, $\phi(t)$. The problem then reduces largely to one of demodulating a phase-modulated (PM) signal. A PLL can be used to form an estimate of the "phase-modulating" component, $\phi(t)$, and the overall phase of the signal, ϕ_{oa}, can be estimated according to

$$\phi_{oa}(t) = \omega t + \hat{\phi}(t) \tag{24.18}$$

Either analog or digital PLLs (DPLLs) can be used, although higher accuracy is attainable with DPLLs. Analog PLLs for demodulating a frequency-modulated (FM) signal are discussed in Chapter 84 and in [12]. The DPLL was developed as an extension of the conventional analog PLL and is therefore similar in structure to its analog counterpart. The DPLL is discussed in [9,10]. The equation to demodulate the digital modulated signal with a first-order DPLL is a simple recursive equation [9].

A block diagram of the DPLL for demodulating a PM signal is shown in Figure 24.12. In this diagram, n represents the discrete-time equivalent of continuous time t. It can be seen that there are

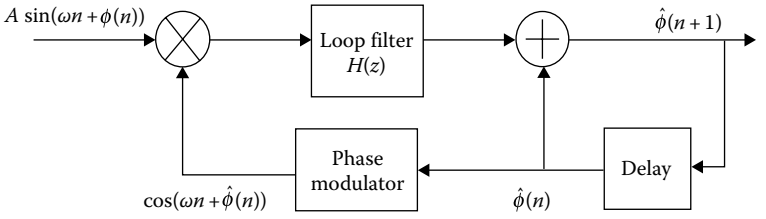


FIGURE 24.12 Block diagram of a DPLL to implement phase demodulation.

strong similarities between Figure 24.12 and the analog PLL-based FM demodulator in Chapter 81. Both have phase comparators (implemented by the multiplier in Figure 24.12), both have loop filters, and both have modulators (either PM or FM) in the feedback loop.

The DPLL can easily be extended to measure the phase of a signal consisting not just of a fundamental component but also of harmonically related components. Details are provided in [11]. The DPLL is near optimal for phase estimation in white Gaussian background noise down to a signal power-to-noise power ratio of about 8 dB [10].

The DPLL will function effectively whether the phase is constant or time varying. Unlike the FFT, the DPLL is a recursive algorithm, with the feedback involved in the recursion creating a vulnerability to quantization errors. However, with proper precautions and long word-lengths, the DPLL will introduce minimal processing error. The main error would then arise from the inaccuracy of the ADC. With appropriate conditioning, one could expect the DPLL to provide accuracies approaching 0.001°.

24.4 Phase-Sensitive Demodulation

It is frequently necessary to track the phase of a carrier that "jitters" in some uncertain manner. This tracking of the carrier phase is necessary for synchronous demodulation schemes, where the phase of the demodulating signal must be made equal to the phase of the carrier. This is briefly re-explained here. Consider, for example, double-sideband (DSB) amplitude modulation. In DSB, the modulated signal is given by $f_s(t) = A[k+\mu m(t)]\cos(\omega_c t)$, where $m(t)$ is the message signal, A is the amplitude of the unmodulated carrier, μ is the modulation index, k is the proportion of modulating signal present in the modulated signal, and $\cos(\omega_c t)$ is the carrier. Demodulation is typically carried out by multiplying $f_s(t)$ by the carrier and then low-pass filtering so that the demodulated signal is given by

$$f_d(t) = \frac{A_c[k + \mu m(t)]}{2} \tag{24.19}$$

However, if because of carrier uncertainty, one multiplies the modulated signal by $\cos(\omega_c t + \phi)$, then the demodulated signal is given by

$$f_d(t) = \frac{A_c[k + \mu m(t)]\cos(\phi)}{2} \tag{24.20}$$

It can be seen from Equation 24.20 that the error in the carrier phase can affect both the amplitude and the sign of the demodulated signal. The phase errors can thus yield substantial errors in system output. The following sections outline important techniques used for tracking the phase of carriers and thus reducing phase errors.

24.4.1 Phase-Locked Loop for Carrier Phase Tracking

The PLL is well known as a means for demodulating FM signals. It is also frequently used for tracking the phase of a carrier in noise, so that a copy of the carrier with correct phase is available for demodulation. This tracking is simple enough if a (noisy) copy of the carrier is directly available; either a digital or analog PLL can be used. In either case, the input can be assumed to have the form $A\sin(\omega t + \phi(t))$, where $\phi(t)$ is the carrier phase. The PLL consists of a multiplier (effectively a phase comparator), a phase modulator, and a loop filter. The design of the loop filter is critical if noise is to be optimally removed. In proper operation, the PLL output will track the phase of the incoming signal (i.e., of the carrier). If a copy of the carrier is not available but needs to be inferred from the modulated signal, the demodulation task is more difficult. DPLLs using a DSP chip can be particularly helpful in this case; the carrier

can be adaptively estimated using intelligent algorithms, with convergence to the "correct" signal being registered when certain desired features of the demodulated signal are observed.

The PLL is quite versatile. It can function in relatively low-noise environments (typically down to about 8 dB SNR). It can be implemented digitally. It can also cope with substantial carrier frequency variations by increasing the order of the loop filter [12] (this is often necessary, for example, in satellite communications because of the Doppler effect). At very low SNR, however, the PLL fails badly. Recent developments in digital signal processing have seen the development of an alternative based on hidden Markov models (HMMs), which will function down to about −5 dB SNR [13]. The HMM method is discussed in the next section.

24.4.2 Hidden Markov Model-Based Carrier Phase Tracker

In the HMM method, the problem of estimating the phase and frequency of a noisy waveform is couched as a "state estimation" problem. The phase of the signal at any point in time can go from 0 to 360°. The 0 to 360° value range is divided into a finite number of intervals or "states," so that the phase at any time occupies a particular (though unknown) state. Similarly, the angular frequency normalized by the sampling frequency at any time in a digital system must be between $-\pi$ and $+\pi$. This value range is also divided into a number of states, so that the frequency at any time has a (hidden or unknown) state associated with it. The frequency is assumed to be a first-order Markov process, and probabilities are assigned to the possibility of the frequency changing from one state to another for successive values of time, that is, frequency transition probabilities are assigned. Large frequency changes are assigned low probabilities, while small changes are assigned high probabilities. The problem of estimating the true phase and frequency states underlying the noisy signal then reduces to one of estimating which states the phase and frequency occupy as time evolves, given the observed noisy signal and the transition probabilities. Computationally efficient optimal algorithms have been developed to estimate these "optimal state sequences" for both the phase and frequency. Details are provided in [13].

24.5 Power Factor

Of particular interest in many applications is the phase angle between the current and voltage of a system. This angle is important because it is a measure of the power that is dissipated in the system. The following paragraphs discuss this angle, its cosine (the system power factor), and its measurement.

In a linear electric circuit that is fed by a current of peak amplitude, I_M, with an angular frequency of ω, the current will have the form, $I_M \cos(\omega t)$. The system voltage will be given by $V_M \cos(\omega t + \phi)$, where V_M is the peak voltage and ϕ is the phase difference between the current and voltage. Then the average power dissipated in the circuit will be given by

$$P_{av} = \frac{1}{2} V_M I_M \cos(\phi) = V_{rms} I_{rms} \cos(\phi) \tag{24.21}$$

where V_{rms} and I_{rms} are the root-mean-square (rms) values of the voltage and current, respectively. The term $\cos(\phi)$ is known as the *power factor*. It may alternatively be expressed as the ratio of real average power to the product of the rms values of voltage and current, respectively:

$$PF = \frac{P_{av}}{V_{rms} I_{rms}} \tag{24.22}$$

The previous expression is, in fact, a general definition of power factor for any current and voltage waveforms. For the special case of sinusoidal voltage and currents, *PF* reduces to $\cos(\phi)$.

There are a number of ways to measure the power factor. One way is to use a wattmeter to measure the real average power and a voltmeter and an ammeter to measure the rms voltage and current, respectively. The power factor is then determined according to Equation 24.22. This is probably the most effective way when the currents and/or voltages are nonsinusoidal. This procedure can easily be implemented with "digital power meters." The power is measured by time-averaging the product of the instantaneous voltage and current, while the rms values are calculated by taking the square root of the time-averaged value of the square of the parameter of interest—current or voltage. Some digital power meters also provide an analysis of the individual harmonics via FFT processing. These meters are accurate and versatile and consequently very popular.

A more direct method is based on the crossed-coil meter, the operation of which was described earlier in this chapter. Note that this meter is a "single-phase meter," which is accurately designed for one frequency only. Errors will occur at other frequencies because of the dependence of the crossed-coil meter method on a constant and known phase angle between the currents in the crossed coils.

With balanced polyphase circuits, it is possible to use a single-phase meter applied to one of the phases. Alternatively, one can use specially designed polyphase meters. In a three-phase meter, for example, one phase is connected to the fixed coil, while the other two phases are connected to the crossed coils on the rotating shaft. The crossed coils are constructed with a 60° angle between them. With four-phase systems, consecutive lines are 90° out of phase. Two of these consecutive lines are connected to the two crossed coils, and the angle between the coils is made equal to 90°.

With unbalanced polyphase circuits amid the presence of harmonics, each of the harmonic components has its own power factor, and so it is likely to be misleading to use a meter that measures a single angle. These methods based on the crossed-coil meter are thus much more limited than their digital counterparts.

24.6 Instrumentation and Components

Table 24.1 lists some ICs and DSP chips that can be used in the various techniques for measuring phase. The list is really only illustrative of what is available. Table 24.2 lists some companies that manufacture these products. An extensive (and indeed rapidly increasing) product range for DSP chip-based products, with details being available from the companies listed in Table 24.2. Table 24.3 lists instruments used for phase measurement. These instruments include CROs, vector voltage meters, vector impedance meters, crossed-coil meters and digital power meters, zero-crossing meters, and phase standards. Again, the table is only representative, as the full range of available instruments is enormous. Addresses of some of the relevant companies are provided in Table 24.4.

TABLE 24.1 ICs Used in Phase Measurement

Function	Designation	Manufacturer
PLL	LM566	National, Motorola, Phillips
PLL	74HC4046	Harris, Motorola
Phase/frequency detector	MC4044P	Motorola
Pair of retriggerable monostables (one-shot)	74HC4538	Motorola, Harris
DSP chip	TMS320C32	Texas Instruments
DSP chip	TMS320C31	Texas Instruments
DSP chip	MC56303	Motorola
DSP chip	ADSP21020	Analog Devices

TABLE 24.2 Companies Making ICs and DSP Chips That Can Be Used for Phase Measurement

Analog Devices, Inc.
One Technology Way
Box 9106
Norwood, MA 02062
Tel: (617) 329-4700

Harris Semiconductor Products Division
P.O. Box 883
Melbourne, FL 37902
Tel: (407) 724-3730

Motorola, Semiconductor Products Sector
3102 N. 56th St.
Phoenix, AZ 85018
Tel: (602) 952-3248

National Semiconductor Corp.
2900 Semiconductor Dr.
P.O. Box 58090
Santa Clara, CA 95052-8090
Tel: (800) 344-4539

Texas Instruments Incorporated
P.O. Box 1443
Houston, Texas 77251-1443
Tel: (800) 344-4539

TABLE 24.3 Instruments for Measuring Phase

Description	Model Number	Manufacturer
CRO	HP54600B	Hewlett-Packard
CRO	HP54602B	Hewlett-Packard
CRO	HP54616	Hewlett-Packard
CRO	TDS220	Tektronix
CRO	TDS510A	Tektronix
Vector signal analyzer	HP89410A	Hewlett-Packard
Vector signal analyzer	HP89440A	Hewlett-Packard
Vector signal analyzer	HP89441A	Hewlett-Packard
Gain/phase impedance meter	HP4193A	Hewlett-Packard
Zero-crossing phase meter	KH6500	Krohn-Hite
Digital power analyzer (with power factor and phase)	Nanovip	Elcontrol
Digital analyzing vector voltmeter	NA2250	North Atlantic Instruments
Digital power analyzer (with power factor and phase, FFT analysis)	3195	Hioki
Crossed-coil meter	246–425G	Crompton Industries
Digital phase standard	5500	Clarke-Hess

TABLE 24.4 Companies Making Instruments for Measuring Phase

Hewlett-Packard Co.
Test and Measurement Sector
P.O. Box 58199
Santa Clara, CA 95052-9943
Tel: (800) 452-4844
www.hp.com

Tektronix Inc. Corporate Offices
26600 SW Parkway
P.O. Box 1000
Wilsonville, OR 97070-1000
Tel: (503) 682-3411, (800) 426-2200
www.tek.com/

Krohn-Hite Corporation
Bodwell St., Avon Industrial Park
Avon, MA
Tel: (508) 580-1660
http://www.krohn-hite.com/

Crompton Instruments
Freebournes Road, Witham
Essex, CM83AH England
Tel: (800) 327-6996
http://www.crompton-instruments.com/

Elcontrol
Via San Lorenzo
1/4 - 40037 Sasso Marconi
Bologna, Italy
Tel: +39 051.6782006
http://www.elcontrol-energy.net/

Hioki
81 Koizumi
Veda, Nagano
386-11 Japan
Tel: +81-268-28-0555
http://www.hioki.com/

Clarke-Hess Comm. Research Corporation
220 W. 19 Street
New York
Tel: (631) 698-3350
http://www.clarke-hess.com/

North Atlantic Industries
110 Wilbur Place, Bohemia, NY 11716
Tel: (631) 567-1100
http://www.naii.com/

References

1. D. Gabor, The theory of communication, *J. Inst. Elec. Eng.,* 93(III), 429–457, 1946.
2. A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.
3. K. C. Pohlmann (ed.), *Advanced Digital Audio*. Carmel, IN: Howard Sams and Co., 1991.
4. D. Rife and R. Boorstyn, Single tone parameter estimation from discrete-time observations, *IEEE Trans. Inf. Theory*, 20, 591–598, 1974.
5. T. Abotzoglou, A fast maximum likelihood algorithm for estimating the frequency of a sinusoid based on Newton's algorithm, *IEEE Trans. Acoust., Speech Signal Process.,* 33, 77–89, 1985.
6. D. McMahon and R. Barrett, ML estimation of the fundamental frequency of a harmonic series, *Proceedings of ISSPA 87*, Brisbane, Queensland, Australia, 1987, pp. 333–336.
7. A. Nehorai and B. Porat, Adaptive comb filtering for harmonic signal enhancement, *IEEE Trans. Acoust., Speech Signal Process.,* 34, 1124–1138, 1986.
8. L. White, An iterative method for exact maximum likelihood estimation of the parameters of a harmonic series, *IEEE Trans. Automat. Control,* 38, 367–370, 1993.
9. C. Kelly and S. Gupta, Discrete-time demodulation of continuous time signals, *IEEE Trans. Inf. Theory,* 18, 488–493, 1972.
10. B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice Hall, 1979.
11. P. Parker and B. Anderson, Frequency tracking of periodic noisy signals, *Signal Processing*, 20(2), 127–152, 1990.
12. R. E. Best, *Phase-Locked Loops: Theory, Design and Applications*. 2nd edn. New York: McGraw-Hill, 1993.
13. L. White, Estimation of the instantaneous frequency of a noisy signal, in B. Boashash (ed.), *Time-Frequency Signal Analysis, Methods and Applications*, Melbourne, Victoria, Australia: Longman-Cheshire, 1992.

Further Readings

Helfrick D. and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques*. Englewood Cliffs, NJ: Prentice Hall, 1990.

Lenk J. D., *Handbook of Oscilloscopes: Theory and Application*. Englewood Cliffs, NJ: Prentice Hall, 1982.

McGraw-Hill, *Encyclopedia of Science and Technology*, 8th edn. New York: McGraw-Hill, 1997.

Taub H. and D. L. Schilling, *Principles of Communication Systems*, 2nd edn. New York: McGraw-Hill, 1986.

25

Energy Measurement

Arnaldo Brandolini
Politecnico di Milano

Alessandro Gandelli
Politecnico di Milano

Energy is one of the most important physical quantities in any branch of science and engineering and especially in electrical engineering. Energy exchange processes lead to the study of electric networks from the physical point of view and allow an in-depth knowledge of power transfer within the electrical world and between electric and other forms of energy.

The definitions of energy and power represent the starting point for any successive study:

1. Energy is the amount of work that the system is capable of doing.
2. Power is the time rate of doing work.

Energy can be mathematically defined as the definite integral of the power over a given time interval Δt.

The power p available in a two-terminal section of an electric circuit is given by the product of the voltage v across the terminals and the current i flowing through the section itself ($p = vi$). The electric energy (E) flowing through the same section is defined by the integral of the power over the observation interval:

$$E(\Delta t) = \int_{t_0}^{\Delta t + t_0} p\,dt \tag{25.1}$$

For this reason, energy measurement is a dynamic measurement, which means it varies with time. The energy measurement unit is the joule (J), but for the electric energy, the watt-hour (W-h) is most common.

The electrostatic energy is defined as the product of the electric charge and the difference of electric potential.

Electricity is generated from different forms of energy (thermal, hydraulic, nuclear, chemical, etc.); after electric transfer and distribution processes, it is converted to other forms of energy. The main feature of electric energy is the simplicity by which one can transfer it over long distances, control the distribution, and measure energy consumption.

25.1 DC Energy Measurements

The simplest way to perform this measurement is to measure voltage and current and then compute the product:

$$E = VI\Delta t \tag{25.2}$$

where Δt is the observation interval measured by means of a chronometer or a time counter.

Note that direct current (dc) systems are limited to a restricted number of applications in power systems, for example, electric traction, electric drives, electrochemical power plants, and HVDC transmission system in limited operating conditions. All these cases, nevertheless, allow energy measurement on either the dc or alternating current (ac) side of the network.

The dc energy measurement has been performed in the past by means of different methodologies and instruments such as electrodynamic measurement devices (Electrodynamics dc Energy Meter) operating as an integrating wattmeter (Figure 25.1). This measuring instrument is built using a small dc motor without iron, whose magnetic field is generated by the line current flowing through a coil arranged as the fixed part of the system. The rotor is connected in series with an additional resistor and is powered by the line voltage (V). Because of the lack of the iron in the magnetic circuit, the rotor magnetic flux ϕ is strictly proportional to the current I.

The rotor current (derived from the line voltage) is

$$I_V = \frac{(V - E)}{R} \tag{25.3}$$

where

$E = k_1 \Gamma \phi$ is the emf induced by the angular speed Γ

R is the total resistance of the voltage circuit

It is possible to make the emf E negligible because of low angular speed Γ, limited amplitude of the flux ϕ, and a significant resistance R. In this way, Equation 25.3 becomes

$$I_V \approx \frac{V}{R} \tag{25.4}$$

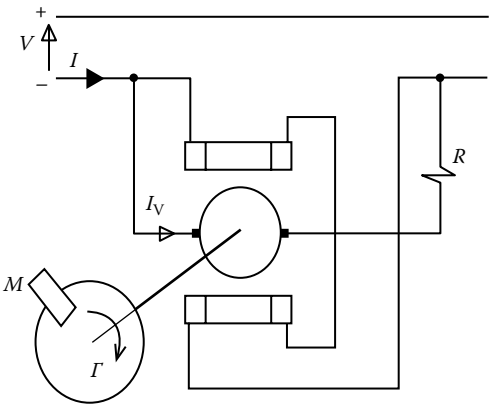


FIGURE 25.1 The electrodynamic dc energy meter. *M*, permanent magnet.

The torque C_m provided by the motor can be written as

$$C_m = k_2 \phi I_v \approx k_3 I \frac{V}{R} = k_4 P \tag{25.5}$$

C_m is therefore approximately proportional to the power P flowing through the line. It is necessary, however, to remember that this torque could create a high angular speed to the rotor, because of constantly incrementing speed. In order to maintain dynamic equilibrium, a simple aluminum disk mounted on the rotor axis and placed in a constant magnetic field provided by a permanent magnet M is added to the dc motor system. In this way, the induced currents in the disk introduce a damped torque proportional to the angular speed Γ, so, at equilibrium, there is a linear dependence of Γ on the power P. Thus,

$$E = \int_{\Delta t} P dt = k_5 \int_{\Delta t} \Gamma dt \tag{25.6}$$

A mechanical counter transfers the rotating motion into a digital representation of the total energy consumed during a specific time interval Δt in the power system.

25.2 AC Induction Energy Meters

The most traditional and widely used ac energy meter is the *induction meter*. This device is built by means of three electric circuits, magnetically coupled, two of them fixed and one rotating around the mechanical axis of the system. Figure 25.2 shows the two fixed circuits, (1) and (2), which are the voltage and the current coils. The third circuit is the rotating disk (3), generally made of aluminum, mounted on a rigid axis (4) transmitting the disk rotation to a mechanical counter (6), which provides the energy display.

The fixed circuits (1) and (2) provide magnetic fluxes interacting with the rotating disk. Fixed circuits (1) and (2) form a C shape and the disk is placed in their iron gaps. Another similar structure, arranged using a permanent magnet (5), is placed over the disk as well. The magnetic fluxes generated by the voltage and current circuits are at the same frequency and are sinusoidal. They induce currents in the

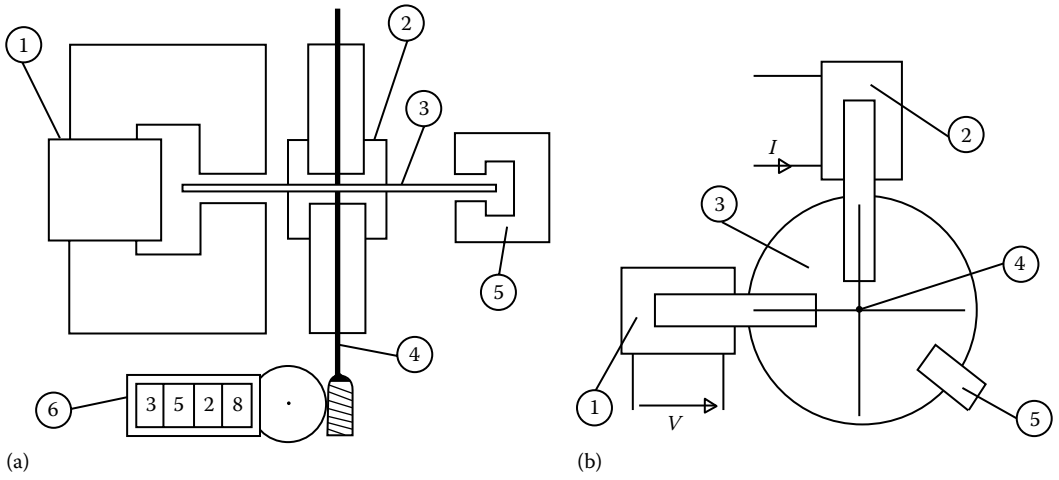


FIGURE 25.2 (a) Side view of an ac induction energy meter: (1) voltage coil and magnetic circuit; (2) current coil and magnetic circuit; (3) aluminum rotating disk; (4) disk axis; (5) permanent magnet; (6) mechanical display. (b) Top view of an ac induction energy meter; (1) voltage coil and magnetic circuit; (2) current coil and magnetic circuit; (3) aluminum rotating disk; (4) disk axis; (5) permanent magnet.

rotating disk that, by means of a cross-interaction with the two generating fluxes, provide mechanical torque acting on the disk. The torque is given by

$$C_m = KVI\sin(\alpha) \tag{25.7}$$

where

C_m is the mechanical torque
K is the system constant
V is the rms of the value of the applied voltage
I is the rms of the value of the applied current
α is the phase angle between the fluxes generated by V and I

The acting torque causes the disk to rotate around its axis. This rotation reaches a dynamic equilibrium by balancing the torque C_m of the voltage and current coils and the reacting torque generated by permanent magnet. The resulting angular speed, Γ, is therefore proportional to the flowing power if

- The angular speed Γ of the disk is much smaller than the voltage and current frequency ω.
- The phase difference between the voltage and current fluxes is equal to $\alpha = \pi - \varphi$, where φ is the phase difference between the voltage and current signals.

The angular speed of the rotating disk can be written as

$$\Gamma = \frac{1}{k}\omega\frac{R_3}{Z_3^2}(M_1 I)\left(\frac{M_2 V}{Z_2}\right)\cos(\phi) = KP \tag{25.8}$$

where

Γ is the angular speed of the rotating circuit (conductor disk), in rad s^{-1}
K is the instrument constant, in rad s^{-1} W^{-1}
P is the mean power in the circuit, in W
$1/k$ is the constant, in Ω V^{-2} s^{-2}
ω is the voltage and current frequency, in rad s^{-1}
R_3 is the equivalent resistance of the rotating disk, relative to the induced current fields, in Ω
Z_3 is the equivalent impedance of the rotating disk, relative to the induced current fields, in Ω
(M_2V/Z_2) is the rms value of the common flux related to the circuits *n*. 1 and 3, in Wb
(M_1I) is the rms value of the common flux related to the circuits *n*. 2 and 3, in Wb
Z_2 is the impedance of the voltage circuit (*n*. 1), in Ω
V is the rms value of the applied voltage, in V
I is the rms value of the applied current, in A
ϕ is the phase difference between current and voltage signals

The integral of Γ over a defined period Δt is proportional (with enough accuracy) to the energy flowing in the power circuit. Thus, it is true that the instrument constant K is strictly related (but not proportional) to the signal frequency ω.

25.3 Static Energy Meters

The development of electronic multipliers led to their use in energy meters that directly multiply voltage by current. In their first version, electronic multipliers used analog components (operational amplifiers, resistors, capacitors, etc.), while recent devices use digital components and programmable logic systems. Voltage and current signals are processed to obtain a signal proportional to the real power flowing into the line. The result is integrated over the observation time in order to calculate the *measured* energy. The devices based on these components are completely static (i.e., they do not have any moving parts).

Moreover, because these electronic components have a frequency range from dc to high frequencies, instruments based on them can be applied to dc, ac, or distorted power systems (some care must be taken in order to provide a correct sampling of signals in all-digital systems).

There are many different prototypes in this class of energy meters. The first realizations were based on analog multipliers and, even if they were not able to replace the traditional induction energy meters, they represented a good solution for all those applications where an increased accuracy was required (up to 0.2%). Now, more sophisticated digital instruments are under design and development, based on dedicated structures mainly implementing digital signal processors (DSPs) as powerful tools for numerical computation and sigma-delta analog-to-digital converters (ADCs) in order to optimize the conversion process.

Many of these instruments can be analyzed by means of the following functional descriptions.

25.3.1 Electronic Energy Meter

Figure 25.3 shows the block diagram of an electronic energy meter. The main feature of this type of instrument is the presence of voltage inputs on both voltage and current channels, because the electronic circuitry accepts only voltage signals. It has negligible current consumption from the system under measurement, due to high input impedance. Moreover, the maximum amplitude level of the input signal must be limited to around 5–15 V. For this reason, the conditioning apparatus must guarantee the correct current-to-voltage transformation and the proper voltage reduction. This type of instrument can work at dc (which omits voltage and current transformers) or ac power systems and can also measure energy from distorted signals.

25.3.2 Conditioning System for DC Electronic Energy Meters

The basic blocks of the conditioning system for a dc energy meter are formed from a voltage divider for the voltage input and a shunt for the current input. After these passive components, two preamplifiers are usually introduced before the processing system. The current preamplifier is very important because

1. The voltage output level of the current shunt is very low, even at full scale ($\leq$1 V).
2. Many times, the current input has to support overloaded signals; the presence of a variable gain amplifier allows acceptable working conditions for the system.
3. It can be used to implement an active filter before signal processing.

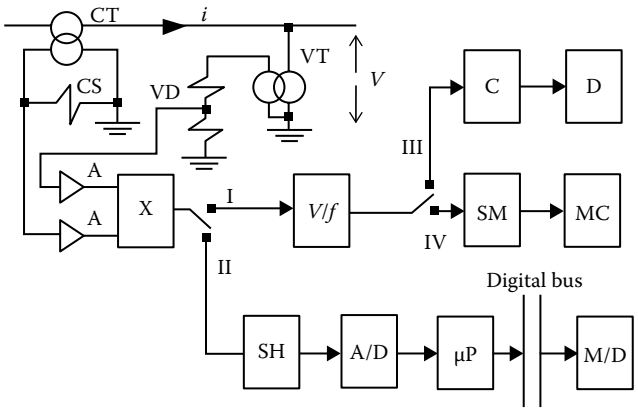


FIGURE 25.3 Electronic energy meter. Mechanical display option (I–IV). Electronic display option (I–III). Electronic display option and digital processing of the power signal (II). CT, current transformer; VT, voltage transformer; CS, current shunt; VD, voltage divider; A, analog signal processing block; X, multiplier; *V*/*f*, voltage-to-frequency converter; SM, step motor; MC, mechanical counter; C, electronic counter; D, display; SH, sample and hold; A/D, analog-to-digital converter; μP, microprocessor (CPU); M/D, memory and display.

25.3.3 Voltage and Current Adapters for AC Electronic Energy Meters

The most common devices to process ac signals for static energy meters are the traditional voltage and current transformers. They must be made with proper components to achieve the right amplitude of the voltage inputs (by nonreactive shunts for the current transformers and nonreactive voltage dividers for the voltage transformers). After the transformers, and related devices, a second block, based on electronic amplifiers, provides the final analog processing of the input signals, as for the dc conditioning systems. It is useful to introduce this second processing element because analog filters are generally required when the input signals need to be digitally processed.

25.3.4 Electronic-Analog Energy Meters with Digital Output

These instruments provide the product of the two input signals (both voltages) through an analog multiplier that evaluates a voltage output proportional to the power of the input signals. This output can be followed by a filtering block.

The output signal is proportional to the instantaneous electric power flowing through the line. To calculate the energy, it is now necessary to complete the process by integrating over the observation time. This last procedure can be performed in two different ways.

23.3.4.1 First Procedure

The power signal at the output of the analog multiplier is applied to the input of a voltage frequency converter. Thus, the power information is converted from a voltage level to the frequency pulse sequence, for which the counting process performs the integration of the power in the observation interval, that is, the measurement of energy.

The final measurement can be performed by means of an electronic counter with digital display or using a dc step motor incrementing the rotor angular position every pulse by a fixed angular increment. The rotor position is shown by a mechanical counter (similar to the system mounted on the induction energy meters) indicating the total number of complete rotations performed by the system, proportional to the energy of the system under measurement. This second arrangement is normally adopted because it allows a permanent record of the energy information, which is not subject to possible lack of electric energy as in the first case.

25.3.4.2 Second Procedure

This arrangement is based on an ADC connected to the output of the analog multiplier. The sampling process is driven by an internal clock. Thus, the ADC provides uniform sampling over the signal period and, under the condition imposed by the sampling theorem, the sum of the samples is proportional to the integral of the power signal, that is, to the energy during the observation interval.

The calculation is performed by means of a dedicated CPU and then the results are sent to the digital memory to be stored and displayed. They can also be used to manage any other automatic processes based on the energy measurement. For this purpose, data are available on a data bus (serial or parallel) connecting the measuring system with other devices.

The sampling process is performed by a sample and hold circuit.

25.3.5 All-Digital Energy Meters

The most advanced solution for energy measurement can be found in all-digital meters (Figure 25.4), where both the voltage and current signals are sampled before any other processing. Thus, the data bus presents the sampled waveforms in digital form, giving the opportunity to perform a wide choice of digital signal processing on the power and energy information. Both sampling devices are driven by a CPU, providing synchronized sampling signals.

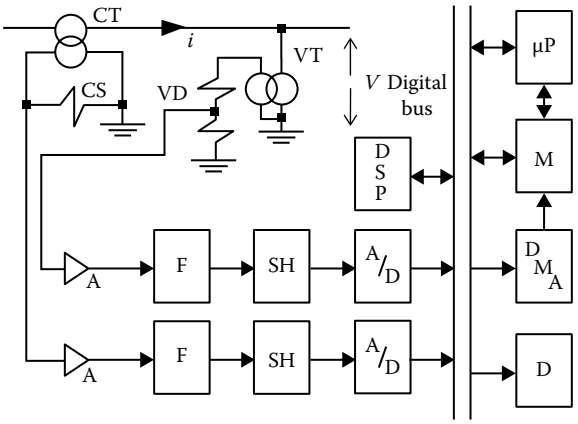


FIGURE 25.4 All-digital energy meter. CT, current transformer; VT, voltage transformer; CS, current shunt; VD, voltage divider; A, analog signal processing block; F, analog electronic filter; SH, sample and hold; A/D, analog-to-digital converter; μP, microprocessor (CPU); M, memory; DSP, digital signal processor; DMA, direct memory access circuit; D, display.

Sometimes, the system is equipped with a DSP capable of providing hardware resources to implement real-time evaluation of complex parameters (i.e., signal transforms) of the signal and energy measurement. Dedicated hardware and software performing instrument testing are also integrated into the meter to complete the device with the most advanced features.

Filters able to meet the sampling theorem requirements, programmable gain amplifiers, and sample and hold circuits generally precede the ADCs.

Data management is arranged in two possible ways: sending the sampled data directly to the processing system for calculations or accessing the memory using DMA procedures, so the data string for a specific time period is first stored and then used for computation of energy and related parameter values. Final results of this computation are then available on the system bus to be sent to the other system resources or to be displayed.

25.4 Accuracy of Energy Meters

Accuracy of energy meters is defined by means of relative parameters (in percent) obtained from a testing process by powering the instrument with a constant (nominal) voltage signal and a variable current signal (5%, 10%, 20%, 50%, 100%, and 120% of the nominal value). The testing procedures are performed by comparing the meter under test with a standard meter (Figure 25.5) or using equivalent methods.

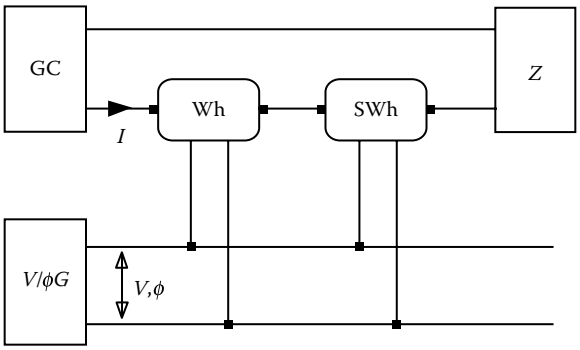


FIGURE 25.5 Testing circuit arrangement to compare an industrial meter (W-h) with a standard energy meter (SW-h). CG, variable-amplitude current generator; $V/\phi G$, variable-amplitude and phase voltage generator; Z, load impedance.

The accuracy of commercial electromechanical (induction) energy meters is generally around 2%. Energy meters with accuracies of 1% have also been built. Electronic energy meters have a better accuracy, generally between 0.5% and 0.2%.

Further Readings

Bosnjakovic P. and B. Djokic, Reactive energy measurement using frequency controlled analog-to-pulse-rate converter, *Arch. Elektrotech.*, 75, 131–135, 1992.

Djokic B., P. Bosnjakovic, and M. Begovic, New method for reactive power and energy measurement, *IEEE Trans. Instrum. Meas.*, 41, 280–285, 1992.

Drisdale C. V. and A. C. Jolley, *Electrical Measuring Instruments*, 2nd edn., New York: John Wiley & Sons, 1952.

Finkelstein L. and K. T. V. Grattan (eds.), *Concise Encyclopedia of Measurement & Instrumentation*, New York: Pergamon Press, 1994.

Harris F. K., *Electrical Measurements*, New York: John Wiley & Sons, 1952.

Kusui S. and T. Nagai, A single-phase 3-wire watt-to-pulse frequency-converter using simple PWM and its accuracy analysis, *IEEE Trans. Instrum. Meas.*, 43, 770–774, 1994.

Montano J. C., A. Lopez, M. Castilla, and J. Gutierrez, DSP-based algorithm for electric-power measurement, *IEE Proc. A Sci. Meas. Technol.*, 140, 485–490, 1993.

Schnell L. (ed.), *Technology of Electrical Measurements,* New York: John Wiley & Sons, 1993.

26

Electrical Conductivity and Resistivity

Michael B. Heaney
Huladyne Research and Consulting

Electrical resistivity is a key physical property of all materials. It is often necessary to accurately measure the resistivity of a given material. The electrical resistivity of different materials at room temperature can vary by over 20 orders of magnitude. No single technique or instrument can measure resistivities over this wide range. This chapter describes a number of different experimental techniques and instruments for measuring resistivities. The emphasis is on explaining how to make practical measurements and avoid common experimental errors. More theoretical and detailed discussions can be found in the sources listed at the end of this chapter.

26.1 Basic Concepts

The electrical resistivity of a material is a number describing how much that material resists the flow of electricity. Resistivity is measured in units of ohmmeters (Ω m). If electricity can flow easily through a material, that material has low resistivity. If electricity has great difficulty flowing through a material, that material has high resistivity. The electrical wires in overhead power lines and buildings are made of copper or aluminum. This is because copper and aluminum are materials with very low resistivities (about 20 nΩ m), allowing electrical power to flow very easily. If these wires were made of high-resistivity material like some types of plastic (which can have resistivities about 1 EΩ m [1×10^{18} Ω m]), very little electrical power would flow.

Electrical resistivity is represented by the Greek letter ρ. Electrical conductivity is represented by the Greek letter σ and is defined as the inverse of the resistivity. This means a high resistivity is the same as a low conductivity and a low resistivity is the same as a high conductivity:

$$\sigma \equiv \frac{1}{\rho} \tag{26.1}$$

This chapter will discuss everything in terms of resistivity, with the understanding that conductivity can be obtained by taking the inverse of resistivity. The electrical resistivity of a material is an intrinsic physical property, independent of the particular size or shape of the sample. This means a thin copper wire in a computer has the same resistivity as the Statue of Liberty, which is also made of copper.

26.2 Simple Model and Theory

Figure 26.1 shows a simple microscopic model of electricity flowing through a material [1]. While this model is oversimplified and incorrect in several ways, it is still a very useful conceptual model for understanding resistivity and making rough estimates of some physical properties. A more correct understanding of the electrical resistivity of materials requires a thorough understanding of quantum mechanics [2].

On a microscopic level, electricity is simply the movement of electrons through a material. The smaller white circle in Figure 26.1 represents one electron flowing through the material. For ease of explanation, only one electron is shown. There are usually many electrons flowing through the material simultaneously. The electron tends to move from the left side of the material to the right side because an external force (represented by the large minus and plus signs) acts on it. This external force could be due to the voltage produced by an electrical power plant or a battery connected to the material. As the electron moves through the material, it collides with the "stationary" atoms of the material, represented by the larger black circles. These collisions tend to slow down the electron. This is analogous to a pinball machine. The electron is like the metal ball rolling from the top to the bottom of a pinball machine, pulled by the force of gravity. The metal ball occasionally hits the pins and slows down. Just like in different pinball machines, the number of collisions the electron has can be very different in different materials. A material that produces many collisions is a high-resistivity material. A material that produces few collisions is a low-resistivity material.

The resistivity of a material can vary greatly at different temperatures. The resistivity of metals usually increases as temperature increases, while the resistivity of semiconductors usually decreases as temperature increases. The resistivity of a material can also depend on the applied magnetic field.

The discussion thus far has assumed that the material being measured is homogeneous and isotropic. Homogeneous means the material properties are the same everywhere in the sample. Isotropic means

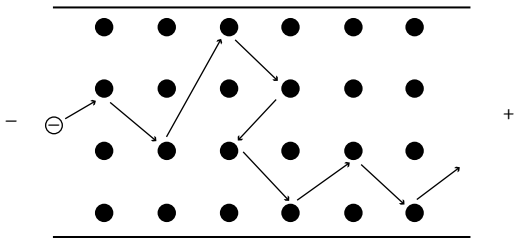


FIGURE 26.1 Simple model of electricity flowing through a material under an applied voltage. The white circle is an electron moving from left to right through the material. The black circles represent the stationary atoms of the material. Collisions between the electron and the atoms slow down the electron, causing electrical resistivity.

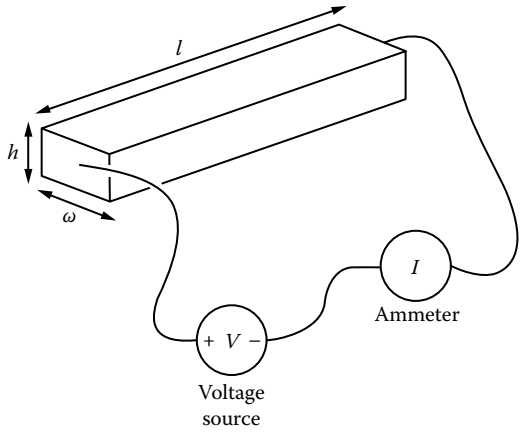


FIGURE 26.2 A two-point technique for measuring the resistivity of a bar of material. The voltage source applies a voltage across the bar, and the ammeter measures the current flowing through the bar.

the material properties are the same in all directions. This is not always a valid assumption. A more exact definition of resistivity is the proportionality coefficient ρ relating a local applied electric field to the resultant current density:

$$E \equiv \rho J \tag{26.2}$$

where

E is the electric field (V m^{-1})
J is the current density (A m^{-2})
ρ is a proportionality coefficient (Ω m)

Equation 26.2 is one form of Ohm's law. Note that E and J are vectors and ρ is, in general, a tensor. This implies that the current does not necessarily flow in the direction of the applied electric field. In this chapter, isotropic and homogeneous materials are assumed, so ρ is a scalar (a single number).

Now consider the bar-shaped sample shown in Figure 26.2. The electric field E is given by the voltage V divided by the distance l over which the voltage is applied:

$$E \equiv \frac{V}{l} \tag{26.3}$$

The current density J is given by the current I, divided by the cross-sectional area A through which the current flows:

$$J \equiv \frac{I}{A} \tag{26.4}$$

where the area A in Figure 26.2 is equal to the width w times the height h. Combining Equations 26.2 through 26.4 and rearranging gives

$$V = \frac{I\rho l}{A} \tag{26.5}$$

Now define a new quantity called "resistance" R with the definition:

$$R \equiv \frac{\rho l}{A} \tag{26.6}$$

Combining Equations 26.5 and 26.6 then gives

$$I = \frac{V}{R} \tag{26.7}$$

where

I is the current in amps (A) flowing through the sample
V is the voltage in volts (V) applied across the sample
R is the resistance in ohms (Ω) of the sample

Equation 26.7 is another form of Ohm's law.

Note that the resistance R can depend on the size and shape of the sample, while ρ is independent of the size or shape of the sample. For example, if the length l of the sample bar is doubled, the resistance will double but the resistivity will remain constant.

The quantitative relationship between the resistivity ρ and the simple microscopic model shown in Figure 26.1 is given by

$$\rho = \frac{m}{ne^2\tau} \tag{26.8a}$$

where

m is the mass of an electron
n is the number of electrons per unit volume carrying current in the material
e is the electric charge on an electron
τ is the average time between collisions of an electron with the stationary atoms of the material

If there were more electrons per unit volume, they could carry more current through the material. This would lower the resistivity. If the electric charge on the electrons were greater, then the applied voltage would pull harder on the electrons, speeding them up. This would lower the resistivity. If the average time between collisions with the stationary atoms were longer, then the electrons could get through the material quicker. This would lower the resistivity. If electrons could be made more massive, they would move slower and take longer to get through the material. This would increase the resistivity.

26.3 Experimental Techniques for Measuring Resistivity

26.3.1 Two-Point Technique

The resistivity of a material can be obtained by measuring the resistance and physical dimensions of a bar of material, as shown in Figure 26.2. In this case, the material is cut into the shape of a rectangular bar of length l, height h, and width w. Copper wires are attached to both ends of the bar. This is called the two-point technique, since wires are attached to the material at two points. A voltage

source applies a voltage V across the bar, causing a current I to flow through the bar. (Alternatively, a current source could force current through the sample bar, while a voltmeter in parallel with the current source measures the voltage induced across the sample bar.) The amount of current I that flows through the bar is measured by the ammeter, which is connected in series with the bar and voltage source. The voltage drop across the ammeter should be negligible. The resistance R of the bar is given by Equation 26.8b:

$$R = \frac{V}{I} \tag{26.8b}$$

where
R is the resistance in Ω
V is the voltage in V
I is the current in A

The physical dimensions can be measured with a ruler, a micrometer, or other appropriate instrument. The two-point resistivity of the material is then

$$\rho \equiv \frac{Rwh}{l} \tag{26.9}$$

where
ρ is the resistivity in Ω m
R is the measured resistance in Ω
w, h, and l are the measured physical dimensions of the sample bar in meters

In practice, measuring resistivity with a two-point technique is often not reliable. There is usually some resistance between the contact wires and the material or in the measuring equipment itself. These additional resistances make the resistivity of the material measure higher than it really is. A second potential problem is modulation of the sample resistivity due to the applied current. This is often a possibility for semiconducting materials. A third problem is that contacts between metal electrodes and a semiconducting sample tend to have other electrical properties that give wrong estimates for the actual sample resistivity. The four-point measurement technique overcomes many of these problems.

26.3.2 Four-Point Technique

Figure 26.3 shows the four-point measurement technique on a bar of material. Four wires are attached to the sample bar as shown. A current source forces a constant current through the ends of the sample bar. A separate ammeter measures the amount of current I passing through the bar. A voltmeter simultaneously measures the voltage V produced across the inner part of the bar. (Alternatively, a voltage source could apply a voltage across the outer contacts, while an ammeter in series with this voltmeter measures the current flowing through the sample bar.)

The four-point resistivity of the material is then

$$\rho = \frac{Vwh}{Il'} \tag{26.10}$$

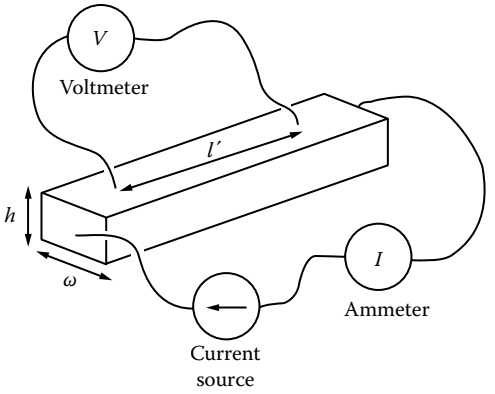


FIGURE 26.3 A four-point technique for measuring the resistivity of a bar of material. The current source forces a current through the bar, which is measured by a separate ammeter. The voltmeter measures the voltage across the middle of the bar.

where

ρ is the resistivity in Ω m
V is the voltage measured by the voltmeter in V
w is the width of the sample bar measured in m
h is the height of the sample bar measured in m
I is the current the ammeter measures flowing through the sample in A
l' is the distance between the two points where the voltmeter wires make contact to the bar, measured in m

Note that the total length l of the bar is not used to calculate the four-point resistivity: the length l' between the two inner contacts is used.

26.3.3 Common Experimental Errors

There are many experimental pitfalls to avoid when making resistivity measurements. The most common sources of error arise from doing a two-point measurement on a material that has any of the contact problems discussed earlier. For this reason, it is advisable to do four-point measurements whenever possible. This section describes experimental techniques to avoid errors in measuring resistivity:

1. The most difficult part of making resistivity measurements is often making good electric contacts to the sample. The general technique for making good electric contacts is to clean the areas of the sample where contacts are to be made with alcohol or an appropriate solvent and then apply the contacts. If this does not work, try scraping the surface with a razor blade where contact is to be made or cutting the sample to expose a fresh surface. Contacts can be made in many ways, such as using alligator clips and silver paint, squeezing a wire against the material, soldering wires to the material, and pressing small pieces of indium or a similar soft metal onto the contact areas. Contacts can age: a good contact can become a bad contact over time. It might be necessary to make fresh contacts to a sample that has aged. There are many complications involved in the electrical properties of contacts. Refer to the sources listed at the end of this chapter for more extensive discussions.
2. The measurement system should be calibrated before measuring any material samples. Calibration procedures are usually described in the equipment manuals.
3. The input resistance (or "impedance") of the voltmeter should be at least 10^5 higher than the resistance of the sample bar. The input impedance is usually listed in the equipment specifications.

Note that some voltmeters and electrometers have sufficiently high impedance between either of the inputs and ground, but not between the two inputs. In this case, it is necessary to use two voltmeters/electrometers (each with one input connected to the ground and the other input connected to the sample bar). Measure the difference between them to obtain the voltage across the sample.

4. The measurement system should be tested before measuring any material samples. First, test "short" with a thick copper wire or sheet in place of the sample. Then, test "open" with nothing in place of the sample. Finally, test with a known, calibrated resistor whose resistance is within an order of magnitude of the sample resistance.
5. The geometry of the sample and electric contacts can be important. Contacts are often made by painting silver paint or applying metal electrodes to the sample. If these contact areas are large or close to each other, this could reduce the accuracy of the resistivity measurement. It is best to make the two voltage contacts in a four-point measurement as small or thin as possible and make the distance between inner electrodes much larger than the sample thickness. This also allows a more accurate estimate of the effective volume of the sample being probed.
6. It is critical that the four contacts to the sample bar in a four-point measurement are completely independent; there should be nothing other than the material of the bar connecting each of the four wires at the bar. For example, when pieces of indium are used to attach wires to a small sample, it is easy to smudge two adjacent indium pieces into one another. Those two contacts are no longer independent and could easily cause errors. Visually inspect the contacts for this condition. If visual inspection is impractical, measure the resistance between the wires going to adjacent contacts. An unusually low resistance might indicate that two contacts are touching each other.
7. The applied voltage or current can cause heating of the material, which can change its resistivity. To avoid this problem, start with very small voltages or currents, and increase until the measured voltages and currents are at least 10 times larger than the random fluctuations of the meters. Then make sure the measured resistance is constant with time: the average resistance should not drift more than 10% in a few minutes.
8. Even if heating of the sample is not a problem, Ohm's law is not always obeyed. Many materials have a resistance that varies as the applied voltage varies, especially at higher voltages. Test for a linear relationship between current and voltage by measuring the resistance at several voltages on both sides of the measurement voltage. Whenever possible, make measurements in the linear (ohmic) region, where resistance is constant as voltage changes.
9. If one or both of the contacts to the voltmeter are bad, the voltmeter may effectively be disconnected from the material. In this situation, the voltmeter might display some random voltage unrelated to the voltage in the material. It might not be obvious that something is wrong, since this random voltage could accidentally appear to be a reasonable value. Check for this by setting the current source to zero amps and seeing if the voltmeter reading drops to zero volts. If it does not, try remaking the two inner contacts.
10. A critical check of a four-point measurement is to reverse the leads and remeasure the resistance. First, turn the current source to zero amps. Without disturbing any of the four contacts at the sample, swap the two sets of wires going to the voltmeter and the current source/ammeter. The two wires that originally plugged into the voltmeter should now plug into one terminal of the current source and one terminal of the ammeter. The two wires that originally plugged into the current source and ammeter should now plug into the voltmeter. Turn the current source on and remeasure the resistance. Note that current is now being forced to flow between the two inner contact points on the sample, while the voltage is being measured between the two outer contacts on the sample. The two measured resistances should be within 10% of each other.
11. The resistivity of some materials can depend on how much light is hitting the material. This is especially a problem with semiconductors. If this is a possibility, try blocking all light from the sample during measurement.

26.3.4 Sheet Resistance Measurements

It is often necessary to measure the resistivities of thin films or sheets of various materials. If the material can be made into the form of a rectangle, then the resistivity can be measured just like the bar samples in Figure 26.2:

$$\rho \equiv \frac{Vwh}{Il} \tag{26.11}$$

where

ρ is the sample resistivity in Ω m
V is the voltage measured by the voltmeter in V
w is the width of the sample measured in m
h is the thickness of the sample measured in m
I is the current the ammeter measures flowing through the sample in A
l is the length of the film measured in m

For the special case of a square film, the width w is equal to the length l and Equation 26.11 becomes

$$\rho \text{ (of square film)} \equiv \frac{Vh}{I} \tag{26.12}$$

The resistivity of a square film of material is called the "sheet resistivity" of the material and is usually represented by the symbol ρ_s. The "sheet resistance" R_s is defined by

$$R_s \equiv R \text{ (of square film)} = \frac{V}{I} \tag{26.13}$$

where

V is the voltage measured by the voltmeter in V
I is the current the ammeter measures flowing through the sample in A

The units for sheet resistance are Ω, but people commonly use the units "Ω per square." The sheet resistance is numerically equal to the measured resistance of a square piece of the material. Note that sheet resistance is independent of the size of the square measured, and it is not necessary to know the film thickness to measure sheet resistance. This makes sheet resistance a useful quantity for comparing different thin films of materials.

It is usually more convenient to measure thin-film samples of arbitrary shape and size. This is usually done by pressing four collinear, equally spaced contacts into a film whose length and width are both much greater than the spacing between contacts. In this situation, the sheet resistance is [3]

$$R_s = 4.532\frac{V}{I} \tag{26.14}$$

where

V is the voltage measured across the two inner contacts
I is the current applied through the two outer contacts

In many practical cases, the size of the thin-film sample will not be much greater than the spacing between the four-point contacts. In other cases, it might be necessary to measure a thin film near a corner or edge. In this situation, use geometric correction factors to accurately estimate the sheet resistance. These correction factors are available for the most commonly encountered sample geometries [3].

26.3.5 Instrumentation for Four-Point Resistivity Measurements

The resistivities of thin films of materials are often measured using commercial four-point probes. These probes generally have four equally spaced, collinear metal points that are pressed against the surface of the film. A current is applied between the outer two points, while the voltage is measured across the inner two points. These probes can also be used to measure the resistivity of bulk samples. Some companies that make probes and systems specifically for four-point resistivity measurements are listed in Table 26.1.

26.3.6 Instrumentation for High-Resistivity Measurements

Many materials such as rocks, plastics, and paper have very high resistivities, up to 1 EΩ m. The techniques described earlier for measuring resistivity are usually not reliable for these materials. In particular, it is often not possible to make a four-point measurement. One problem is that high voltages are needed to get any measurable current flowing through these materials. A second problem is that very long time constants prevent making steady-state measurements. A third problem is that the surfaces of these materials can often have significantly lower resistivity than the bulk, due to defects or other contamination. Measurements using the techniques described earlier then give falsely low values for the bulk resistivity. The best way to measure the resistivity of these materials is to use a specialized commercial instrument. These are designed to separate out the bulk resistivity from the surface resistivity and to minimize the many other problems encountered when measuring very high resistivities. Table 26.2 lists some companies that make high-resistivity measurement systems.

26.3.7 van der Pauw Technique

The four-point measurement technique described earlier has assumed the material sample has the shape of a rectangular thin film or a bar. There is a more general four-point resistivity measurement technique that allows measurements on samples of arbitrary shape, with no need to measure all the physical dimensions of the sample. This is the van der Pauw technique [4]. There are four conditions that must be satisfied to use this technique:

1. The sample must have a flat shape of uniform thickness.
2. The sample must not have any isolated holes.
3. The sample must be homogeneous and isotropic.
4. All four contacts must be located at the edges of the sample.

In addition to these four conditions, the area of contact of any individual contact should be at least an order of magnitude smaller than the area of the entire sample. For small samples, this might not be possible or practical. If sufficiently small contacts are not achievable, it is still possible to do accurate van der Pauw resistivity measurements, using geometric correction factors to account for the finite size of the contacts. See Ref. [5] for further details.

The inset illustration of Figure 26.4 illustrates one possible sample measurement geometry. A more common geometry is to attach four contacts to the four corners of a square-shaped sheet of the material.

The procedure for doing a van der Pauw measurement is as follows:

1. Define a resistance $R_{ij,kl} \equiv V_{kl}/I_{ij,}$, where $V_{kl} \equiv V_k - V_l$ is the voltage between points k and l and I_{ij} is the current flowing from contact i to contact j.
2. Measure the resistances $R_{21,34}$ and $R_{32,41}$. Define $R_>$ as the greater of these two resistances and $R_<$ as the lesser of these two resistances.
3. Calculate the ratio $R_>/R_<$ as the corresponding value of the function $f(R_>/R_<)$ from Figure 26.4. Be careful to use the appropriate horizontal scale!

TABLE 26.1 Companies That Make Four-Point Resistivity Measurement Probes and Systems

Company and Comments
Creative Design Engineering, Inc. 20565 Elves Drive Cupertino, CA 95014 Tel: (408) 736-7273 Fax: (408) 738-3912 Creative Design Engineering makes manual and automatic four-point resistivity systems specially designed for both small and large semiconductor wafers
Four Dimensions, Inc. 3138 Diablo Ave. Hayward, CA 94545 Tel: (510) 782-1843 Fax: (510-786-9321 http://www.4dimensions.com Four Dimensions makes a variety of manual and automatic four-point probe systems for measurement of resistivity and resistivity mapping of flat samples such as semiconductor wafers
Hewlett-Packard Company Test and Measurement Organization 5301 Stevens Creek Blvd. Santa Clara, CA 95052-8059 Tel: (800) 452-4844 Fax: (303) 754-4801 http://www.hp.com Hewlett-Packard makes a variety of high-quality instruments useful for four-point measurements
Jandel Engineering, Ltd. Grand Union House Leighton Road Linslade, Leighton Buzzard LU7 7LA United Kingdom Tel: (01525)-378554 Fax: (01525)-381945 http://www.getnet.com/~bridge/jandel.html Jandel makes four-point probes useful for flat samples such as semiconductor wafers. They will build custom four-point probes for your particular needs. They also make a combined constant current source and digital voltmeter for resistivity measurements
Keithley Instruments, Inc. 28775 Aurora Road Cleveland, OH 44139-1891 Tel: (440) 248-0400 Fax: (440) 248-6168 http://www.keithley.com Keithley makes a wide variety of four-point measurement systems. They also have useful, free literature detailing techniques for making accurate resistivity measurements

TABLE 26.1 (continued) Companies That Make Four-Point Resistivity Measurement Probes and Systems

Company and Comments
KLA-Tencor Corp. 1 Technology Drive Milpitas, CA 95035 Tel: (408) 875-3000 Fax: (408) 875-3030 http://www.kla-tencor.com KLA-Tencor makes automated sheet resistance mapping systems designed for semiconductor wafers
Lucas-Signatone Corp. 393-J Tomkins Ct. Gilroy, CA 95020 Tel: (408) 848-2851 Fax: (408) 848-5763 http://www.signatone.com Signatone makes four-point resistivity measurement systems and a variety of four-point probe heads. They make a high-temperature, four-point probe head for temperatures up to 670 K
Miller Design and Equipment, Inc. 2231-C Fortune Drive San Jose, CA 95131-1806 Tel: (408) 434-9544 Fax: (408) 943-1491 Miller Design makes semiautomatic resistivity probe systems, designed for semiconductor wafers
Mitsubishi Chemicals Corp./Yuka Denshi Co., Ltd. Kyodo Bldg., 1-5 Nihonbashi Muromachi 4-chome Chuo-ku, Tokyo 103 Japan Tel: 03-3270-5033 Fax: 03-3270-5036 Yuka Denshi makes a low-resistivity meter and a variety of four-point probe heads
MMR Technologies, Inc. 1400 North Shoreline Blvd., # A5 Mountain View, CA 94043 Tel: (650) 962-9620 Fax: (650) 962-9647 http://www.mmr.com MMR makes systems for four-point resistivity, Hall mobility, and Seebeck potential measurements over the temperature range 80–400 K
Napson Corporation Momose Bldg. 7F 2-3-6 Kameido Koto-ku Tokyo 136 Japan Napson Corporation manufactures high-performance sheet resistance measuring systems for thin films on semiconductor or flat panel, Solar Cell, and resistivity systems for silicon wafers

(*continued*)

TABLE 26.1 (continued) Companies That Make Four-Point Resistivity Measurement Probes and Systems

Company and Comments
QuadTech, Inc. 100 Nickerson Rd., Suite 3 Marlborough, MA 01752-9605 Tel: (800) 253-1230 Fax: (508) 485-0295 http://www.quadtechinc.com QuadTech makes a four-point ohmmeter capable of measuring resistances from 1 μΩ to 2 MΩ
Quantum Design 11578 Sorrento Valley Rd. San Diego, CA 92121-1311 Tel: (800) 289-6996 Fax: (619) 481-7410 http://www.quandsn.com Quantum Design makes an automated system for measuring four-point resistivity, Hall mobility, and other properties over the temperature range 2–400 K in magnetic fields up to 14 T

4. Calculate the resistivity ρ_a using

$$\rho_a = \frac{\pi d(R_> + R_<) f(R_>/R_<)}{\ln 4} \tag{26.15}$$

where
ρ_a is the resistivity in Ω m
d is the thickness of the sample in m
resistances $R_>$ and $R_<$ are measured in Ω
ln 4 is approximately 1.3863

It is not necessary to measure the width or length of the sample.

5. Switch the leads to measure $R_{43,12}$ and $R_{14,23}$. Repeat steps 3 and 4 to calculate ρ_b using these new values for $R_>$ and $R_<$. If the two resistivities ρ_a and ρ_b are not within 10% of each other, either the contacts are bad or the sample is too nonuniform to measure reliably. Try making new contacts. If the two resistivities ρ_a and ρ_b are within 10% of each other, the best estimate of the material resistivity ρ is the average:

$$\rho = \frac{(\rho_a + \rho_b)}{2} \tag{26.16}$$

Note: The function $f(R_>/R_<)$ plotted in Figure 26.4 is defined by the transcendental equation:

$$f\left(\frac{R_>}{R_<}\right) \equiv \frac{-\ln 4(R_>/R_<)}{[1 + (R_>/R_<)\ln\{1 - 4^{-[(1+R_>/R_<)f]^{-1}}\}]} \tag{26.17}$$

TABLE 26.2 Companies That Make High-Resistivity Measurement Probes and Systems

Company and Comments
Agilent Technologies Test and Measurement Organization 5301 Stevens Creek Blvd. Santa Clara, CA 95052-8059 Tel: (800) 452-4844 Fax: (303) 754-4801 http://www.hp.com Hewlett-Packard makes high-resistance meters and specially designed resistivity test chambers
Keithley Instruments, Inc. 28775 Aurora Road Cleveland, OH 44139-1891 Tel: (440) 248-0400 Fax: (440) 248-6168 http://www.keithley.com Keithley makes special meters and resistivity test chambers for measuring high resistivities. They also have useful, free literature detailing techniques for making accurate resistivity measurements
Mitsubishi Chemicals Corp./Yuka Denshi Co., Ltd. Kyodo Bldg., 1-5 Nihonbashi Muromachi 4-chome Chuo-ku, Tokyo 103 Japan Tel: 03-3270-5033 Fax: 03-3270-5036 Yuka Denshi makes high-resistance meters and a variety of probes and resistivity test chambers
Monroe Electronics, Inc. 100 Housel Avenue Lyndonville, New York 14098 Tel: (800) 821-6001 Fax: (716) 765-9330 http://www.monroe-electronics.com Monroe Electronics makes portable and handheld instruments for measuring surface resistivity, designed for testing antistatic materials
QuadTech, Inc 100 Nickerson Rd. Suite 3 Marlborough, MA 01752-9605 Tel: (800) 253-1230 Fax: (508) 485-0295 http://www.quadtechinc.com QuadTech makes a high-resistance ohmmeter

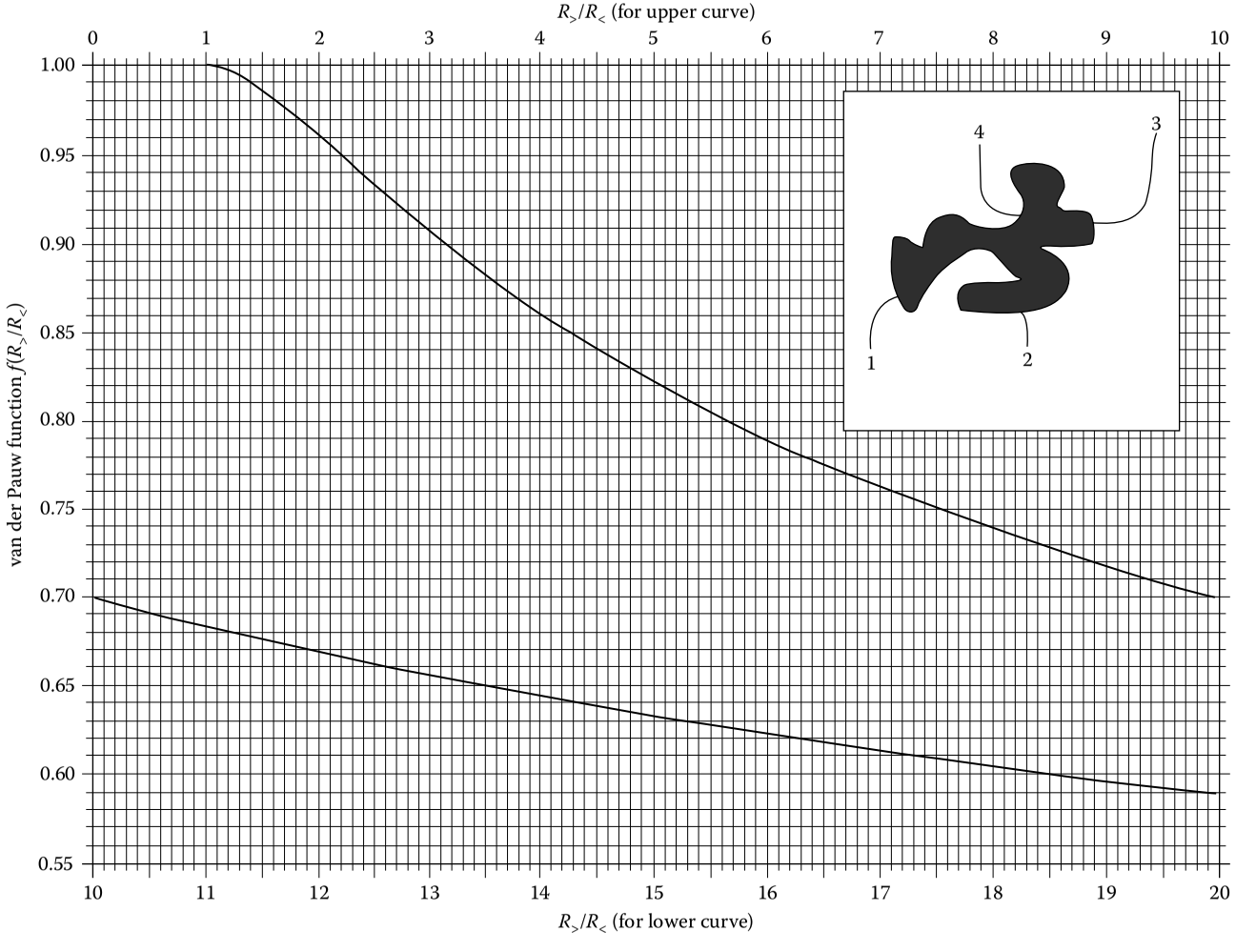


FIGURE 26.4 The van der Pauw technique. The inset shows one of the possible measurement geometries. The graph shows the function $f(R_>/R_<)$ needed to find the resistivity in Equation 26.15.

Defining Terms

Conductance: The inverse of resistance.

Conductivity: The inverse of resistivity.

Contact resistance: The resistance between the surface of a material and the electric contact made to the surface.

Four-point technique: A method for measuring the resistivity of a material, using four electric contacts to the material, which avoids many contact resistance problems.

Resistance: The physical property of a particular piece of a material, quantifying the ease with which electricity can flow through it. Resistance will depend on the size and shape of the piece of material.

Resistivity: The intrinsic physical property of a material quantifying the ease with which electricity can flow through it. Resistivity will not depend on the size and shape of the piece of material. Higher resistivity means the flow of electricity is more difficult.

Sheet resistance: The resistance of a square thin film or sheet of material.

Two-point technique: A method for measuring the resistivity of a material, using two electric contacts to the material.

van der Pauw technique: A method of measuring the four-point resistivity of an arbitrarily shaped material sample.

Acknowledgments

I thank Alison Breeze, John Clark, Kirsten R. Daehler, James M. E. Harper, Linda D. B. Kiss, Heidi Pan, and Shukri Souri for many useful suggestions.

References

1. P. Drude, Zur elektronentheorie der metalle, *Annalen der Physik*, 1, 566–613, 1900; 3, 369–402, 1900. See Ref. [2] for modern discussions of the Drude model and electrical conductivity.
2. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Philadelphia, PA: Saunders College, 1976; C. Kittel, *Introduction to Solid State Physics*, 7th edn., New York: John Wiley & Sons, 1996.
3. L. B. Valdes, Resistivity measurements on germanium for transistors, *Proc. IRE*, 42, 420–427, 1954.
4. L. J. van der Pauw, A method of measuring specific resistivity and Hall effect of discs of arbitrary shape, *Philips Res. Rep.*, 13, 1–9, 1958.
5. R. Chwang, B. J. Smith, and C. R. Crowell, Contact size effects on the van der Pauw method for resistivity and Hall coefficient measurement, *Solid-State Electron.*, 17, 1217–1227, 1974.

Further Readings

Look D. C., Bulk and contact electrical properties by the magneto-transmission-line method: Application to GaAs, *Solid-State Electron.*, 30, 615–618, 1987.

Look D. C., *Electrical Characterization of GaAs Materials and Devices*, New York: John Wiley & Sons, 1989.

Maissel L. I., Electrical properties of metallic thin films (13-1 to 13-33), in L. I. Maissel and R. Glang (eds.), *Handbook of Thin Film Technology*, San Francisco, CA: McGraw-Hill, 1970.

Wieder H. H., *Laboratory Notes on Electrical and Galvanomagnetic Measurements*, New York: Elsevier, 1979.

27

Charge Measurement

Saps Buchman
Stanford University

John T. Mester
Stanford University

T.J. Sumner
Imperial College

Electric charge, a basic property of elementary particles, is defined by convention as negative for the electron and positive for the proton. In 1910, Robert Andrews Millikan (1868–1953) demonstrated the quantization and determined the value of the elementary charge by measuring the motion of small charged droplets in an adjustable electric field. The SI unit of charge, the *coulomb* (C), is defined in terms of base SI units as

$$1 \text{ coulomb} = 1 \text{ ampere} \times 1 \text{ second} \tag{27.1a}$$

In terms of fundamental physical constants, the coulomb is measured in units of the elementary charge e:

$$1\text{C} = 1.60217733 \times 10^{19} e \tag{27.1b}$$

where the relative uncertainty in the value of the elementary charge is 0.30 ppm [1].

Charge measurement is widely used in electronics, physics, radiology, and light and particle detection, as well as in technologies involving charged particles or droplets (e.g., toners used in copiers). Measuring charge is also the method of choice for determining the average value for small and/or noisy electric currents by utilizing time integration. The two standard classes of charge-measurement devices are the electrostatic voltmeters and the charge amplifiers.

Electrostatic instruments function by measuring the mechanical displacement caused by the deflecting torques produced by electric fields on charged conductors [2,3]. Electrostatic voltmeters also serve as charge-measurement devices, using the fact that charge is a function of voltage and instrument capacitance. This class of instruments can be optimized for a very wide range of measurements, from about 100 V to 100 kV full scale, with custom devices capable of measuring voltages in excess of 200 kV. The accuracy of electrostatic voltmeters is about 1% of full scale, with typical time constants of about 3 s. Their insulation resistance is between 10^{10} and 10^{15} Ω, with instrument capacitances in the range of 1–500 pF. Figure 27.1 gives a schematic representation of several types of electrostatic voltmeters.

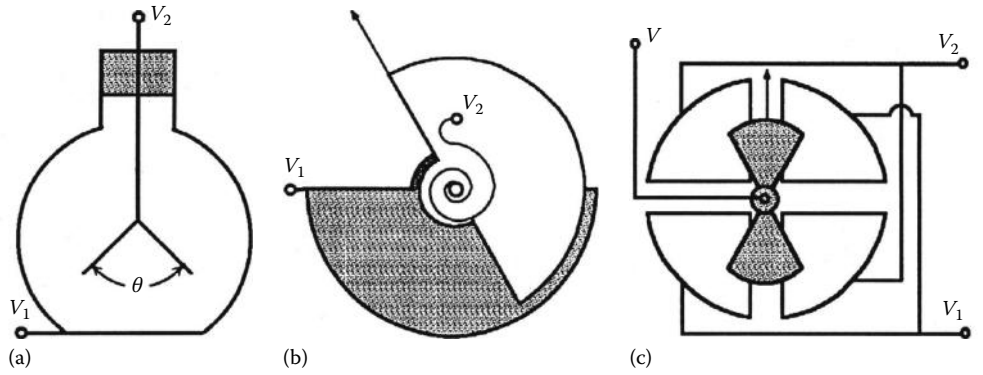


FIGURE 27.1 Examples of the repulsion, attraction, and symmetrical mechanical configurations of electrostatic voltmeters: (a) gold-leaf electroscope, (b) schematic representation of an attraction electrostatic voltmeter, and (c) a symmetrical quadrant electrostatic voltmeter. (From Harris, F., *Electrical Measurements*, John Wiley & Sons, New York, 1952.)

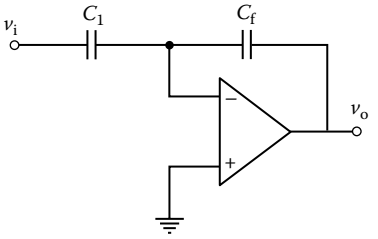


FIGURE 27.2 Basic concept of the charge amplifier. The output voltage is $v_o = C_1/C_f \times v_i$.

Modern electronic instruments have replaced in great measure the *electrostatic voltmeters* as devices of choice for the measurement of charge. The charge amplifier is used for the measurement of charge or charge variation [4]. Figure 27.2 shows the basic configuration of the charge amplifier. The equality of charges on C_1 and C_f results in

$$v_0 = \frac{C_1}{C_f} v_i \quad \text{or} \quad \Delta v_0 = \frac{C_1}{C_f} \Delta v_i \tag{27.2}$$

This same measurement principle is realized in the *electrometer*. The charge, Q, to be measured is transferred to the capacitor, C, and the value, V, of the voltage across the capacitor is measured: $Q = CV$. Figure 27.3 shows the block diagram for the typical digital electrometer [5]. Charge is measured in the coulomb mode, in which a capacitor C_f is connected across the operational amplifier, resulting in the input capacitance AC_f. Typical gain A for these amplifiers is in the range 10^4–10^6, making AC_f very large and thus eliminating the problem of complete charge transfer to the input capacitor of the coulombmeter. Electrometers have input resistances in the range 10^{14}–10^{16} Ω, resulting in very long time constants and thus minimizing the discharging of the capacitor. Typical leakage currents are from 5×10^{-14} to 5×10^{-16}A, again minimizing the variation in the charge. In the coulombmeter mode, electrometers can measure charges as low as 10^{-15} C and currents as low as 10^{-17} A.

Errors in charge-measurement instruments are caused by extraneous currents [5]. These currents are generated as thermal noise in the shunt resistance, by resistive leakage and by triboelectric, piezoelectric, pyroelectric, electrochemical, and dielectric absorption effects. The coulombmeter

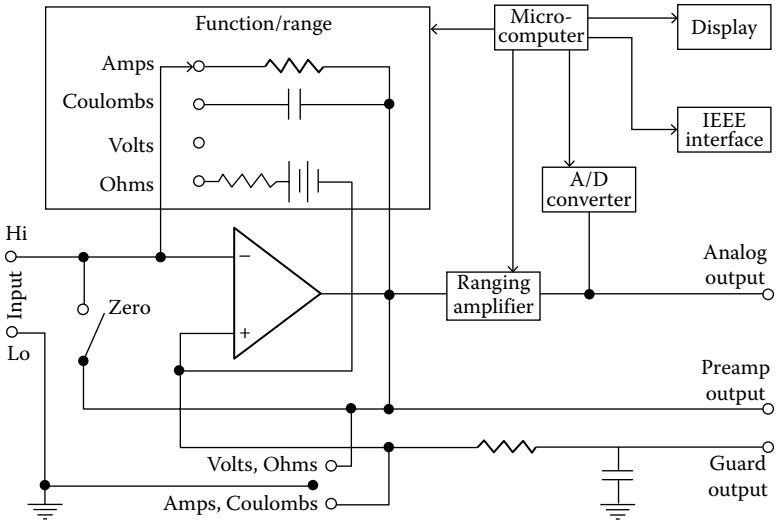


FIGURE 27.3 Conceptual block diagram of the digital electrometer. In the coulomb function, the charge to be determined is transferred to the corresponding capacitor, and the voltage across this capacitor is measured.

function of the electrometers does not use internal resistors, thus eliminating this thermal noise source. Triboelectric charges due to friction between conductors and insulators can be minimized by using low-noise triaxial cables and by reducing mechanical vibrations in the instrument. Under mechanical stress, certain insulators will generate electric charge due to piezoelectric effects. Judicious choices of materials and reduction of stress and mechanical motion can significantly reduce this effect.

Trace chemicals in the circuitry can give rise to electrochemical currents. It is therefore important to thoroughly clean and dry chemicals of all sensitive circuitry. Variations in voltages applied across insulators cause the separation and recombination of charges and thus give rise to dielectric absorption parasitic currents. The solution is to limit the voltages applied to insulators used for high-sensitivity charge measurements to less than about 5 V.

Dielectric materials used in sensitive charge-measurement experiments should be selected for their high resistivity (low-resistive leakage), low water absorptivity, and minimal piezoelectric, pyroelectric, triboelectric, and dielectric absorption effects. Sapphire and polyethylene are two examples of suitable materials. *Guarding* is used to minimize both shunt currents and errors associated with the capacitance of cables and connectors. The block diagram in Figure 27.3 shows typical guarding arrangements for modern electrometers.

27.1 Electrostatic Voltmeters

Electrostatic voltmeters and the more sensitive mechanical electrometers use an indicator to read out the position of a variable capacitor. Depending on their mechanical configuration, the electrostatic voltmeters can be categorized into three types: repulsion, attraction, and symmetrical [2,3]. The moving system in the high-sensitivity instruments is suspended from a torsion filament or pivoted in precision bearings to increase ruggedness. A wide variety of arrangements are used for the capacitive elements, including parallel plates, concentric cylinders, and hinged plates. Motion damping of the moving parts is provided by air or liquid damping vanes or by eddy current damping.

One of the oldest devices used to measure charge is the *gold-leaf electroscope*, shown in Figure 27.1a. Thin leaves of gold are suspended from a conductive contact that leads out of a protective case through an insulator. As charge applied to the contact is transferred to the leaves, the leaves separate by a certain angle,

the mutual repulsion being balanced by gravity. In principle, this device can also be used to measure the voltage difference between the contact electrode and the outer case, assuming the capacitance as a function of leaf separation angle is known. The electroscope is an example of a repulsion-type device, as is the Braun voltmeter in which the electroscope leaves are replaced by a balanced needle [2]. The delicate nature and low accuracy of this class of instruments limit their use in precise quantitative measurement applications.

An example of an attraction electrostatic voltmeter used for portable applications is shown in Figure 27.1b. While the fixed sector disk is held at V_1, the signal V_2 is applied to the movable sector disk through a coil spring that supplies the balancing torque. Opposite charges on the capacitor cause the movable plate to rotate until the electric attraction torque is balanced by the spring. If a voltage $V = V_1 - V_2$ is applied, the electric torque is given by [3]

$$\tau = \frac{dU}{d\theta} = \frac{d(1/2(CV)^2)}{d\theta} = \frac{1}{2}V^2\frac{dC}{d\theta} \tag{27.3}$$

The balancing spring torque is proportional to angular displacement, so the angle at equilibrium is given by

$$\frac{1}{2}V^2\frac{dC}{d\theta} = K\theta \tag{27.4}$$

Since the rotation is proportional to V^2, such an instrument can be used to measure ac voltages as well.

Symmetrical instruments are used for high-sensitivity, low-voltage measurements. The voltage is applied to a mobile element positioned between a symmetrical arrangement of positive and negative electrodes. Common mode displacement errors are thus reduced, and the measurement accuracy increased. One of the first devices sensitive enough to be called an "electrometer" was the quadrant electrometer shown schematically in Figure 27.1c. As a voltage difference, $V_1 - V_2$, is applied across the quadrant pairs, the indicator is attracted toward one pair and repelled by the other. The indicator is suspended by a wire allowing the stiffness of the suspension to be controlled by the potential V, so that the displacement is given by [2]

$$\theta = K\left[(V_1 - V_2)\left(V - \frac{1}{2}(V_1 - V_2)\right)\right] \tag{27.5}$$

where K is the unloaded spring constant of the suspension.

The advantage of electrostatic instruments is that the only currents they draw at dc are the leakage current and the current needed to charge up the capacitive elements. High-performance symmetrical electrostatic instruments have leakage resistances in excess of 10^{16} Ω, sensitivities of better than 10 μV, and capacitances of 10–100 pF. They are capable of measuring charges as small as 10^{-16} C and are sensitive to charge variations of 10^{-19} C.

Historically, as stated earlier, the symmetrical electrostatic voltmeters have been called "electrometers." Note that this can give rise to some confusion, as the term *electrometer* is presently also used for the electronic electrometer. This is a high-performance dc multimeter with special input characteristics and high sensitivity, capable of measuring voltage, current, resistance, and charge.

Modern noncontacting electrostatic voltmeters have been designed for voltage measurements up to the 100 kV range. An advantage of these instruments is that no physical or electric contact is required between the instrument and test surface, ensuring that no charge transfer takes place. Kawamura et al. [6] report the design of an attraction-type device that uses a strain gage to determine the displacement of a movable plate electrode. Hsu and Muller [7] have constructed a micromechanical shutter to modulate the capacitance between the detector electrode and the potential surface to be measured. Trek Inc. [8] electrostatic voltmeters achieve a modulated capacitance to the test surface by electromechanically vibrating the detector electrode. Horenstein [9], Gunther [10], and MacDonald and Fallone [11] have employed noncontacting electrostatic voltmeters to determine the charge distributions on semiconductor and insulator surfaces. Tables 27.1 and 27.2 contain a selection of available commercial devices and manufacturers.

TABLE 27.1 Instruments Used in Charge Measurement Applications

Instrument Manufacturer	Model #	Description
Advantest	TR8652	Electrometer
	R8340/8340A	Electrometer
	R8240	Digital electrometer
	TR8601	Micro current meter
	TR8641	Picoammeter
Amptek	A101	Charge preamplifier
	A111	Charge preamplifier
	A203	Charge preamplifier/shaper
	A225	Charge preamplifier/shaper
	A250	Charge preamplifier
EIS	ESH1-33	Electrostatic voltmeter
	ESD1-11	Electrostatic voltmeter
	CRV	Electrostatic peak voltmeter
Jennings	J-1005	RF kilovoltmeter
Keithley	610C	Electrometer
	614	Digital electrometer
	617	Programmable electrometer
	642	Digital electrometer
	6512	Electrometer
	6517	Electrometer
Kistler	5011B	Charge amplifier
	5995	Charge amplifier
	5395A	Charge calibrator
Monroe	168-3	Electrostatic voltmeter
	174-1	Electrostatic voltmeter
	244AL	Electrostatic millivoltmeter
	253-1	Nanocoulombmeter/Faraday cup
Nuclear Associates	37-720FW	Digital electrometer for dosimetry
Trek	320B	Electrostatic voltmeter
	341	Electrostatic voltmeter
	344	Electrostatic voltmeter
	362A	Electrostatic voltmeter
	368	Electrostatic voltmeter

[a] Available in a range of specifications.

TABLE 27.2 Instrument Manufacturers

Advantest Corporation
Shinjuku NS Building 4-1
Nishi-Shinjuku 2-Chome, Shinjuku-ku
Tokyo 163-08 Japan

Amptek Inc.
6 De Angelo Drive
Bedford, MA 01730

Electrical Instrument Service Inc.
Sensitive Research Instruments
25 Dock St.
Mount Vernon, New York 10550

Jennings Technology Corporation
970 McLaughlin Ave.
San Jose, CA 95122

Keithley Instruments Inc.
28775 Aurora Road
Cleveland, OH 44139

Kistler Instrumente AG
CH-8408 Winterthur, Switzerland

Monroe Electronics
100 Housel Ave.
Lyndonville, New York 14098

Nuclear Associates
Div. of Victoreen, Inc.
100 Voice Rd.
P.O. Box 349
Carle Place, New York 11514-0349

Trek Incorporated
3922 Salt Works Rd.
P.O. Box 728
Medina, New York 14103

27.2 Charge Amplifiers

The conversion of a charge, Q, into a measurement voltage involves at some stage the transfer of that charge onto a reference capacitor, C_r. The voltage, V_r, developed across the capacitor gives a measure of the charge as $Q = V_r/C_r$. There are two basic amplifier configurations for carrying out such measurements using the reference capacitor in either a *shunt* or *feedback* arrangement.

27.2.1 Shunt Amplifiers

Figure 27.4 shows a typical circuit in which the reference capacitor is used in a shunt mode. In this example, it is assumed that the charge that is to be measured is the result of the integrated current delivered by a time-dependent current source, $i(t)$. With the measurement circuit disconnected (switch in position s2), the charge on the source capacitor, C_s, at time τ will be $Q = \int_0^C i(t)dt$ (assuming Q starts from zero at $t = 0$), and the output voltage, V_o will be zero, as the input voltage to the (ideal) operational amplifier is zero. On closing the switch in position s1, the charge on C_s will then be shared between it and C_r and

$$V_o = \left(\frac{R_1 + R_2}{R_2}\right)\frac{Q}{C_s + C_r} \tag{27.6}$$

In order to accurately relate the output voltage to the charge Q, not only does the gain of the noninverting amplifier and the reference capacitance need to be known, which is relatively straightforward, but it is also necessary to know the source capacitance. This is not always easy to determine. The effect of any uncertainty in the value of C_s can be reduced by increasing the value of the reference capacitor to the point where it dominates the total capacitance. However, in so doing, the output voltage is also reduced and the measurement becomes more difficult. The dependence of the measurement on C_s is one of the main limitations to this simple method of charge measurement. In addition, any leakage currents into the input of the operational amplifier, through the capacitors, or back into the source circuitry during the measurement period will affect the result. For the most accurate measurements of low charge levels, *feedback amplifiers* are more commonly used.

27.2.2 Feedback Amplifiers

Figure 27.5 shows a circuit where the reference capacitor now provides the feedback path around the operational amplifier. The output voltage from this configuration for a given charge Q transfers from the source is then

$$V_o = \frac{AQ}{C_s + AC_r} \tag{27.7}$$

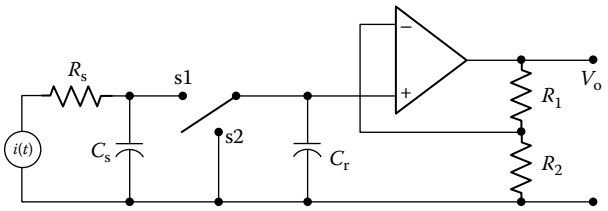


FIGURE 27.4 Schematic representation of a charge amplifier using a shunt reference capacitor. With the switch in position sl, the measurement circuit is connected and the charge is proportional to the output voltage and to the sum $C_s + C_r$. Note the significant sensitivity to C_s.

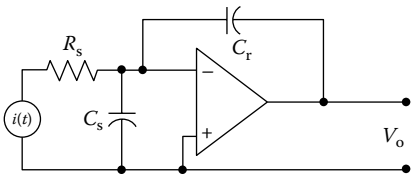


FIGURE 27.5 Schematic representation of a charge amplifier with reference feedback capacitor. The charge is proportional to the output voltage and to the sum $C_s + AC_r$, where A is the amplifier gain. Note the reduced sensitivity to C_s.

where A is the open-loop gain of the operational amplifier. For most situations, $AC_r > C_s$ and the charge measurement becomes independent of the source capacitance. In addition, the inverting input to the operational amplifier is kept close to ground potential, reducing the magnitude of leakage currents in that part of the circuit. However, in contrast to these two benefits is the new problem that the input bias current for the operational amplifier is integrated by the feedback capacitor, producing a continual drift in the output voltage. Several solutions have been used to overcome this problem, including the use of a parallel feedback resistor, R_f, which suppresses the integrating behavior at low frequencies (periods longer than R_fC_r), balancing the bias current with another externally provided current and incorporating a reset switch that discharges the circuit each time the output voltage ramps beyond a set trigger level.

The sensitivity of feedback amplifiers depends on the noise sources operating within any specific application. The most impressive performance is obtained by amplifiers integrated into charge-coupled device (CCD) chips that can, under the right operational conditions, provide sensitivities measured in terms of a few electron charges. To illustrate the important parameters involved in the design of ultralow-noise charge preamplifiers for CCD-type applications, consider the circuit shown in Figure 27.6. The source (detector) is now shown as a biased photodiode, which is assumed to be producing individual bursts of charge each time a photon (e.g., an x-ray) interacts in it. In this example, the photodiode is coupled to the amplifier using a large value capacitor, C_c. This blocks the direct current path from the diode bias supply, V_b, but provides a low-impedance path for the short-duration charge deposits. The preamplifier is a variant on that shown in Figure 27.5, in which there is now a parallel feedback resistor to provide baseline restoration on long time scales and an field effect transistor (FET) transistor to reduce the effect of the operational amplifier input bias current by virtue of its high current gain factor, β.

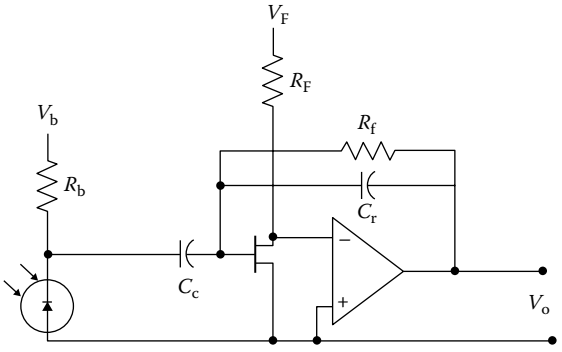


FIGURE 27.6 Typical ultralow-noise charge preamplifier configuration for charge pulse readout from ionization-type radiation detectors (e.g., x-ray detection using photodiodes or CCDs). The large capacitor C_c provides a low-impedance path for the short charge deposit pulses, while the parallel feedback resistor provides baseline restoration on long time scales. An FET transistor reduces the effect of the operational amplifier input bias current.

In practice, the dominant noise contributions in most applications of this type come from Johnson (current) noise in the bias and feedback resistors, shot noise on the photodiode bias current, voltage noise across the FET, and finally the inevitable 1/*f* component. The two resistors effectively feed thermal current noise into the input of the integrator. Similarly, the shot noise associated with the photodiode bias current feeds into the input. Together, these components are known as *parallel* noise and the total parallel noise charge, q_p, is given by

$$q_p = \sqrt{\left(\frac{4kT}{R_f + R_b} + 2eI_b(T)\right)\frac{1}{\Delta B}} \tag{27.8}$$

where

k is the Boltzmann's constant
e is the charge on the electron
T is the absolute temperature
ΔB is the bandwidth associated with the measurement that will depend on the details of subsequent shaping amplifier stages [12]

Voltage noise across the FET (and hence operational amplifier inputs) will arise from junction noise in the FET itself and from Johnson noise in its bias resistor, R_F. In practice, the FET junction noise usually dominates, in which case this *series* noise, q_s, contribution is given by

$$q_s = \sqrt{\varepsilon_n^2 C_{in}^2 \Delta B} \tag{27.9}$$

where

ε_n is the junction voltage noise for the FET in $V\sqrt{Hz^{-1}}$
C_{in} is the total capacitance seen at the gate of the FET

This will include both the source capacitance and the gate capacitance of the FET and any stray capacitance. The total noise is then the quadrature sum of Equations 27.8 and 27.9. The different dependencies on the bandwidth for Equations 27.8 and 27.9 imply there will be some optimum bandwidth for the measurement and this will depend on the relative contributions from each. 1/*f* noise manifests itself as a bandwidth-independent term that again must be added in quadrature. The Johnson noise associated with the resistors and FET junction will show a temperature dependence decreasing with $\sqrt{T}$. For the FET, this reduction does not continue indefinitely and there is usually an optimum temperature for the FET around 100 K. Photodiode bias currents also fall with decreasing temperature, and, for silicon devices, this is about a factor of 2 for every 10 K drop in temperature. Most ultralow-noise applications thus operate at reduced temperature, at least for the sensitive components. Bias resistors and feedback resistors are kept as high as possible (typically >100 MΩ), and FETs are specially selected for low junction voltage noise (typically 1 $nV\sqrt{Hz^{-1}}$). Ideally, photodiode capacitances should be kept as low as possible, and there is also interplay between the FET junction noise, ε_n, and the FET gate capacitance that is affected by altering the FET bias current, which can be used to fine-tune the series noise component.

Finally, there is another noise component that can often be critical and difficult to deal with. This is from microphonics. There are two effects. First, the feedback reference capacitor is typically made as small as possible (<1 pF) to reduce the effect of noise in the following shaping amplifier stages. This makes it sensitive to any stray capacitances and if there are vibrations in the system that alter the local geometry, then this can change the feedback capacitance that changes the "gain" of the preamplifier. Second, the photodiode will be operating with some applied bias voltage (often several tens of volts), and any change in its apparent capacitance through mechanical movement of components will result in

charge being moved around. These charge movements will be sensed by the charge amplifier. Ultralow-noise applications that ignore mechanical stability in their design phase are in peril. Tables 27.1 and 27.2 contain a selection of available commercial devices and manufacturers.

27.3 Applications

The Millikan technique of measuring charges on particles suspended in electric fields continues to be developed for various applications. Kutsuwada et al. [13] use a modified Millikan experiment to measure the charge distribution on electrophotographic toner particles of various sizes. They show a comparison of the results obtained using an ordinary Millikan apparatus and a modified system in which an additional ac electrode has been inserted in the hyperbolic quadrupole electrode assembly. The two methods agree to within a small multiplicative calibration factor. A different method of measuring the charge of toner particles uses the q/d meter [14] in which q and d refer, respectively, to the charge and the diameters of the particles. In the q/d meter, the charged particles are transported horizontally in a steady laminar air flow and move vertically in an electric field until deposited on a registration electrode. The position at which the particle is deposited on the registration electrode defines the charge-to-diameter ratio. The size of the deposited particle is then measured, thus completing the determination of the charge distribution for various particle sizes.

The experiments searching for fractional charges [15] make use of superconducting niobium spheres 0.25 mm in diameter, suspended in vacuum at 4.2 K in a magnetic field. The vertical position of the spheres is modulated by an alternating electric field and measured with an ultrasensitive magnetometer. Positrons and electrons generated by radioactive sources are used to cancel all integer charges on the spheres. Fractional charges are detected and measured as that residual charging of the niobium spheres that cannot be neutralized by the integral charges from the radioactive sources. Although this experiment is sensitive to about 0.01 electron charges (10^{-21} C), it has produced no conclusive evidence of fractional charges.

A similar approach, using force modulation, is used for the noncontact measurement of charge on gyroscopes [16]. Out-of-phase equal forces are applied to an electrostatically suspended gyroscope at a frequency well within the suspension control bandwidth. The charge of the gyroscope is then proportional to modulation frequency component of the suspension control effort. The sensitivity of this method is about 10^{-12} C gyroscope charge, limited by the allowable modulation force and position sensor noise.

Noncontact measurement of charge on liquid drops in a microgravity environment can be also performed using field mill instruments [17]. A mechanical chopper is used to modulate the electric field induced by the spherical charge on a grounded sensing plate. The resulting alternating current from sensor to ground is a measure of the charge on the drop. In a refinement of this method, the modulation is achieved by varying the distance between charge and sensor. The authors [18] claim that this system provides an increase in sensitivity of two to three orders of magnitude over the original technique.

Optical sensors based on the Pockels effect [19] are used to measure the space-charge field in gaseous dielectrics. The Pockels effect involves the change in the birefringence of certain crystalline materials on the application of an electric field. In the measuring system, a circularly polarized beam is detected after passing through the Pockels sensor and a polarizing plate. The detected intensity varies linearly with the intensity of the electric field applied to the sensor. This system is capable of performing a vector measurement of the electric field produced by the space charge, determining both intensity and direction.

Optical methods of charge measurement are also used for particles whose physical structure depends on their charge. An example is the degree of dissociation of the end groups of polystyrene particles in colloidal solutions [20]. In this application, the intensities of the Raman scattering spectrum lines depend on the degree of dissociation of the polystyrene end groups and thus determine the charge of these particles.

A widely used application of charge measurement is as an integral element of radiation dosimetry. Radiation is passed through ionization chambers, where the generated ions are collected and the charge

measured with electrometers. Ionization chambers used in dosimetry for radiation therapy have typical sensitivities in the range 0.02–0.2 nC R^{-1}. Coulombmeter electrometers with sensitivities of 1–10 pC are therefore required for this application.

Defining Terms

Charge, also **electric charge:** A basic property of elementary particles, defined by convention as negative for the electron and positive for the proton. The SI unit of charge is the coulomb (C), defined as 1 ampere (A) × 1 second (s).

Charge amplifier: Charge-measurement instrument. The charge is transferred to a reference capacitor and the resulting voltage across the capacitor is measured. Shunt and feedback versions of the charge amplifier have the reference capacitor used in shunt and feedback mode, respectively.

Electrometer: *Historic usage,* type of electrostatic voltmeter. *Modern usage*, electronic electrometer, a high-performance dc multimeter with special input characteristics and high sensitivity, capable of measuring voltage, current, resistance, and charge.

Electrostatic instrument: An instrument that functions by measuring the mechanical displacement or strain caused by electric fields.

Electrostatic voltmeter: Electrostatic instrument used to measure charge. The charge is determined as a function of voltage and instrument capacitance.

References

1. E. Cohen and B. Taylor, The 1986 adjustment of the fundamental physical constants, *Rev. Mod. Phys*., 59(4), 1121–1148, 1987.
2. F. Harris, *Electrical Measurements*, New York: John Wiley & Sons, 1952.
3. W. Michels, *Electrical Measurements and Their Applications*, Princeton, NJ: D. Van Nostrand, 1969.
4. A. J. Diefenderfer, *Principles of Electronic Instrumentation*, 3rd edn., Philadelphia, PA: Saunders College Publishing, 1994.
5. Keithley Instruments, Inc., *Low Level Measurements*, 4th edn., Cleveland, OH: Keithley Instruments, Inc., 1993.
6. K. Kawamura, S. Sakamoto, and F. Noto, Design and development of new electrostatic voltmeter using strain gauge, *IEEE Trans. Ind. Appl*., 25, 563–568, 1989.
7. C. Hsu and R. Muller, Micromechanical electrostatic voltmeter, *Transducers '91. International Conference on Solid-State Sensors and Actuators*, San Francisco, CA, 1991, pp. 659–662.
8. Technical Information, and Downloads, Trek Incorporated, Medina, NY. http://www.trekinc.com/, accessed on August 18, 2013.
9. M. M. Horenstein, Measuring surface charge with a noncontacting voltmeter, *Proceedings of the IEEE Industry Applications Society Annual Meeting*, *IAS'93*, Toronto, Ontario, Canada, vol. 3, 1993, pp. 1811–1816.
10. P. Gunther, Determination of charge density and charge centroid location in electrets with semiconducting substrates, *IEEE Trans. Electr. Insul*., 27, 698–701, 1992.
11. B. MacDonald and B. Fallone, Surface-charge distributions on radiation-charged electrets. *7th International Symposium on Electrets (ISE 7)*. Berlin, Germany, 1991, pp. 798–803.
12. V. Radeka, Signal, noise and resolution in position-sensitive detectors, *IEEE Trans. Nucl. Sci*., 21, 51–64, 1974; *20th Nuclear Science Symposium, 5th Nuclear Power Systems Symposium*, San Francisco, CA.
13. N. Kutsuwada, T. Shohdohji, N. Okada, H. Izawa, T. Sugai, Y. Nakamura, and T. Murata, Measurement of electric charge of electrophotographic toner, *J. Imaging Sci. Technol*., 37(5), 480–484, 1993.
14. M. Mehlin and R. M. Hess, Electrical charge measurement of toner particles using the *q/d* meter, *J. Imaging Sci. Technol*., 36, 142–150, 1992.

15. A. F. Hebard, G. S. LaRue, J. D. Phillips, and C. R. Fisel, *Search for Fractional Charge in Near Zero*: *New Frontiers of Physics*, New York: W. H. Freeman and Company, 1988, p. 511.
16. S. Buchman, T. Quinn, G. M. Keiser, D. Gill, and T. J. Sumner, Charge measurement and control for the gravity probe B gyroscopes, *Rev. Sci. Instrum*., 66, 120–129, 1995.
17. M. N. Horenstein, Peak sampled vibrating-reed for the measurement of electric fields in the presence of space charge, *Rev. Sci. Instrum*., 54, 591–593, 1983.
18. K.-C. Lin and T. G. Wang, Noncontact charge measurement, *Rev. Sci. Instrum*., 63, 2040–2043, 1992.
19. K. Hidaka, Progress in Japan of space charge field measurement in gaseous dielectrics using a Pockels sensor, *IEEE Electr. Insul. Mag*., 12(1), 17–28, 1996.
20. R. Kesavamoorthy, T. Sakuntala, and A. K. Arora, In situ measurement of charge on polystyrene particles in colloidal suspension, *Meas. Sci. Technol*., 1(5), 440–445, 1990.

28

Capacitance and Capacitance Measurements

Halit Eren
Curtin University

28.1 Theory of Capacitance

A capacitor is a system of two conducting electrodes separated by a dielectric material. The electrodes have equal and opposite charges. The capacitance C of a capacitor is equal to the ratio of the absolute value of the charge Q to the absolute value of the voltage between the electrodes, which can be expressed as

$$C = \frac{Q}{V} \tag{28.1}$$

where
C is the capacitance in farads (F)
Q is the charge in coulomb (C)
V is the voltage (V)

The unit of capacitance, the farad, is a large unit; practical capacitors have capacitances in microfarads (μF or $l0^{-6}$ F), nanofarads (nF or 10^{-9} F), and picofarads (pF or 10^{-12} F). The conversions of these units are shown in Table 28.1.

The capacitance C depends on the size and shape of charged bodies and their relative positions; examples are shown in Table 28.2. Generally, capacitance is inherent wherever an electrostatic field is present. In many electronic systems, it is necessary to deal not only with the capacitances that are designed consciously within the circuits but also with the unwanted interference and the stray capacitances that are introduced externally or internally at various stages of the circuits. For example, some sensors operate

TABLE 28.1 Capacitance Unit Conversions

Microfarads	Nanofarads	Picofarads
l0^{-6} F	10^{-9} F	10^{-12} F
0.000001 μF	0.001 nF	1.0 pF
0.001 μF	1.0 nF	1,000 pF
1.0 μF	1000 nF	1,000,000 pF

on capacitance principles, giving useful signals; in others, capacitance is inherent but undesirable. In many cases, cables and external circuits introduce additional capacitances that need to be accounted for the desirable operation of the system. In these cases, Table 28.2 is useful to identify and analyze possible sources of capacitances where charged bodies are involved.

The capacitance of structure can be determined by solving Laplace's equations $\nabla^2 V(x,y,z) = 0$ with appropriate boundary conditions. The boundary conditions specify the electrode voltages V_1 and V_2 of the plates. Laplace's equation yields to V and the electric field $E(x,y,z) = -\nabla V(x,y,z)$ between the electrodes. The charge of each electrode can also be obtained by integration of the flux density over the surface are of each electrode, as

$$Q = \int \varepsilon(x,y,z)E(x,y,z)dA \tag{28.2}$$

If the capacitor is made from two parallel plates, as shown in Figure 28.1, the capacitance value in terms of dimensions can be expressed by

$$C = \frac{\varepsilon A}{d} = \varepsilon_r \varepsilon_0 \frac{A}{d} \tag{28.3}$$

where

ε is the dielectric constant or permittivity of material
ε_r is the relative dielectric constant (in air, $\varepsilon_r = 1$)
ε_0 is the dielectric constant of vacuum (8.854188×10^{-12} F m^{-1})
d is the distance of the plates in m
A is the effective area of two parallel plates in m^2

In arriving to Equation 28.3, the fringe field is neglected for small distances, d, between the plates.

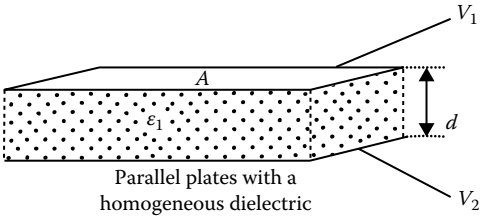
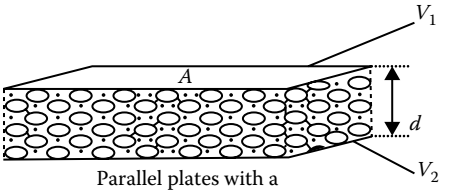
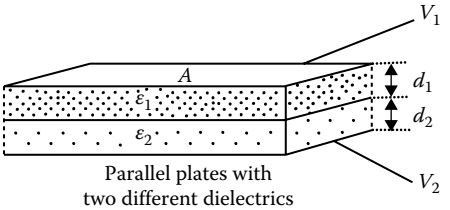
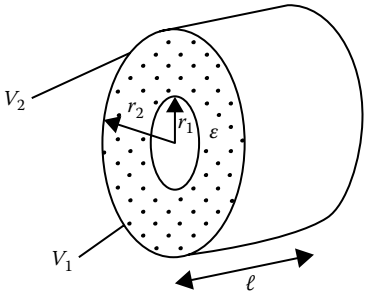
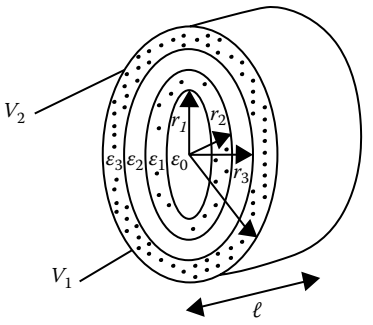
Capacitances can also be expressed in terms of dielectric properties currents and voltages. Suppose that a uniform dielectric between two parallel plates has a resistance, R, which can be written as

$$R = \frac{d\rho}{A} \tag{28.4}$$

where ρ is the specific resistance of the dielectric in Ω m. Then, using Equation 28.3 and eliminating d and A gives

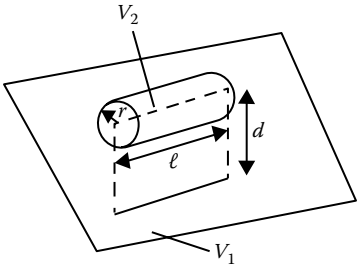
$$C = \frac{\varepsilon\rho}{R} \tag{28.5}$$

TABLE 28.2 Capacitances of Various Electrode Systems

 Parallel plates with a homogeneous dielectric	$C = \dfrac{\varepsilon A}{d}$
 Parallel plates with a two constituent dielectric	$C = \dfrac{(\varepsilon_1 r + \varepsilon_2)A}{(1+r)d}$, where r value is the volumetric ratio
 Parallel plates with two different dielectrics	$C = \dfrac{\varepsilon A}{(d_1/\varepsilon_1) + (d_2/\varepsilon_2)}$
 Coaxial cylinders with single dielectric	$C = \dfrac{4\pi\varepsilon\ell}{\ln(r_2/r_1)}$
 Coaxial cylinders with three dielectrics	$C = \dfrac{2\pi\varepsilon_0\ell}{\ln[(r_2/r_1)^{\varepsilon_0/\varepsilon_1} \times (r_3/r_2)^{\varepsilon_0/\varepsilon_2} \times (r_4/r_3)^{\varepsilon_0/\varepsilon_3}]}$ $= \left(\dfrac{1}{C_1} + \dfrac{1}{C_2} + \dfrac{1}{C_3}\right)^{-1}$

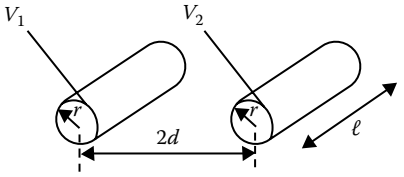
(*continued*)

TABLE 28.2 (continued) Capacitances of Various Electrode Systems



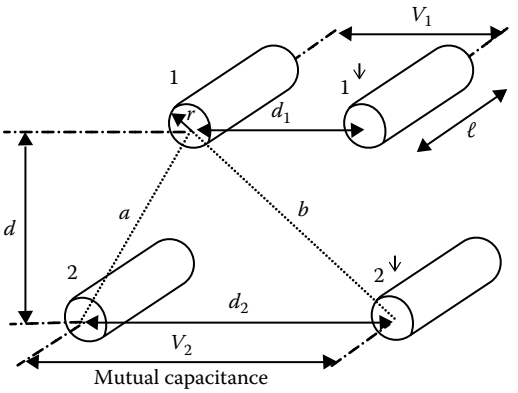
Cylinder parallel with plate

$$C = \frac{\pi\varepsilon\ell}{\ln\left[(d/r)+\sqrt{(d/r)^2-1}\right]} = \frac{2\pi\varepsilon\ell}{\cosh^{-1}(d/r)}$$



Parallel cylinders with equal radius

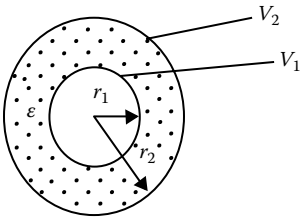
$$C = \frac{\pi\varepsilon\ell}{\ln\left[(d/r)+\sqrt{(d/r)^2-1}\right]} = \frac{\pi\varepsilon\ell}{\cosh^{-1}(d/r)}$$



Mutual capacitance

$$C_{11} = \frac{\pi\varepsilon\ell}{\ln(d_1/r)} + \frac{\pi\varepsilon}{2}\frac{\ln(b/a)}{[\ln(a/r)]\ln(b/r)}$$

$$C_{21} = C_{12} = \frac{\pi\varepsilon\ell}{2}\frac{\ln(b/a)}{[\ln(a/r)]\ln(b/r)}$$



Concentric spheres

$$C = \frac{4\pi\varepsilon r_1 r_2}{r_2 - r_1}$$

A voltage V across the capacitor causes a leakage current $I_l = V/R$ such that

$$C = \frac{\varepsilon\rho I_1}{V} \tag{28.6}$$

This indicates that the leakage current of a capacitor is simply proportional to its capacitance value.

As seen in Equations 28.3 and 28.4, the value of the capacitance is proportional to the permittivity of the dielectric material used. In the construction of capacitors, the permittivity of commonly used materials is given in Table 28.3.

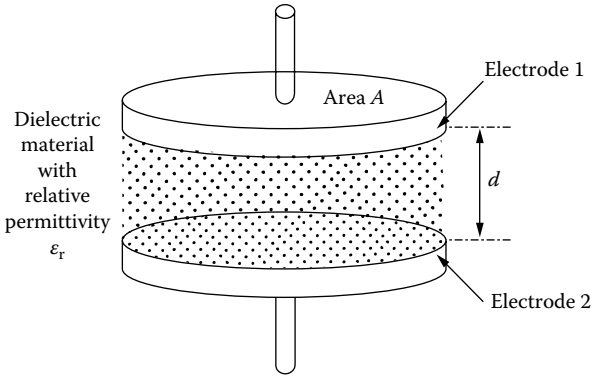


FIGURE 28.1 A basic capacitor made from two parallel plates. The capacitance between two charged bodies depends on the permittivity of the medium in between, the distance between the bodies, and the effective area. Capacitance can also be expressed in terms of the absolute values of the charge and the absolute values of the voltages between the bodies.

TABLE 28.3 Permittivity (Dielectric Constants of Materials Used in Capacitors)

Material	Permittivity
Vacuum	1.0
Air	1.0006
Teflon	2.1
Polyethylene, etc.	2.0–3.0
Impregnated paper	4.0–6.0
Glass and mica	4.0–7.0
Ceramic (low *K*)	≤20.0
Ceramic (medium *K*)	80.0–100.0
Ceramic (high *K*)	≥1000.0

28.2 Construction and Types of Capacitors

Commonly used fixed capacitors are constructed with dielectric materials of air, paper, mica, polymers, and ceramics. A comprehensive list of common capacitors and their characteristics are given in Tables 28.4 and 28.5. Variable capacitors are generally made with air or ceramic-dielectric materials. The capacitors used in electronic circuits can be classified as low-loss, medium-loss, and high-tolerance capacitors.

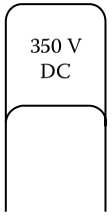
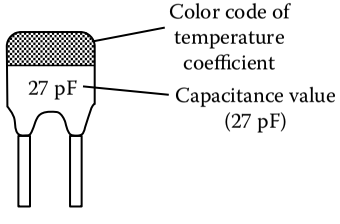
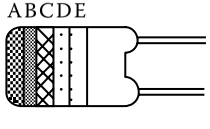
- *Low-loss capacitors* such as mica, glass, low-loss ceramic, and low-loss plastic film capacitors generally have good capacitance stability. These capacitors are expensive and often selected in precision applications, for example, telecommunication filters.
- *Medium-loss capacitors* have medium stability in a wide range of ac and dc applications. These are paper, plastic film, and medium-loss ceramic capacitors. Their applications include coupling, decoupling, bypass, energy storage, and some power electronic applications (e.g., motor starter, lighting, power line applications, and interference suppressions).
- *High-tolerance capacitors* such as aluminum and tantalum electrolytic capacitors deliver high capacitances. Although these capacitors are relatively larger in dimension, they are reliable and have longer service lives. They are used in polarized voltage applications, radios, televisions, consumer goods, as well as military equipment, and harsh industrial environmental applications.

TABLE 28.4 Characteristics of Common Capacitors

Capacitor Types	Range	Tolerance (%)	Voltage Range (V)	Temperature Range (°C)	Temperature Coefficient (ppm/°C)	Frequency Range	Permittivity ($\varepsilon/\varepsilon_0$ C²/N m²)	Dielectric Strength (C/C_0)	Dissipation Factor (%)	Insulation Resistance (MΩ/μF)	Typical Average Failure Rates (Fail per 10^6 h)
Mica, glass, porcelain, and Teflon											0.0133
Mica	5 pF–0.01 μF	5	100–600	−55/125	−50	100 Hz–10 GHz	7.0	1000	0.001	2.5×10^4	
Glass	5–1000 pF	5	100–600	−55/125	40	100 Hz–10 GHz	6.6	2500	0.001	10^6	
Porcelain	100 pF–0.1 μF	5	50–400	−55/125	120				0.1	5×10^5	
Teflon	1000 pF–2 μF	10	50–200	−70/250	−200				0.04	5×10^6	
Ceramic											
Low loss	100 pF–1 μF	10	50–400	−55/125	±30	100 Hz–10 GHz	5.7	200–300	0.02	5×10^3	0.11–0.008
Disk											
Multilayer											
Plate											
High permittivity	10 pF–1 μF		50–30,000	−55/125		1 kHz–1 GHz	1000–7000	100			
Disk											
Multilayer											
Plate											
Paper											
Paper	0.1–10 μF	10	200–1,600	−55/125	±800	D.C.–1 MHz	4.5	500–1000	1.0	5×10^3	0.002
Metalized paper											
Plastic											
Kapton	1000 pF–1 μF	10		−55/220	100				0.3	10^5	0.05
Polyester/ Mylar	1000 pF–50 μF	10	50–600	−55/125	400	D.C.–10 GHz	2.3	1000	0.75	10^5	

Parylene	5000 pF–1 μF	10		−55/125	±100	D.C.–10 GHz		1000	0.1	10^5	
Polysulfone	1000 pF–1 μF	5		−55/150	80				0.3	10^5	
Polycarbonate											
Axial	100 pF–30 μF	10	50–800	−55/125	±100	D.C.–10 GHz	2.8		0.2	5×10^5	
Can											
Radial											
Polypropylene											
Axial	100 pF–50 μF	10	100–800	−55/105	−200	D.C.–10 GHz			0.2	10^5	
Can											
Radial											
Polystyrene											
Axial	10 pF–2.7 μF	10	100–600	−55/85	−100	D.C.–10 GHz	10		0.05	10^6	
Can											
Radial											
Electrolytic and solid											0.04
Aluminum	0.1–1.6 μF	−10/100	3–600	−40/85	2500		8–10		10	100	
Tantalum											
Axial	0.1–1000 μF	−10/100	6–100	−55/85	800	D.C.–1 kHz	25–27		4.0	20	
Can											
Radial											
TiTiN Film	10–200 pF	10	6–30	−55/125	100				0.01	10^6	
Oil	0.1–20 μF		200–10,000			D.C.–1 MHz	2.4	1000	0.5		
Air/vacuum	1–100 pF		2,000–3,600				1.0				

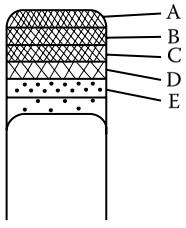
TABLE 28.5 Capacitor Specifications and Applications

Capacitor Types	Typical Commercial Specifications: Voltage A.C.V	Capacitance F	Tolerance%	Applications	Samples/Code
Mica, glass, porcelain, and Teflon				High temperature, low absorption, good in RF applications, and circuit requiring long-term stability	
Mica	350	2.2–1000 p	1		
Glass					
Porcelain					
Teflon					
Ceramic				Active filters, and high-density PCB applications, power-tuned circuits, coupling and decoupling of high-frequency circuits (PCB versions available)	
Low loss disk	100	1.8–470 p	±2		
Multilayer plate	63/50	10–1 μ	10		
	100	390–4700 p	10		
High permittivity disk					
Multilayer plate					
Paper				General-purpose motor applications	
Metalized paper					

Color	α ppm/°C
Red	100
Purple	
Black	0
Brown	−33
Red	−75
Orange	−150
Yellow	−220
Green	−330
Blue	−470
Purple	−750

Color	Significant Figures A	B	Multiplier pF C	Tolerance% D	Voltage V E
Black	—	0	1	±20	
Brown	1	1	10		125
Red	2	2	10^2		160
Orange	3	3	10^3		250

Plastic				Filters, timing, and other high-stability applications, high quality, small low TC, tuned circuits, timing networks, stable oscillator circuits, resonance circuits and other high-performance pulse-handling applications, phase shifting, pulse applications
Kapton	63	0.1–68 μ	10	
Polyester/ Mylar				
Parylene				
Polysulfone				
Polycarbonate				
Axial	280	1–10 p	10	
Can	100	0.1–68 μ	10	
Radial				
Polypropylene				
Axial	63	100–2200 p	5	
Can	400	150–1000 p	1	
Radial	1000	1–470 n	20	
Polystyrene				
Axial	450	47–680 p	1	
Can	1000	1–470 n	20	
Radial	160	10–10,000 p	2.5	

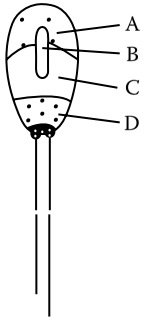


Color	Significant Figures		Multiplier pF	Tolerance%	Voltage V(D.C.)
	A	B	C	D	E
Black	0	0	1	±20	—
Brown	1	1	10		100
Red	2	2	10^2		250
Orange	3	3	10^3		—
Yellow	4	4	10^4		400
Green	5	5	10^5		—
Blue	6	6	—		630
Purple	7	7	—		—
Gray	8	8	10^{-2}		—
White	9	9	10^{-1}	±10	—

(*continued*)

TABLE 28.5 (continued) Capacitor Specifications and Applications

Capacitor Types	Typical Commercial Specifications: Voltage A.C.V	Capacitance F	Tolerance%	Applications
Electrolytic and solid				General-purpose to high-performance applications, power-supply filters, motor capacitors, switching circuits, high-voltage filter transmitters, and long-life applications
Aluminum	25	680–6800 p	20	
Tantalum				
Axial	6.3	6.8–150 μ	20	
Can	35	6.8–150 μ	20	
Radial	16	2.2–68 μ	20	



Samples/Code

Color	Significant figures A	Significant figures B	Multiplier μF C	Voltage V D
Black	—	0	1	10
Brown	1	1	—	1.6
Red	2	2	—	4
Orange	3	3	—	40
Yellow	4	4	—	6.3
Green	5	5	—	16
Blue	6	6	—	—
Purple	7	7	10^{-3}	—
Gray	8	8	10^{-2}	25
White	9	9	10^{-1}	2.5

There are also specially designed capacitors (e.g., mica, glass, oil, gas, and vacuum). These capacitors are used particularly in high-voltage (35 kV) and high-current (200 A) applications.

In the majority of cases, the manufacturing process of the capacitors begins by forming one plate using metallization of one side of a flexible dielectric film. A foil such as aluminum is used as the other plate. The film/foil combination is rolled on a suitable core with alternate layers slightly extended and then heat treated. In some cases, two-foil layers are divided by a dielectric film or paper impregnated with oil.

Generally, capacitors are two-terminal devices with one electrode as the ground terminal. However, if both terminals are separated from the common terminal, the additional capacitances between ground and electrodes might have to be taken into account. Usually, capacitance between electrodes and ground are small compared to the dominant capacitance between plates. Three-terminal capacitances exist and are manufactured in many different ways, as illustrated in Table 28.2.

As far as construction materials and construction techniques are concerned, the capacitors can broadly be classified as electrolytic, ceramic, paper, polymer, mica, variable capacitors, or integrated circuit (IC) capacitors.

Paper capacitors: Usually, paper capacitors are made with thin (5–50 μm in thickness) wood pulp. A number of sheets are used together to eliminate possible chemical and fibrous defects that may exist in each sheet. The paper sheets are placed between thin aluminum foils and convolutedly wound, as shown in Figure 28.2. The moisture of the paper is removed at high-temperature vacuum drying method before the capacitor is vacuum impregnated with oil, paraffin, or wax. The losses and self-inductance in these capacitors are sizeable and frequency dependent. The applications are usually restricted to low frequencies and high voltages. When impregnated with silicone oil, they can withstand voltages up to 300 kV.

Electrolytic capacitors: This describes any capacitor in which the dielectric layer is formed by an electrolytic method. The electrolytic capacitors in dry foil form may be similar in construction to the paper film capacitors; that is, two-foil layers separated by an impregnated electrolyte spacer and are rolled together. In this case, one of the plates is formed using metallization of one side of a flexible dielectric film. A foil (e.g., aluminum) is used as the plate. The capacitor is then hermetically sealed in aluminum or plastic can, as shown in Figure 28.3. These capacitors can be divided into two main subgroups.

Tantalum electrolytic: The anode consists of sintered tantalum powder and the dielectric is Ta_2O_5, which has a high value of ε_r. A semiconductor layer MnO_2 surrounds the dielectric. The cathode made from graphite is deposited around MnO_2 before the capacitor is sealed. The form of a tantalum electrolytic capacitor includes a porous anode slug to obtain a large active surface. These capacitors are highly stable and reliable, with good temperature ranges, and are suitable for high-frequency applications.

Aluminum electrolytic capacitors: Aluminum foil is oxidized on one side as $A1_2O_3$. The oxide layer is the dielectric having a thickness of about 0.1 μm and high electric field strength (7×10^5 V mm^{-1}). A second

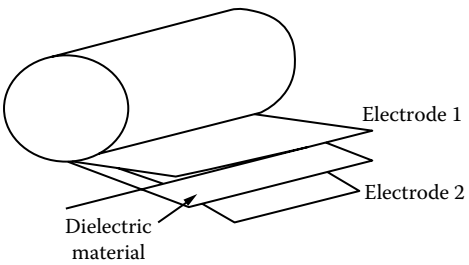


FIGURE 28.2 Construction of a typical capacitor. Dielectric material sheets are placed between electrode foils and convolutedly wound. The moisture of the dielectric material is removed at high temperatures by vacuum drying technique before the capacitor is impregnated with oil, paraffin, or wax.

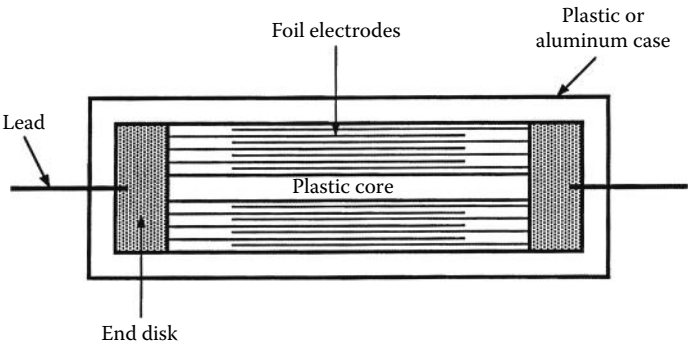


FIGURE 28.3 Construction of an electrolytic capacitor. The two-foil layer electrodes are separated by an impregnated electrolyte paper spacer and rolled together on a plastic core. Usually, a flexible metalized dielectric film is used as one of the plates and a common foil is used for the other. The capacitor is then hermetically sealed in aluminum or plastic can.

layer acting as the cathode, made from etched Al foil, is inserted. The two layers are separated by a spacer when the layers are rolled and mounted.

Electrolytic capacitors must be handled with caution, since in these capacitors, the electrolytic can be easily polarized. That is, the anode should always be positive with respect to the cathode. If not connected correctly, hydrogen gas will form; this damages the dielectric layer, causing a high leakage current or blowup. These capacitors can be manufactured in values up to 1.0 F. They are used in not so critical applications such as coupling, bypass, and filtering. However, they are not useful at frequencies above 1 kHz.

Ceramic and glass capacitors: The dielectric is a ceramic material with deposited metals. They are usually rod or disk shaped. They have good temperature characteristics and are suitable for high-frequency applications. There are many different types, such as the following: (1) *Low K ceramic*: These capacitors are made with materials that contain a large fraction of titanium dioxide (TiO_2). The relative permittivity of these materials varies from 10 to 500, with negative temperature coefficient. The dielectric is TiO_2 + MgO + SiO_2, suitable in high-frequency applications in filters, tuned circuits, coupling and bypass circuits, etc. (2) *High K ceramic*: The dielectric contains a large fraction of barium titanate, $BaTiO_3$, mixed with $PbTiO_3$ or $PbZrO_3$ giving relative permittivity of 250–10,000. They have high losses and also have high-voltage time dependence with poor stability. (3) *Miniature ceramic capacitors*: These are used in critical high-frequency applications. They are made in the ranges of 0.25 pF to 1 nF. (4) *Dielectric ceramic capacitors*: The material is a semiconducting ceramic with deposited metals on both sides. This arrangement results in two depletion layers that make up the very thin dielectric. In this way, high capacitances can be obtained. Due to thin depletion layers, only small dc voltages are allowed. They are used in small and lightweight equipment such as hearing aids.

Glass capacitors are made with glass dielectric materials. The properties of glass dielectrics are similar to ceramics.

Polymer capacitors: Various polymers, such as polycarbonate, polystyrol, polystyrene, polyethylene, and polypropylene, are used as the dielectric material. The construction is similar to that of paper capacitors. Polystyrene capacitors, in particular, are very stable and are virtually frequency independent. They have low voltage ratings and are used in transistorized applications as tuning capacitors and standard capacitors.

Mica capacitors: A thin layer of mica, usually muscovite mica ($\geq$0.003 mm), is stapled with Cu foil or coated with a layer of deposited silver. They are then vacuum impregnated and coated with epoxy. The field strength of these capacitors is very high (10^5 V mm^{-1}) and resistivity $\rho = 10^6$–10^{15} Ω m. These capacitors are available in values from 1.0 pF to several microfarads for high-voltage (from 100 to 2000 V) and high-frequency applications. They have tolerances between ±20% and ±0.5%.

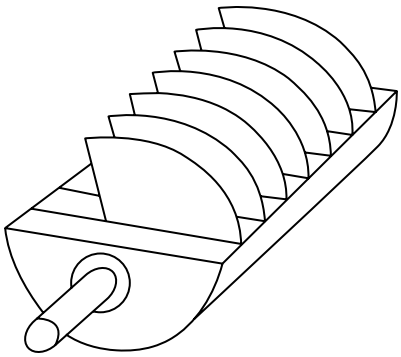


FIGURE 28.4 A variable capacitor consists of two assemblies of spaced plates positioned together by insulation members such that one set of plates can be rotated. The majority of variable capacitors have air as the dielectric. They are commonly used in adjustment of resonant frequencies of tuned circuits in receivers and transmitters. By shaping the plates suitably, they can be made to be linear or logarithmic.

Variable capacitors: These capacitors usually have air as the dielectric and consist of two assemblies of spaced plates positioned together by insulation members such that one set of plates can be rotated. A typical example of variable capacitors is given in Figure 28.4. Their main use is the adjustment of resonant frequency of tuned circuits in receivers and transmitters, filters, etc. By shaping the plates, various types of capacitances can be obtained, such as *linear capacitance*, in which capacitance changes as a linear function of rotation, and *logarithmic capacitance*.

Variable capacitors can be grouped as precision types, general-purpose types, transmitter types, trimmer types, and special types such as the phase shifters.

Precision-type variable capacitors are used in bridges, resonant circuits, and many other instrumentation systems. The capacitance swing can be from 100 to 5000 pF. They have excellent long-term stability with very tight tolerances.

General-purpose type variable capacitors are used as tuning capacitors in radio and other broadcasting devices. They are available in many laws such as straight-line frequency and straight-line wavelength. The normal capacitance swing is from 400 to 500 pF. In some cases, a swing range of 10–600 pF is available.

Transmitter-type variable capacitors are similar to general-purpose variable capacitors, but they are specially designed for high-voltage operations. The vanes are rounded and spaced wider to avoid flashover and excessive current leakages. The swing of these capacitors can go from few picofarads up to 1000 pF. In some cases, oil filling or compressed gases are used to increase operating voltages and capacitances.

Trimmer capacitors are used for coil trimming at intermediate radio frequencies. They can be air-spaced rotary types (2–100 pF), compression types (1.5–2000 pF), ceramic-dielectric rotary types (5–100 pF), and tubular types (up to 3 pF).

Sometimes, special type variable capacitors are produced for particular applications, such as differential and phase shift capacitors in radar systems. They are used for accurate measurement of time intervals, high-speed scanning circuits, transmitters and receivers, etc.

Integrated circuit capacitors: IC capacitors are designed and constructed for use in microelectronic circuits. They include miniature ceramic capacitors, tantalum oxide solid capacitors, and tantalum electrolyte solid capacitors. The ceramic and tantalum oxide IC capacitors are encapsulated and are fitted with endcaps for direct surface mounting onto circuit boards. The beam-leaded tantalum electrolytic IC capacitors are usually attached by pressure bonding. Typical values of these capacitors are 1 pF to 27 nF for temperature compensating ceramic, 100–3000 pF for tantalum oxide, 390 pF to 0.47 μF for general-purpose ceramic, and 0.1–10 μF for tantalum electrolyte. Operating voltages ranges from 25 to 200 V for ceramic, 12–35 V for tantalum electrolyte, and 12–25 V for tantalum oxide.

IC capacitors are made mostly within MOS IC technology since it is easier to deposit tantalum and other materials. The plates of IC capacitors are generally formed by two heavily doped polysilicon layers, formed on a thick layer of oxide. The dielectric is usually made from a thin layer of silicon oxide. These capacitors are temperature stable, with a temperature coefficient of about 20 ppm/°C. In addition, IC capacitive sensors can easily be constructed by incorporating dielectrics sensitive to physical variables. Usually, the metallization layer formed on top of the dielectric forms a shape to provide access to measured physical variable to the dielectric.

Voltage variable capacitors: These capacitors make use of the capacitive effect of the reversed-biased p–n junction diode. By applying different reverse-bias voltages to the diode, the capacitance can be changed. Hence, the name varicap or varactor diodes are given to these devices. Varactors are designed to provide various capacitance ranges from a few picofarads to more than 100 pF. For improved performances, it is also possible to make use of high-speed switching silicon diodes as voltage variable capacitors. However, they are limited by the very low maximum capacitance availability. Typical applications of these varactor diodes are in the tuning circuits in radio frequency receivers. Present-day varactor diodes operate into the microwave part of the spectrum. These devices are quite efficient as frequency multipliers at power levels as large as 25 W. The efficiency of a correctly designed varactor multiplier can exceed 50% in most cases. It is also worth noting that some Zener diodes and selected silicon power-supply rectifier diodes can work effectively as varactors at frequencies as high as 144 MHz. In the case of the Zener diode, it should be operated below its reverse breakdown voltage.

28.3 Characteristics of Capacitors

Capacitors are characterized by dielectric properties, breakdown voltages, temperature coefficients, insulation resistances, frequency and impedances, power dissipation and quality factors, reliability, and aging properties. Typical characteristics of common capacitors are given in Tables 28.4 and 28.5. Some of the important characteristics will be explained next.

Dielectric properties: Dielectrics of capacitors can be made from polar or nonpolar materials. Polar materials have dipolar characteristics; that is, they consist of molecules whose ends are oppositely charged. This polarization causes oscillations of the dipoles at certain frequencies, resulting in high losses.

Capacitor properties are largely determined by the dielectric properties. For example, the losses in the capacitors occur due to the current leakage and the dielectric absorption. These losses are frequency dependent, as typified by Figure 28.5.

The dielectric absorption introduces a time lag during the charging and discharging of the capacitor, thus reducing the capacitance values at high frequencies and causing unwanted time delays in circuits. The leakage current, on the other hand, prevents indefinite storage of energy in the capacitor.

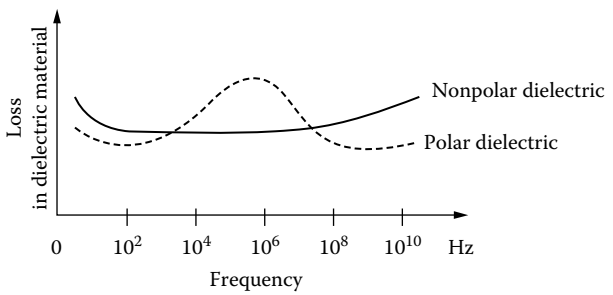


FIGURE 28.5 The frequency dependence of dielectric losses. The dielectric material of capacitors can be polar or nonpolar. In polar materials, polarization causes oscillations at certain frequencies, resulting in high losses at those frequencies. The dielectric losses introduce a time lag during the charging and discharging of the capacitors, thus reducing the capacitance values.

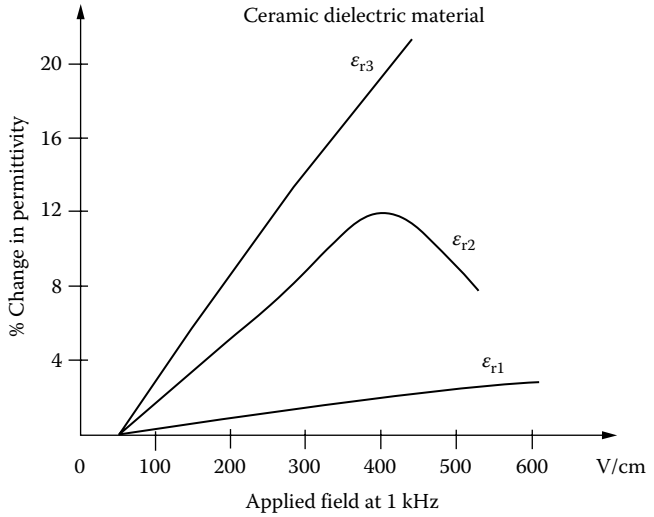


FIGURE 28.6 Changes in relative permittivity versus field strength. The dielectric strength depends on the temperature, frequency, and applied voltage. Increases in the applied voltage cause higher changes in the dielectric strength. If the capacitor is subjected to high operating voltages, the electric field in the dielectric exceeds the breakdown value that can damage the dielectric permanently.

An associated parameter to leakage currents is the leakage resistance, which is measured in megohms, but usually expressed in megohm–microfarads or ohms–farads. The leakage resistance and capacitance introduces time constants that can vary from a few days for polystyrene to several seconds in some electrolytic capacitors. It is important to mention that the leakage current does not only depend on the properties of the dielectric materials but also depends on the construction and structural integrity of capacitors. This is particularly true for capacitors having values less than 0.1 μF, having very thin dielectric materials between the electrodes.

Breakdown voltage: If the capacitor is subjected to high operating voltages, the electric field in the dielectric exceeds the breakdown value, which damages the dielectric permanently. The dielectric strength, which is the ability to withstand high voltages without changing properties, depends on the temperature, frequency, and the applied voltage. An example of this dependence on the applied voltage is given in Figure 28.6. It is commonly known that the use of capacitors below their rated values increases the reliability and the expected lifetime. The standard voltage ratings of most capacitors are quoted by the manufacturers as 50, 100, 200, 400, and 600 V. Tantalum and electrolytic capacitors have ratings of 6, 10, 12, 15, 20, 25, 35, 50, 75, 100 V, and higher.

Usually, values for surge voltages are given to indicate the ability of capacitors to withstand high transients. Typically, the surge voltage for electrolytic capacitors is 10% above the rated voltage, 50% for aluminum capacitors, and about 250% for ceramic and mica capacitors.

The rated reverse voltages of electrolytic capacitors are limited to 1.5 V and, in some cases, to 15% of the rated forward voltages.

Temperature coefficient: The temperature characteristics of capacitors largely depend on the temperature properties of the dielectric materials used, as given in Figure 28.7. The temperature coefficients of glass, Teflon, mica, and polycarbonate are very small, whereas in ceramic capacitors, they can relatively be much higher.

Insulation resistance: The insulation resistance of capacitors is important for effective operations in many circuits since the insulation resistance is very susceptible to temperature and humidity. For example, unsealed capacitors show large and rapid changes against temperature and humidity,

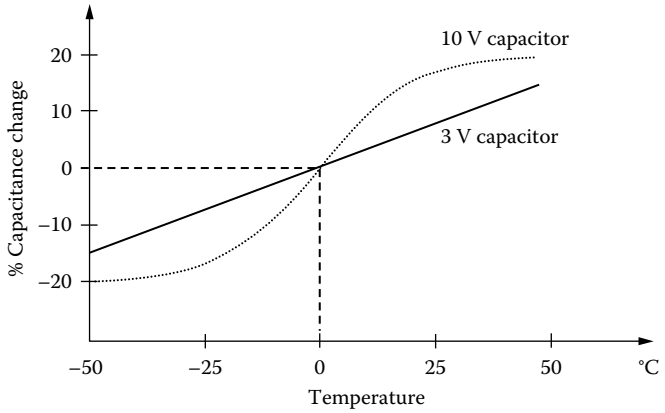


FIGURE 28.7 The temperature dependence of capacitors. The temperature characteristics of capacitors are largely dependent on the temperature properties of the dielectric materials used. The variations in capacitances due to temperature are also dependent on the type of capacitor and the operational voltages. The temperature coefficient of glass, Teflon, mica, and polycarbonate are very small and relatively high in ceramic capacitors.

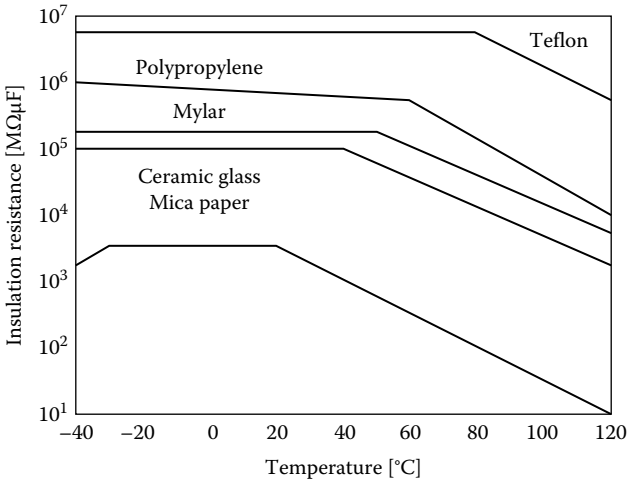


FIGURE 28.8 Temperature dependence of insulation resistance. The insulation resistance of many capacitors is not affected at low temperatures. However, under high-temperature conditions, the change in insulation resistance can be approximated by an exponential relation. The insulation resistance is also susceptible to variations in humidity.

which can make them unsuitable for many circuits. For most capacitors, the change in insulation resistance is an exponential function of temperature ($R_{T1} = R_{T2}e^{K(T1-T2)}$). This can be substantial in high-temperature operations. The temperature dependence of insulation resistance of common capacitors is shown in Figure 28.8.

Frequency and impedance: Practical capacitors have increases in losses at very low and at very high frequencies. At low frequencies, the circuit becomes entirely resistive and the dc leakage current becomes effective. At very high frequencies, the current flowing through the capacitor become important and the dielectric losses increase. Approximate ranges of useable frequencies of capacitors are provided in Tables 28.4 and 28.5.

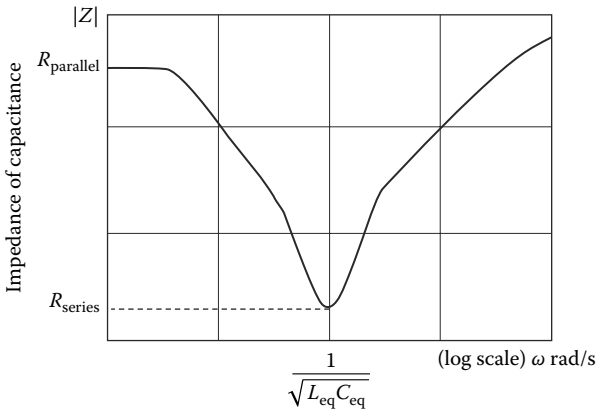


FIGURE 28.9 Frequency and impedance relation of capacitors. The losses and inherent inductance affects the ideal operation of capacitors; hence, the impedance becomes a function of the frequency. Depending on the construction, all capacitors will resonate at some frequency.

An ideal capacitor should have an entirely negative reactance, but losses and inherent inductance prevent ideal operation. Depending on the construction, capacitors will resonate at certain frequencies due to unavoidable construction-based inductances. A typical impedance characteristic of a capacitor is depicted in Figure 28.9.

Power–dissipation and quality factors: Ideally, a capacitor should store energy without dissipating any power. However, due to equivalent resistances, R_{eq}, real power will be dissipated. Therefore, the power factor of a capacitor can be expressed as

$$\text{PF} = \cos\theta = \frac{R_{eq}}{Z_{eq}} \tag{28.7}$$

where
θ is the phase angle
Z_{eq} is the equivalent total impedance

An important characteristic, the *dissipation factor* of capacitors, is expressed as

$$\text{DF} = \tan\delta = \frac{R_{eq}}{X_{eq}} \tag{28.8}$$

where
δ is the angle of loss
X_{eq} is the equivalent reactance

The dissipation factor depends on the frequency. Capacitors are designed such that this dependence is minimal. The measurement of dissipation factor δ is made at 1 kHz and 1.0 Vrms applied to the capacitor. A typical dissipation factor curve is depicted in Figure 28.10.

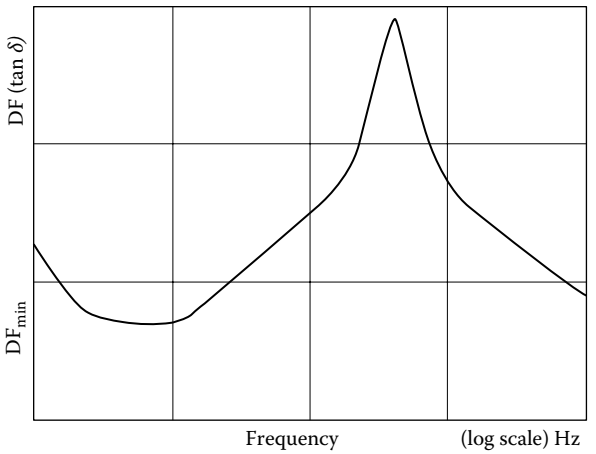


FIGURE 28.10 Power–dissipation factors. In ideal operations, capacitors should store energy without dissipating power. Nevertheless, due to resistances, some power will be dissipated. The dissipation depends on the frequency. The standard measurement of dissipation factor δ is determined by applying a voltage of 1.0 Vrms at 1 kHz.

28.4 Selection, Reliability, and Standards of Capacitors

28.4.1 Capacitor Selection

Experience shows that a substantial part of component failures in electronic equipment is due to capacitors. The major cause of this can be attributed to the improper selection and inappropriate applications of capacitors. The following factors are therefore important criteria in the selection of capacitors in circuit applications: (1) the capacitance *values and tolerances* are determined by operating frequencies and by requirements of timing, energy storage, phase shifting, coupling, or decoupling; (2) the *voltages* are determined by the type and nature of the source, ac, dc, transient, surges, and ripples; (3) the *stability* is determined by operating conditions like temperature, humidity, shock, vibration, and life expectancy; (4) the *electric properties* are determined by life expectancy, leakage current, dissipation factor, impedance, and self-resonant frequency; (5) the mechanical properties are determined by the types and construction, for example, size, configuration, and packaging; and (6) the cost is determined by the types and physical dimensions of capacitors and the required tolerance.

28.4.2 Capacitor Reliability

Some of the common causes of capacitor failure are due to voltage and current overloads, high temperature, humidity, shock, vibration, pressure, frequency effects, and aging. The voltage overload produces an excessive electric field in the dielectric that results in the breakdown or total destruction of the dielectric. The electric current overload caused by rapid voltage variations results in current transients. If these currents are of sufficient amplitude and duration, the dielectric can be deformed and damaged, resulting in drastic changes in capacitance values, and thus leading to malfunctioning of the equipment. The high temperatures are mainly due to voltage and current overloads. Overheating and high temperatures accelerate the dielectric aging. This causes the plastic film to be brittle and also introduces cracks in the hermetic seals. The moisture and humidity due to severe operating environments can cause corrosions, reducing the dielectric strength, and lowering the insulation resistance. The mechanical effects are mainly the pressure, variation, shock, and stress, which can cause mechanical damages of seals that

TABLE 28.6 Standard Capacitors and Tolerances

Value (pF, nF, μF) Tolerance	5%	10%	20%
10	x	x	x
11	x		
12	x		
13	x		
15	x	x	x
16	x		
18	x	x	
20	x		
22	x	x	x
24	x		
27	x	x	
30	x		
33	x	x	x
36	x		
39	x		
43	x		
47	x	x	x
51	x		
56	x	x	x
62	x		
68	x	x	x
75	x		
82	x		
91	x		

result in electric failures. Aging deteriorates the insulation resistance and affects the dielectric strength. The aging is usually determined by shelf life; information about aging is supplied by the manufacturers.

28.4.3 Capacitor Standard Values and Tolerances

General-purpose capacitor values tend to be grouped close to each other in a bimodal distribution manner within their tolerance values. Usually, the tolerances of standard capacitors are 5%, 10%, and 20% of their rated values, as shown in Table 28.6. Nevertheless, tolerances of precision capacitors are much tighter—in the range of 0.25%, 0.5%, 0.625%, 1%, 2%, and 3% of the rated values. Undoubtedly, these capacitors are much more expensive than the general-purpose range.

For capacitors in the small pF range, the standard tolerances are ±1.5, ±1, ±0.5, ±0.25, and ±0.1 pF. Usually, low tolerance ranges are achieved by selecting good-quality materials and elaborate manufacturing.

Standard capacitors are constructed from interleaved metal plates using air as the dielectric material. The area of the plates and distance between them are determined and constructed with precision. National Bureau of Standards maintains a bank of primary standard air capacitors that can be used to calibrate the secondary and working standards in laboratory and industrial application. The working standards of small capacitances are obtained by using air capacitors, whereas working standards of larger capacitances are made from solid dielectric materials. Usually, silver-mica capacitors are selected as working standards since they are very stable, exhibit low dissipation factors and small temperature coefficients, and aging effects are excellent. They are available in decade mounted forms with multiplication factors of 10.

28.5 Capacitors in Circuits

The capacitor is used as a two-terminal element in electric circuits with the current–voltage relationship given by

$$i(t) = \frac{C\,dv(t)}{dt} \tag{28.9}$$

where C is the capacitance.

The circuit representation is given in Figure 28.11a.

From the $v(t)$, $i(t)$ relationship, the instantaneous power of this element can then be given by

$$p(t) = v(t)i(t) \tag{28.10}$$

The stored energy in the capacitor at time t s can be calculated as

$$w(t) = \int Cv(t)\left\{\frac{dv(t)}{dt}\right\}dt$$

$$= \frac{[Cv^2(t)]}{2} \tag{28.11}$$

If the voltage across the capacitor is time varying, the energy stored in the capacitor in a time interval $t_1 - t_2$ can be calculated from Equation 28.11 as

$$W = \left(\frac{1}{2}\right)C[v^2(t_2) - v^2(t_1)] \tag{28.12}$$

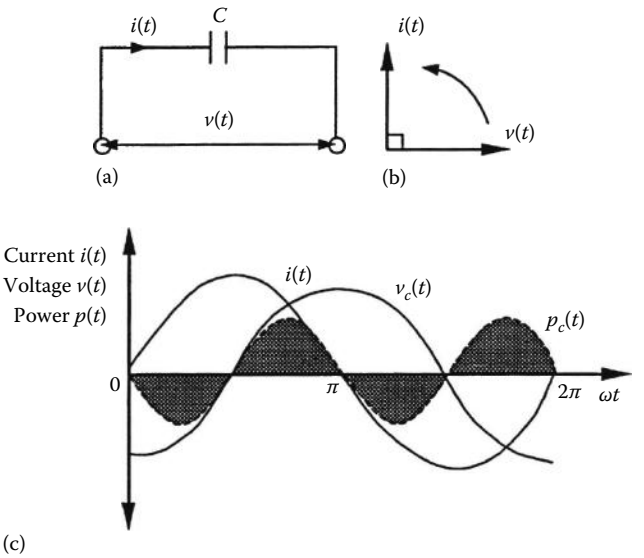


FIGURE 28.11 A capacitor as a two-terminal circuit element: (a) connection of a capacitor in electric circuits, (b) voltage–current relationship under sinusoidal operations, and (c) the power, voltage, and current relationships. The power has positive and negative values with twice the frequency of the applied voltage.

Although the voltage assumes different values in time interval $t_1 - t_2$, if the initial and final voltages are equal (e.g., $v(t_1) = v(t_2)$), then the net energy stored in the capacitor will be equal to zero. This implies that the energy stored in the capacitor is returned to the circuit during that time interval; that is, the capacitor transforms the energy without dissipation. The stored energy in the capacitor is associated with the electrostatic field and, in the absence of the electrostatic field, the energy will be zero.

From the voltage–current relationship in Equation 28.9, it can be seen that a capacitor is a passive element. That is, if the voltage is known, the current can be immediately determined by differentiating the voltage. Conversely, if the current is known, then the voltage can be determined by integration. If the voltage $v(t)$ is considered as an input variable, and the current $i(t)$ as an output variable, then the behavior of the current in a certain time interval is completely determined by the behavior of the voltage in this interval. In this case, the solution of the differential equation has the forced component only. Here, the particular solution of the voltage–current differential equation coincides with a full solution and the Laplace transform gives the relationship

$$I(s) = sCV(s) \tag{28.13}$$

From this, the input–output relationship yields the impedance of the capacitor as

$$Z(s) = \frac{1}{sC}$$

or (28.14)

$$Y(s) = sC$$

Under the stationary conditions, $s \to 0$, $Z \to \infty$, and $Y \to 0$.

In the sinusoidal condition, $s = j\omega = 2\pi f$, where f is the frequency and, hence,

$$Z(j\omega) = \frac{1}{\{j\omega C\}} = \frac{-j}{\{\omega \mathrm{C}\}} \tag{28.15}$$

and

$$Y(j\omega) = j\omega C \tag{28.16}$$

The capacitor can then be characterized under sinusoidal conditions, by a reactance of $X_C = 1/\omega C$, measured in ohms with the current leading the voltage by 90°, as shown in the phasor diagram in Figure 28.11b.

In sinusoidal operations, the instantaneous power $p(t) = v(t)i(t)$ can be calculated as

$$v(t) = V_{\max} \cos \omega t = \sqrt{2} V \cos \omega t \tag{28.17}$$

Using the relationship given by Equation 28.9, the current can be written as

$$i(t) = \frac{C\,dv}{dt} = -\omega C \sqrt{2} V \sin \omega t \tag{28.18}$$

giving

$$p(t) = v(t)i(t) = -2\omega CV^2 \sin\omega t \cos\omega t = \frac{V^2 \sin 2\omega t}{X_C} \tag{28.19}$$

This indicates that the average power is zero because of the $\sin 2\omega t$ term, but there is a periodic storage and return of energy and the amplitude of that power is V^2/X_C. The power, voltage, and current relationships in a capacitor are given in Figure 28.11c.

28.5.1 Series and Parallel Connections

The formulae for series and parallel connection of capacitors can be obtained from the impedances of such connections as shown in Figure 28.12a and b. For the series connection, the impedances are added such that

$$\frac{1}{sC} = \frac{1}{sC_1} + \frac{1}{sC_2} + \cdots + \frac{1}{sC_n} \tag{28.20}$$

where C_1, C_2, …, C_n are the capacitance of the capacitors connected in series.

Thus, the equivalent value of the capacitances in series connection can be expressed by

$$C = \left\{ \frac{1}{C_1} + \frac{1}{C_2} + \cdots + \frac{1}{C_n} \right\}^{-1} \tag{28.21}$$

The final capacitance value will always be smaller than the smallest value.

In a similar way, the equivalent capacitance of parallel connections can be expressed as

$$C = C_1 + C_2 + \cdots + C_n \tag{28.22}$$

and the final value of C is always larger than the largest capacitance in the circuit.

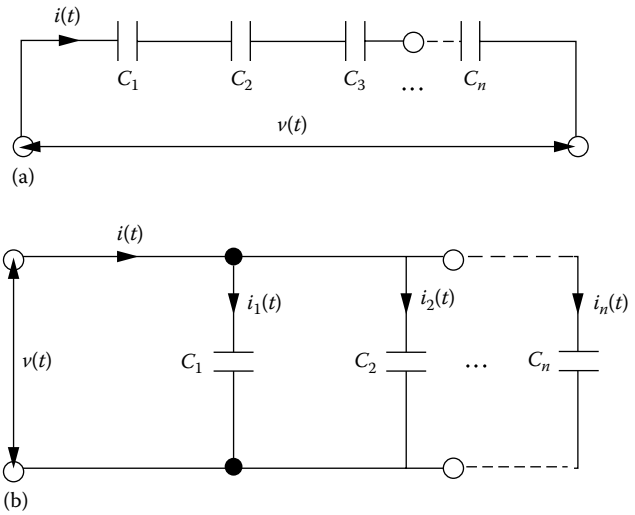


FIGURE 28.12 Series and parallel connection of capacitors: (a) series connection and (b) parallel connection. In the series connection, the final capacitance value will always be smaller than the smallest value of the capacitor in the circuit, whereas in parallel connection, the final value is greater than the largest capacitance.

28.5.2 Distributed Capacitance Circuits

Capacitance is inherent whenever an electric potential exists between two conducting bodies, albeit in electric circuits, communication equipment, electronic circuit boards, computer devices, and so on. For example in electric circuits, its effect will be most noticeable in the coils and in the transmission lines. In the case of coils, there are small capacitances between adjacent turns, between turns that are not adjacent, between terminal leads, between turns and ground, etc. Each of the various capacitances associated with the coil stores a quantity of electrostatic energy that is determined by the capacitance involved and the fraction of the total coil voltage that appears across it. The total effect is that the numerous small coil capacitances can be replaced by a single capacitor of appropriate size shunted across the coil terminals. This equivalent capacitance is called either the *distributed capacitance* or the *self-capacitance* of the coil, and it causes the coil to show parallel resonance effects under some conditions. In the case of a mismatched or unterminated transmission line, the distributed capacitance, together with the inductive effects, will create a phase difference between the voltage and current in the line. This phase difference depends on the type of termination and the electric length of the line and, as a result, the input impedance of the line can effectively be an equivalent capacitor when its electric length is less than a quarter wavelength for an open-circuit termination or between a quarter wavelength and half a wavelength for a short-circuit termination.

28.5.3 Capacitor Equivalent Circuits

The electric equivalent circuit of a capacitor consists of a pure capacitance (C_p), plate inductances (L_1, L_2), plate resistances (R_1, R_2), and a parallel resistance R_p that represents the resistance of the dielectric or leakage resistance, as shown in Figure 28.13. The capacitors that have high leakage currents flowing through the dielectric have relatively low R_p values. Very low leakage currents are represented by extremely large R_p values. Examples of these two extremes are electrolytic capacitors that have high leakage current (low R_p) and plastic film capacitors, which have very low leakage current (high R_p). Typically, an electrolytic capacitor might easily have several microamperes of leakage current ($R_p < 1$ MΩ), while a plastic film capacitor could have a resistance greater than 100,000 MΩ

It is usual to represent a low leakage capacitor (high R_p) by a series resistance–capacitance (*RC*) circuit, while those with high leakage (low R_p) are represented by parallel *RC* circuits. However, when a capacitor is measured in terms of the series *C* and *R* quantities, it is desirable to resolve them into parallel equivalent circuit quantities. This is because the (parallel) leakage resistance best represents the quality of the capacitor dielectric.

28.5.4 Capacitive Bridges and Measurement of Capacitance

Bridges are used to make precise measurements of unknown capacitances and associated losses in terms of known external capacitances and resistances. The most commonly used bridges are series *RC* bridge, parallel *RC* bridge, Wien bridge, and Schering bridge.

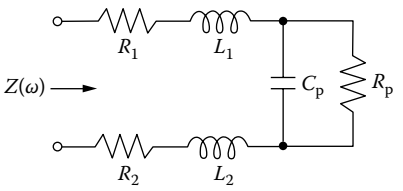


FIGURE 28.13 The capacitor equivalent circuit. A practical capacitor has resistance and inductances. Often, the electric equivalent circuit of a capacitor can be simplified by a pure capacitance C_p and a parallel resistance R_p by neglecting resistances R_1, R_2 and inductances L_1, L_2. In low leakage capacitors where R_p is high, the equivalent circuit can be represented by a series *RC* circuit.

28.5.4.1 Series Resistance–Capacitance Bridge

Figure 28.14 is a series *RC* bridge, which is used for the comparison of a known capacitance with an unknown capacitance. The unknown capacitance is represented by C_x and R_x. A standard adjustable resistance R_1 is connected in series with a standard capacitor C_1. The voltage drop across R_1 balances the resistive voltage drop when the bridge is balanced. The additional resistor in series with C_x increases the total resistive component, so that small values of R_1 will not be required to achieve balance. Generally, the bridge balance is most easily achieved when capacitive branches have substantial resistive components. To obtain balance, R_1 and either R_3 or R_4 are adjusted alternately. This type of bridge is found to be most suitable for capacitors with a high-resistance dielectric and hence very low leakage currents.

At balance,

$$Z_1 Z_3 = Z_2 Z_x \tag{28.23}$$

Substituting impedance values gives

$$\left(R_1 - \frac{j}{\omega C_1}\right) R_3 = \left(R_x - \frac{j}{\omega C_2}\right) R_2 \tag{28.24}$$

Equating the real terms gives

$$R_x = \frac{R_1 R_3}{R_2} \tag{28.25}$$

and equating *imaginary* terms gives

$$C_x = \frac{C_1 R_2}{R_3} \tag{28.26}$$

An improved version of the series *RC* bridge is the *substitution bridge*, which is particularly useful to determine the values of capacitances at radio frequencies. In this case, a series-connected *RC* bridge is balanced by disconnecting the unknown capacitance C_x and resistance R_x and replacing it by an adjustable standard capacitor C_s and adjustable resistor R_s. After having obtained the balance position, the unknown capacitance and resistance C_x an R_x are connected in parallel to the capacitor C_s.

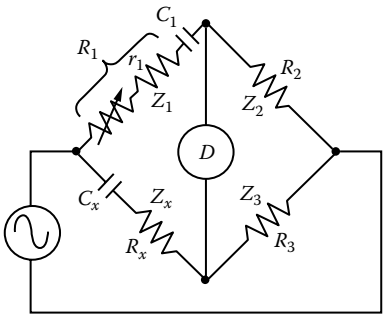


FIGURE 28.14 A series *RC* bridge. In these bridges, the unknown capacitance is compared with a known capacitance. The voltage drop across R_1 balances the resistive voltage drop in branch Z_2 when the bridge is balanced. The bridge balance is most easily achieved when capacitive branches have substantial resistive components. The resistors R_1 and either R_3 or R_4 are adjusted alternately to obtain the balance. This type of bridge is found to be most suitable for capacitors with high-resistance dielectrics; hence, they have very low leakage currents.

The capacitor C_s and resistor R_s are adjusted again for the rebalance of the bridge. The changes in the ΔC_s and ΔR_s lead to unknown values as

$$C_x = \Delta C_s \quad \text{and} \quad R_x = \Delta R_s \left(\frac{C_{s1}}{C_x} \right)^2 \tag{28.27}$$

where C_{s1} is the value of C_s in the initial balance condition.

28.5.4.2 Parallel Resistance–Capacitance Bridge

Figure 28.15 illustrates a parallel *RC* bridge. In this case, the unknown capacitance is represented by its parallel equivalent circuit C_x in parallel with R_x. The Z_2 and Z_3 impedances are pure resistors with either or both being adjustable. The Z_1 is balanced by a standard capacitor C_1 in parallel with an adjustable resistor R_1. The bridge balance is achieved by adjustment of R_1 and either R_2 or R_3. The parallel *RC* bridge is found to be most suitable for capacitors with a low-resistance dielectric, hence relatively high leakage currents. At balance,

$$\frac{1}{((1/R_1) + j\omega C_1)} R_3 = \frac{1}{((1/R_x) + j\omega C_x)} R_2 \tag{28.28}$$

Equating *real* terms gives

$$R_x = \frac{R_3 R_1}{R_2} \tag{28.29}$$

and equating *imaginary* terms gives

$$C_x = \frac{C_1 R_2}{R_3} \tag{28.30}$$

28.5.4.3 Wien Bridge

Figure 28.16 shows a Wien bridge. This is a special resistance-ratio bridge that permits two capacitances to be compared once all the resistances of the bridge are known. At balance, it can be proven that the unknown resistance and the capacitance are

$$R_x = \frac{R_3(1 + \omega^2 R_1^2 C_1^2)}{\omega^2 R_1 R_2 C_1^2} \tag{28.31}$$

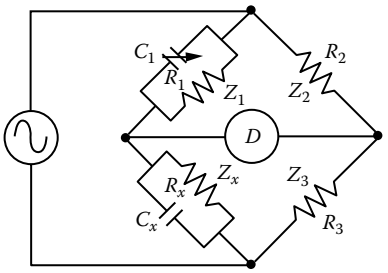


FIGURE 28.15 A parallel *RC* bridge. The unknown capacitance is represented by its parallel equivalent circuit, C_x in parallel with R_x. The bridge balance is achieved by adjustment of R_1 and either R_3 or R_4. The parallel *RC* bridge is found to be most suitable for capacitors with a low-resistance dielectric, hence relatively high leakage currents.

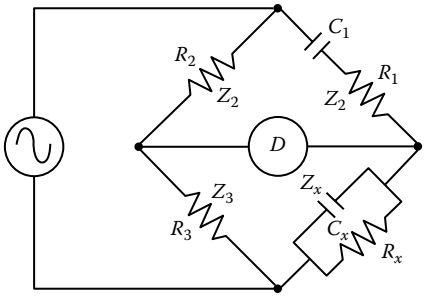


FIGURE 28.16 The Wien Bridge. This bridge is used to compare two capacitors directly. It finds applications particularly in determining the frequency in *RC* oscillators. In some cases, capacitors C_1 and C_x are made equal and ganged together so that the frequency at which the null occurs varies linearly with the capacitance.

and

$$C_x = \frac{C_1 R_2}{[R_3(1+\omega^2 R_1^2 C_1^2)]} \tag{28.32}$$

It can also be shown that

$$\omega^2 = \frac{1}{R_1 C_1 R_x C_x} \tag{28.33}$$

As indicated in Equation 28.33, the Wien Bridge has an important application in determining the frequency in *RC* oscillators. In frequency meters, C_1 and C_x are made equal, and the two capacitors are ganged together so that the frequency at which the null occurs varies linearly with capacitances.

28.5.4.4 Schering Bridge

Figure 28.17 illustrates the configuration of the Schering Bridge. This bridge is used for measuring the capacitance, the dissipation factors, and the loss angles. The unknown capacitance is directly proportional to the known capacitance C_3. That is, the bridge equations are

$$C_x = \frac{C_3 R_2}{R_1} \quad \text{and} \quad R_x = \frac{C_2 R_1}{C_3} \tag{28.34}$$

Usually, R_2 and R_3 are fixed and C_2 and C_3 are made variable. Schering bridges are frequently used in high-voltage applications with high-voltage capacitor C_3. They are also used as high-frequency bridges since the use of two variable adjustment capacitors are convenient for precise balancing.

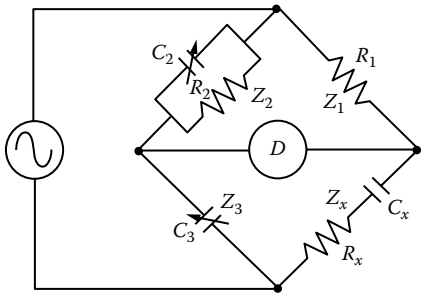


FIGURE 28.17 The Schering Bridge. This bridge is particularly useful for measuring the capacitance, associated dissipation factors, and the loss angles. The unknown capacitance is directly proportional to the known capacitance C_1. The Schering bridge is frequently used as a high-voltage bridge with a high-voltage capacitor as C_1.

Partial List of Manufacturers and Suppliers

Capacitor & Components LLC
2451 McMullen Booth Rd. Suite 248, Clearwater, FL 33759
Phone: 727-939-1932
Fax: 727-939-1900
www.capacitorsupplier.com/index.htm

Captor Corp.
5040 S. County Rd. 25-A, Tipp City, OH 45371
Phone: 937-667-8484
Fax: 937-667-5133
www.captorcorp.com

Condenser Products, A Custom Capacitors Inc. Co.
2131 Broad St., US Hwy. 41 S. Brooksville, FL 34604
Phone: 352-796-3561, 888-598-0957 (toll free)
Fax: 352-799-0221
www.condenser.com/

Core Electronics
5001 Belle Ave, Cypress, CA 90630
Phone: 714-484-0132
Fax: 714-484-0377
www.corelectronics.com

Cougar Electronics Corp.
10 Lyman St. New Haven, CT 06511
Phone: 203-562-6545, 866-497-6682 (toll free)
Fax: 203-562-6748
www.cougarelectronics.com/

Custom Electronics, Inc.
87 Browne St. Oneonta, NY 13820
Phone: 607-432-3880, 877-735-9234 (toll free)
Fax: 607-432-3913
www.customelec.com

Daitron
27520 S.W. 95th Ave. Wilsonville, OR 97070-3500
Phone: 503-682-7560, 888-324-8766 (toll free)
Fax: 503-682-2861
www.daitron.com

Dearborn Electronics, Inc.
1221 N. Highway 17 Longwood, FL 32750
Phone: 407-695-6562
www.dearbornelectronics.com

ebm-papst, Inc.
100 Hyde Rd. Farmington, CT 06034
Phone: 860-674-1515
Fax: 860-674-8536
www.ebmpapst.us

Electronic Expediters, Inc.
3700 Via Pescador, Camarillo, CA 93012
Phone: 805-987-7171
Fax: 805-987-3344
www.expediters.com/EParts.jsp

Electronic Film Capacitors
41 Interstate Ln, Waterbury, CT 06705
Phone: 203-755-5629
Fax: 203-755-0659
www.filmcapacitors.com

Electronic Supply Chain Solutions, Inc.
3023 Eastland Blvd., Bldg. H-109, Clearwater, FL 33761
Phone: 727-723-8255
Fax: 727-723-9246
electronicsupplychainsolutions.com/

GE Energy
8800 E. 63rd St, Raytown, MO 64133
Phone: 800-391-0696
www.ge-energy.com/prod_serv/products/mem…

Grainger Industrial Supply
100 Grainger Pkwy., Lake Forest, IL 60045
Phone: 847-535-1000, 888-361-8649 (toll free)
Fax: 847-535-9123
www.grainger.com/thomasnet

High Energy Corp.
Lower Valley Rd., P.O. Box 308, Parkesburg, PA 19365
Phone: 610-593-2800
Fax: 610-593-2985
www.highenergycorp.com

Hindley Electronics, Inc.
715 Innovation Dr., Bowling Green, OH 43402
Phone: 419-352-7110, 888-352-7110 (toll free)
Fax: 419-352-0628
www.hindleyelectronics.com/

Illinois Capacitor, Inc.
3757 W. Touhy Ave., Lincolnwood, IL 60712
Phone: 847-675-1760
Fax: 847-673-2850
www.illinoiscapacitor.com

Keytroller, Inc.
3907 W. Martin Luther King Blvd., Tampa, FL 33614
Phone: 813-877-4500
Fax: 813-871-6250
www.keytroller.com

Knight Electronics International, Inc.
31 Myrtle St., Manhasset, NY 11030
Phone: 516-365-3830
Fax: 516-365-3830
knightelectronicsinternational.com/index…

Mag Cap Engineering, LLC
222 Bolivar St., Canton, MA 02021
Phone: 781-821-2300, 866-340-2300 (toll free)
Fax: 781-821-2111
www.magcap.com

Marsh Electronics, Inc.
1563 S. 101st St., Milwaukee, WI 53214
Phone: 414-475-6000, 800-236-8327 (toll free)
Fax: 414-771-2847
www.marshelectronics.com

Meco Motors
22L Commerce Rd. Fairfield, NJ 07004
Phone: 877-308-3728 (toll free)
Fax: 973-808-6670
mecomotors.rtrk.com/?scid=2099227&rl_alt…

Mini-Systems, Inc.
P.O. Box 69, North Attleboro, MA 02761
Phone: 508-226-2111
Fax: 508-226-2211
www.mini-systemsinc.com

Myron Zucker, Inc.
36825 Metro Ct., Sterling Heights, MI 48312
Phone: 586-979-9955, 800-245-0583 (toll free)
Fax: 586-979-9484
www.myronzucker.com/

OptiXtal, Inc.
1901 S. 54th St., Philadelphia, PA 19143
Phone: 267-257-8139
Fax: 215-729-1380
www.optixtal.com

Oren Elliott Products, Inc.
128 W. Vine St., Edgerton, OH 43517-9600
Phone: 419-298-2306, 888-771-8290 (toll free)
Fax: 419-298-3545
www.orenelliottproducts.com

Oriental Motor U.S.A. Corp.
1001 Knox St., Torrance, CA 90502
Phone: 310-715-3301, 800-418-7903 (toll free)
Fax: 310-225-2594, 800-309-7999 (toll free)
www.orientalmotor.com

Oxley, Inc.
31 Business Park Dr., Branford, CT 06405-2977
Phone: 203-488-1033
Fax: 203-481-6971
www.oxleygroup.com/

Radio Research Instrument Co.
584 N. Main St. Waterbury, CT 06704-3506
Phone: 203-753-5840
Fax: 203-754-2567
www.radiores.com

Schneider Electric
1300 South Wolfe Rd., Des Plaines, IL 60018
Phone: 847-441-2526
Fax: 847-441-2522
www.schneider-electric.com/site/home/ind…

Smith & Loveless, Inc.
14040 Santa Fe Trail Dr., Lenexa, KS 66215
Phone: 913-888-5201, 800-898-9122 (toll free)
Fax: 913-888-2173
www.smithandloveless.com

Spectrum Advanced Specialty Products
8031 Avonia Rd., Fairview, PA 16415
Phone: 814-474-2207
Fax: 814-474-2208
www.specemc.com/

STK Electronics, Inc.
2747 Us Route 20, Cazenovia, NY 13035-9481
Phone: 315-655-2530
Fax: 315-655-3862
www.stkelectronics.com

Tecate Group
7520 Mission Valley Rd. San Diego, CA 92108-4400
Phone: 619-398-9700
Fax: 619-398-9777
www.tecategroup.com

TDK Components USA Inc.
1-T TDK Blvd., Peachtree City, GA 30269
Phone: 770-631-0410
Fax: 770-487-1460
www.tdk.com

TriCom, Inc.
359 E. Madison Ave., Clifton Heights, PA 19018
Phone: 610-259-7400, 800-347-8066 (toll free)
Fax: 610-259-4785
www.tricom-sales.com

Vectronix
105 Enterprise Dr Ste F, Cumming, GA 30040
Phone: 866-244-7778 (toll free)
Fax: 770-205-7848
www.look4ic.com

World Electronics
37 Hanover Pl., Glen Rock, NJ 07452-2705
Phone: 201-670-1177
Fax: 201-447-1633
www.worldtubes.com

Bibliography

Capacitor Types, Radio Electronics.com, http://www.radio-electronics.com/info/data/capacitor/capacitor_types.php, accessed on September 11, 2011.

Laughton, M.A. and Warne, D.F. *Electrical Engineers Reference Book*, 6th edn., Elsevier, Oxford, U.K., 2003.

Lin, D.Y., Wu, J.D., Chang, Y.J., and Wu, S.J., Practical and simple circuitry for the measurement of small capacitance, *Review of Scientific Instruments*, 78(1), 14703-1-5, 2007.

Magin, R.L., Fractional order models of capacitor and dielectric behavior, *Critical Review in Biomedical Engineering*, 32(2), 124–153, 2004.

Mukherjee, S.P., Borland, W., and Suh, S., High capacitance density ceramic thin films on metal foils for embedded capacitor applications: A review, *Proceedings—Electrochemical Society*, PV 2005–13, 2006.

Schulz, A.L., *Capacitors: Theory, Types and Applications (Electronic Engineering Development)*, Nova Science Publication, New York, 2010.

Sharma, P. and Bhatti, T.S., A review on electrochemical double-layer capacitors, *Energy Conversion and Management*, 51(12), 2901–2912, 2010.

Sinclair, I., *Practical Electronics Handbook*, 5th edn., Newness, Oxford, U.K., 2001.

Types of Capacitors, Electronics tutorials, http://www.electronics-tutorials.ws/capacitor/cap_2.html, accessed on September 15, 2011.

Wang, Y. and Lee, L., A digitally programmable capacitance standard, *Review of Scientific Instruments*, 75(4), 1158–1160, 2004.

29

Permeability and Hysteresis Measurement

Jeff P. Anderson
LTV Steel Corporation

Richard J. Blotzer
LTV Steel Corporation

Magnetic fields are typically conceptualized with so-called flux lines or lines of force. When such flux lines encounter any sort of matter, an interaction takes place in which the number of flux lines is either increased or decreased. The original magnetic field therefore becomes amplified or diminished in the body of matter as a result of the interaction. This is true whether the matter is a typical "magnetic" material, such as iron or nickel, or a so-called "nonmagnetic" material, such as copper or air.

The *magnetic permeability* of a substance is a numerical description of the extent to which that substance interacts with an applied magnetic field. In other words, permeability refers to the degree to which a substance can be magnetized.

Different substances possess varying degrees of magnetization. The aforementioned examples of strongly magnetic materials have the ability to strengthen an applied magnetic field by a factor of several thousand. Such highly magnetizable materials are called *ferromagnetic*. Certain other substances, such as Al, only marginally increase an applied magnetic field. Such weakly magnetizable materials are called *paramagnetic*. Still other substances, such as Cu and the rare gases, slightly weaken an applied magnetic field. Such "negatively magnetizable" substances are called *diamagnetic*.

In common parlance, diamagnetic and paramagnetic substances are often called *nonmagnetic*. However, as detailed in the following text, all substances are magnetic to some extent. Only empty space is truly nonmagnetic.

The term *hysteresis* has been used to describe many instances where an effect lags behind the cause. However, Ewing was apparently the first to use the term in science [1] when he applied it to the particular magnetic phenomenon displayed by ferromagnetic materials. Magnetic hysteresis occurs during the cyclical magnetization of a ferromagnet. The magnetization path created while increasing an externally applied field is not retraced on subsequent decrease (and even reversal) of the field. Some magnetization, known as *remanence*, remains in the ferromagnet after the external field has been removed. This remanence, if appreciable, allows for the permanent magnetization observed in common bar magnets.

TABLE 29.1 Conversion Factors between the mks and cgs Systems for Important Quantities in Magnetism

Quantity	mks	cgs
H, applied field	A/m	$=4\pi \times 10^{-3}$ Oe
B, flux density	Wb/m^2	$=10^4$ G
M, magnetization	Wb/m^2	$=10^4/4\pi$ emu/cm^3
κ, susceptibility	$Wb/(A \cdot m)$	$=16\ \pi^2 \times l0^{-7}$ emu/Oe $\cdot$ cm^3

29.1 Definition of Permeability

Let an externally applied field be described by the vector quantity H. This field may be produced by a solenoid or an electromagnet. Regardless of its source, H has units of ampere turns per meter (A/m). On passing through a body of interest, H magnetizes the body to a degree, M, formally defined as the magnetic moment per unit volume. The units of M are usually webers per square meter. A secondary coil (and associated electronics) is typically used to measure the combined effects of the applied field and the body's magnetization. This sum total flux-per-unit-area (flux density) is known as the induction, B, which typically has units of Wb/m^2, commonly referred to as a tesla (T). Because H, M, and B are usually parallel to one another, the vector notation can be dropped, so that

$$B = \mu_0 H + M \tag{29.1}$$

where μ_0 is the permeability of free space ($47\pi \times 10^{-7}$ Wb/A $\cdot$ m).

The absolute permeability, μ, of a magnetized body is defined as the induction achieved for a given strength of applied field, or

$$\mu = \frac{B}{H} \tag{29.2}$$

Often, the absolute permeability is normalized by μ_0 to result in the relative permeability, $\mu_r = (\mu/\mu_0)$. This relative permeability is numerically equal and physically equivalent to the cgs version of permeability. This, unfortunately, is still in common usage and often expressed in units of gauss per oersted (G/Oe), although the cgs permeability is actually dimensionless. Conversion factors between the mks and cgs systems are listed in Table 29.1 for the important quantities encountered.

29.2 Types of Material Magnetization

All substances fall into one of three magnetic groups: diamagnetic, paramagnetic, or ferromagnetic. Two important subclasses, antiferromagnetic and ferrimagnetic, will not be included here. The interested reader can find numerous discussions of these subclasses; for example, see [1].

29.2.1 Diamagnetism

Diamagnetic and paramagnetic (see next section) substances are usually characterized by their magnetic susceptibility rather than permeability. Susceptibility is derived by combining Equations 29.1 and 29.2, namely,

$$\mu_r = 1 + \frac{M}{\mu_0 H} = 1 + \frac{\kappa}{\mu_0} \tag{29.3}$$

where κ is the susceptibility with units of Wb/A $\cdot$ m. This so-called *volume susceptibility* is often converted to a mass susceptibility (χ) or a molar susceptibility (χ_M). Values for the latter are readily available for many pure substances and compounds [2]. Susceptibility is also often called "intrinsic permeability" [3].

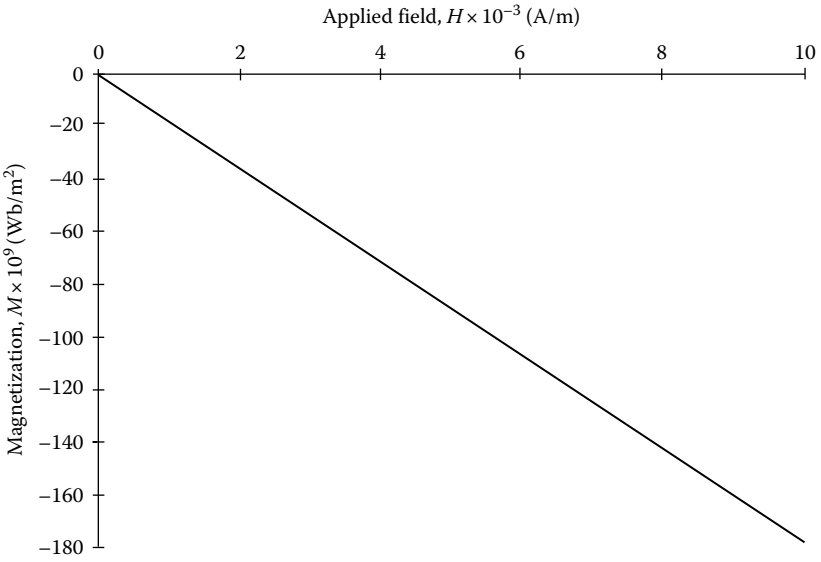


FIGURE 29.1 For diamagnetic substances, magnetization M is small and opposite the applied field H as in this schematic example for graphite ($\kappa = -1.78 \times 10^{-11}$ Wb/A · m).

In any atom, the orbiting and spinning electrons behave like tiny current loops. As with any charge in motion, a magnetic moment is associated with each electron. The strength of the moment is typically expressed in units of Bohr magnetons.

Diamagnetism represents the special case in which the moments contributed by all electrons cancel. The atom as a whole possesses a net zero magnetic moment. An applied field, however, can induce a moment in the diamagnetic material, and the induced moment opposes the applied field. The magnetization, M, in Equation 29.3 is therefore antiparallel to the applied field, H, and the susceptibility, κ, is negative. For diamagnetic materials, $\mu < 1$. Figure 29.1 shows a schematic M vs. H curve for graphite with $\kappa = -1.78 \times 10^{-11}$ Wb/A · m. Note that κ is a constant up to very high applied field values.

29.2.2 Paramagnetism

In a paramagnetic substance, the individual electronic moments do not cancel and the atom possesses a net nonzero moment. In an applied field, the weak diamagnetic response is dominated by the atom's tendency to align its moment parallel with the applied field's direction. Paramagnetic materials have relatively small positive values for κ, and $\mu > 1$.

Thermal energy retards a paramagnet's ability to align with an applied field. Over a considerable range of applied field and temperature, the paramagnetic susceptibility is constant. However, with very high applied fields and low temperatures, a paramagnetic material can be made to approach saturation—the condition of complete alignment with the field. This is illustrated in Figure 29.2 for potassium chromium alum, a paramagnetic salt. Even at a temperature as low as 1.30 K, an applied field in excess of about 3.8×10^6 A/m is necessary to approach saturation. (Note in Figure 29.2, that 1 Bohr magneton = 9.27×10^{-24} J/T.)

29.2.3 Ferromagnetism

Ferromagnetic substances are actually a subclass of paramagnetic substances. In both cases, the individual electronic moments do not cancel, and the atom has a net nonzero magnetic moment that tends to align itself parallel to an applied field. However, a ferromagnet is much less affected by the randomizing

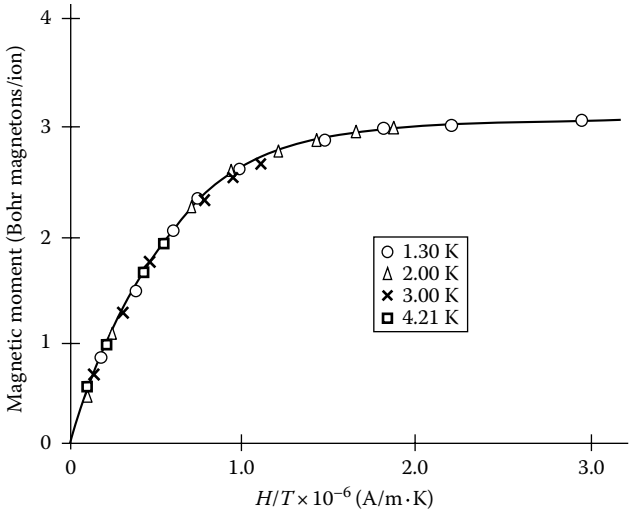


FIGURE 29.2 For paramagnetic substances, the susceptibility is constant over a wide range of applied field and temperature. However, at very high H and low T, saturation can be approached, as in this example for potassium chromium alum. (After Henry, W.E., *Phys. Rev.*, 88, 559, 1952.)

action of thermal energy compared to a paramagnet. This is because the individual atomic moments of a ferromagnet are coupled in rigid parallelism, even in the absence of an applied field.

With no applied field, a demagnetized ferromagnet is comprised of several magnetic domains. Within each domain, the individual atomic moments are parallel to one another, or coupled, and the domain has a net nonzero magnetization. However, the direction of this magnetization is generally opposed by a neighboring domain. The vector sum of all magnetizations among the domains is zero. This condition is called the *state of spontaneous magnetization*.

With an increasing applied field, domains with favorable magnetization directions, relative to the applied field direction, grow at the expense of the less favorably oriented domains. This process is schematically illustrated in Figure 29.3. The exchange forces responsible for the ferromagnetic coupling are explained by Heisenberg's quantum mechanical model [4]. Above a critical temperature known as the Curie point, the exchange forces disappear and the formerly ferromagnetic material behaves exactly like a paramagnet.

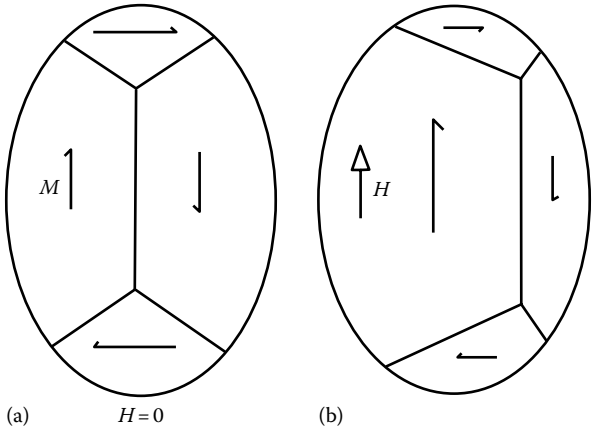


FIGURE 29.3 With no applied field, (a) a ferromagnet assumes spontaneous magnetization. With an applied field, (b) domains favorably oriented with H grow at the expense of other domains.

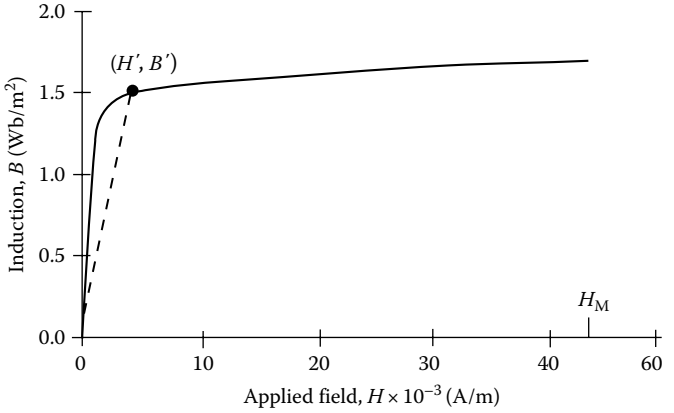


FIGURE 29.4 Magnetization (*B*–*H*) curve for a typical soft ferromagnet. Permeability at point (H', B') is the slope of the dashed line.

During magnetization, ferromagnets show very different characteristics from diamagnets and paramagnets. Figure 29.4 is a so-called *B*–*H* curve for a typical soft ferromagnet. Note that *B* is no longer linear with *H* except in the very low- and very high-field regions. Because of this, the permeability μ for a ferromagnet must always be specified at a certain applied field, *H*, or, more commonly, a certain achieved induction, *B*. Note that μ is the slope of the line connecting a point of interest on the *B*–*H* curve to the origin. It is not the slope of the curve itself, although this value, *dB*/*dH*, is called the *differential permeability* [3].

Another ferromagnetic characteristic evident in Figure 29.4 is *saturation*. Once the applied field has exceeded a certain (and relatively low) value, the slope of the magnetization curve assumes a constant value of unity. At this point, *M* in Equation 29.1 has reached a maximum value, and the ferromagnet is said to be saturated. For all practical purposes, all magnetic moments in the ferromagnet are aligned with the applied field at saturation. This maximum magnetization is often called the saturation induction, B_s [5]. Note that B_s is an intrinsic property—it does not include the applied field in its value.

29.3 Definition of Hysteresis

If *H* is decreased from H_M in Figure 29.4, *B* does not follow the original magnetization path in reverse. Even if *H* is repeatedly cycled from H_M to $-H_M$, *B* will follow a path on increasing *H* that is different from decreasing *H*. The cyclical *B*-*H* relationship for a typical soft ferromagnet is displayed by the hysteresis loops in Figure 29.5. Two loops are included: a minor loop inside a major loop generated by independent measurements. The two differ in the value of maximum applied field: H_M. for the minor loop was below saturation, while H_M for the major loop was near saturation. Both loops are symmetrical about the origin as a point of inversion since in each case $H_M = |-H_M|$.

Notice for the minor loop that when the applied field is reduced from H_M. to 0, the induction does not also decrease to zero. Instead, the induction assumes the value B_r, known as the *residual induction*. If the peak applied field exceeds the point of saturation, as for the major loop in Figure 29.5, B_r assumes a maximum value known as the *retentivity*, B_{rs}.

Note that B_r and B_{rs} are short-lived quantities observable only during cyclical magnetization conditions. When the applied field is removed, B_r rapidly decays to a value B_d, known as the *remanent induction*. B_d is a measure of the permanent magnetization of the ferromagnet. If the maximum applied field was in excess of saturation, B_{rs} rapidly decays to a maximum value of permanent magnetization, or *remanence*, B_{dm}.

The minor loop in Figure 29.5 shows that in order to reduce the induction *B* to zero, a reverse applied field, H_c, is needed. This is known as the *coercive force*. If the maximum applied field was in excess of saturation, the coercive force assumes a maximum value, H_{cs}, known as the *coercivity*. Note that H_c and H_{cs} are usually expressed as positive quantities, although they are negative fields relative to H'_M and H''_M.

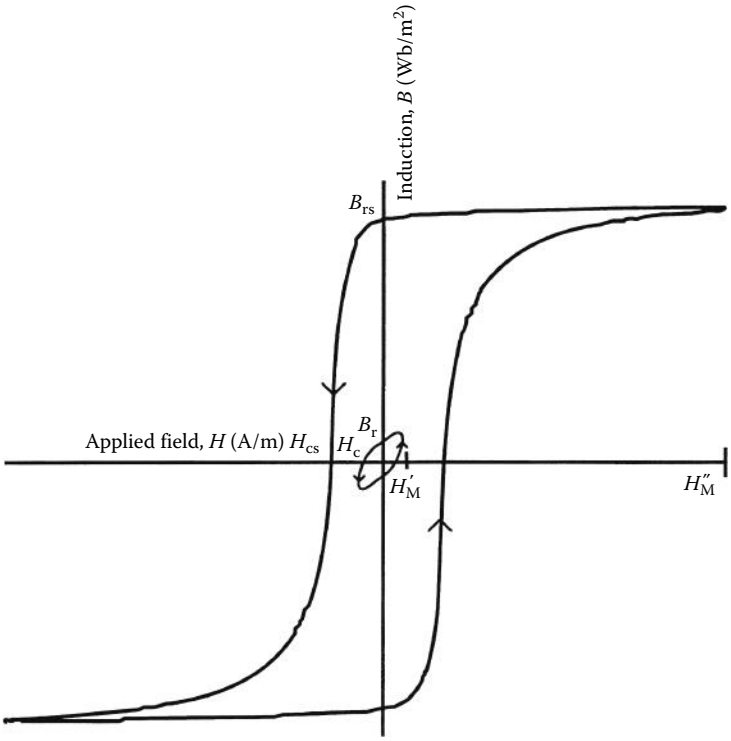


FIGURE 29.5 Major and minor dc hysteresis loops for a typical soft ferromagnet. Labeled points of interest are described in the text.

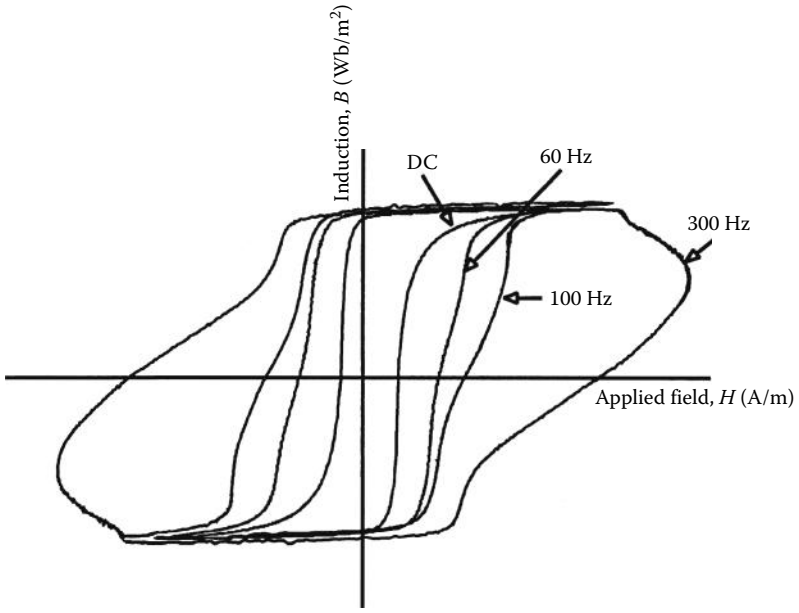


FIGURE 29.6 With increasing test frequency, coercivity for a soft ferromagnet also increases.

The hysteresis loops in Figure 29.5 are known as *direct current (dc) loops*. Typical sweep times for such loops range from 10 to 120 s. At faster sweep times, the coercivity will show frequency dependence, as shown experimentally in Figure 29.6. For soft magnetic materials, this dependence can be influenced by the metallurgical condition of the ferromagnet [6].

29.4 Core Loss

During alternating current (ac) magnetization, some of the input energy is converted to heat in ferromagnetic materials. This heat energy is called *core loss* and is classically comprised of three parts. The first, *hysteresis loss*, P_h, is proportional to the ac frequency, f, and the area of the (slow-sweep) dc hysteresis loop:

$$P_h = kf \int BdH \tag{29.4}$$

The second part is the *loss due to eddy current formation*, P_e. In magnetic testing of flat-rolled strips (e.g., the Epstein test; see next section), this core loss component is classically expressed as

$$P_e = \frac{(\pi Bft)^2}{6d\rho} \tag{29.5}$$

where

B is the peak induction
t is the strip thickness
d is the material density
ρ is the material resistivity

The sum total $P_h + P_e$ almost never equals the observed total core loss, P_t. The discrepancy chiefly originates from the assumptions made in the derivation of Equation 29.5. To account for the additional observed loss, an anomalous loss term, P_a, has often been included, so that

$$P_t = P_h + P_e + P_a \tag{29.6}$$

29.5 Measurement Methods

Reference [3] is a good source for the various accepted test methods for permeability and hysteresis in diamagnetic, paramagnetic, and ferromagnetic materials. Unfortunately, only a few of the instruments described there are available commercially. Examples of these are listed in Table 29.2.

The instruments in Table 29.2 include hysteresigraphs (LDJ models 3500, 5600, and 5500) and vibrating sample magnetometers (VSMs) (LDJ and Lakeshore Cryotronics VSM models). Also included are two Donart models of Epstein testers. The Epstein test is commonly used to characterize flat-rolled soft ferromagnets such as silicon electrical steels in sheet form. A recent alternative to the Epstein test is the single-sheet test method. The Soken instrument in Table 29.2 is an example of such a tester. This method requires much less sample volume than the Epstein test. It can also accommodate irregular sample geometries. However, the Soken instrument is not yet accepted by the American Society for Testing and Materials (ASTM) for reasons explained in the next section.

Note that all instruments in Table 29.2 can measure permeability (or susceptibility), but not all can provide hysteresis loop measurements. Diamagnetic and paramagnetic materials generally require VSM instruments unless one is prepared to construct their own specialty apparatus [3]. All instruments in Table 29.2 can measure ferromagnetic materials, although only a few can accommodate hard (i.e., permanently magnetizable) ferromagnets.

TABLE 29.2 Commercially Available Instruments for Measurement of Permeability and Hysteresis

Manufacturer	Model	Power	Material Type[a]	Ferromagnetic Type	Hysteresis Loop?	Core Loss?
LDJ	3500/5600	ac/dc	F	Soft and hard	Y	Y
Troy, MI	5500	dc	F	Soft and hard	Y	N
(810) 528-2202	VSM	dc	D, P, F	Soft and hard	Y	N
Lakeshore Cryotronics	VSM	dc	D, P, F	Soft and hard	Y	N
Westerville, OH	Susceptometer	ac	F	Soft and hard	N	N
(614) 891-2243	Magnetometer	dc	F	Soft and hard	N	N
Donart Electronics	3401	dc	F	Soft	Y	N
Pittsburgh, PA	MS-2	ac	F	Soft	N	Y
(412) 796-5941						
Soken/Magnetech	DAC-BHW-2	ac	F	Soft	Y	Y
Racine, WI						
(501) 922-6899						

[a] D, diamagnetic; P, paramagnetic; F, ferromagnetic.

29.6 Validity of Measurements

For a ferromagnet under sinusoidal ac magnetization, the induction will show a waveform distortion (i.e., B is nonsinusoidal) once H_M exceeds the knee of the B–H curve in Figure 29.4. Brailsford [7] has discussed such waveform distortion in detail. With one exception, all ac instruments in Table 29.2 determine H from its sinusoidal waveform and B from its distorted waveform.

The single exception is the Soken instrument. Here, feedback amplification is employed to deliberately distort the H waveform in a way necessary to force a sinusoidal B waveform. In general, this will result in a smaller measured value for permeability compared to the case where feedback amplification is not used. Some suggest this to be the more precise method for permeability measurement, but the use of feedback amplification has prevented instruments such as the Soken from achieving ASTM acceptance to date.

Defining Terms

Hysteresis: A ferromagnetic phenomenon in which the magnetic induction B is out of phase with the magnetic driving force H.

Permeability: The extent to which a material can be magnetized.

References

1. Nocola Spaldin, Magnetic Materials: Fundamentals and Device Applications, Cambridge University Press, Cambridge, UK, 2003.
2. D. R. Lide (ed.), *CRC Handbook of Chemistry and Physics*, Boca Raton, FL: CRC Press, 1992–1993.
3. Anonymous, *Annual Book of ASTM Standards*, Philadelphia, PA: ASTM, 3.04, 1995.
4. C. Kittel, *Introduction to Solid State Physics*, 5th edn., New York: John Wiley & Sons, 1976.
5. Anonymous, *Annual Book of ASTM Standards*, Philadelphia, PA: ASTM, 3.04, p. 11, 1995.
6. R. Thomas and G. Morgan, *Proceedings of the Eleventh Annual Conference on Properties and Applications of Magnetic Materials*, Chicago, IL, 1992.
7. A.V. Narlikar, *Frontiers in Magnetic Materials*, Springer, Heidelberg, 2010.

30

Inductance Measurement

Michal Szyper
AGH University of Science and Technology and University of Mining and Metallurgy

Inductance is a property of all electrical conductors. It is a measurable quantity; therefore, it has a physical dimension, a measurement unit (the henry), as well as reference standards. It is found as *self-inductance L* of a single conductor and as *mutual inductance M* in the case of two or more conductors. Inductors can vary in construction, but windings in the form of coils are the most frequent. In this case, they have specific geometrical parameters: surface *A*, length *l*, and a number of turns of windings *N*. The path of magnetic flux can be specially shaped by magnetic cores. Figure 30.1 shows examples of different inductive elements of electrical circuits.

Figure 30.1a and b present windings with self-inductance made as a coreless coil (a) and wound on a ferromagnetic core that is the concentrator of the magnetic field (b). A transformer loaded by impedance Z_L and an electromagnet loaded by impedance related to eddy currents in the metal board, shown in Figure 30.1c and d, will have not only self-inductance of windings but also mutual inductance between windings (c) or between winding and eddy currents (d). Self-inductances and mutual inductances can be detected in busbars of electric power stations as shown in Figure 30.1e and also on tracks on printed circuit boards as in Figure 30.1f. Couplings between currents and electromagnetic fields can be made deliberately but can also be harmful, for example, due to energy losses caused by eddy currents or due to electromagnetic disturbances induced in the tracks of printed circuit boards. Inductors made as windings on ferromagnetic cores have numerous other applications: in electrical machines, transformers, reactors, relays, instrument transformers, as well as inductive sensors, magnetic field detectors, radio receiver antennas, components of power supplies for electronic systems, and other.

Inductive elements are characterized, apart from the inductance *L* and *M*, by the resistance *R* of their windings, which also represents active power losses in the core, and by the capacitance *C* between their windings. Inductive elements can be therefore described by the impedance *Z* (two-terminal network) or transmittance *T* (four-terminal network) whose numerical values are determined experimentally from the voltage and current measurements or by comparison with reference impedances. The models' parameters: self-inductances, mutual inductances, resistances, and capacitances are computed from the measurement results, based on a chosen circuit model (equivalent circuit).

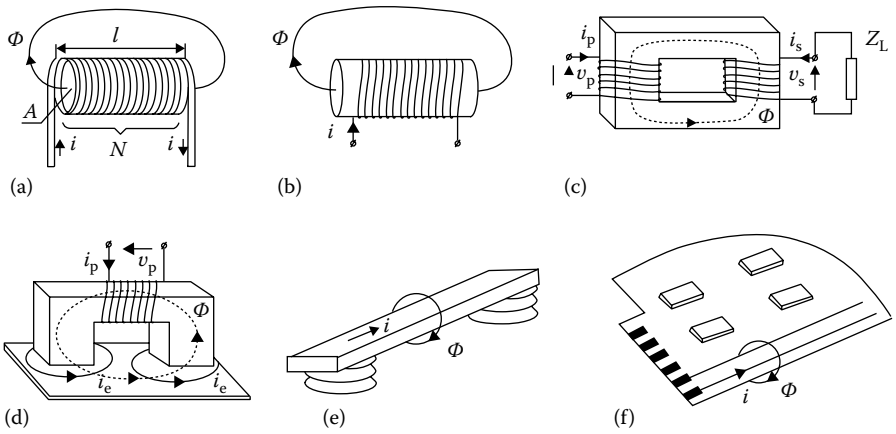


FIGURE 30.1 Examples of different inductive elements: (a) coreless coil, (b) coil with ferromagnetic concentrating core, (c) transformer, (d) electromagnet, (e) element of electrical power station busbars, and (f) printed circuit board with conductive tracks.

30.1 Definitions of Inductance

Inductance is a physical quantity measured in electric circuits in which current is changing in time. A time-varying electric current $i(t)$ produces a time-varying magnetic flux $\Phi(t)$ around the current-carrying conductors; thus, the current and the magnetic field are mutually coupled. The result of this coupling is the electromotive force, that is, the electric voltage $v_L(t)$ induced in these conductors according to Maxwell's law (Dorf and Svoboda 2003). The value of the voltage $v_L(t)$ and the form of its variability in time are thus dependent on the magnetic flux $\Phi(t)$ coupled with the current $i(t)$. In the simple case of a single current-carrying conductor placed in air, whose resistances and capacitances have been neglected as very small, these relations are

$$\Phi(t) = Li(t) \tag{30.1}$$

$$v_L(t) = \frac{d\Phi(t)}{dt} = L\frac{di(t)}{dt} \tag{30.2}$$

where L is the self-inductance of the conductor. The form of the Equation 30.2 and assumption of placing the conductor in air mean that the inductance L is ideal and it does not depend on neither the current nor time. In other words, regarding the inductance L of a current-carrying conductor, the electric circuit comprising this conductor is linear and stationary.

Using the Equation 30.2, the unit of self-inductance one henry (1 H) has been defined as a physical dimension of this quantity in the SI system (Dyer 2001, Tumanski 2006):

One henry (1 H) is the self-inductance of an electric conductor in which an electromotive force of one volt (1 V) is induced when the current in this conductor is uniformly changing at the rate one ampere (1 A) per second (1 s).

Apart from the self-inductance L of a single current-carrying conductor, also the mutual inductance M of two (or a larger number of) conductors is defined. Similarly, as in the case of the self-inductance, the mutual inductance is defined as the effect of coupling the windings by a magnetic flux passing through both windings and produced by the currents: primary $i_p(t)$ and secondary $i_s(t)$ in the primary and the secondary winding, respectively, as in Figure 30.1c. If we make the simplifying assumption about a linear relation between the flux $\Phi(t)$ and the current $i_p(t)$, assume the current $i_s(t) = 0$

(i.e., the secondary circuit is open), and neglect the resistances and capacitances, the mutual inductance between the primary and secondary winding can be defined as

$$v_s(t) = N_s \frac{d\Phi(t)}{dt} = k\sqrt{L_p L_s}\,\frac{di_p(t)}{dt} = M\frac{di_p(t)}{dt} \tag{30.3}$$

where $-1 \le k \le 1$ is the coupling coefficient between the primary and secondary winding. Its value depends on the windings design and, if the inductor has a magnetic core, on the core properties, and its sign depends on the windings direction. Self-inductances of the primary and secondary winding are denoted L_p and L_s, respectively, and the number of turns of the secondary winding is denoted N_s. The value $k \neq \pm 1$ means that the primary and secondary windings are not fully magnetically coupled; thus, a portion of the flux leaks out. Ferromagnetic cores significantly reduce the phenomenon of magnetic flux leakage but they can be a cause of the nonlinear relation between currents and the magnetic flux. The unit of mutual inductance is also the henry (1 H), defined similarly to the self-inductance unit (Dyer 2001, Tumanski 2006).

Since the use of a current test signal $i(t)$ or $i_p(t)$ uniformly (i.e., linearly) changing with time is inconvenient, usually sinusoidal test signals are employed for the inductance measurement: a current source signal $i(t) = I_m \sin\omega t$ or a voltage source signal $v(t) = V_m \sin\omega t$. For such case, and under formerly made simplifying assumptions, Equations 30.1 through 30.3 take the form

$$\Phi(t) = LI_m \sin\omega t \tag{30.4}$$

$$v_L(t) = \omega L I_m \cos\omega t = X_L(\omega) I_m \cos\omega t = V_{Lm} \cos\omega t \tag{30.5}$$

$$v_s(t) = \omega M I_{pm} \cos\omega t = T_{ps}(\omega) I_{pm} \cos\omega t = V_{sm} \cos\omega t \tag{30.6}$$

The reactance of a winding with self-inductance L, that is, a two-terminal network is denoted $X_L(\omega) = \omega L$. The *transmittance* between the secondary winding voltage $v_s(t)$ and the current $i_p(t)$ in the primary winding with mutual inductance M, that is, a four-terminal network, is denoted $T_{ps} = \omega M$.

Equations 30.1 through 30.6 indicate the method for measuring ideal self- and mutual inductances. It consists of determining the voltage and current signal values: the current signal rate of change and the voltage steady-state value for the first case, or the voltage and current amplitudes and frequency for the second case, and calculating the inductance from the quotient of these values. In such simple case, the use of a current source and an oscilloscope is sufficient. However, as mentioned in the introduction, inductive elements in electric circuits are characterized by resistances and capacitances, and relations between their currents and magnetic fluxes can be nonlinear. Expressions that represent actual properties of inductive elements are termed their models, and graphical representations of models are called equivalent circuits.

30.2 Models and Equivalent Circuits of Inductive Elements

Possibilities and methods of measuring the inductance of inductive elements depend on the form and properties of mathematical models we assign to them. Like in a case of any measurement, we only can measure phenomena, which we are able to describe mathematically. The representation error of actual properties of inductive elements models is always limited and its value can be selected. We choose the representation error value depending on the purpose, for which the model is intended, and then we measure its variables (currents and voltages) and calculate coefficients (inductances, resistances, and capacitances), that is, the values of the equivalent circuit components. The chosen model should take into account physical phenomena occurring in the measured inductive element that are necessary for obtaining the required accuracy of the measurement. In real inductive elements, aside from the

coupling between the current in conductors and the magnetic flux produced by this current and the voltage induction at the ends of these conductors, the following phenomena occur:

- Heat dissipation in resistances of the inductive element conductors.
- Heat dissipation in magnetic cores, on which windings having an inductance are wound, and in conductive elements in which eddy currents flow.
- The nonlinear and ambiguous relation between the current in a winding wound on a magnetic core and the magnetic flux produced in the core by this current termed the magnetization curve and the phenomenon of magnetic hysteresis.
- The phenomenon of electrical energy storing in stray (i.e., spatially distributed) self-capacitances of current-carrying conductors, for example, of individual turns of a winding with respect to each other or to shields and adjacent structures.
- The phenomenon of coupling of the magnetic flux produced the current in conductors with eddy currents induced by this current in adjacent conductive elements, like metallic components of sensors, shields, and metallic structures.
- The phenomenon of the increase in a conductor resistance with the frequency of the current in this conductor, termed the skin effect.
- The influence of external magnetic fields (i.e., not being produced by the current flowing in the windings) on the total magnetic flux coupled with these windings and, consequently, on the additive voltage from electromagnetic interferences (normal) induced by external magnetic fields in these windings.
- Spatial heterogeneity of magnetic flux density in magnetic core volume due to material properties and the core design (Craigen 2010).

These phenomena may, depending on the measurement conditions, considerably influence the measurement results and, consequently, their accuracy, that is, the inductance measurement errors. Therefore, if this influence is significant, they should be taken into account in the inductive elements models.

The following can be used as the models of inductive elements:

1. Operator algebraic equations, usually in the complex notation with the Fourier operator ($j\omega$) having the form of linear algebraic relations between the variables $I(j\omega)$, $V(j\omega)$, and $\Phi(j\omega)$ with coefficients $L(\omega)$, $M(\omega)$, $C(\omega)$, and $R(\omega)$ occurring in the determined impedances $Z(j\omega)$ or transmittances $T(j\omega)$ (Dyer 2001)
2. Linear or nonlinear differential equations and stationary or nonstationary differential equations with variables $i(t)$, $v(t)$, and $\Phi(t)$ and coefficients L, M, C, and R, or the corresponding integral operations (Dorf and Svoboda 2003)

A graphical representation of these models is equivalent circuits that have the form of electric circuits comprising L,C,R elements and represent the aforementioned physical phenomena with a given error.

Figure 30.2 shows basic equivalent circuits of typical inductive elements used in electric circuits. They are categorized according to the nature of their inductance (the self-inductance L and mutual inductance M), the frequency band (low frequencies LF, radio frequencies RF) and the presence (PMC) or absence (AMC) of ferromagnetic core. Equivalent circuits of coreless inductive elements for low frequencies (LF) comprise exclusively inductances and resistances connected in series, whereas for radio frequencies (RF), they also include capacitances connected in parallel with inductances and resistances. In the case of inductive elements having a magnetic core, their models and equivalent circuits are usually linearized. The method consists of determining the equivalent circuit parameters for the steady-state (e.g., nominal) values of the current amplitude and frequency. In such case, the magnetizing inductance $L\mu$ and, usually much smaller, leakage inductances L_d, L_{pd}, and L_{sd} are determined, as well as the resistances that represent active power losses in the core R_{Fe} and in windings R_{Cu}, R_{pCu}, and R_{sCu}.

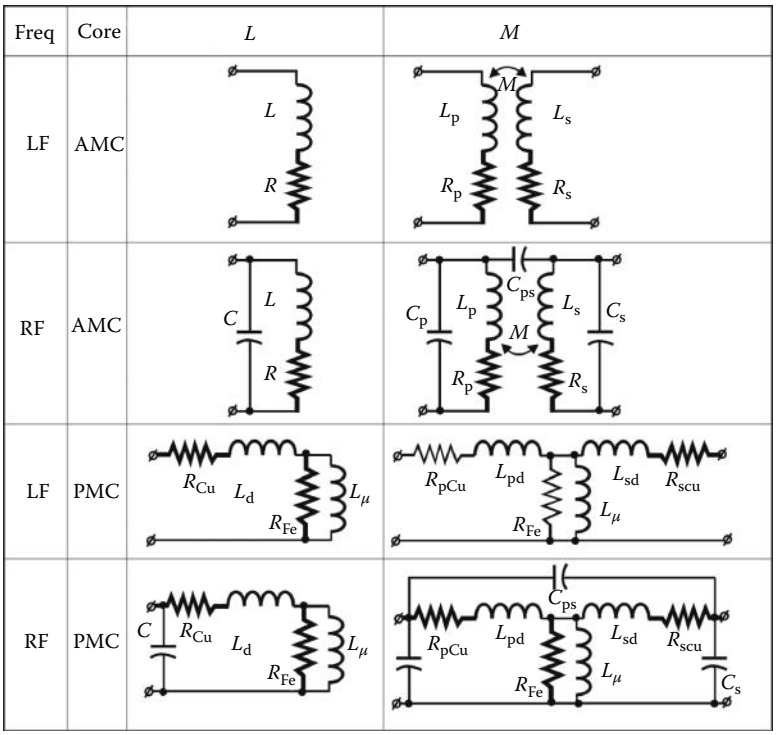


FIGURE 30.2 Basic equivalent circuits of inductive elements with self-inductance L, mutual inductance M for low (LF) and high (HF) frequencies, without ferromagnetic cores (AMC), and with ferromagnetic cores (PMC).

Since the calculation of *LCR* components' values in a complex equivalent circuit is complicated, appropriate computational tools, designed for solving electric circuits, are employed for this purpose.

In the complex number notation, the impedances and transmittances take the form

$$Z = R_Z + jX_Z = Z_m \exp(j\phi_Z), \quad Z_m = \sqrt{R_Z^2 + X_Z^2}, \quad \phi_Z = \arctan\frac{X_Z}{R_Z} \tag{30.7}$$

$$T = R_T + jX_T = T_m \exp(j\phi_T), \quad T_m = \sqrt{R_T^2 + X_T^2}, \quad \phi_T = \arctan\frac{X_T}{R_T} \tag{30.8}$$

The components R_Z, X_Z, and R_T, X_T of the impedance Z and the transmittance T are the algebraic functions of the determined values of *LCR* elements of equivalent circuits. Thus, by measuring the real and imaginary components, or moduli and phase angles of an impedance or transmittance, and comparing them with complex components of the functions representing equivalent circuits, we obtain algebraic equations systems whose solutions are the sought *LCR* elements values. For example, in the equivalent circuit denoted (L, LF, AMC) in Figure 30.2, from the measured impedance $Z = R_Z + jX_Z$, component values can be calculated: $L = \omega^{-1}X_Z$ and $R_L = R_Z$.

Linearization of models of inductive elements with ferromagnetic cores can however be the cause of considerably large errors in determination of inductance, whereas taking into consideration nonlinear relations between the measured inductances and currents requires the use of more complex models, for example, in the form of differential equation representing the inductive element with the voltage test signal $v(t)$:

$$L(i(t))\frac{di(t)}{dt} + i(t)\frac{dL(i(t))}{dt} + R_L(i(t))i(t) = v(t) \tag{30.9}$$

whose equivalent circuit is denoted (L, LF, PMC) in Figure 30.2. The measured values of the equivalent current-dependent self-inductance and the equivalent resistance are denoted $L(i(t))$ and $R_L(i(t))$, respectively. Steady-state $L(i)$ values can be determined from the relation $L(i) \approx \Delta\Phi(i)/\Delta i$ by measuring the magnetizing curve $\Phi(i)$. A shape of this curve depends not only on the amplitude of the current $i(t)$ but also on its frequency ω. The value of the resistance $R_L(i(t))$ also depends on the current amplitude and frequency. Where a dc component of the current in the inductive element (bias current) is applied, the inductance and resistance values depend also on this component value. Therefore, application of nonlinear models of inductive elements for determination of their variable coefficients is a difficult measuring task, usually solved by means of specialized measuring instruments (Agilent Handbook 2009). The variable coefficients of models having the form of Equation 30.9 can be determined by employing computational methods for model identification (Utsun 2009), for example, the tuned model method that is based on minimizing the difference between responses of the measured inductive element and its model to identical excitation with variable in time test signal.

30.3 Measurement Methods

Impedance (or transmittance) measurement methods for inductive elements used for determination of their inductance can be divided into three basic groups:

1. Current–voltage methods based on determination of the impedance (transmittance) and calculation of their equivalent circuit *LCR* components values from the measured current and voltage signal parameters, for example, amplitudes and phase angles in the case of sinusoidal signals, or from the measured parameters of the test signals of a different shape, for example, a step signal
2. Bridge and differential methods based on comparison of sinusoidal current and voltage signals of the measured (unknown) and the reference (known) impedances until a state of the signals equality is reached, that is, the approximate equality of the compared impedances with a given (small) error
3. Resonant methods based on the series or parallel connection of the measured inductive impedance with a variable (controlled) capacitive impedance, tuning resulting circuit to resonance at the given frequency of the excitation signal, and determination of the inductive reactance value from relations describing the state of resonance

30.3.1 Current–Voltage Methods

Current–voltage methods are presently a wide group of measurement methods employed in inductance measurements of all types of inductive elements. In these methods, periodic excitation signals are used within the frequency ranges from ca. 10 Hz to ca. 100 kHz (LF) and from ca. 100 kHz to ca. 1 GHz (RF), as well as in the range from ca. 1 GHz to ca. 100 GHz. Also, the current and voltage values are selected from a wide range, according to both the application and operating conditions of these elements. The accuracy of current–voltage methods is usually not very high and is contained within ca. 1% to ca. 10% of the measured impedance value. Simple current–voltage methods consist of the use of sinusoidal excitation and measuring the amplitudes or rms values and phase angles between currents and voltages in the measured inductance elements. In the case of inductive elements with large currents, and at low frequencies, the rms voltmeters and ammeters are used, as well as wattmeters applied for the measurement of active power losses in the measured element (Figure 30.3a). The "wattmeter" method can be used for measuring the equivalent inductances of nonlinear elements with cores, particularly those operated at the power network frequency (50 or 60 Hz). In the case of approximately linear elements and for higher frequencies, oscilloscopes are used in the *X*–*Y* mode, with both axes scaled in the current and voltage values, to allow determination of amplitudes of these signals and phase angles between them (Figure 30.3b). The orthogonal composition of the measured current and voltage signals, termed the

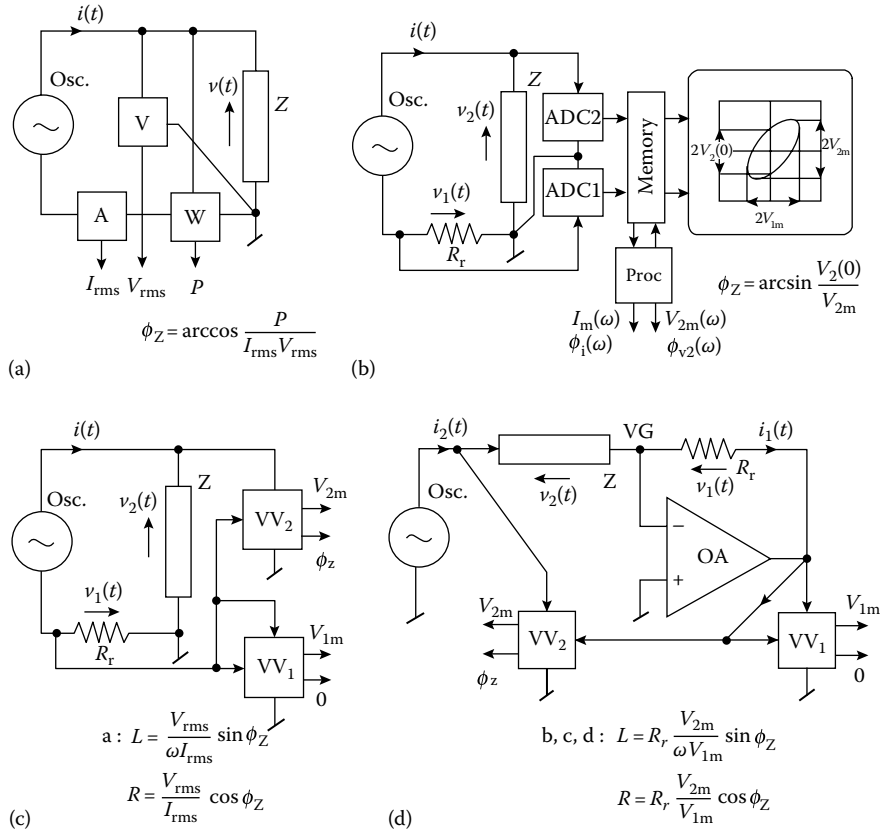


FIGURE 30.3 Circuit diagram for impedance measurement by current and voltage method: (a) the wattmeter method, (b) the oscilloscope method, (c) the concept of vector voltmeters method, and (d) the current-to-voltage conversion method using vector voltmeters.

"ellipse" method, allows observing the effects of the measured impedance nonlinearity, but the condition of its correct application is that transmittances of both oscilloscope channels are identical. Instead of using the "ellipse" method, the current and voltage waveforms can be stored in the oscilloscope memory using sufficiently fast sampling AD converters. On the stored signals, mathematical operations can be performed, for example, determination of their amplitude and phase spectra over the specified frequency band within an appropriate time window, from which the impedances or transmittances can be computed (NI Developer Zone [tutorial]).

Other conceptions of measuring systems employing current–voltage methods utilizing vector voltmeters are shown in Figure 30.3c and d. They consist of determining the modules and phase angles of the measured impedances or transmittances from the quotient of the voltage and current complex values from Equations 30.10 and 30.11 and computing the self- or mutual inductances basing on the assumed equivalent circuits:

$$Z = \frac{v_2}{i} = R_r \frac{v_2}{v_1} = R_r \frac{V_{2m} \exp(j\phi_2)}{V_{1m} \exp(j\phi_1)} = R_r \frac{V_{2m}}{V_{1m}} \exp j(\phi_2 - \phi_1) = Z_m \exp(j\phi_Z) \tag{30.10}$$

$$T = \frac{v_s}{i_p} = R_r \frac{v_s}{v_{pr}} = R_r \frac{V_{sm} \exp(j\phi_s)}{V_{prm} \exp(j\phi_{pr})} = R_r \frac{V_{sm}}{V_{prm}} \exp j(\phi_s - \phi_{pr}) = T_m \exp(j\phi_T) \tag{30.11}$$

The amplitudes of voltages across the measured impedance and transmittance are denoted V_{2m} and V_{sm}, the amplitudes of voltages proportional to the impedance and transmittance currents are denoted V_{1m} and V_{prm}, and Φ_2, Φ_1, Φ_s, Φ_{pr} denote the phase angles of these signals, respectively. The results of voltage measurements using vector voltmeters are amplitudes and phase angles with respect to the reference voltage phase angle. Since Equations 30.10 and 30.11 allow computing the differences between the results of the phase angle measurements, it is convenient to take one of the measured voltage signals as the reference voltage, with respect to which the other voltage phase angles will be determined as the value of the measured angle difference. In Figure 30.3c and d, it has been assumed that the reference voltage signal is the voltage across the resistor R_r. This voltage is proportional to the current but can be measured in different ways. The measuring system, shown in Figure 30.3c as the simplified conception of the measurement method, requires an appropriate choice of the reference point ("ground") common for both voltmeters, which may be difficult due to common-mode interference. The measuring system with current-to-voltage conversion shown in Figure 30.3d utilizes precise, wideband operational amplifiers (OAs). The node denoted VG is called "virtual ground point."

The vector voltmeter VV_1 measures the amplitude of the voltage drop $V_{1m} = R_r I_m$ caused by the test current $i_1(t) = i_2(t)$, and the result of phase angle measurement is $\Phi_1 = 0$. The vector voltmeter VV_2 measures the amplitude V_{2m} of the voltage drop $v_2(t)$ across the impedance Z and the phase angle $\Phi_2 = \Phi_Z$ between this voltage and the current. If the measurements are conducted at the radio frequencies, the capacitances of inductive elements should be taken into account, that is, the equivalent circuits denoted (L, RF, AMC) or (M, RF, AMC) in Figure 30.2 should be used. The measurements should be performed for the number of frequencies corresponding to the number of the necessary algebraic equations. The measuring system conception shown in Figure 30.3d is currently one of the most commonly employed to inductance measurement method.

A block diagram illustrating the principle of the vector voltmeter is shown in Figure 30.4a. The system consists of the multiplier or gated synchronous phase-sensitive detector (PSD) (Dyer 2001, Tumanski 2006) of the measured voltage v with the switching system of the phase of the reference voltage v_{dn}, integrator, digital voltmeter, and processor. The principle of vector voltmeter operation is based on determination of the magnitude V_m and phase angle ϕ of the measured voltage v in reference to voltage v_1, which is proportional to the current i. Assume that voltages v and v_{dn} are in the following forms:

$$v = V_m \sin(\omega t + \phi) = V_m(\sin\omega t \cos\phi + \cos\omega t \sin\phi) \tag{30.12}$$

$$v_{dn} = V_{md} \sin(\omega t + n\,\pi/2), \quad n = 0, 1, 2, 3 \tag{30.13}$$

Phase angle $n\pi/2$ of voltage v_{dn} can take values from the set $\{0, \pi/2, \pi, 3/2\,\pi\}$ by choosing the number n that gives the possibility of detecting the phase angle ϕ in all four quadrants of the Cartesian coordinate

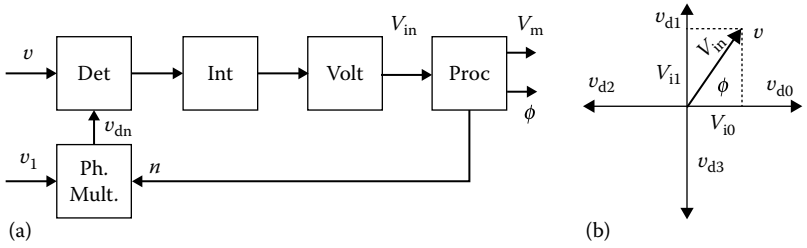


FIGURE 30.4 Block diagram (a) and phasor diagram (b) illustrating the principle of operation of a vector voltmeter. Block abbreviations: "Det," phase-sensitive amplitude detector; "Int," integrator; "Volt," voltmeter; "Proc," processor; and "Ph. mult," controlled phase multiplexer.

system, as is shown in Figure 30.4b. A multiplying synchronous phase detector multiplies voltages v and v_{dn}, and the integrator averages the multiplication result during time T_i:

$$V_{in} = \frac{1}{T_i}\int_0^{T_i} v\, v_{dn}\, dt \tag{30.14}$$

Averaging time $T_i = k\ T$, $k = 1, 2, 3, \ldots$ is a multiple of the period T of the measured voltage. From Equations 30.12 through 30.14, an example for $0 \le \phi \le \pi/2$ (e.g., $n = 0$ and $n = 1$), a pair of numbers is obtained:

$$V_{i0} = 0.5 V_m V_{md} \cos\phi, \quad V_{i1} = 0.5 V_m V_{md} \sin\phi \tag{30.15}$$

which are the values of the Cartesian coordinates of the measured voltage v. The module and phase angle of voltage are calculated from

$$V_m = \frac{2}{V_{md}}\sqrt{V_{i0}^2 + V_{i0}^2}, \quad \phi = \text{arctg}\frac{V_{i1}}{V_{i0}} \tag{30.16}$$

Both coordinates of the measured voltage v can be calculated in a similar way in the remaining quadrants of the Cartesian coordinate system. A vector voltmeter determines the measured voltage as vector (phasor) by measurement of its magnitude and angle as shown in Figure 30.4.

The current and voltage method of impedance or transmittance measurement is based on the measurement of voltages $v_1 = i\ R_r$ and v_2 or $v_{pr} = iR_r$ and v_s and the use of Equation 30.10 or 30.11. Calculation of the voltage measurement results and control of number n is performed by a processor. Errors in phase angle measurement of impedance when using vector voltmeters are within the range of 1% to 10% or between 10^{-5} and 10^{-3} rad, respectively, for a frequency equal to 1 GHz (information from websites). Advanced instruments for impedance and inductance measurements utilizing the vector current–voltage method are known as vector impedance meter or vector network analyzers.

Another method of comparative current/voltage type is a modernized version of the "three-voltmeter" method (Sydenham 1983). A diagram of a measurement system illustrating the principle of the method is shown in Figure 30.5a and b.

The method is based on the properties of an OA, in which output voltage v_2 is proportional to input voltage v_1 and to the ratio of the reference resistance R_r to measured impedance Z. The phasor

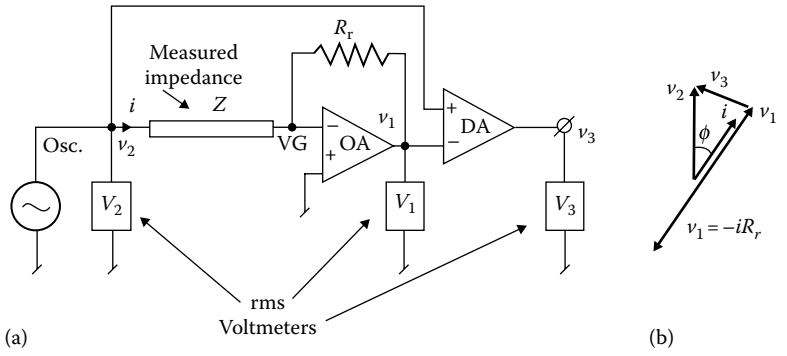


FIGURE 30.5 Block diagram (a) and phasor diagram (b) of the "three-voltmeter" method. Operational and differential amplifiers are represented by blocks "OA" and "DA."

difference v_3 of voltages v_1 and v_2 can be obtained using the differential amplifier (DA). The three voltages (as in Figure 30.5b) can be used for the module Z_m and phase ϕ calculation using relations

$$V_2 = i R_r = -\frac{R_r}{Z} v_1 \rightarrow Z_m = \frac{V_1}{V_2} R_r \tag{30.17}$$

$$v_3 = v_1 - v_2 \rightarrow \phi = \arccos \frac{V_1^2 + V_2^2 - V_3^2}{2V_1V_2} \tag{30.18}$$

where V_1, V_2, V_3 are the results of rms voltage measurements in the circuit. The advantage of the method lies in limiting the influence of stray capacitances as a result of attaching one of the terminals of the measured impedance to a point of "virtual ground." However, in order to obtain small measurement errors, particularly at high frequencies, wideband OAs and wideband rms voltmeters must be used.

Total errors of the inductance measurement, obtained by the current–voltage methods, depend on voltmeter errors, the amplifiers and resistance R_r errors, system factors (residual and leakage inductances, resistances, and capacitances), and the quality of approximation of the measured impedances by the equivalent circuit.

30.3.2 Bridge Methods

There is a variety of bridge methods for measuring inductances. Principles of operation and circuit diagrams of bridges are described in Helfrick and Cooper (2009), Dyer (2001), Tumanski (2006), and Sydenham (1983). The most common ac bridges for inductance measurements and the formulae for calculating measurement results are shown in Figure 30.6.

The procedure referred to as *bridge balancing* is based on a proper selection of the reference values of the bridge so as to reduce the differential voltage to zero (as referred to the output signal of the balance indicator). It can be done manually or automatically.

The condition of the balanced bridge $v_0 = 0$ leads to the following relation between the impedances of the bridge branches; one of them (e.g., Z_1) is the measured impedance:

$$Z_1Z_3 = Z_2Z_4 \rightarrow (R_1 + jX_1)(R_3 + jX_3) = (R_2 + jX_2)(R_4 + jX_4) \tag{30.19}$$

Putting Equation 30.19 into complex form and using expressions for the impedances of each branch, two algebraic equations are obtained by comparing the real and imaginary components. They are used to determine the values of the equivalent circuit elements of the measured impedance. In the simplest case, they are L and R elements connected in series. More complicated equivalent circuits need more equations to determine the equivalent circuit parameters. Additional equations can be obtained from measurements made at different frequencies.

In self-balancing bridges, vector voltmeters preceded by an amplifier of the out-of-balance voltage v_0 are used as "zero" detectors. The detector output is coupled with variable standard bridge components.

The Maxwell Wien bridge shown in Figure 30.6a is one of the most popular ac bridges. Its range of measurement values is large and the relative error of measurement is about 0.1% of the measured value. It is used in the wide frequency band of 20 Hz–1 MHz. The bridge is balanced by varying the R_2 and R_3 resistors or by varying R_3 and capacitor C_3. Some difficulties can be expected when balancing a bridge with inductors with high time constants.

The Hay bridge presented in Figure 30.6b is also used for measurement of inductors, particularly those with high time constants. The balance conditions of the bridge depend on the frequency value, so the frequency should be kept constant during the measurements, and the bridge supply voltage should

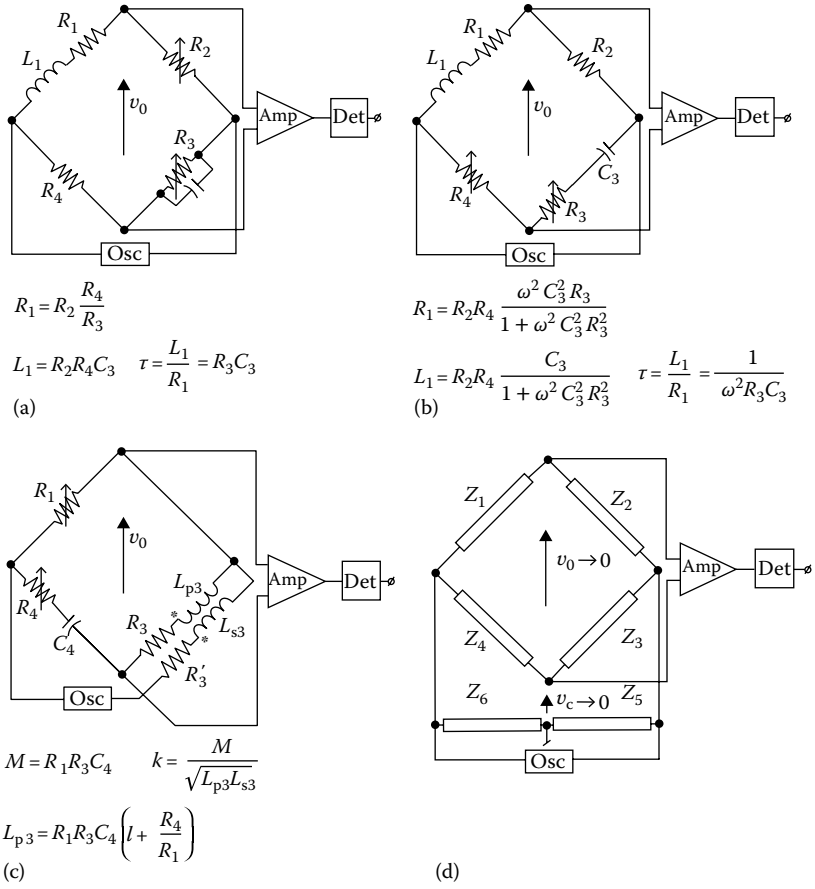


FIGURE 30.6 Bridge circuits used for inductance measurements: (a) Maxwell Wien bridge, (b) Hay bridge, (c) Carey Foster bridge, and (d) ac bridge with Wagner branch. Block abbreviations: Osc, oscillator; Amp, amplifier; Det, amplitude detector.

be free from harmonic distortions. The dependence of bridge balance conditions on frequency also limits the measurement ranges. The bridge is balanced by varying R_3 and R_4 resistors and by switching capacitor C_3.

Mutual inductance M of two windings with self-inductances L_p and L_s can be determined by Maxwell Wien or Hay bridges. For this, two inductance measurements have to be made for two possible combinations of the series connection of both coupled windings: one of them for the corresponding directions of the windings, and one for the opposite directions. Two values of inductances L_1 and L_2 are obtained as the result of the measurements:

$$L_1 = L_p + L_s + 2M, \quad L_2 = L_p + L_s - 2M \tag{30.20}$$

Mutual inductance is calculated from

$$M = 0.25(L_1 - L_2) \tag{30.21}$$

The Carey Foster bridge described in Figure 30.6c is used for mutual inductance measurement. The self-inductances of the primary and secondary windings can be determined by two consecutive measurements. The expressions presented in Figure 30.6c yield the magnetic coupling coefficient k.

The bridge can be used in a wide frequency range. The bridge can be balanced by varying R_1 and R_4 resistances and switching the remaining elements.

For correct measurements with ac bridges, it is essential to minimize the influence of harmful couplings among the bridge elements and connection wires, both between each other and to the environment. Elimination of parallel (common) and series (normal, electromagnetic) voltage distortions is necessary for high measurement resolution. These couplings are produced by the capacitances, inductances, and parasitic resistances of bridge elements to the environment and among themselves. Because of their appearance, they are called stray couplings. Series (electromagnetic, normal) voltage distortions are induced in the bridge circuit by varying common electromagnetic fields. Parallel (common) voltage distortions are caused by potential differences between the reference point of the supply voltage and the points of the out-of-balance voltage detector.

Magnetic shields applied to connection wires and bridge-balancing elements are the basic means of minimizing the influence of parasitic couplings and voltage distortions (Sydenham 1983). All the shields should be connected as a "star" connection, that is, at one point, and connected to one reference ("ground") point of the system. For these reasons, amplifiers with symmetric inputs are frequently used in an out-of-balance voltage detector of ac bridges, as they reject parallel voltage distortions well.

When each of the four nodes of the bridge has different stray impedances to the reference ground, an additional circuit called a Wagner branch is used (see Figure 30.6d). By varying impedances Z_5 and Z_6 in this circuit, voltage v_c can be reduced to zero; by varying the other impedances, voltage v_0 can also be reduced to zero. This way, the bridge becomes symmetrical in relation to the reference ground point and the influence of the stray impedances is minimized.

The total error of the inductance measurement results (when using bridge methods) depends on the following factors: the accuracy of the standards used as the bridge elements, mainly standard resistors and capacitors; errors of determining the frequency of the bridge-supplying voltage (if it appears in the expressions for the measured values); errors of the resolution of the zero detection systems (errors of state of balance); errors caused by the influence of residual and stray inductances; resistances and capacitances of the bridge elements and wiring; and the quality of approximation of the measured impedances in the equivalent circuit.

The errors of equivalent resistance measurements of inductive elements using bridge methods are higher than the errors of inductance measurements. The number of various existing ac bridge systems is very high. Often, the bridge system is built as a universal system that can be configured for different applications by switching elements. One of the designs of such a system is described in Helfrick and Cooper (2009). The example of an automatic bridge is given in Dyer (2001).

30.3.3 Differential Methods

Differential methods (Sydenham 1983) can be used to build fully automatic digital impedance meters (of real and imaginary components or module and phase components) that can measure inductive impedances. Differential methods are characterized by small errors, high resolution, and a wide frequency band and often utilize precise digital control and digital processing of the measurement results. The principle of differential methods is presented through the example of a measuring system with an inductive voltage divider (IVD) and a magnetic current comparator (MCC) (Figure 30.7). An IVD is a precise voltage transformer with several secondary winding taps that can be used to vary the secondary voltage in precisely known steps (Sydenham 1983). By combining several IVDs in parallel, it is possible to obtain a precise voltage division, usually in the decade system. The primary winding is supplied from a sinusoidal voltage source. An MCC is a precise differential transformer with two primary windings and a single secondary winding. The primary windings are connected in a differential way; that is, the magnetic fluxes produced by the currents in these windings subtract. The output voltage of the secondary winding depends on the current difference in the primary windings. MCCs are characterized by very high resolution and low error but are expensive. In systems in common use, precise control of

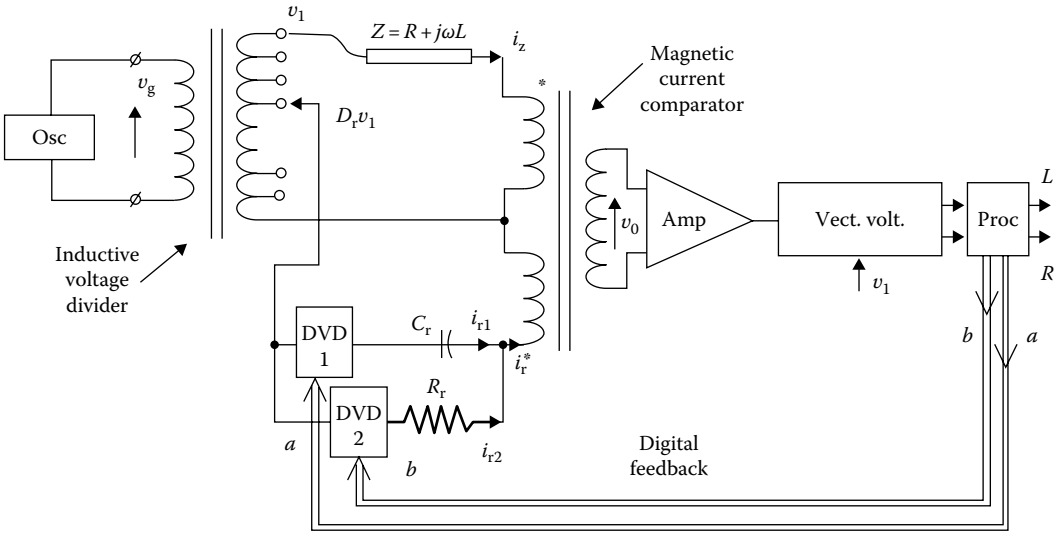


FIGURE 30.7 Scheme of the differential method. Block abbreviations: Osc, oscillator; Amp, amplifier; Vect. voltm, vector voltmeter; Proc, processor; DVD, digital voltage divider.

voltages (IVD) is provided by digitally controlled (sign and values) digital voltage dividers (DVDs), and the magnetic comparator is replaced by a DA.

The algorithms that enable calculation of L and R element values of the series equivalent circuit of an inductive impedance Z in a digital processor result from the mathematical model described in Figure 30.7. When the system is in a state of equilibrium, the following relations occur:

$$V_0 = 0 \rightarrow i_z - i_r = 0 \tag{30.22}$$

$$i_z = D_1 v_2 \frac{1}{R + j\omega L}, \quad i_r = v_r \left(\frac{b}{R_r} - aj\omega C_r \right) \tag{30.23}$$

where $0 < D_r \leq 1$ is the coefficient of v_1 voltage division and the values a and b are equivalent to the binary signals used by the processor to control DVD. Multiple digital-to-analog converters (DACs) are used as digitally controlled voltage dividers. They multiply voltage $D_r v_1$ by negative numbers ($-a$), which is needed in the case of using a standard capacitor C_r for measurements of inductance. After substituting Equation 30.23 into Equation 30.22 and equating the real and imaginary parts, the following formulae are obtained:

$$L = \frac{aR_r^2 C_r}{D_r \left(b^2 + a\omega^2 R_r^2 C_r^2 \right)} \tag{30.24}$$

$$R = \frac{bR_r}{D_r \left(b^2 + a\omega^2 R_r^2 C_r^2 \right)} \tag{30.25}$$

The length N of the code words $\{a_n\} \leftrightarrow a$ and $\{b_n\} \leftrightarrow b$, $n = 1, 2, \ldots, N$ determines the range and resolution of the measurement system; that is, the highest measurable inductance and resistance values and the lowest detectable values. The range can be chosen automatically by changing the division

coefficients D_r. Achieved accuracy is better than 0.1% in a very large range of impedances and in a sufficiently large frequency range. Measurements can be periodically repeated and their results can be stored and processed.

30.3.4 Resonance Methods

Resonance methods are a third group of methods of measuring inductance. They are based on application of a series or parallel resonance *LC* circuits as elements of either a bridge circuit or a two-port (four terminals) "T"-type network. Examples of both circuit applications are presented in Figure 30.8. In the bridge circuit shown in Figure 30.8a, which contains a series resonance circuit, the resonance state is obtained by varying the capacitor C_r, and then the bridge is balanced ($v_0 = 0$) using the bridge resistors. From the resonance and balance conditions, the following expressions are obtained:

$$L = \frac{1}{\omega^2 C_r}, \quad R = R_2 \frac{R_4}{R_3} \tag{30.26}$$

To calculate the values of the *LR* elements of the series equivalent circuit of the measured impedance, it is necessary to measure (or know) the angular frequency ω of the supply voltage. The frequency band is limited by the influence of unknown interturn capacitance value. In the "shunted T" network presented in Figure 30.8b, the state of balance (i.e., the minimal voltage v_0 value) is achieved by tuning the multiloop *LCR* circuit to parallel resonance. The circuit analysis (Sydenham 1983) is based on the "star-delta" transformation of the $C_rR_rC_r$ element loop and leads to the relations for *L* and *R* values:

$$L = \frac{2}{\omega^2 C_r}, \quad R = \frac{1}{\omega^2 C_r^2 R_r} \tag{30.27}$$

According to Sydenham (1983), a "double T" network can be used for inductive impedance measurements at high frequencies (up to 100 MHz).

In resonance methods, the quality factor *Q* and dissipation factor D are defined for the equivalent circuit of inductive elements. In the case of the equivalent circuit (L, LF, NMC) with inductance and resistance, it can be as follows:

$$Q = \frac{1}{D} = \frac{\omega L}{R} = \tau\omega \tag{30.28}$$

where parameter τ is a time constant. Connecting a variable capacitance *C* in parallel with this circuit and tuning the circuit to resonance, for $R \ll \omega L$ (e.g., $Q \gg 1$), the measured inductance can be computed from the measured resonant frequency in Equation 30.26.

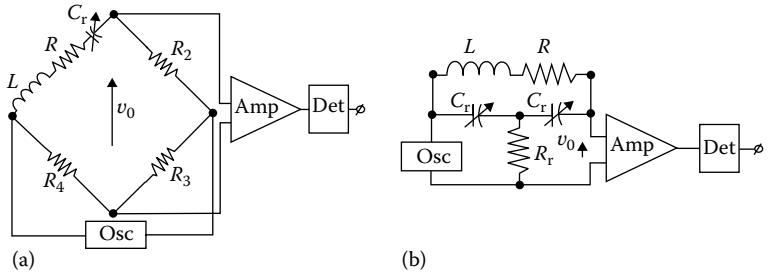


FIGURE 30.8 Circuits diagrams applied in resonance methods: (a) bridge circuit with series resonance circuit and (b) two-port "shunted T" type with parallel resonance circuit. Block abbreviations: Osc, oscillator; Amp, amplifier; Det, amplitude detector.

30.4 Instrumentation

Instruments used for inductance measurements can be categorized into several groups according to their application and metrological properties, mainly the test signal frequency band and the accuracy of measurement. The largest group are instruments for test measurements of *LCR* elements: *LCR* meters and *LCR* bridges, designed as built-in or self-contained digital multimeters, operating in the low and radio frequency band (Waltrip et al. 2005). The second group are advanced vector impedance meters, operating in the low and radio frequency band, and vector network analyzers, used for inductance measurements at high frequencies (Blaschke and Victory 2007). Into the third group can be classified instruments intended for building dedicated measuring systems, for example, for research purposes: vector voltmeters, selective signal detectors, precision differential systems (Dyer 2001), reference *LCR* elements (e.g., quantum Hall resistors [Bohacek 2004]) (NIST Calibration Services: Impedance Measurement), and sources of digitally synthesized analogue signals. Professional design of measuring instruments for inductance measurements reduces the influence of factors that give rise to measurement errors: residual and stray *LCR* parameters, inductive and capacitive couplings, normal-mode (electromagnetic) and common-mode interference, as well as other factors (Agilent Handbook 2009, Callegaro 2009). Operating manuals for such instruments provide necessary guidelines on measurement methods and conditions. More information can be found on the following websites:

http://www.agilent.com/products
http://www.anritsu.com/products
http://www.bkprecision.com/products
http://www.hioki.com/products
http://www.quadtech.com/products
http://www.rhode-schwarz.com/products
http://www.wayne-kerr.com/products.

References

Agilent Technologies, *Agilent Impedance Measurement Handbook*. 2009. *A Guide to Measurement Technology and Techniques*, 4th edn., Santa Clara, California.

Blaschke, V. and Victory, J. 2007. Accurate De-embedding Technique for Scalable Inductor Models, *IEEE International Conference on Microelectronic Test Structures*, Tokyo, Japan, pp. 248–252.

Bohacek, J. 2004. A QHE-based system for calibrating impedance standards, *IEEE Trans. Instr. Meas.*, 53(4):977–980.

Callegaro, L. 2009. The metrology of electrical impedance at high frequency: A review, *Meas. Sci. Technol.* 20(022002):1–14.

Craigen, D. 2010. The problem of coil inductance with nonlinear core material, http://www.integratedsoft.com/News/RandDnews.aspx?Article=Coil-Inductance-with-Nonlinear-Core-Material (May 21, 2009).

Dorf, R. C. and Svoboda, J. A. 2003. *Introduction to Electric Circuits*, 6th edn., New York: John Wiley & Sons.

Dyer, S. A. 2001. *Survey of Instrumentation and Measurement*, New York: Wiley-IEEE.

Helfrick, A. D. *Modern Electronic Instrumentation and Measurement Techniques*, Dorling Kindrsley Pvt Ltd, India, 2005.

National Instruments, Capacitance/Inductance Measurement, http://zone.ni.com/devzone/cda/tut/p/id/3078, accessed on August 18, 2013.

NIST Calibration Services: Impedance Measurements, http://ts.nist.gov/MeasurementServices/Calibrations/Impedance.cfm, accessed on August 18, 2013.

Sydenham, P. H. (ed.), 1983. *Handbook of Measurement Science, Vol.* 2, *Practical Fundamentals,* New York: John Wiley & Sons.

Tumanski, S. 2006. *Principles of Electrical Measurement*, New York: Taylor & Francis Group.

Utsun, O. 2009. Measurement and real-time modelling of inductance and flux linkage in switched reluctance motors, *IEEE Trans. Magnetic*, 45(12):5376–5382.

Waltrip, B., Avramov-Zamurovic, and S. Koffman, A. Inductance measurement using an LCR meter and a current transformer interface, *IEEE Conf.: IMTC 2005*, 1:1005–1007.

31

Immittance Measurement

Achim Dreher
German Aerospace Center

Electronic circuits consist of numerous elements that can be lumped, distributed, or a combination of both. The components are regarded as *lumped* if their size is much smaller than the signal wavelength. This condition holds for resistors, inductors, capacitors, transformers, diodes, transistors, or similar devices operating in printed circuits at frequencies up to a few hundred megahertz or even higher in small integrated circuits. In the microwave or millimeter-wave region, the elements and their connecting transmission lines must be considered as *distributed*. While in lumped circuits a change of voltage or current at one single point immediately affects these quantities at all other points, in distributed circuits the propagation properties now have to be taken into account. The same holds for long connecting cables even at lower frequencies.

To describe the effect of any element within an electronic circuit or of the connection of different circuits, the *immittance* is used as a characteristic quantity. It simply provides a relation of sinusoidal voltage and current at the terminals of the element as a function of frequency. In practice, the single elements are interconnected to networks and the immittance also characterizes arbitrarily complicated networks considered as one port. Moreover, the elements themselves are not ideal. A resistor, for example, made of wound resistive wire, has parasitic components such as capacitance and inductance of winding and terminals. It must be represented by an equivalent circuit forming a complex network [1].

The word "immittance" was proposed by Bode [2] and is a combination of the words "impedance" and the reverse quantity called "admittance." These terms do not only occur in electrodynamics but wherever wave propagation takes place—in acoustics as well as in elasticity. The emphasis of this chapter is on lumped electronic networks and guided electromagnetic waves. Readers interested in more general propagation and scattering phenomena are referred to [3].

31.1 Definitions

Assume a stable linear and time-invariant (LTI) network with only one port. Linearity and time independence are generally met for combinations of passive elements but also for active devices with small-signal driving under constant physical conditions (temperature, humidity, dimensions, etc.). In the steady state, a voltage $v(t) = V_m \cos(\omega t + \varphi_v)$ with amplitude V_m varying harmonically with the angular frequency $\omega = 2\pi f$ that is applied to the terminal then only produces voltages and currents of the same frequency within the network (Figure 31.1). Using complex notation

$$v(t) = \mathrm{Re}\left\{Ve^{j\omega t}\right\} \quad \text{with } V = V_m e^{j\varphi_v} \tag{31.1}$$

the current flowing into the network is given by

$$i(t) = I_m \cos(\omega t + \varphi_i) = \mathrm{Re}\{Ie^{j\omega t}\} \quad \text{with } I = I_m e^{j\varphi_i} \tag{31.2}$$

The phasors V and I are time independent and can be represented in the complex plane (Figure 31.2). Relating voltage and current at the terminal, the network is uniquely described by means of a complex frequency-dependent quantity, the impedance Z:

$$Z = \frac{V}{I} = \frac{V_m}{I_m} e^{j(\varphi_v - \varphi_i)} = |Z|\, e^{j\varphi_z} \tag{31.3}$$

For a given frequency, an arbitrarily complex network within a circuit thus can be replaced by a single element without changing the electrical properties at the terminals. Sometimes it is more convenient to use the inverse of Z, the admittance Y:

$$Y = \frac{1}{Z} = \frac{1}{V} = |Y|\, e^{j\varphi_y} \quad \text{with } \varphi_y = \varphi_i - \varphi_v = -\varphi_z \tag{31.4}$$

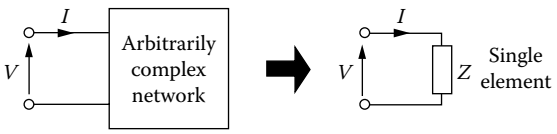


FIGURE 31.1 An arbitrarily complex network can be replaced by its impedance for a given frequency without changing the electrical properties at the terminal.

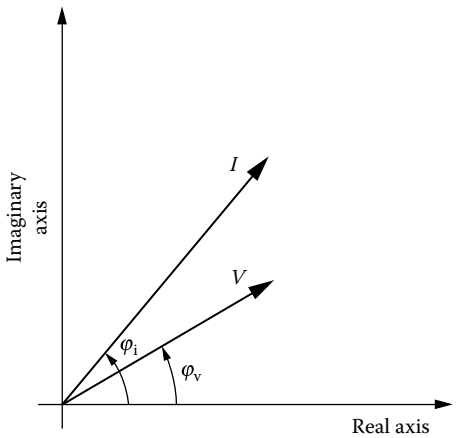


FIGURE 31.2 Voltage and current phasors in the complex plane.

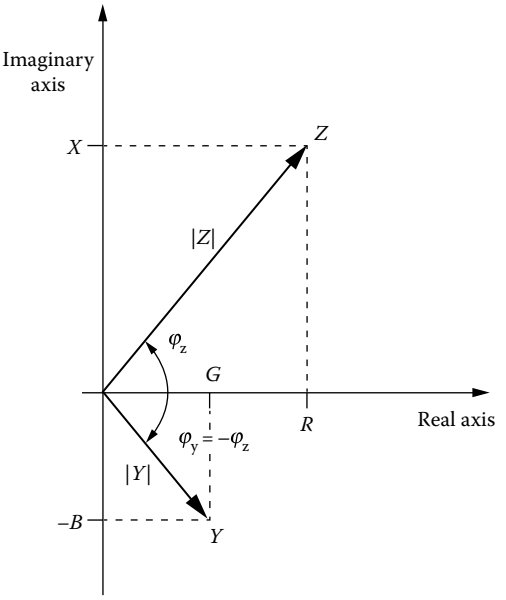


FIGURE 31.3 Representation of impedance and admittance in the complex plane showing the relations between rectangular and polar coordinates. Note that the units are different for each vector.

Both quantities are combined to form the word "immittance." Figure 31.3 shows their representation in the complex plane. Equations 31.3 and 31.4 give the definition in polar coordinates. In data sheets, they are often written as

$$|Z| \angle \varphi_z, \quad |Y| \angle \varphi_y \tag{31.5}$$

Using Euler's identity $e^{j\varphi} = \cos\varphi + j\sin\varphi$, one obtains in rectangular coordinates:

$$\begin{aligned} Z &= |Z|\cos\varphi_z + j|Z|\sin\varphi_z = R + jX \\ Y &= |Y|\cos\varphi_y + j|Y|\sin\varphi_y = G + jB \end{aligned} \tag{31.6}$$

From Figure 31.3, the following relations between rectangular and polar coordinate representation can be deduced immediately:

$$\begin{aligned} R &= |Z|\cos\varphi_z \;\; |Z| = \sqrt{R^2 + X^2} \quad G = |Y|\cos\varphi_y \quad |Y| = \sqrt{G^2 + B^2} \\ X &= |Z|\sin\varphi_z \;\; \varphi_z = \tan^{-1}\left(\frac{X}{R}\right) \quad B = |Y|\sin\varphi_y \quad \varphi_y = \tan^{-1}\left(\frac{B}{G}\right) \end{aligned} \tag{31.7}$$

The real parts are the resistance R and the conductance G. They indicate the losses within the network. The imaginary parts, which are termed reactance X and susceptance B, respectively, are a measure of the reactive energy stored in the network during one period. In general, all these quantities are frequency dependent.

Note that the correct sign of the imaginary parts must be used: the angle φ is in the range of $-180° < \varphi \leq 180°$ and $\varphi < 0$ always corresponds to $X, B < 0$.

For elements with low losses, the loss angle δ or loss factor D is often given instead of the phases φ_z and φ_y. They are always positive small quantities and tend to 0 for a lossless device:

$$D = \tan\delta = \tan\left(\frac{\pi}{2} - |\varphi_z|\right) = \tan\left(\frac{\pi}{2} - |\varphi_y|\right) = \left|\frac{R}{X}\right| = \left|\frac{G}{B}\right| \tag{31.8}$$

The inverse quantity is the quality factor $Q = 1/D$. It involves a ratio of stored electric energy to power dissipated:

$$Q = 2\pi \frac{\text{Energy stored}}{\text{Energy loss per period}} \tag{31.9}$$

A high Q indicates a nearly pure reactive component.

In high-power electronics, it is necessary to reduce losses on transmission lines and therefore avoid currents associated with reactive load components. To obtain a criterion for the application and efficiency of compensation techniques, a power factor is defined. From complex power representation,

$$P = VI^{\star} = |P|(\cos\varphi + j\sin\varphi) \tag{31.10}$$

(the asterisk, *, indicates the conjugate complex number) follows from Equations 31.3 and 31.4:

$$P = |I|^2 \quad Z = |I|^2|Z|(\cos\varphi_z + j\sin\varphi_z) = |V|^2 \quad Y^{\star} = |V|^2|Y|(\cos\varphi_y - j\sin\varphi_y) \tag{31.11}$$

and since the effective power is given by the real part of P,

$$P_{\text{eff}} = \text{Re}\{P\} = |P|\cos\varphi \tag{31.12}$$

the power factor is

$$\cos\varphi = \cos\varphi_z = \cos\varphi_y \tag{31.13}$$

In general, root mean square (rms) values are used for the phasors. Otherwise, a factor 1/2 has to be taken into account in Equations 31.10 and 31.11, since $|P| = 1/2V_mI_m$ for sinusoidal quantities.

It can also be seen from Equations 31.10 and 31.11 that the immittances are directly related to the apparent power:

$$|P| = |V||I| = |I|^2|Z| = |V|^2|Y| \tag{31.14}$$

31.2 Ideal Lumped Components

The immittances of the fundamental passive circuit elements are derived from their instantaneous voltage current relations using Equations 31.1 through 31.4 and the differentiation rules.

31.2.1 Resistances

From Equation 31.15

$$v(t) = Ri(t) \tag{31.15}$$

it follows $V = RI$ and thus $Z = R$ or $Y = G$. The immittance of a resistance is real and identical to its dc resistance or conductance.

31.2.2 Inductances

Voltage and current are related via the differential equation

$$v(t) = L\frac{\mathrm{d}i(t)}{\mathrm{d}t} \tag{31.16}$$

with inductance L, from which follows that $V = j\omega LI$ and

$$Z = j\omega L = jX_\mathrm{L}, \quad Y = \frac{1}{j\omega L} = -j\frac{1}{X_\mathrm{L}} = -jB_\mathrm{L} \tag{31.17}$$

31.2.3 Capacitances

From Equation 31.18

$$i(t) = C\frac{\mathrm{d}v(t)}{\mathrm{d}t} \tag{31.18}$$

with capacitance C, it follows that $I = j\omega CV$ and

$$Y = j\omega C = jB_\mathrm{C}, \quad Z = \frac{1}{j\omega C} = -j\frac{1}{B_\mathrm{C}} = -jX_\mathrm{C} \tag{31.19}$$

The immittance of ideal inductors and capacitors is purely imaginary with different signs according to the phase shift of ±90° between voltages and currents. A general element or network is therefore called inductive or capacitive at a given frequency corresponding to the sign of the imaginary part of its impedance. Note, however, that the frequency dependence can be much more complicated than for these ideal elements and the impedance can even change several times between capacitive and inductive characteristic.

31.3 Distributed Elements

At high frequencies, the size of the elements may no longer be small compared to the signal wavelength. Propagation effects must then be taken into account and the components can no longer be described by means of simple lumped equivalent circuits. If at all possible, they are replaced by transmission line circuits, which are easier to characterize; they realize the required electrical properties more exactly within a defined frequency range.

31.3.1 Transmission Lines

Assuming a simplifying transverse electromagnetic wave (TEM mode) with no field components in the propagation direction, voltages and currents can be uniquely defined and are given as solutions of the corresponding wave equations [4]:

$$\frac{\mathrm{d}^2V}{\mathrm{d}z^2} - \gamma^2 V = 0, \quad \frac{\mathrm{d}^2I}{\mathrm{d}z^2} - \gamma^2 I = 0 \tag{31.20}$$

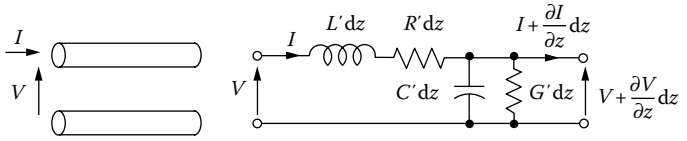


FIGURE 31.4 Equivalent circuit of a differential length of transmission line. The wave equations can be obtained by simply applying Kirchhoff's laws to voltages and currents.

They vary along the line in the z-direction according to

$$V(z) = V_0^+ e^{-\gamma z} + V_0^- e^{\gamma z}, \quad I(z) = I_0^+ e^{-\gamma z} + I_0^- e^{\gamma z} \tag{31.21}$$

These solutions are sums of forward ($e^{-\gamma z}$) and backward ($e^{\gamma z}$) traveling waves with amplitudes V_0^+, I_0^+ and V_0^-, I_0^- and a propagation constant

$$\gamma = \sqrt{(R' + j\omega L')(G' + j\omega C')} \tag{31.22}$$

The equivalent circuit of the transmission line is shown in Figure 31.4. The energy storage in the electric field is accounted for by the distributed shunt capacitance C' per unit length, while the effect of the magnetic field is represented by the series inductance L' per unit length. The series resistance R' per unit length and the shunt conductance G' per unit length represent the power losses in the conductors and in the dielectric, respectively. The amplitudes of voltage and current are related by means of the characteristic impedance Z_0:

$$Z_0 = \frac{V^+}{I^+} = -\frac{V^-}{I^-} = \sqrt{\frac{R' + j\omega L'}{G' + j\omega C'}} \tag{31.23}$$

Of special interest for the use within a network is the input impedance Z_{in} of the transmission line. It depends also on the termination Z_{L} at the other end of the line. For a transmission line of length l, it is given by

$$Z_{\text{in}} = Z_0 \frac{Z_{\text{L}} + Z_0 \tanh \gamma l}{Z_0 + Z_{\text{L}} \tanh \gamma l} \tag{31.24}$$

that is, a transmission line transforms the impedance Z_{L} into Z_{in} at the input.

A quantity more suitable to wave propagation and measurement at high frequencies is the reflection coefficient Γ. It is defined by the relation of the voltages associated with forward and backward traveling waves. At the end of the line, using $V(l) = Z_{\text{L}} I(l)$, one finds

$$\Gamma = \frac{V_0^- e^{\gamma l}}{V_0^+ e^{-\gamma l}} = \frac{Z_{\text{L}} - Z_0}{Z_{\text{L}} + Z_0} \tag{31.25}$$

For devices that support quasi or strong non-TEM waves like microstrip lines, hollow waveguides, and dielectric and optical waveguides, a voltage cannot be uniquely defined. That is why several definitions of the characteristic impedance Z_0 exist [5].

31.4 Interconnections and Graphical Representations

Since Kirchhoff's laws for voltages and currents also hold for complex quantities, the rules for series and parallel connections of resistances and susceptances in the dc case apply as well for immittances.
Series connection:

$$Z = \sum_i Z_i \quad \frac{1}{Y} = \sum_i \frac{1}{Y_i} \tag{31.26}$$

Parallel connection:

$$Y = \sum_i Y_i \quad \frac{1}{Z} = \sum_i \frac{1}{Z_i} \tag{31.27}$$

As an example, consider a simplified equivalent circuit of a wire-wound resistor with the nominal value R_0 including parasitic elements that represent the effects of windings and terminal (Figure 31.5). Gradually using the rules for series and parallel connection and the impedances for inductances (Equation 31.17) and capacitances (Equation 31.19), the impedance of the real resistor with parasitic elements as given leads to

$$Z = \frac{R_0 + j\omega L}{1 - \omega^2 LC + j\omega R_0} \tag{31.28}$$

The magnitude and phase of Z/R_0 as a function of ω/ω_0 are shown in Figure 31.6 with $\omega_0 = 1/\sqrt{LC}$ as the resonant frequency defined by the parasitic elements. The network is inductive for low ($\varphi_z > 0$) and capacitive for high frequencies. An alternative representation is to plot real and imaginary parts in the impedance plane with the frequency as parameter as indicated by the labels (Figure 31.7). This version, called the *locus*, is very suitable to see immittance changes caused by parameters like frequency or adjustable elements within the network. Note that both real and imaginary parts are parameter dependent and vary with frequency.

In high-frequency applications, one obtains the impedance more easily from the reflection coefficient. Rewriting Equation 31.25 in the form

$$\Gamma = \frac{\bar{Z}_L - 1}{\bar{Z}_L + 1} \quad \text{with} \quad \bar{Z}_L = \frac{Z_L}{Z_0} \tag{31.29}$$

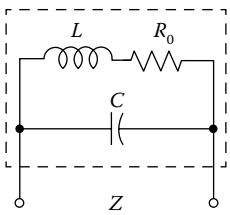


FIGURE 31.5 The simple equivalent circuit of a wire-wound resistor with nominal value R_0, inductance of the winding L, and capacitance of winding and terminal C. It is valid for a wide frequency range.

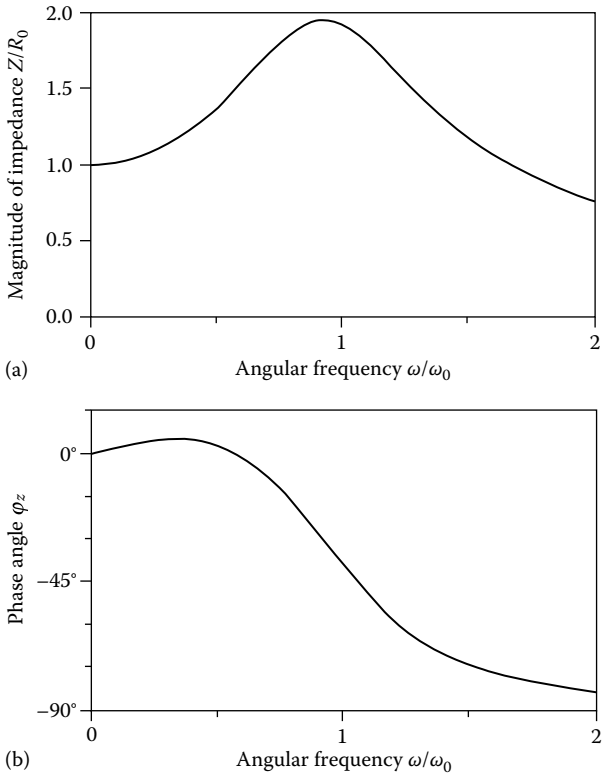


FIGURE 31.6 Normalized magnitude (a) and phase (b) of the impedance of a wire-wound resistor varying with frequency. ω_0 is the resonant frequency defined by the parasitic elements.

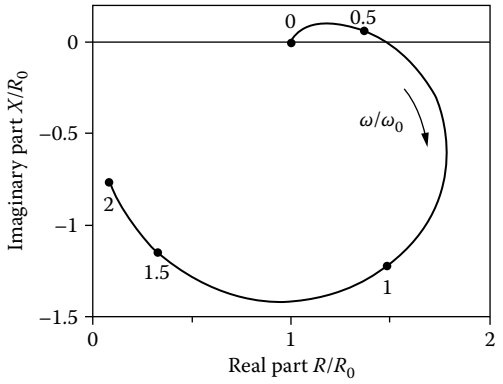


FIGURE 31.7 Normalized impedance of a wire-wound resistor in the complex plane. The arrow indicates the direction of increasing frequency.

defines a transformation of which the graphical representation has been called the Smith chart (Figure 31.8). It can be regarded as two coordinate systems lying one on top of the other. The reflection coefficient is given in polar coordinates around the center; the circles give the real and imaginary part of the associated impedance. The Smith chart is very useful for solving transmission line and waveguide impedance matching problems [6].

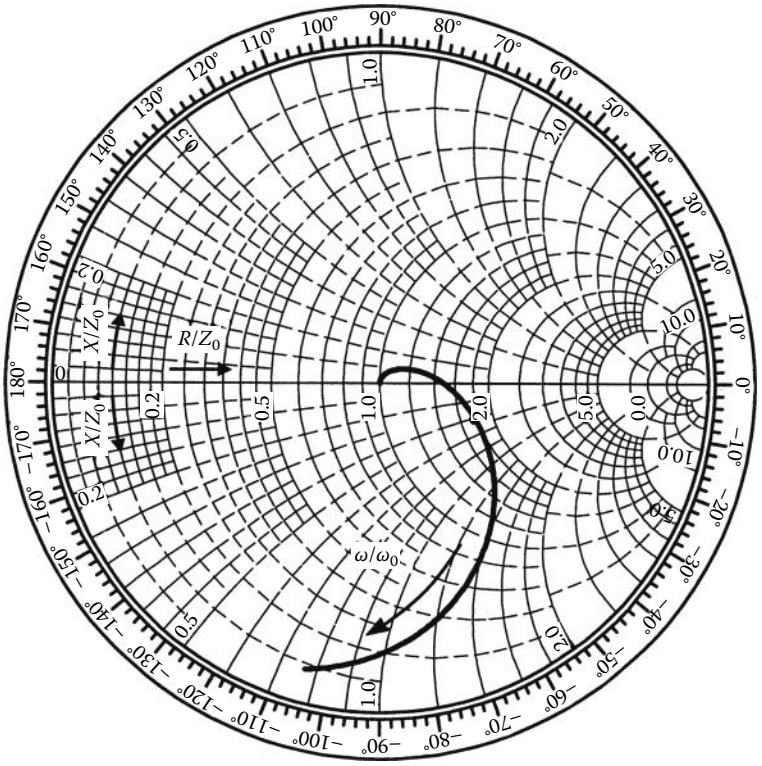


FIGURE 31.8 Smith chart representation of the impedance of a wire-wound resistor.

31.5 Measurement Techniques

Since immittances are complex quantities, one must determine two parameters: magnitude and phase or real and imaginary part, described as vector measurements. There exist several techniques depending on frequency range and required accuracy [7].

31.5.1 Current–Voltage Methods

A simple way to measure immittances follows directly from the defining Equation 31.3. Applying a well-known sinusoidal voltage to the terminal and measuring magnitude and phase of the current give the desired quantity (Figure 31.1). However, the internal impedance Z_A of the ammeter should be known exactly and the unknown impedance is then given by

$$Z = \frac{V}{I} - Z_A \tag{31.30}$$

In practical applications, impedances below 1000 Ω are measured by passing a predetermined current through the unknown device and measuring the voltage across it. Phase angle information is obtained by comparing the relative phase between voltage and current by means of a phase detector.

A variant on this method using only the better practicable voltage measurements is shown in Figure 31.9. The accurately known resistor R must be small compared to Z_x and to the internal resistance of V_2. One finds that

$$Z_x = \left(\frac{V_1}{V_2} - 1\right)R, \quad \text{or} \quad Z_x \approx \frac{V_1}{V_2}R \quad \text{if } R \ll |Z_x| \tag{31.31}$$

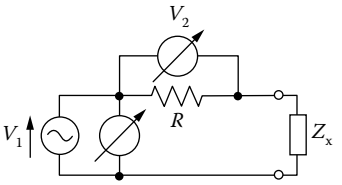


FIGURE 31.9 Determination of impedance Z_x by phase-sensitive voltage measurements, only using a well-known resistor R.

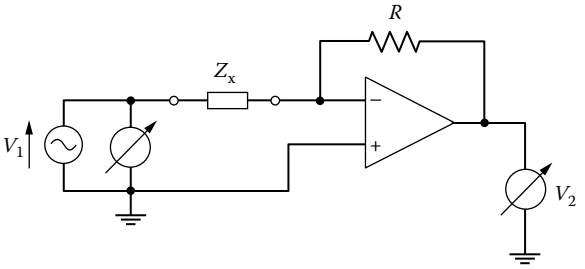


FIGURE 31.10 Impedance measurements with an inverting operational amplifier circuit. Its advantages are high input and low output resistance.

The measurement can be enhanced using an operational amplifier with high input and low output resistance in an inverting circuit (Figure 31.10). The unknown is then given by

$$Z_x = -\frac{V_1}{V_2}R \tag{31.32}$$

Practical implementations use operational amplifiers as part of an autobalancing bridge; see [7].

31.5.2 Bridge Methods

Alternating current bridges are low-cost standard laboratory devices to measure impedances over a wide frequency range from dc up to 300 MHz with very high precision (Figure 31.11). A comprehensive survey is given in [1]. Their main advantage is that only a zero indicator in the diagonal branch is necessary. In this case, the internal impedance does not influence the accuracy and the null can be detected with a high-sensitivity ac galvanometer as well as with headphones in the audio frequency range.

If the bridge is balanced, the unknown immittance is given by

$$Z_x = \frac{Z_1}{Z_3}Z_2 \quad \text{or} \quad Y_x = \frac{Z_3}{Z_1}Y_2 \tag{31.33}$$

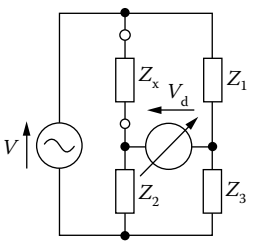


FIGURE 31.11 Impedance measurements by bridge methods. The bridge is balanced when the voltage V_d across the diagonal branch is adjusted to zero by tuning Z_1, Z_2, or Z_3.

Since the quantities are complex, Equation 31.33 involves the adjustment of two parameters: magnitude and phase

$$|Z_x| = \left|\frac{Z_1}{Z_3}\right| |Z_2|, \quad \varphi_x = \varphi_1 - \varphi_3 + \varphi_2 \tag{31.34}$$

or real and imaginary parts, respectively.

An important property of an impedance bridge is the sensitivity ε

$$\varepsilon = \left|\frac{\partial V_d}{\partial Z_x}\right| = V \frac{Z_2}{(Z_2 + Z_x)^2} \tag{31.35}$$

or (independent of Z_x)

$$\varepsilon = V \frac{Z_3^2}{Z_2 (Z_1 + Z_3)^2} \tag{31.36}$$

in the vicinity of zero crossing when the bridge is balanced.

The precision of the measurement not only depends on the exact zero adjustment, which can be enhanced by choosing the elements and the voltage according to Equation 31.36 to obtain a high sensitivity, but also on the realization of $Z_1 \ldots Z_3$. Mostly, these are connections of resistors and capacitors. Inductors are avoided because they always have a resistive component and it is difficult and expensive to manufacture inductors with exactly defined and reproducible electrical properties. There exist various types of bridges depending on how the elements are designed and interconnected. To choose the correct configuration, it must be known whether the unknown impedance is capacitive or inductive; otherwise, a zero adjustment is not always possible since the balancing condition cannot be fulfilled. Bridges are therefore principally used to measure capacitances and inductances as well as loss and quality factors of capacitors and coils. Since magnitude and phase conditions must be matched simultaneously, two elements must be tuned. To obtain a wide measurement range, the variable elements are designed as combinations of switchable and tunable capacitors and resistors. The sensitivity of the zero indicators can be changed for global search and final adjustment. Unfortunately, magnitude and phase cannot be adjusted independently of each other. If the balancing is performed by hand, a suitable strategy is to search the minimum voltage by tuning each element successively.

Frequently used bridges are the Wheatstone bridge (Figure 31.12) for the measurement of lossy capacitances and the Hay bridge (Figure 31.13) to determine inductivity and quality factor of coils. Because of its symmetrical structure, the balancing condition for the Wheatstone bridge is simply

$$R_x = \alpha R_1, \quad C_x = \alpha C_1, \quad \alpha = \frac{R_3}{R_2} \tag{31.37}$$

which is independent of frequency.

The measurement of a coil with the Hay bridge requires that

$$R_x + j\omega L_x = \frac{j\omega C_3 R_1 R_2}{1 + j\omega R_3 C_3} = \frac{\omega^2 C_3^2 R_1 R_2 R_3}{1 + (\omega R_3 C_3)^2} + j\omega \frac{C_3 R_1 R_2}{1 + (\omega R_3 C_3)^2} \tag{31.38}$$

from which the quality factor is obtained as

$$Q = \frac{\omega L_x}{R_x} = \frac{1}{\omega R_3 C_3} \tag{31.39}$$

The inductance of high-Q coils can be determined frequency independent since

$$L_x \approx R_1 R_2 C_3 \quad \text{if } (\omega R_3 C_3)^2 \ll 1 \tag{31.40}$$

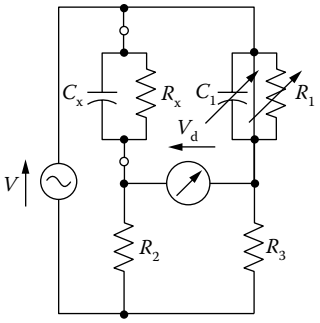


FIGURE 31.12 The Wheatstone bridge for the capacitance and dissipation factor measurement of capacitors. The balancing condition is frequency independent. The resistor R_1 and the capacitor C_1 must be tuned successively until the bridge is balanced.

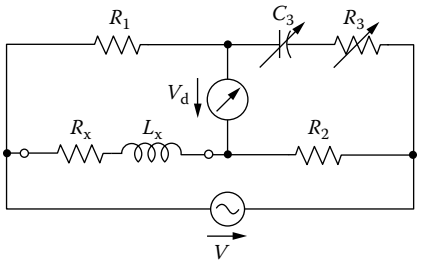


FIGURE 31.13 The Hay bridge for the measurement of the inductance and the quality factor of coils. If Q is sufficiently high, the inductance can be determined nearly frequency independent.

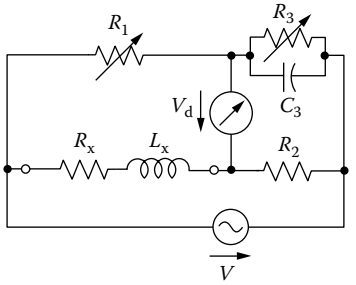


FIGURE 31.14 The Maxwell bridge with simple and frequency-independent balancing conditions. Despite these advantages, it is not recommended for high-Q coils because of a very large R_1 value.

A very interesting alternative is the Maxwell bridge (Figure 31.14), since it requires only resistors as variable elements, which can be manufactured with high precision. The balancing is frequency independent and leads to

$$R_x = \frac{R_1 R_2}{R_3}, \quad L_x = R_1 R_2 C_3, \quad Q = \omega C R_1 \tag{31.41}$$

Nevertheless, the Hay bridge is preferred for high-Q coils, because a very large value of R_1 is required for the Maxwell bridge leading to a disadvantageous balancing [8].

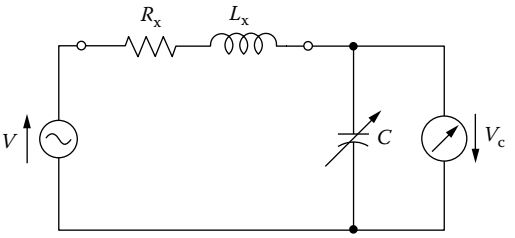


FIGURE 31.15 Coil as part of a resonance circuit to determine inductance and quality factor. The capacitor C is tuned to maximum voltage V_C.

31.5.3 Resonant Method

Using the coil as part of a resonance circuit as in Figure 31.15 and tuning C to maximum voltage, the quality factor can be measured directly as

$$Q = \left| \frac{V_{C,\max}}{V} \right| = \frac{1}{\omega R_x C} \tag{31.42}$$

The unknown inductance is then obtained from the test frequency by means of the resonance condition

$$L_x = \frac{1}{\omega^2 C} \tag{31.43}$$

If a capacitor with sufficiently low losses is used, Q values as high as 1000 can be measured.

31.5.4 Network Analysis Methods

31.5.4.1 Frequency Domain

In the case of distributed elements, measurements of currents and voltages depend on the position and are often not directly applicable to high-frequency devices like waveguides or microstrip lines. For that reason the determination of immittances is derived from measuring the reflection coefficient. Equation 31.24 shows the importance of defining a proper measurement plane. This is the cross section of the line or waveguide perpendicular to the direction of propagation at a definite length l_0, where the reflection coefficient has to be measured. It can then be transformed along the line toward load or source using this relation or the Smith chart. Exact microwave measurements are very sophisticated and need a lot of practical experience. Further details can be found in the literature [5,9–11].

Automated and precise immittance measurements over a wide frequency range are best carried out with a vector network analyzer [10]. Unfortunately, this is also the most expensive method. The principle of measurement is shown in Figure 31.16. A power divider splits the incident signal into a transmitted and a

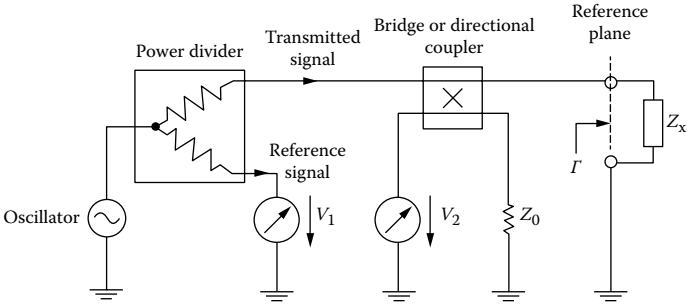


FIGURE 31.16 Schematic of network analyzer measurements. The voltage ratio V_2/V_1 of reflected wave and reference signal is proportional to the reflection coefficient Γ. The impedance Z_x can then be computed.

reference part. The directional bridge or coupler separates forward and backward traveling waves, and the reflected signal now appears in the branch with the phase-sensitive voltmeter V_2. Using a bridge with impedances matched to the line ($Z_1 = Z_2 = Z_3 = Z_0$), the voltage in the diagonal branch is given by (Figure 31.11)

$$V_d = \frac{Z_x - Z_0}{2(Z_x + Z_0)} V = V_2 \tag{31.44}$$

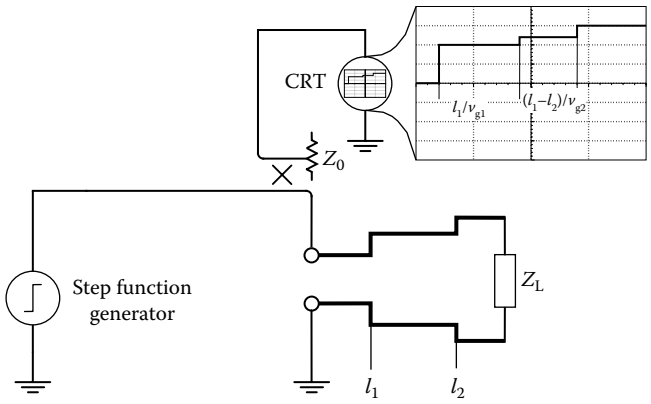


FIGURE 31.17 Detection and measurement of impedance steps on a line or waveguide with a TDR. Since phase information is usually not available, the characteristics and magnitudes of the impedances can only be estimated. Notice that the group or signal velocity v_g varies from step to step.

TABLE 31.1 Instruments for Immittance Measurements

Manufacturer	Model Number	Description
Agilent Technologies	4285A	Precision LCR Meter 75 kHz–30 MHz
Agilent Technologies	4287A	RF LCR Meter 1 MHz–3 GHz
Agilent Technologies	4294A	Precision Impedance Analyzer 40 Hz–110 MHz
Agilent Technologies	E4980A	Precision LCR Meter 20 Hz–2 MHz
Agilent Technologies	E5061B	Network Analyzer 5 Hz–3 GHz
Agilent Technologies	N9923A	FieldFox Handheld RF Vector Analyzer 2 MHz–4/6 GHz
Agilent Technologies	PNA Series	Vector Network Analyzers 300 kHz–1.05 THz
Anritsu	ME7828A	Vector Network Analyzer 70 kHz–110 GHz
Fluke	PM 6303A	Automatic RCL Meter
Fluke	PM 6306/6304	Automatic RCL Meter
Keithley	3321	LCZ Meter, 4 Test Frequencies to 100 kHz
Keithley	3322	LCZ Meter, 11 Test Frequencies to 100 kHz
Keithley	3330	LCZ Meter 20 Hz–200 kHz
QuadTech	1715 CE	LCR Meter 4 Test Frequencies to 10 kHz
QuadTech	1730 T	LCR Meter 7 Test Frequencies to 100 kHz
QuadTech	1910/1920 CE	LCR Meter 20 Hz–1 MHz
QuadTech	7600 Plus	LCR Meter 10 Hz–2 MHz
Rohde & Schwarz	ZVA Series	Vector Network Analyzer 300 kHz–80 GHz
SST	SR715/720	LCR Meter 10 kHz/100 kHz
Tinsley Precision Instruments	QJ Series	Portable Bridges (Wheatstone, Kelvin)
TTi	LCR400	Precision LCR Meter
Voltech	ATi	Precision Transformer Tester

and thus the reflection coefficient

$$\Gamma = \frac{Z_x - Z_0}{Z_x + Z_0} = \alpha \frac{V_2}{V_1} \tag{31.45}$$

is directly proportional to the voltage ratio.

Network analyzers use an automatic error correction to eliminate the effect of internal and external couplers and junctions. Because of that, a calibration procedure with standard terminations is necessary. These terminations must be manufactured very precisely, since they define the measurement plane and determine the overall measurement error.

31.5.4.2 Time Domain

It is often necessary to locate an impedance step, whether to find out the distance of a cable defect or to track down the origin of reflections within a connection. To this end, high-performance vector network analyzers have a Fourier transform procedure. But there also exist cheaper time-domain reflectometers (TDRs) [12]. They use an incident step or impulse signal (Figure 31.17) and the reflected signal is

TABLE 31.2 Companies Producing Immittance Measurement Equipment

Agilent Technologies
5301 Stevens Creek Blvd.
Santa Clara, CA 95051
(800) 829-4444
www.agilent.com

Fluke Corporation
6920 Seaway Boulevard
Everett, WA 98203
(800) 44-FLUKE (443-5853)
www.fluke.com

QuadTech
734 Forest St.
Suite 500
Marlborough, MA 01752
(800) 253-1230
www.quadtech.com

SRS Stanford Research Systems, Inc.
1290-D Reamwood Ave.
Sunnyvale, CA 94089
(408) 744-9040
www.thinksrs.com

TTi (Thurlby Thandar Instruments Ltd.)
Glebe Road
Huntingdon
Cambridgeshire PE29 7DR
United Kingdom
+44-1480-412451
www.tti-test.com

Anritsu Company
Americas Sales Region Headquarters
1155 East Collins Blvd., Suite 100
Richardson, TX 75081
(800) ANRITSU (267-4878)
www.anritsu.com

Keithley Instruments, Inc.
28775 Aurora Road
Cleveland, OH 44139
(800) 552-1115
www.keithley.com

Rohde & Schwarz, Inc.
8661A Robert Fulton Drive
Columbia, MD 21046-2265
(410) 910-7800
www.rohde-schwarz.com/USA

Tinsley Precision Instruments
2 Gatton Park Business Centre
Wells Place
Redhill
Surrey RH1 3LG
+44-1737-649300
www.tinsley.co.uk

Voltech Instruments, Inc.
12851 Banyan Creek Drive
Fort Myers, FL 33908
(239) 437-0494
www.voltech.com

separated by means of a directional coupler and displayed on an oscilloscope in the time domain. From the shape of the signal, the impedance step can be localized by means of the time delay:

$$l = \frac{1}{2} v_g t \tag{31.46}$$

with v_g as signal or group velocity on the line varying from step to step. Characteristic and magnitude of the impedance can only be estimated, since phase information is usually not available. TDR measurements are restricted to the localization of impedance steps and not to be recommended for exact measurements. Moreover, additional pulse deformations occur in dispersive waveguides.

31.6 Instrumentation and Manufacturers

A broad range of instrumentation for measuring immittance is available. Some of these instruments are included in Table 31.1. Table 31.2 provides the names and addresses of some companies that produce immittance-measuring instrumentation.

Defining Terms

Admittance (*Y*): The reciprocal of impedance.

Immittance: A response function for which one variable is a voltage and the other a current. Immittance is a general term for both impedance and admittance, used where the distinction is irrelevant.

Impedance (*Z*): The ratio of the phasor equivalent of a steady-state sine-wave voltage to the phasor equivalent of a steady-state sine-wave current. The real part is the *resistance*; the imaginary part is the *reactance*.

Phasor: A complex number, associated with sinusoidally varying electrical quantities, such that the absolute value (modulus) of the complex number corresponds to either the peak amplitude or root mean square (rms) value of the quantity and the phase (argument) to the phase angle at zero time. The term "phasor" can also be applied to impedance and related complex quantities that are not time dependent.

Reflection coefficient: At a given frequency, at a given point, and for a given mode of propagation, the ratio of voltage, current, or power of the reflected wave to the corresponding quantity of the incident wave.

References

1. B. M. Oliver and J. M. Cage, *Electronic Measurements and Instrumentation*, New York: McGraw-Hill, 1971.
2. H. W. Bode, *Network Analysis and Feedback Amplifier Design*, Princeton, NJ: Van Nostrand, 1959.
3. A. T. de Hoop, *Handbook of Radiation and Scattering of Waves*, London, U.K.: Academic Press, 1995.
4. R. E. Collin, *Foundations for Microwave Engineering*, New York: IEEE Press, 2001.
5. P. I. Somlo and J. D. Hunter, *Microwave Impedance Measurement*, London, U.K.: Peter Peregrinus, 1985.
6. R. L. Thomas, *A Practical Introduction to Impedance Matching*, Dedham, MA: Artech House, 1976.
7. M. Honda, *The Impedance Measurement Handbook*, Yokogawa, Japan: Hewlett-Packard, 1989.
8. David M. Pozar, *Microwave Engineering*, 5th Edn., John Wiley & Sons, New Jersey, 2004.
9. T. S. Laverghetta, *Modern Microwave Measurements and Techniques*, Norwood, MA: Artech House, 1988.
10. J. P. Dunsmore, *Handbook of Microwave Component Measurement: With Advanced VNA Techniques*, Wiley, Chichester, West Essex, U.K., 2012.
11. T. S. Laverghetta, *Handbook of Microwave Testing*, Dedham, MA: Artech House, 1981.
12. E. K. Miller (ed.), *Time-Domain Measurements in Electromagnetics*, New York: Van Nostrand Reinhold, 1986.

Further Information

Anonymous, *Low Level Measurements Handbook*, Cleveland, OH: Keithley Instruments, Inc., 2004.

Bobrow L. S., *Fundamentals of Electrical Engineering*, New York: Oxford University Press, 1996.

Fujimoto H., G. Mana, and K. Nakayama, A universal geometry for calculable frequency-response coefficient of LCR standards and new 10-MHz resistance and 1.6-MHz quadrature-bridge systems, *IEEE Trans. Instrum. Meas*., 56(2), 221–225, April 2007.

Morris A. S., *Measurement and Instrumentation Principles*, Oxford, U.K.: Butterworth-Heinemann, 2001.

Roadstrum W. H. and D. H. Wolaver, *Electrical Engineering for all Engineers*, New York: John Wiley & Sons, 1994.

Suzuki K., A new self-calibration method of an LCR meter for RF resistance calibration using capacitance standards, *IEEE Trans. Instrum. Meas*., 58(4), 993 –996, April 2009.

Whitczak M., R. Rybski, and J. Kaczmarek, Impedance measurement with the D-optimum experimental conditions, *IEEE Trans. Instrum. Meas*., 58(8), 2535–2543, August 2009.

32

Distortion Measurement

Michael F. Toner
Nortel Networks

Gordon W. Roberts
McGill University

A sine-wave signal will have only a single-frequency component in its spectrum, that is, the frequency of the tone. However, if the sine wave is transmitted through a system (such as an amplifier) having some nonlinearity, then the signal emerging from the output of the system will no longer be a pure sine wave. That is, the output signal will be a distorted representation of the input signal. Since only a pure sine wave can have a single component in its frequency spectrum, this situation implies that the output must have other frequencies in its spectral composition. In the case of *harmonic distortion*, the frequency spectrum of the distorted signal will consist of the fundamental (which is the same frequency as the input sine wave) plus harmonic frequency components that are at integer multiples of the fundamental frequency. Taken together, these will form a Fourier representation of the distorted output signal. This phenomenon can be described mathematically. Refer to Figure 32.1, which depicts a sine-wave input signal $x(t)$ at frequency f_1, applied to the input of a system $A(x)$, which has an output $y(t)$. Assume that system $A(x)$ has some nonlinearity. If the nonlinearity is severe enough, then the output $y(t)$ might have excessive harmonic distortion such that its shape no longer resembles the input sine wave. Consider the example where the system $A(x)$ is an audio amplifier and $x(t)$ is a voice signal. Severe distortion can result in a situation where the output signal $y(t)$ does not represent intelligible speech. The *total harmonic distortion* (THD) is a figure of merit that is indicative of the quality with which the system $A(x)$ can reproduce an input signal $x(t)$. The output signal $y(t)$ can be expressed as

$$y(t) = a_0 + \sum_{k=1}^{N} a_k \cos(2\pi k f_1 t + \theta_k) \tag{32.1}$$

where

a_k, $k = 0, 1, \ldots, N$ are the magnitudes of the Fourier coefficients
θ_k, $k = 0, 1, \ldots, N$ are the corresponding phases

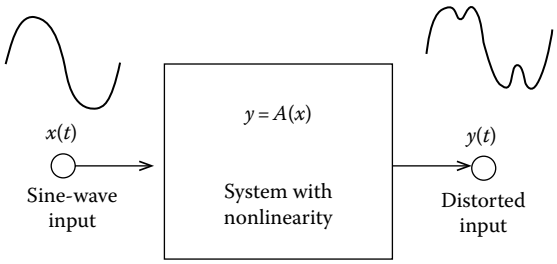


FIGURE 32.1 Any system with nonlinearity gives rise to distortion.

The THD is defined as the percentage ratio of the rms voltage of all harmonics components above the fundamental frequency to the rms voltage of the fundamental. Mathematically, the definition is written as

$$\text{THD} = \frac{\sqrt{\sum_{k=2}^{N} a_k^2}}{a_1} \times 100\% \tag{32.2}$$

If the system has good linearity (which implies low distortion), then the THD will be a smaller number than that for a system having poorer linearity (higher distortion). To provide the reader with some feeling for the order of magnitude of a realistic THD, a reasonable audio amplifier for an intercom system might have a THD of about 2% or less, while a high-quality sound system might have a THD of 0.01% or less. For the THD to be meaningful, the bandwidth of the system must be such that the fundamental and the harmonics will lie within the passband. Therefore, the THD is usually used in relation to low-pass systems or band-pass systems with a wide bandwidth. For example, an audio amplifier might have a 20 Hz to 20 kHz bandwidth, which means that a 1 kHz input sine wave could give rise to distortion up to the 20th harmonic (i.e., 20 kHz), which can lie within the passband of the amplifier. On the other hand, a sine wave applied to the input of a narrow-band system such as a radio frequency amplifier will give rise to harmonic frequencies that are outside the bandwidth of the amplifier. These kinds of narrow-band systems are best measured using *intermodulation distortion*, which is treated elsewhere in this handbook. For the rest of the discussion at hand, consider the example system illustrated in Figure 32.1 that shows an amplifier system $A(x)$ that is intended to be linear but has some undesired nonlinearities. Obviously, if a linear amplifier is the design objective, then the THD should be minimized.

32.1 Mathematical Background

Let $y = A(x)$ represent the input–output transfer characteristic of the system $A(x)$ in Figure 32.1 containing the nonlinearity. Expanding into a power series yields

$$A(x) = \sum_{k=0}^{\infty} c_k x^k = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots \tag{32.3}$$

Let the input to the system be $x = \cos(2\pi f_0 t)$. Then the output will be

$$y = A(x) = c_0 + c_1 A_0 \cos(2\pi f_0 t) + c_2 A_0^2 \cos^2(2\pi f_0 t) + c_3 A_0^3 \cos^3(2\pi f_0 t) + \cdots \tag{32.4}$$

This can be simplified using the trigonometric relationships

$$\cos^2(\theta) = \frac{1}{2} - \frac{1}{2}\cos(2\theta) \tag{32.5}$$

$$\cos^3(\theta) = \frac{3}{4}\cos(\theta) + \frac{1}{4}\cos(3\theta) \tag{32.6}$$

$$\cos^4(\theta) = \frac{1}{8} - \frac{1}{2}\cos(2\theta) + \frac{1}{8}\cos(4\theta) \tag{32.7}$$

$$\cos^5(\theta) = \frac{5}{8}\cos(\theta) + \frac{5}{16}\cos(3\theta) + \frac{1}{16}\cos(5\theta) \tag{32.8}$$

and so on. Performing the appropriate substitutions and collecting terms results in an expression for the distorted signal $y(t)$ that is of the form shown in Equation 32.1. The THD can then be computed from Equation 32.2.

Closer inspection of Equations 32.6 and 32.8 reveals that a cosine wave raised to an odd power gives rise to only the fundamental and odd harmonics, with the highest harmonic corresponding to the highest power. A similar phenomenon is observed for a cosine raised to even powers; however, the result is only a dc component and even harmonics without any fundamental component. In fact, any nonlinear system that possesses an odd input–output transfer characteristic $A(x)$ (i.e., the function $A(x)$ is such that $-A(x) = A(-x)$) will give rise to odd harmonics only. Consider Figure 32.2, which illustrates an example of two-sided symmetrical clipping. It is an odd function. The application of a sine wave to its input will result in a waveform similar to that shown in Figure 32.3, which has only odd harmonics as shown in Figure 32.4. The majority of physical systems are neither odd nor even. (An even function is one that has the property $A(x) = A(-x)$; e.g., a full-wave rectifier.) Consider the enhancement N-type metal oxide semiconductors (NMOS) transistor illustrated in Figure 32.5, which

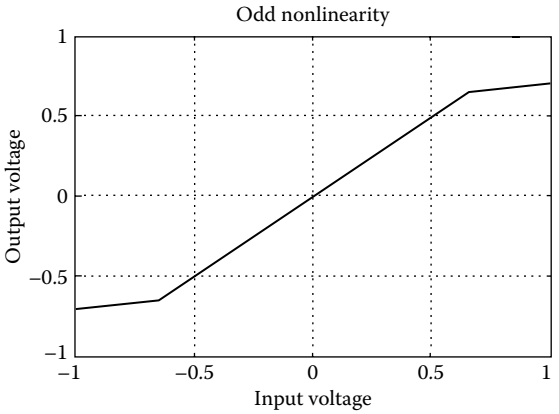


FIGURE 32.2 An off nonlinearity with $f(-x) = -f(x)$.

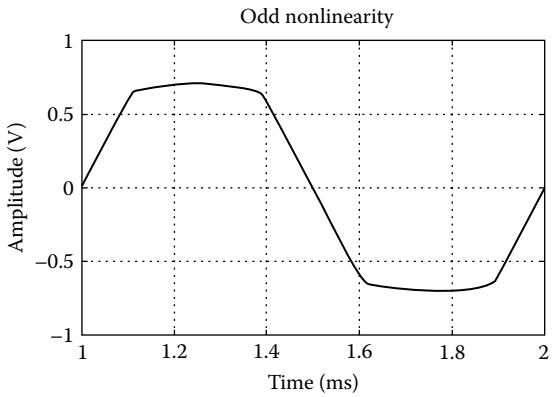


FIGURE 32.3 Distortion due to symmetrical two-sided clipping.

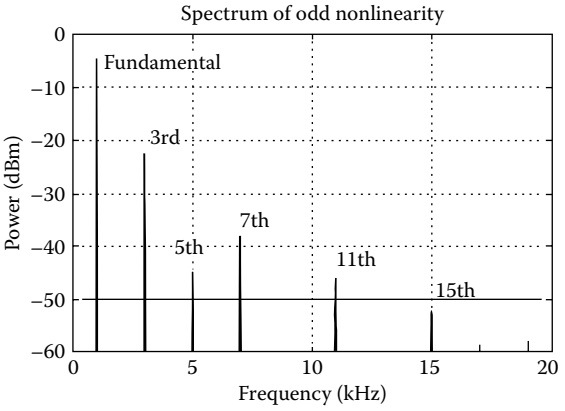


FIGURE 32.4 Frequency spectrum of signal distorted by symmetrical two-sided clipping.

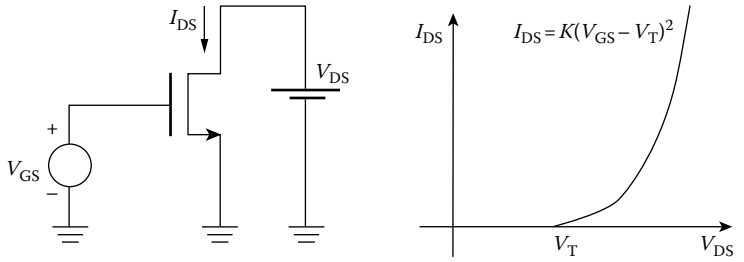


FIGURE 32.5 NMOS enhancement transistor is actually a nonlinear device. It is neither odd nor even in the strict sense.

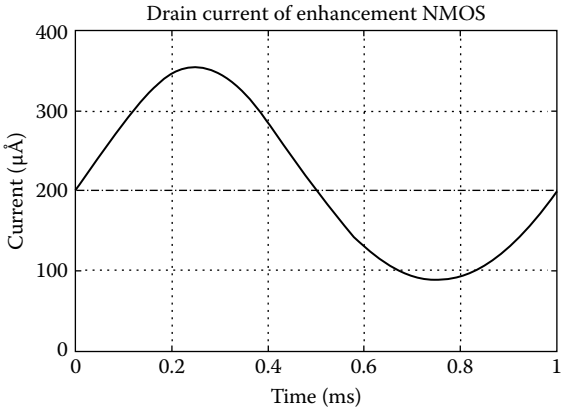


FIGURE 32.6 Showing how the drain current of the enhancement NMOS device is distorted.

has the square-law characteristic shown. Assume that the voltage V_{GS} consists of a dc bias plus a sine wave such that V_{GS} is always more positive than V_T (the threshold voltage). Then the current flowing in the drain of this NMOS transistor could have the appearance shown in Figure 32.6. It is observed that the drain current is distorted, since the positive-going side has a greater swing than the negative-going side. The equation for the drain current can be derived mathematically as follows. A MOS transistor operating in its saturation region can be approximated as a square-law device:

$$I_{DS} = \frac{\mu C_{ox}}{2}\frac{W}{L}(V_{GS} - V_T)^2 \tag{32.9}$$

If the gate of the *n*-channel enhancement MOSFET is driven by a voltage source consisting of a sine-wave generator in series with a dc bias, that is,

$$V_{GS} = V_B + A_0 \sin(2\pi f_0 t) \tag{32.10}$$

then the current in the drain can be written as

$$I_{DS} = \frac{\mu C_{ox}}{2}\frac{W}{L}\{[V_B + A_0 \sin(2\pi f_0 t)] - V_T\}^2 \tag{32.11}$$

Expanding and using the trigonometric relationship

$$\sin^2(\theta) = \frac{1}{2} - \frac{1}{2}\sin\left(2\theta + \frac{\pi}{2}\right) \tag{32.12}$$

Equation 32.11 can be rewritten as

$$I_{DS} = \frac{\mu C_{ox}}{2}\frac{W}{L}\left[(V_B - V_T)^2 + \frac{A_0^2}{2} + 2(V_B - V_T)A_0 \sin(2\pi f_0 t) + \frac{A_0^2}{2}\sin(4\pi f_0 t) + \frac{\pi}{2}\right] \tag{32.13}$$

which clearly shows the dc bias, the fundamental, and the second harmonic that are visible in the spectrum of the drain current I_{DS} in Figure 32.7. There is one odd harmonic (i.e., the fundamental) and two even harmonics (strictly counting the dc component and the second harmonic). This particular transfer characteristic is neither odd nor even. Finally, for an ideal square-law characteristic, the second harmonic is the highest frequency component generated in response to a sine-wave input. Another example of a transfer characteristic that is neither odd nor even is the single-sided clipping as shown in Figure 32.8, which gives rise to the distortion of Figure 32.9. One last example of an odd input–output transfer characteristic is symmetrical crossover distortion as depicted in Figure 32.10. The distorted output in response to a 1 kHz sine-wave input is shown in Figure 32.11. The spectrum of the output signal is shown in Figure 32.12. Note that only odd harmonics have been generated.

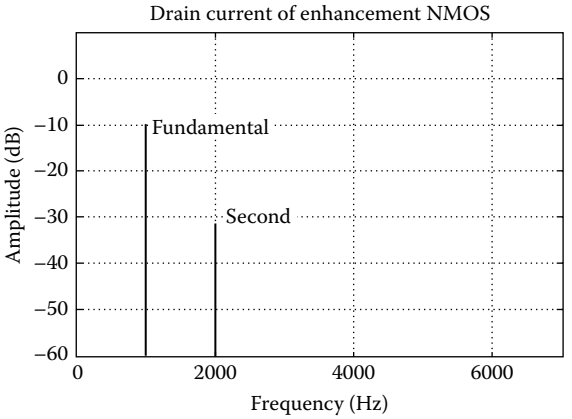


FIGURE 32.7 The drain current contains a dc bias, the fundamental, and the second harmonic only, for an ideal device.

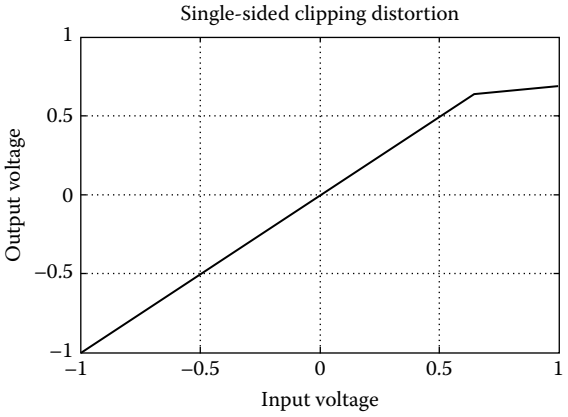


FIGURE 32.8 Single-sided clipping is neither even nor odd.

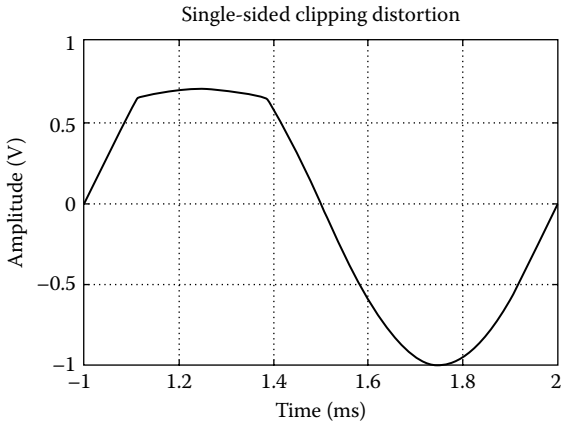


FIGURE 32.9 Distortion due to single-sided clipping.

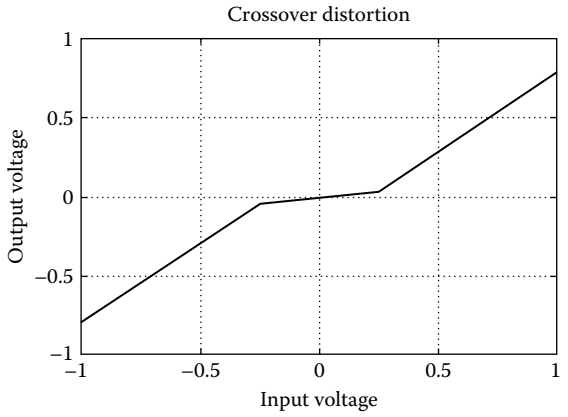


FIGURE 32.10 Symmetrical crossover distortion is odd.

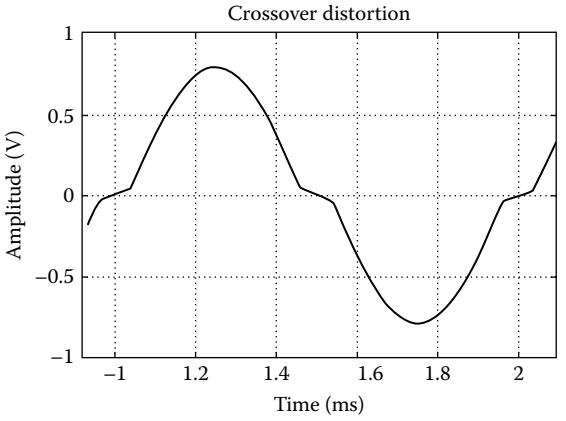


FIGURE 32.11 An example of crossover distortion.

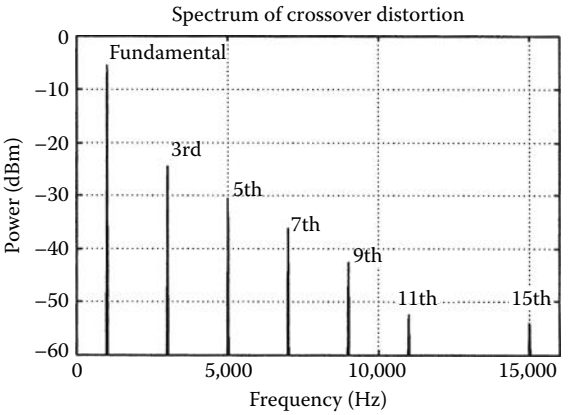


FIGURE 32.12 Symmetrical crossover distortion gives rise to odd harmonics.

To round out the discussion, consider a mathematical example wherein the harmonics are derived algebraically. Consider an input–output transfer function $f(x) = c_1x + c_3x^3 + c_5x^5$ that has only odd powers of x. If the input is a cosine $x = A_0\cos(2\pi f_0 t)$, then the output will be of the form

$$y(t) = f(x) = c_1A_0\cos(2\pi f_0 t) + c_3A_0^3\cos^3(2\pi f_0 t) + c_5A_0^5\cos^5(2\pi f_0 t) \tag{32.14}$$

This can be simplified using the trigonometric relationships given in Equations 32.5 through 32.8 with the following result:

$$y(t) = \left(c_1A_0 + \frac{3c_3A_0^3}{4} + \frac{5c_5A_0^5}{8}\right)\cos(2\pi f_0 t) + \left(\frac{c_3A_0^3}{4} + \frac{5c_5A_0^5}{16}\right)\cos(2\pi 3 f_0 t) + c_5A_0^5\cos^5(2\pi 5 f_0 t) \tag{32.15}$$

Clearly, only the fundamental plus the third and fifth harmonics are present. Should the exercise be repeated for an input–output transfer function consisting of only even powers of x, then only a dc offset plus even harmonics (not the fundamental) would be present in the output.

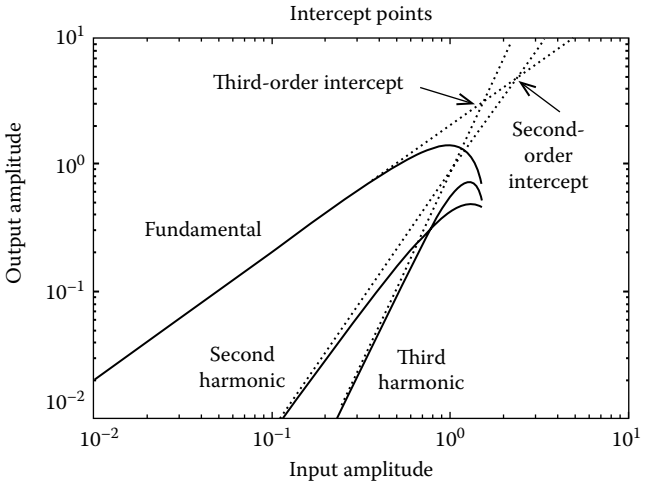


FIGURE 32.13 An example showing the IP2 and IP3 for a hypothetical system. Both axes are plotted on a logarithmic scale.

32.2 Intercept Points

It is often desirable to visualize how the various harmonics increase or decrease as the amplitude of the input sine wave $x(t)$ is changed. Consider the example of a signal applied to a nonlinear system $A(x)$ having single-sided clipping distortion as shown in Figure 32.8. The clipping becomes more severe as the amplitude of the input signal $x(t)$ increases in amplitude, so the distortion of the output signal $y(t)$ becomes worse. The *intercept point* (IP) is used to provide a figure of merit to quantify this phenomenon. Consider Figure 32.13, which shows an example of the power levels of the first three harmonics of the distorted output $y(t)$ of a hypothetical system $A(x)$ in response to a sine-wave input $x(t)$. It is convenient to plot both axes on a log scale. It can be seen that the power in the harmonics increases more quickly than the power in the fundamental. This is consistent with the observation of how clipping becomes worse as the amplitude increases. It is also consistent with the observation that, in the equations earlier, the higher harmonics will rapidly become more prominent because they are proportional to higher exponential powers of the input signal amplitude. The IP for a particular harmonic is the power level where the extrapolated line for that harmonic intersects with the extrapolated line for the fundamental. The second-order intercept is often abbreviated IP2, the third-order intercept abbreviated IP3, etc.

32.3 Measurement of the THD

32.3.1 Classical Method

The traditional method of measuring THD is shown in Figure 32.14. A sine-wave test stimulus $x(t)$ is applied to the input of the system $A(x)$ under test. The system output $y(t)$ is fed through a band-pass filter tuned to the frequency of the input stimulus to extract the signal. Its power p_1 can be measured

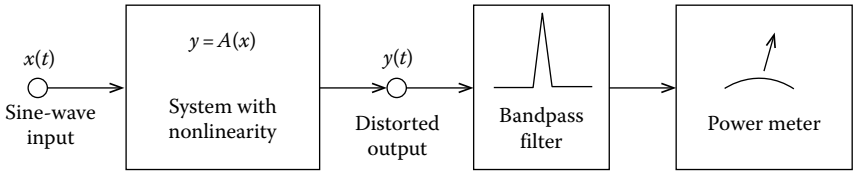


FIGURE 32.14 Illustrating the classical method of measuring THD.

with a power meter. The band-pass filter is then tuned to each of the desired harmonics in turn and the measurement is repeated to determine the required p_i. The THD is then calculated from

$$\text{THD} = \sqrt{\frac{\sum_{k=2}^{N} p_k}{p_1}} \times 100\% \tag{32.16}$$

In the case of an ordinary audio amplifier, nearly all of the power in the distorted output signal is contained in the first 10 or 11 harmonics. However, in more specialized applications, a much larger number of harmonics might need to be considered.

32.3.2 Spectrum Analyzer Method

THD measurements are often made with a spectrum analyzer using the setup shown in Figure 32.15. The readings for the power levels of each of the desired harmonic components in the frequency spectrum of the distorted signal $y(t)$ are collected from the spectrum analyzer, usually in units of dB. They are converted to linear units by means of the relationship

$$a_i = 10^{r_i/20} \tag{32.17}$$

where r_i is the reading for the ith component in dB. The THD is then computed from Equation 32.2. The spectrum analyzer method can be considered as an extension of the classical method described earlier, except that the spectrum analyzer itself is replacing both the band-pass filter and the power meter.

32.3.3 DSP Method

Digital signal processing (DSP) techniques have recently become popular for use in THD measurement. In this method, the distorted output $y(t)$ is digitized by a precision A/D converter and the samples are stored in the computer's memory as shown in Figure 32.16. One assumes that the samples have been collected with a uniform sample period T_s and that appropriate precautions have been taken with regard

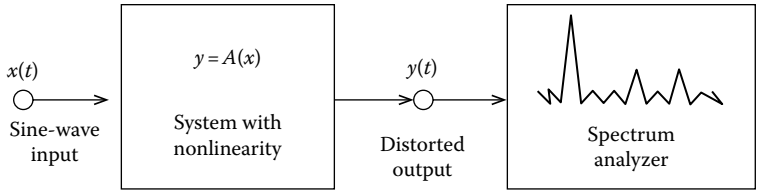


FIGURE 32.15 Illustrating measurement of THD using a spectrum analyzer.

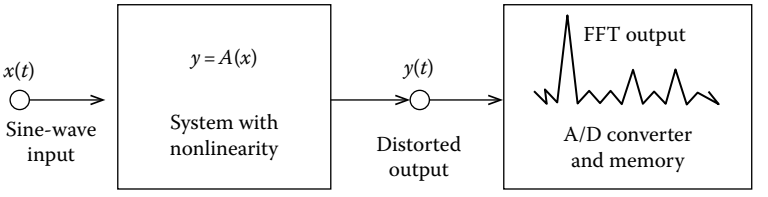


FIGURE 32.16 Illustrating the measurement of THD using FFT.

to the Nyquist criterion and aliasing. Let $y(n)$ refer to the nth stored sample. A fast Fourier transform (FFT) is executed on the stored data using the relationship

$$Y(k) = \sum_{n=0}^{N-1} y(n) e^{-j(2\pi/N)kn} \tag{32.18}$$

where N is the number of samples that have been captured. The frequency of the input test stimulus is chosen such that the sampling is coherent. Coherency in this context means that if N samples have been captured, then the input test stimulus is made to be a harmonic of the primitive frequency f_P, which is defined as

$$f_P = \frac{f_s}{N} = \frac{1}{T_s N} \tag{32.19}$$

One can view the primitive frequency f_P as the frequency of a sinusoidal signal whose period is exactly equal to the time interval formed by the N-points. Thus, the frequency of the test stimulus can be written as

$$f_0 = M \times f_P = M \times \frac{f_s}{N} = \frac{M}{N} \times f_s \tag{32.20}$$

where M and N are integers. To maximize the information content collected by a set of N-points, M and N are selected so that they have no common factors, that is, relatively prime. This ensures that every sample is taken at a different point on the periodic waveform. An example is provided in Figure 32.17, where $M = 3$ and $N = 32$. The FFT is executed on the distorted signal as per Equation 32.18, and then the THD is computed from

$$\text{THD} = \frac{\sqrt{\sum_{k=2}^{N} |Y(k \times M)|^2}}{|Y(M)|} \times 100\% \tag{32.21}$$

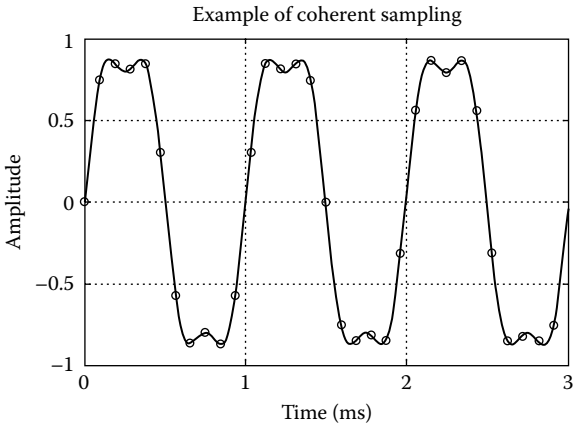


FIGURE 32.17 With coherent sampling, each sample occurs on a unique point of the signal.

32.4 Conclusions

The THD is a figure of merit for the quality of the transmission of a signal through a system having some nonlinearity. Its causes and some methods of measuring it have been discussed. Some simple mathematical examples have been presented. However, in real-world systems, it is generally quite difficult to extract all of the parameters c_k in the transfer characteristic. The examples were intended merely to assist the reader's understanding of the relationship between even and odd functions and the harmonics that arise in response to them.

Defining Terms

Distortion: The effect of corrupting a signal with undesired frequency components.

Fundamental: The lowest frequency component of a signal other than zero frequency.

Harmonic: Any frequency component of a signal that is an integer multiple of the fundamental frequency.

Nonlinearity: The deviation from the ideal of the transfer function, resulting in such effects as clipping or saturation of the signal.

Total harmonic distortion (THD): A numerical figure of merit of the quality of transmission of a signal, defined as the ratio of the power in all the harmonics to the power in the fundamental.

Further Information

Mahoney M., *DSP-Based Testing of Analog and Mixed-Signal Circuits*, Los Alamos, CA: EEE Computer Society Press, 1987.

Pederson D. O. and K. Mayaram, *Analog Integrated Circuits for Communications*, New York: Kluwer Academic Press, 1991.

33

Noise Measurement

W. Marshall Leach, Jr.
Georgia Institute of Technology

This chapter describes the principal sources of electric noise and discusses methods for the measurement of noise. The notations for voltages and currents correspond to the following conventions: dc quantities are indicated by an uppercase letter with uppercase subscripts, for example, V_{BE}. Instantaneous small-signal ac quantities are indicated by a lowercase letter with lowercase subscripts, for example, v_n. The mean-square value of a variable is denoted by a bar over the square

of the variable, for example, $\overline{v_n^2}$, where the bar indicates an arithmetic average of an ensemble of functions. The root-mean-square or rms value is the square root of the mean-square value. Phasors are indicated by an uppercase letter and lower case subscripts, for example, V_n. Circuit symbols for independent sources are circular, symbols for controlled sources are diamond shaped, and symbols for noise sources are square. In the evaluation of noise equations, Boltzmann's constant is $k = 1.38 \times 10^{-23}$ J K^{-1} and the electronic charge is $q = 1.60 \times 10^{-19}$ C. The standard temperature is denoted by T_0 and is taken to be $T_0 = 290$ K. For this value, $4kT_0 = 1.60 \times 10^{-20}$ J and the thermal voltage is $V_T = kT_0/q = 0.025$ V.

33.1 Thermal Noise

Thermal noise or *Johnson noise* is generated by the random collision of charge carriers with a lattice under conditions of thermal equilibrium [1–7]. Thermal noise in a resistor can be modeled by a series voltage source or a parallel current source having the mean-square values:

$$\overline{v_t^2} = 4kTR\Delta f \tag{33.1}$$

$$\overline{i_t^2} = \frac{4kT\Delta f}{R} \tag{33.2}$$

where R is the resistance and Δf is the bandwidth in hertz (Hz) over which the noise is measured. The equation for $\overline{v_t^2}$ is commonly referred to as the *Nyquist formula*. Thermal noise in resistors is independent of the resistor composition.

The *crest factor* for thermal noise is the ratio of the peak value to the rms value. A common definition for the peak value is the level that is exceeded 0.01% of the time. The amplitude distribution of thermal noise is modeled by a Gaussian or normal probability density function. The probability that the instantaneous value exceeds 4 times the rms value is approximately 0.01%. Thus, the crest factor is approximately 4.

33.2 Spectral Density

The *spectral density* of a noise signal is defined as the mean-square value per unit bandwidth. For the thermal noise generated by a resistor, the voltage and current spectral densities, respectively, are given by

$$S_v(f) = 4kTR \tag{33.3}$$

$$S_i(f) = \frac{4kT}{R} \tag{33.4}$$

Because these are independent of frequency, thermal noise is said to have a uniform or flat distribution. Such noise is also called *white noise*. It is called this by analogy to white light, which also has a flat spectral density in the optical band.

33.3 Fluctuation–Dissipation Theorem

Consider any system in thermal equilibrium with its surroundings. If there is a mechanism for energy in a particular mode to leak out of that mode to the surroundings in the form of heat, then energy can leak back into that mode from the surrounding heat by the same mechanism. The fluctuation

dissipation theorem of quantum mechanics states that the average energy flow in each direction is the same. Otherwise, the system would not be in equilibrium.

Mathematically, the fluctuation dissipation theorem states, in general, that the generalized mean-square force $\overline{\mathfrak{F}^2}$ acting on a system in the frequency band from f_1 to f_2 is given by

$$\overline{\mathfrak{F}^2} = 4kT \int_{f_1}^{f_2} \mathrm{Re}[Z(f)]\mathrm{d}f \tag{33.5}$$

where

Re $[Z(f)]$ is the real part of the system impedance $Z(f)$
f is the frequency in hertz (Hz)

For a mechanical system, the generalized force is the mechanical force on the system and the impedance is force divided by velocity. For an electric system, the generalized force is the voltage and the impedance is the ratio of voltage to current.

Equation 33.1 is a statement of the fluctuation dissipation theorem for a resistor. The theorem can be used to calculate the mean-square thermal noise voltage generated by any two-terminal network containing resistors, capacitors, and inductors. Let $Z(f)$ be the complex impedance of the network. The mean-square open-circuit thermal noise voltage is given by

$$\overline{v_\mathrm{t}^2} = 4kT \int_{f_1}^{f_2} \mathrm{Re}[Z(f)]\mathrm{d}f \simeq 4kT\ \mathrm{Re}[Z(f)]\Delta f \tag{33.6}$$

where $\Delta f = f_2 - f_1$ and the approximation holds if Re $[Z(f)]$ is approximately constant over the band.

33.4 Equivalent Noise Resistance and Conductance

A mean-square noise voltage can be represented in terms of an *equivalent noise resistance* [8]. Let $\overline{v_\mathrm{n}^2}$ be the mean-square noise voltage in the band Δf. The noise resistance R_n is defined as the value of a resistor at the standard temperature T_0 = 290 K that generates the same noise. It is given by

$$R_\mathrm{n} = \frac{\overline{v_\mathrm{n}^2}}{4kT_0\Delta f} \tag{33.7}$$

A mean-square noise current can be represented in terms of an *equivalent noise conductance*. Let $\overline{i_\mathrm{n}^2}$ be the mean-square noise current in the band Δf. The noise conductance G_n is defined as the value of a conductance at the standard temperature that generates the same noise. It is given by

$$G_\mathrm{n} = \frac{\overline{i_0^2}}{4kT_0\Delta f} \tag{33.8}$$

33.5 Shot Noise

Shot noise is caused by the random emission of electrons and by the random passage of charge carriers across potential barriers [1–7]. The shot noise generated in a device is modeled by a parallel noise current source. The mean-square shot noise current in the frequency band Δf is given by

$$\overline{i_{\text{sh}}^2} = 2qI\Delta f \tag{33.9}$$

where I is the dc current through the device. This equation is commonly referred to as the *Schottky formula*. Like thermal noise, shot noise is white noise and has a crest formula of approximately 4.

33.6 Flicker Noise

The imperfect contact between two conducting materials causes the conductivity to fluctuate in the presence of a dc current [1–7]. This phenomenon generates what is called *flicker noise* or *contact noise*. It is modeled by a noise current source in parallel with the device. The mean-square flicker noise current in the frequency band Δf is given by

$$\overline{i_{\text{f}}^2} = \frac{K_{\text{f}} I^m \Delta f}{f^n} \tag{33.10}$$

where

K_{f} is the flicker noise coefficient
I is the dc current
m is the flicker noise exponent, and $n \approx 1$

Other names for flicker noise are *1/f noise* (read "one-over-f-noise"), *low-frequency noise*, and *pink noise*. The latter comes from the optical analog of pink light, which has a spectral density that increases at lower frequencies.

33.7 Excess Noise

In resistors, flicker noise is caused by the variable contact between particles of the resistive material and is called *excess noise*. Metal film resistors generate the least excess noise; carbon composition resistors generate the most, with carbon film resistors lying between the two. In modeling excess noise, the flicker noise exponent has the value $m = 2$. The mean-square excess noise current is often written as a function of the *noise index NI* as follows:

$$\overline{i_{\text{ex}}^2} = \frac{10^{NI/10}}{10^{12}\ln 10} \times \frac{I^2 \Delta f}{f} \tag{33.11}$$

where I is the dc current through the resistor. The noise index is defined as the number of μA of excess noise current in each decade of frequency per A of dc current through the resistor. An equivalent definition is the number of μV of excess noise voltage in each decade of frequency per volt of dc drop across the resistor. In this case, the mean-square excess noise voltage generated by the resistor is given by

$$\overline{v_{\text{ex}}^2} = \frac{10^{NI/10}}{10^{12}\ln 10} \times \frac{V^2 \Delta f}{f} \tag{33.12}$$

where $V = IR$ is the dc voltage across the resistor.

33.8 Burst Noise

Burst noise or *popcorn noise* is caused by a metallic impurity in a *pn* junction [4]. When amplified and reproduced by a loudspeaker, it sounds like corn popping. When viewed on an oscilloscope, it appears as fixed amplitude pulses of randomly varying width and repetition rate. The rate can vary from less than one pulse per minute to several hundred pulses per second. Typically, the amplitude of burst noise is 2–100 times that of the background thermal noise.

33.9 Partition Noise

Partition noise occurs when the charge carriers in a current have the possibility of dividing between two or more paths. The noise is generated in the resulting components of the current by the statistical process of partition [9]. Partition noise occurs in bipolar junction transistors (BJTs) where the current flowing from the emitter into the base can take one of two paths. The recombination of injected carriers in the base region corresponds to the current flow in one path. This current flows in the external base lead. The current carried to the collector corresponds to the current flow in the second path. Because the emitter current exhibits full shot noise, the base and collector currents also exhibit full shot noise. However, the base and collector noise currents are correlated because they have equal and opposite partition components. Partition noise in the BJT can be accounted for if all shot noise is referred to two uncorrelated shot noise current sources, one from base to emitter and the other form collector to emitter [10]. This noise model of the BJT is described here.

33.10 Generation–Recombination Noise

Generation–recombination noise is a semiconductor that is generated by the random fluctuation of free-carrier densities caused by spontaneous fluctuations in the generation, recombination, and trapping rates [7]. In BJTs, it occurs in the base region at low temperatures. The generation–recombination gives rise to fluctuations in the base-spreading resistances that are converted into a noise voltage due to the flow of a base current. In junction FETs, it occurs in the channel at low temperatures. Generation–recombination causes fluctuations of the carrier density in the channel, which gives rise to a noise voltage when a drain current flows. In silicon junction FETs, the effect occurs below 100 K. In germanium junction FETs, it occurs at lower temperatures. The effect does not occur in metal–oxide–semiconductor field-effect transistors (MOSFETs).

33.11 Noise Bandwidth

The *noise bandwidth* of a filter is the bandwidth of an ideal filter having a constant passband gain that passes the same rms noise voltage, where the input signal is white noise [1–7]. The filter and the ideal filter are assumed to have the same gains. Let $A_v(f)$ be the complex voltage gain transfer function of a filter, where f is the frequency in Hz. Its noise bandwidth B_n is given by

$$B_n = \frac{1}{A_{vo}^2} \int_0^\infty |A_v(f)|^2 \, df \tag{33.13}$$

where A_{vo} is the maximum value of $|A_v(f)|$. For a white noise input voltage with the spectral density $S_v(f)$, the mean-square noise voltage at the filter output is

$$\overline{v_{no}^2} = A_{vo}^2 S_v(f) B_n$$

TABLE 33.1 Noise Bandwidth B_n of Low-Pass Filters

Number of Poles	Slope (dB/dec)	Real Pole (B_n)		Butterworth (B_n)
1	20	$1.571\,f_0$	$1.571\,f_3$	$1.571\,f_3$
2	40	$0.785\,f_0$	$1.220\,f_3$	$1.111\,f_3$
3	60	$0.589\,f_0$	$1.155\,f_3$	$1.042\,f_3$
4	80	$0.491\,f_0$	$1.129\,f_3$	$1.026\,f_3$
5	100	$0.420\,f_0$	$1.114\,f_3$	$1.017\,f_3$

Two classes of low-pass filters are commonly used in making noise measurements. The first has n real poles, all with the same frequency, having the magnitude-squared transfer function given by

$$|A_v(f)|^2 = \frac{A_{vo}^2}{\left[1+(f/f_0)^2\right]^n} \tag{33.14}$$

where f_0 is the pole frequency. The second is an n-pole Butterworth filter having the magnitude-squared transfer function given by

$$|A_v(f)|^2 = \frac{A_{vo}^2}{1+(f/f_3)^{2n}} \tag{33.15}$$

where f_3 is the –3 dB frequency. Table 33.1 gives the noise bandwidths as a function of the number of poles n for $1 \le n \le 5$. For the real-pole filter, B_n is given as a function of both f_0 and f_3. For the Butterworth filter, B_n is given as a function of f_3.

33.12 Noise Bandwidth Measurement

The noise bandwidth of a filter can be measured with a white noise source with a known voltage spectral density $S_v(f)$. Let $\overline{v_n^2}$ be the mean-square noise output voltage from the filter when it is driven by the noise source. The noise bandwidth is given by

$$B_n = \frac{\overline{v_o^2}}{A_{vo}^2 S_v(f)} \tag{33.16}$$

If the spectral density of the source is not known, the noise bandwidth can be determined if another filter with a known noise bandwidth is available. With both filters driven simultaneously, let $\overline{v_{o1}^2}$ and $\overline{v_{o2}^2}$ be the two mean-square noise output voltages, B_{n1} and B_{n2} the two noise bandwidths, and A_{vo1} and A_{vo2} the two maximum gain magnitudes. The noise bandwidth B_{n2} is given by

$$B_{n2} = B_{n1}\left(\frac{\overline{v_{o2}^2}}{\overline{v_{o1}^2}}\right)\left(\frac{A_{vo1}}{A_{vo2}}\right)^2 \tag{33.17}$$

The white noise source should have output impedance that is low enough so that the loading effect of the filters does not change the spectral density of the source.

33.13 Spot Noise

Spot noise is the rms noise in a band divided by the square root of the noise bandwidth. For a noise voltage, it has the units $V/\sqrt{Hz}$, which is read "volts per root Hz." For a noise current, the units are $A/\sqrt{Hz}$. For white noise, the spot noise in any band is equal to the square root of the spectral density. Spot noise measurements are usually made with a band-pass filter having a bandwidth that is small enough so that the input spectral density is approximately constant over the filter bandwidth. The spot-noise voltage at a filter output is given by $\sqrt{(\overline{v_{no}^2}/B_n)}$, where $\overline{v_{no}^2}$ is the mean-square noise output voltage and B_n is the filter noise bandwidth. The spot-noise voltage at the filter input is obtained by dividing the output voltage by A_{vo}, where A_{vo} is the maximum value of $|A_v(f)|$.

A filter that is often used for spot-noise measurements is a second-order band-pass filter. The noise bandwidth is given by $B_n = \pi B_3/2$, where B_3 is the −3 dB bandwidth. A single-pole high-pass filter having a pole frequency f_1 cascaded with a single-pole low-pass filter having a pole frequency f_2 is a special case of band-pass filter having two real poles. Its noise bandwidth is given by $B_n = \pi(f_1 + f_2)/2$. The −3 dB bandwidth in this case is $f_1 + f_2$, not $f_2 - f_1$.

33.14 Addition of Noise Voltages

Consider the instantaneous voltage $v = v_n + i_n R$, where v_n is a noise voltage and i_n is a noise current. The mean-square voltage is calculated as follows:

$$\overline{v^2} = \overline{(v_n + i_n R)^2} = \overline{v_n^2} + 2\rho\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}R + \overline{i_n^2}R^2 \tag{33.18}$$

where ρ is the *correlation coefficient* defined by

$$\rho = \frac{\overline{v_n i_n}}{\sqrt{\overline{v_n^2}}\sqrt{\overline{v_n^2}}} \tag{33.19}$$

For the case $\rho = 0$, the sources are said to be uncorrelated or independent. It can be shown that $-1 \le \rho \le 1$.

In ac analysis, noise signals are often represented by phasors. The square magnitude of the phasor represents the mean-square value at the frequency of analysis. Consider the phasor voltage $V = V_n + I_n Z$, where V_n is a noise phasor voltage, I_n is a noise phasor current, and $Z = R + jX$ is a complex impedance. The mean-square voltage is given by

$$\begin{aligned}\overline{v^2} &= \overline{(V_n + I_n R)(V_n^* + I_n^* Z_n^*)} \\ &= \overline{v_n^2} + 2\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}\,\mathrm{Re}(\gamma Z^*) + \overline{i_n^2}\,|Z|^2\end{aligned} \tag{33.20}$$

where

* denotes the complex conjugate

γ is the complex *correlation coefficient* defined by

$$\gamma = \gamma_r + j\gamma_i = \frac{\overline{V_n I_n^*}}{\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}} \tag{33.21}$$

It can be shown that $|\gamma| \le 1$.

Noise equations derived by phasor analysis can be converted easily into equations for real signals. However, the procedure generally cannot be done in reverse. For this reason, noise formulas derived by phasor analysis are the more general form.

33.15 Correlation Impedance and Admittance

The *correlation impedance* Z_γ and *correlation admittance* Y_γ between a noise phasor voltage V_n and a noise phasor current I_n are defined by [8]

$$Z_\gamma = R_\gamma + jX_\gamma = \gamma\sqrt{\frac{\overline{v_n^2}}{\overline{i_n^2}}} \tag{33.22}$$

$$Y_\gamma = G_\gamma + jB_\gamma = \gamma^*\sqrt{\frac{\overline{i_n^2}}{\overline{v_n^2}}} \tag{33.23}$$

where

$\overline{v_n^2}$ is the mean-square value of V_n
$\overline{i_n^2}$ is the mean-square value of I_n
γ is the complex correlation coefficient between V_n and I_n

With these definitions, it follows that $\overline{V_n I_n^*} = \overline{i_n^2} Z_\gamma = \overline{v_n^2}\ Y_\gamma^*$.

33.16 $v_n - i_n$ Amplifier Noise Model

The noise generated by an amplifier can be modeled by referring all internal noise sources to the input [1–7,11]. In order for the noise sources to be independent of the source impedance, two sources are required—a series voltage source and a shunt current source. In general, the sources are correlated.

Figure 33.1 shows the amplifier noise model, where V_s is the source voltage, $Z_s = R_s + jX_s$ is the source impedance, V_{ts} is the thermal noise voltage generated by R_s, $A_v = V_o/V_i$ is the complex voltage gain, and Z_i is the input impedance. The output voltage is given by

$$V_o = \frac{A_v Z_i}{Z_s + Z_i}(V_s + V_{ts} + V_n + I_n Z_s) \tag{33.24}$$

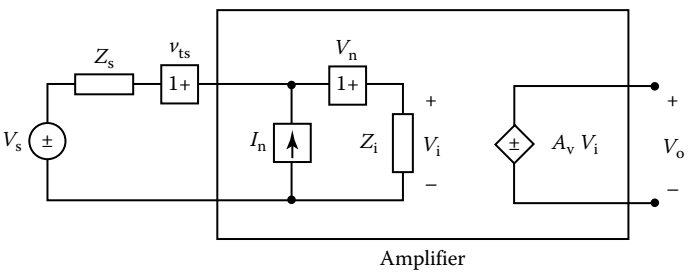


FIGURE 33.1 Amplifier $v_n - i_n$ noise model.

The *equivalent noise input voltage* is the voltage in series with V_s that generates the same noise voltage at the output as all noise sources in the circuit. It is given by $V_{ni} = V_{ts} + V_n + I_n Z_S$. The mean-square value is

$$\overline{v_{ni}^2} = 4kTR_sB_n + \overline{v_n^2} + 2\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}\,\mathrm{Re}(\gamma Z_s^*) + \overline{i_n^2}\,|Z_s|^2 \tag{33.25}$$

where

B_n is the amplifier noise bandwidth

γ is the complex correlation between V_n and I_n

For $|Z_s|$ very small, $\overline{v_{ni}^2} \cong \overline{v_n^2}$ and γ is not important. Similarly, for $|Z_s|$ very large, $\overline{v_{ni}^2} \cong \overline{i_n^2}|Z_s|^2$ and γ is again not important.

When the source is represented by a Norton equivalent consisting of a source current i_s in parallel with source admittance $Y_s = G_s + jB_s$, the noise referred to the input must be represented by an *equivalent noise input current*. The mean-square value is given by

$$\overline{i_{ni}^2} = 4kTG_sB_n + \overline{v_n^2}\,|Y_s|^2 + 2\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}\,\mathrm{Re}(\gamma Y_s) + \overline{i_n^2} \tag{33.26}$$

33.17 Measuring $\overline{V_{ni}^2}$, $\overline{V_n^2}$, and $\overline{i_n^2}$

For a given Z_s, the mean-square equivalent noise input voltage can be measured by setting $V_s = 0$ and measuring the mean-square noise output voltage $\overline{v_{no}^2}$. It follows that $\overline{v_{ni}^2}$ is given by

$$\overline{v_{ni}^2} = \frac{\overline{v_{no}^2}}{|A_v|^2} \times \left|1 + \frac{Z_s}{Z_i}\right|^2 \tag{33.27}$$

To measure $\overline{v_n^2}$ and $\overline{i_n^2}$, $\overline{v_{no}^2}$ is measured with $Z_s = 0$ and with Z_s replaced by a large-value resistor. It follows that $\overline{v_n^2}$ and $\overline{i_n^2}$ are then given by

$$\overline{v_n^2} = \frac{\overline{v_{no}^2}}{|A_v|^2} \quad \text{for } Z_s = 0 \tag{33.28}$$

$$\overline{i_n^2} = \left|\frac{1}{R_s} + \frac{1}{Z_i}\right|^2 \frac{\overline{v_{no}^2}}{|A_v|^2} \quad \text{for } Z_s = R_s \text{ and } R_s \text{ large} \tag{33.29}$$

33.18 Noise Temperature

The internal noise generated by an amplifier can be expressed as an equivalent *input-termination noise temperature* [12]. This is the temperature of the source resistance that generates a thermal noise voltage equal to the internal noise generated in the amplifier when referred to the input. The noise temperature T_n is given by

$$T_n = \frac{\overline{v_{ni}^2}}{4kR_sB_n} - T \tag{33.30}$$

where

$\overline{v_{ni}^2}$ is the mean-square equivalent input noise voltage in the band B_n

R_s is the real part of the source output impedance

T is the temperature of R_s

33.19 Noise Reduction with a Transformer

Let a transformer be connected between the source and the amplifier in Figure 33.1. Let n be the transformer turns ratio, R_1 the primary resistance, and R_2 the secondary resistance. The equivalent noise input voltage in series with the source voltage V_s has the mean-square value:

$$\overline{v_{ni}^2} = 4kT\left(R_s + R_1 + \frac{R_2}{n^2}\right)\Delta f + \frac{\overline{v_n^2}}{n^2}$$
$$+2\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}\,\mathrm{Re}\left[\gamma\left(Z_s^* + R_1 + \frac{R_2}{n^2}\right)\right] + n^2\overline{i_n^2}\left|Z_s + R_1 + \frac{R_2}{n^2}\right|^2 \tag{33.31}$$

In general, $R_2/R_1 \propto n$, which makes it difficult to specify the value of n that minimizes $\overline{v_{ni}^2}$. For $R_1 + R_2/n^2 \ll |Z_s|$ it is minimized when

$$n^2 = \frac{1}{|Z_s|}\sqrt{\frac{\overline{v_n^2}}{\overline{i_n^2}}} \tag{33.32}$$

33.20 Signal-to-Noise Ratio

The *signal-to-noise ratio* of an amplifier is defined by

$$\mathrm{SNR} = \frac{\overline{v_{so}^2}}{\overline{v_{no}^2}} \tag{33.33}$$

where

$\overline{v_{so}^2}$ is the mean-square signal output voltage
$\overline{v_{no}^2}$ is the mean-square noise output voltage

The SNR is often expressed in dB with the equation $\mathrm{SNR} = 10\log(\overline{v_s^2}/\overline{v_{no}^2})$. In measuring the SNR, a filter should be used to limit the bandwidth of the output noise to the signal bandwidth of interest. An alternative definition of the SNR that is useful in making calculations is

$$\mathrm{SNR} = \frac{\overline{v_s^2}}{\overline{v_{ni}^2}} \tag{33.34}$$

where

$\overline{v_s^2}$ is the mean-square signal input voltage
$\overline{v_{ni}^2}$ is the mean-square equivalent noise input voltage

33.21 Noise Factor and Noise Figure

The *noise factor F* of an amplifier is defined by [1–8]

$$F = \frac{\overline{v_{no}^2}}{\overline{v_{nos}^2}} \tag{33.35}$$

where

$\overline{v_{no}^2}$ is the mean-square noise output voltage with the source voltage zeroed
$\overline{v_{nos}^2}$ is the mean-square noise output voltage considering the only source of noise to be the thermal noise generated by the source resistance R_s

The *noise figure* is the noise factor expressed in dB and is given by

$$\mathrm{NF} = 10\log F \tag{33.36}$$

For the $v_n - i_n$ amplifier noise model, the noise factor is given by

$$F = \frac{\overline{v_{ni}^2}}{4kTR_sB_n} = 1 + \frac{\overline{v_n^2} + 2\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}\,\mathrm{Re}(\gamma Z_s^*) + \overline{i_n^2}|Z_s|^2}{4kTR_sB_n} \tag{33.37}$$

where B_n is the amplifier noise bandwidth. The value of Z_s that minimizes the noise figure is called the *optimum source impedance* and is given by

$$Z_{so} = R_{so} + jX_{so} = \left[\sqrt{1-\gamma_i^2} - j\gamma_i\right]\sqrt{\frac{\overline{v_n^2}}{\overline{i_n^2}}} \tag{33.38}$$

where $\gamma_i = \mathrm{Im}(\gamma)$. The corresponding value of F is denoted by F_0 and is given by

$$F_0 = 1 + \frac{\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}}{2kTB_n}\left(\gamma_r + j\sqrt{1-\gamma_i^2}\right) \tag{33.39}$$

It follows that F can be expressed in terms of F_0 as follows:

$$F = F_0 + \frac{G_n}{R_{ns}}\left[(R_s - R_{so})^2 + (X_s - X_{so})^2\right] \tag{33.40}$$

where G_n is the noise conductance of I_n and R_{ns} is the noise resistance of the source. These are given by

$$G_n = \frac{\overline{i_n^2}}{4kT_0B_n} \tag{33.41}$$

$$R_{ns} = \frac{\overline{v_{ts}^2}}{4kT_0B_n} = \frac{T\ R_s}{T_0} \tag{33.42}$$

When the source is represented by a Norton equivalent consisting of a source current i_s in parallel with a source admittance $Y_s = G_s + jB_s$, the *optimum source admittance* is given by

$$Y_{so} = G_{so} + jB_{so} = \left[\sqrt{1-\gamma_i^2} + j\gamma_i\right]\sqrt{\frac{\overline{i_n^2}}{\overline{v_n^2}}} \tag{33.43}$$

which is the reciprocal of Z_{so}. The noise factor can be written as

$$F = F_0 + \frac{R_n}{G_{ns}}\left[(G_s - G_{so})^2 + (B_s - B_{so})^2\right] \tag{33.44}$$

where

R_n is the noise resistance of V_n
G_{ns} is the noise conductance of the source

These are given by

$$R_n = \frac{\overline{v_n^2}}{4kT_0B_n} \tag{33.45}$$

$$G_{ns} = \frac{\overline{i_{ts}^2}}{4kT_0B_n} = \frac{TG_s}{T_0} \tag{33.46}$$

33.22 Noise Factor Measurement

The noise factor can be measured with a calibrated white noise source driving the amplifier. The source output impedance must equal to the value of Z_s for which F is to be measured. The source temperature must be the standard temperature T_0. First, measure the amplifier noise output voltage over the band of interest with the source voltage set to zero. For the amplifier model of Figure 33.1, the mean-square value of this voltage is given by

$$\overline{v_{nol}^2} = \left|\frac{A_{vo}Z_i}{Z_s + Z_i}\right|^2 \left[4kT_0R_sB_n + \overline{v_n^2} + 2\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}\,\mathrm{Re}(\gamma Z_s^*) + \overline{i_n^2}\,|\,Z_s\,|^2\right] \tag{33.47}$$

The source noise voltage is then increased until the output voltage increases by a factor r. The new mean-square output voltage can be written as

$$r^2\overline{v_{nol}^2} = \left|\frac{A_{vo}Z_i}{Z_s + Z_i}\right|^2 \left[(S_v(f) + 4kT_0R_s)B_n + \overline{v_n^2} + 2\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}\,\mathrm{Re}(\gamma Z_s^*) + \overline{i_n^2}|Z_s|^2\right] \tag{33.48}$$

where $S_v(f)$ is the open-circuit voltage spectral density of the white noise source.

The aforementioned two equations can be solved for F to obtain

$$F = \frac{S_s(f)}{(r^2 - 1)4kT_0R_s} \tag{33.49}$$

A common value for r is $\sqrt{2}$. The gain and noise bandwidth of the amplifier are not needed for the calculation. If a resistive voltage divider is used between the noise source and the amplifier to attenuate the input signal, the source spectral density $S_s(f)$ is calculated or measured at the attenuator output with it disconnected from the amplifier input.

If the noise bandwidth of the amplifier is known, its noise factor can be determined by measuring the mean-square noise output voltage $\overline{v_{no}^2}$ with the source voltage set to zero. The noise factor is given by

$$F = \left|1 + \frac{Z_s}{Z_i}\right|^2 \frac{\overline{v_{no}^2}}{4kT_0R_sB_nA_{vo}^2} \tag{33.50}$$

This expression is often used with $B_n = \pi B_3/2$, where B_3 is the −3 dB bandwidth. Unless the amplifier has a first-order low-pass or a second-order band-pass frequency response characteristic, this is only an approximation.

33.23 Junction Diode Noise Model

When forward biased, a diode generates both shot noise and flicker noise [1–7]. The noise is modeled by a parallel current source having the mean-square value

$$\overline{i_n^2} = 2qI\Delta f + \frac{K_f I^m \Delta f}{f} \tag{33.51}$$

where I is the dc diode current. A plot of $\overline{i_n^2}$ versus f for a constant Δf exhibits a slope of –10 dB/decade for low frequencies and a slope of zero for higher frequencies.

Diodes are often used as noise sources in circuits. Specially processed zener diodes are marketed as solid-state noise diodes. The noise mechanism in these is called *avalanche noise*, which is associated with the diode reverse breakdown current. For a given breakdown current, avalanche noise is much greater than shot noise in the same current.

33.24 BJT Noise Model

Figure 33.2a shows the BJT noise model [1–7]. The base-spreading resistance r_x is modeled as an external resistor; v_{tx} is the thermal noise generated by r_x; i_{shb} and i_{fb}, respectively, are the shot noise and flicker noise in the base bias current I_B; and i_{shc} is the shot noise in the collector bias current I_C. The sources have the mean-square values of

$$\overline{v_{tx}^2} = 4kTr_x\Delta f \tag{33.52}$$

$$\overline{i_{shb}^2} = 2qI_B\Delta f \tag{33.53}$$

$$\overline{i_{fb}^2} = \frac{K_f I_B^m \Delta f}{f} \tag{33.54}$$

$$\overline{i_{shc}^2} = 2qI_C\Delta f \tag{33.55}$$

Let the resistances to signal ground seen looking out of the base and the emitter, respectively, be denoted by R_1 and R_2. The mean-square equivalent noise input voltages in series with the base or the emitter that generates the same collector noise current is given by

$$\overline{v_{ni}^2} = 4kT(R_1 + r_x + R_2)\Delta f + \left(2qI_B\Delta f + \frac{K_f I_B \Delta f}{f}\right)(R_1 + r_x + R_2)^2$$

$$+2qI_C\Delta f\left(\frac{R_1 + r_x + R_2}{\beta} + \frac{V_T}{I_C}\right)^2 \tag{33.56}$$

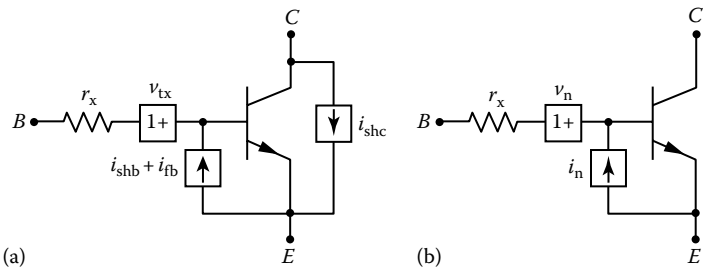


FIGURE 33.2 (a) BJT noise model and (b) BJT $v_n - i_n$ noise model.

At frequencies where flicker noise can be neglected, the value of I_C that minimizes $\overline{v_{ni}^2}$ is called the *optimum collector current*. It is given by

$$I_{C_{opt}} = \frac{V_T}{R_1 + r_x + R_2} \times \frac{\beta}{\sqrt{1+\beta}} \tag{33.57}$$

The corresponding value of $\overline{v_{ni}^2}$ is given by

$$\overline{v_{ni_{min}}^2} = 4kT(R_1 + r_x + R_2)\Delta f \times \frac{\sqrt{1+\beta}}{\sqrt{1+\beta}-1} \tag{33.58}$$

If N identical BJTs that are identically biased are connected in parallel, the equivalent noise input voltage is given by Equation 33.56 with r_x replaced with r_x/N, I_B replaced with NI_B, and I_C replaced with NI_C. In this case, R_1 and R_2, respectively, are the resistances to signal ground seen looking out of the parallel connected bases and the parallel connected emitters. For N fixed, the value of I_C that minimizes $\overline{v_{no}^2}$ is given by Equation 33.57 with R_1 replaced with NR_1 and R_2 replaced with NR_2. The corresponding value of $\overline{v_{ni_{min}}^2}$ is given by Equation 33.58 with r_x replaced with r_x/N. It follows that parallel connection of BJTs can be used to reduce the thermal noise of r_x, provided the devices are optimally biased.

The BJT $v_n - i_n$ noise model is given in Figure 33.2b, where r_x is a noiseless resistor, for its thermal noise is included in v_n. The mean-square values of v_n and i_n and the correlation coefficient are given by

$$\overline{v_n^2} = 4kTr_x\Delta f + 2kT\frac{V_T}{I_C}\Delta f \tag{33.59}$$

$$\overline{i_n^2} = 2qI_B\Delta f + \frac{K_f I_B^m \Delta f}{f} + \frac{2qI_C\Delta f}{\beta^2} \tag{33.60}$$

$$\rho = \frac{2kT\Delta f}{\beta\sqrt{\overline{v_n^2}}\sqrt{\overline{i_n^2}}} \tag{33.61}$$

where $\beta = I_C/I_B$ is the current gain. An alternative model puts r_x inside the BJT. For this model, the expressions for i_n and ρ are more complicated than the ones given here.

The $v_n - i_n$ noise model of Figure 33.1 does not have a noiseless resistor in series with its input. Before formulae that are derived for this model are applied to the BJT model of Figure 33.2b, the formulae must be modified to account for the noiseless r_x. For the common-emitter amplifier, for example, the source resistance R_s would be replaced in the expression for $\overline{v_{ni}^2}$ by $R_s + r_x$ in all occurrences except in terms that represent the thermal noise of R_s.

The value of the base-spreading resistance r_x depends on the method used to measure it. For noise calculations, r_x should be measured with a noise technique. A test circuit for measuring r_x is shown in Figure 33.3. The emitter bias current I_E and the collector bias voltage V_C are given by

$$I_E = \frac{-V_{BE} - V_{EE}}{R_E} \tag{33.62}$$

$$V_C = V_{CC} - \alpha I_E R_C \tag{33.63}$$

where $\alpha = \beta/(1 + \beta)$. Capacitors C_1 and C_2 should satisfy $C_1 >> I_E/(2\pi f V_T)$ and $C_2 >> 1/(2\pi f R_C)$, where f is the lowest frequency of interest. To minimize the noise contributed by R_C, R_F, and the op-amp, a

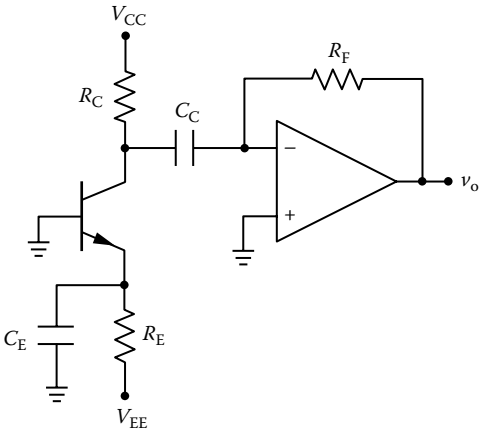


FIGURE 33.3 Test circuit for measuring r_x of a BJT.

low-noise op-amp should be used and R_F should be much larger than R_C. The power supply rails must be properly decoupled to minimize power supply noise.

To prevent flicker noise from affecting the data, the op-amp output voltage must be measured over a noise bandwidth where the spectral density is white. Denote the mean-square op-amp output voltage over the band B_n with the BJT in the circuit by $\overline{v_{no1}^2}$. Denote the mean-square voltage over the band B_n with the BJT removed by $\overline{v_{no2}^2}$. The base-spreading resistance r_x can be obtained by solving

$$r_x^2\left[\frac{A}{\beta^2}-2qI_BB_n\right]+r_x\left[\frac{2AV_T}{\beta I_C}-4kTB_n\right]+\frac{AV_T^2}{I_C^2}=0 \tag{33.64}$$

where

$$A=\frac{\overline{v_{no1}^2}-\overline{v_{no2}^2}}{R_F^2}-2qI_CB_n \tag{33.65}$$

The test circuit of Figure 33.3 can be used to measure the flicker noise coefficient K_f and the flicker noise exponent m. The plot of $(\overline{v_{no1}^2}-\overline{v_{no2}^2})$ versus frequency for a constant noise bandwidth Δf must be obtained, for example, with a signal analyzer. In the white noise range, the slope of the plot is zero. In the flicker noise range, the slope is −10 dB per decade. The lower frequency at which $(\overline{v_{no1}^2}-\overline{v_{no2}^2})$ is 3 dB greater than its value in the white noise range is the flicker noise corner frequency f_f. It can be shown that

$$K_fI_B^m=\frac{\left(\overline{v_{no1}^2}-\overline{v_{no2}^2}\right)f_f}{2R_F^2\Delta f}\times\left(\frac{r_x}{\beta}+\frac{V_T}{I_C}\right)^2 \tag{33.66}$$

By repeating the measurements for at least two values of I_C, this equation can be used to solve for both K_f and m. Unless I_C is large, the r_x/β term can usually be neglected compared to the V_T/I_C term. The value of the flicker noise exponent is usually in the range $1 < m < 3$ but is often taken as unity. If it is assumed that $m = 1$, the value of K_f can be calculated by making the measurements with only one value of I_C.

33.25 FET Noise Model

Figure 33.4a shows the MOSFET noise equivalent circuit, where i_{td} is the channel thermal noise current and i_{fd} is the channel flicker noise current [1–7]. The mean-square values of these currents are given by

$$\overline{i_{td}^2} = \frac{8kT\Delta f}{3g_m} \tag{33.67}$$

$$\overline{i_{fd}^2} = \frac{K_f \Delta f}{4KfL^2C_{ox}} \tag{33.68}$$

where

K is the transconductance parameter
$g_m = 2\sqrt{KI_D}$ is the transconductance
L is the effective length of the channel
C_{ox} is the gate oxide capacitance per unit area

Let the resistances to signal ground seen looking out of the gate and the source, respectively, be denoted by R_1 and R_2. The mean-square equivalent noise input voltage in series with either the gate or the source that generates the same drain noise current is given by

$$\overline{v_{ni}^2} = 4kT(R_1 + R_2)\Delta f + \frac{4kT\Delta f}{3\sqrt{KI_D}} + \frac{K_f \Delta f}{4KL^2C_{ox}f} \tag{33.69}$$

where it is assumed that the MOSFET bulk is connected to its source in the ac. If N identical MOSFETs that are identically biased are connected in parallel, the equivalent noise input voltage is given by Equation 33.69 with the exception that the second and third terms are divided by N.

The noise sources in Figure 33.4a can be reflected into a single source in series with the gate. The circuit is shown in Figure 33.4b. The mean-square value of v_n is given by

$$\overline{v_n^2} = \frac{8kT\Delta f}{3g_m} + \frac{K_f \Delta f}{4KfL^2C_{ox}} \tag{33.70}$$

The FET flicker noise coefficient K_f can be measured by replacing the BJT in Figure 33.3 with the FET. On a plot of $(\overline{v_{no1}^2} - \overline{v_{no2}^2})$ as a function of frequency for a constant noise bandwidth, the flicker noise corner frequency f_f is the lower frequency at which $(\overline{v_{no1}^2} - \overline{v_{no2}^2})$ is up 3 dB above the white noise level. A signal analyzer can be used to display this plot. The flicker noise coefficient is given by

$$K_f = \frac{32kTKf_fL^2C_{ox}}{3g_m} \tag{33.71}$$

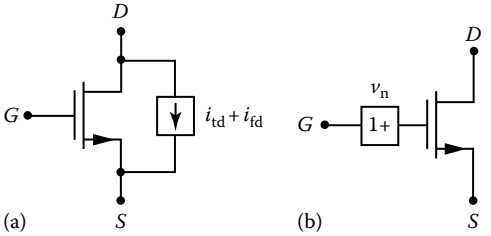


FIGURE 33.4 FET noise model: (a) current noise model and (b) voltage noise model.

The MOSFET circuits and equations also apply to the junction FET with the exception that the L^2 and C_{ox} terms are omitted from the formulae. This assumes that the junction FET gate-to-channel junction is reverse biased, which is the usual case. Otherwise, shot noise in the gate current must be modeled.

33.26 Operational Amplifier Noise Models

Variations of the $v_n - i_n$ amplifier noise model are used in specifying op-amp noise performance. The three most common models are given in Figure 33.5. In Figure 33.5b and c, v_n can be placed in series with either input [4,6]. In general, the sources in each model are correlated. In making calculations that use specified op-amp noise data, it is important to use the noise model for which the data apply.

33.27 Photodiode Detector Noise Model

Figure 33.6a shows the circuit symbol of a photodiode detector [4]. When reverse biased by a dc source, an incident light signal causes a signal current to flow in the diode. The diode small-signal noise model is shown in Figure 33.6b, where i_s is the signal current (proportional to the incident

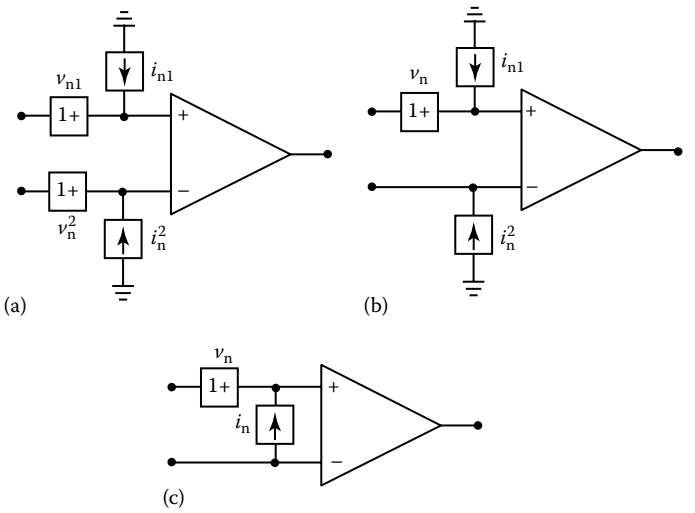


FIGURE 33.5 Op-amp noise models: (a) multiple noise currents and multiple noise voltage representation, (b) multiple noise currents and single voltage representation, and (c) single noise current and single noise voltage representation.

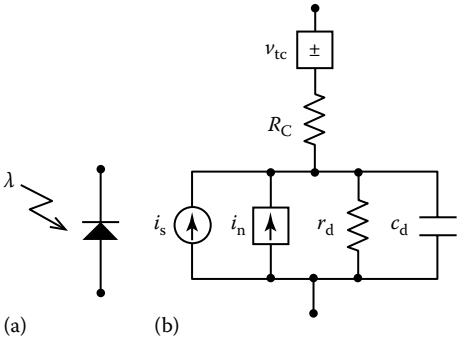


FIGURE 33.6 Photodiode noise model (a) photodiode symbol and (b) small-signal noise model of photodiode.

light intensity), i_n is the diode noise current, r_d is the small-signal resistance of the reverse-biased junction, c_d is the small-signal junction capacitance, r_c is the cell resistance (typically <50 Ω), and v_{tc} is the thermal noise voltage generated by r_c. The noise current i_n consists of three components: shot noise i_{sh}, flicker noise i_f, and carrier generation–recombination noise i_{gr}. The first three have the mean-square values

$$\overline{v_{tc}^2} = 4kTr_c\Delta f \tag{33.72}$$

$$\overline{i_{sh}^2} = 2qI_D\Delta f \tag{33.73}$$

$$\overline{i_f^2} = \frac{K_f I_D^m \Delta f}{f} \tag{33.74}$$

where I_D is the reverse-biased diode current. The carrier generation–recombination noise has a white spectral density up to a frequency determined by the carrier lifetime. Because the detector has a large output resistance, it should be used with amplifiers that exhibit a low input current noise.

33.28 Piezoelectric Transducer Noise Model

Figure 33.7a shows the circuit symbol of a piezoelectric transducer [4]. This transducer generates an electric voltage when a mechanical force is applied between two of its surfaces. An approximate equivalent circuit that is valid for frequencies near the transducer mechanical resonance is shown in Figure 33.7b. In this circuit, C_e represents the transducer electric capacitance, while C_s, L_s, and R_s are chosen to have a resonant frequency and quality factor numerically equal to those of the transducer mechanical resonance. The source v_s represents the signal voltage, which is proportional to the applied force. The source v_{ts} represents the thermal noise generated by R_s. It has a mean-square value of

$$\overline{v_{ts}^2} = 4kTR_s\Delta f \tag{33.75}$$

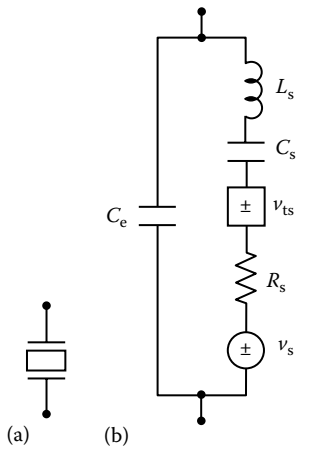


FIGURE 33.7 (a) Piezoelectric transducer symbol and (b) noise model of piezoelectric transducer.

This noise component is negligible in most applications.

The piezoelectric transducer has two resonant frequencies: short-circuit resonant frequency f_{sc} and an open-circuit resonant frequency f_{oc} given by

$$f_{sc} = \frac{1}{2\pi\sqrt{L_s C_s}} \tag{33.76}$$

$$f_{oc} = \frac{1}{2\pi\sqrt{L_1 C_1}} \tag{33.77}$$

where $C_1 = C_s C_e/(C_s + C_e)$. It is normally operated at the open-circuit resonant frequency where the transducer output impedance is very high. For this reason, it should be used with amplifiers that exhibit a low input current noise.

33.29 Parametric Amplifiers

A *parametric amplifier* is an amplifier that uses a time-varying reactance to produce amplification [13]. In low-noise microwave parametric amplifiers, a reverse-biased *pn* junction diode is used to realize a variable capacitance. Such diodes are called *varactors*, for variable reactance. The depletion capacitance of the reverse-biased junction is varied by simultaneously applying a signal current and a pump current at different frequencies. The nonlinear capacitance causes frequency mixing to occur between the signal frequency and the pump frequency. When the power generated by the frequency mixing exceeds the signal input power, the diode appears to have a negative resistance and signal amplification occurs. The only noise that is generated is the thermal noise of the effective series resistance of the diode, which is very small.

Figure 33.8 shows block diagram of a typical parametric amplifier. The varactor diode is placed in a resonant cavity. A circulator is used to isolate the diode from the input and output circuits. A pump signal is applied to the diode to cause its capacitance to vary at the pump frequency. The filter isolates the pump signal from the output circuit. The idler circuit is a resonant cavity that is coupled to the diode cavity to reduce the phase sensitivity. Let the signal frequency be f_s, the pump frequency be f_p, and the resonant frequency of the idler cavity be f_i. In cases where $f_p = f_s + f_i$, the varying capacitance of the diode looks like a negative resistance and the signal is amplified. If $f_i = f_s$, the amplifier is called a *degenerate* amplifier. This is the simplest form of the parametric amplifier and it requires the lowest

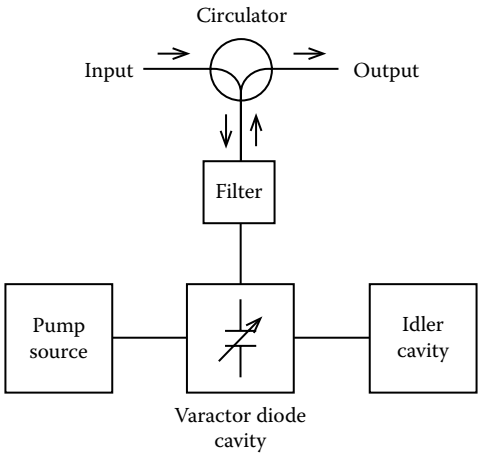


FIGURE 33.8 Block diagram of a typical parametric amplifier.

pump frequency and power to operate. For the *nondegenerate* amplifier, $f_p > 2f_s$. In both cases, the input and the output are at the same frequency. In the *upconverter* amplifier, $f_p = f_i - f_s$ and $f_p > 2f_s$. In this case, the varying capacitance of the diode looks like a positive resistance and the signal frequency output is not amplified. However, there is an output at the idler frequency that is amplified. Thus, the output frequency is higher than the input frequency. The conversion gain can be as high as the ratio of the output frequency to the input frequency.

33.30 Measuring Noise

A typical setup for measuring noise is shown in Figure 33.9. To prevent measurement errors caused by signals coupling in through the ground and power supply leads, the circuit under test and the test set must be properly grounded and good power supply decoupling must be used [6]. For measurement schemes requiring a sine-wave source, an internally shielded oscillator is preferred over a function generator. This is because function generators can introduce errors caused by radiated signals and signals coupled through the ground system.

When making measurements on a high-gain circuit, the input signal must often be attenuated. Attenuators that are built into sources might not be adequately shielded, so that errors can be introduced by radiated signals. These problems can be minimized if a shielded external attenuator is used between the source and the circuit under test. Such an attenuator is illustrated in Figure 33.9. When a high attenuation is required, a multistage attenuator is preferred. For proper frequency response, the attenuator might require frequency compensation [4]. Unless the load impedance on the attenuator is large compared to its output impedance, both the attenuation and the frequency compensation can be a function of the load impedance.

Figure 33.9 shows source impedance Z_s in series with the input to the circuit under test. This impedance is in series with the output impedance of the attenuator. It must be chosen so that the circuit under test has the desired source impedance termination for the noise measurements.

Because noise signals are small, a low-noise amplifier is often required to boost the noise level sufficiently so that it can be measured. Such an amplifier is shown in Figure 33.9. The noise generated by the amplifier will add to the measured noise. To correct for this, first measure the mean-square noise voltage with the amplifier input terminated in the output impedance of the circuit under test. Then subtract this from the

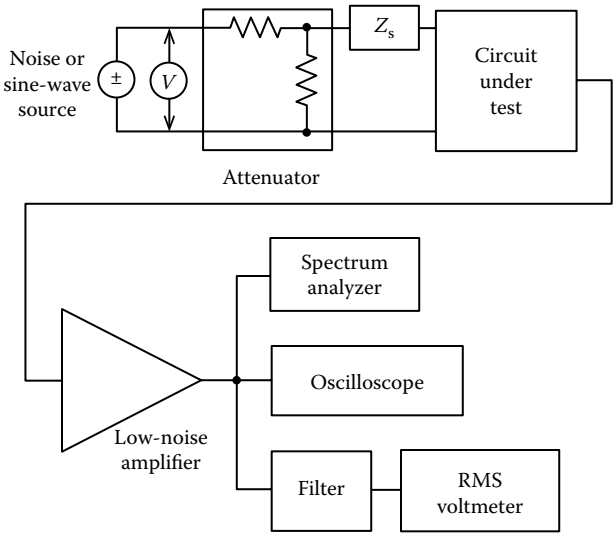


FIGURE 33.9 Noise measuring setup.

measured mean-square noise voltage with the circuit under test driving the amplifier. The difference is the mean-square noise due to the circuit. Ideally, the amplifier should have no effect on the measured noise.

The noise voltage over a band can be measured with either a spectrum analyzer or with a filter having a known noise bandwidth and a voltmeter. The noise can be referred to the input of the circuit under test by dividing by the total gain between its input and the measuring device. The measuring voltmeter should have a bandwidth that is at least 10 times the noise bandwidth of the filter. The *voltmeter crest factor* is the ratio of the peak input voltage to the full-scale rms meter reading at which the internal meter circuits overload. For a sine-wave signal, the minimum voltmeter crest factor is $\sqrt{2}$. For noise measurements, a higher crest factor is required. For Gaussian noise, a crest factor of 3 gives an error less than 1.5%. A crest factor of 4 gives an error less than 0.5%. To avoid overload on noise peaks caused by an inadequate crest factor, measurements should be made on the lower one-third to one-half of the voltmeter scale.

A true rms voltmeter is preferred over one that responds to the average rectified value of the input voltage but is calibrated to read rms. When the latter type of voltmeter is used to measure noise, the reading will be low. For Gaussian noise, the reading can be corrected by multiplying the measured voltage by 1.13. Noise voltages measured with a spectrum analyzer must also be corrected by the same factor if the spectrum analyzer responds to the average rectified value of the input voltage but is calibrated to read rms.

Noise measurements with a spectrum analyzer require knowledge of the noise bandwidth of the instrument. For a conventional analyzer, the bandwidth is proportional to frequency. When white noise is analyzed, the display exhibits a slope of +10 dB per decade. However, the measured voltage level at any frequency divided by the square root of the noise bandwidth of the analyzer is a constant equal to the spot-noise value of the input voltage at that frequency. Band-pass filters that have a bandwidth proportional to the center frequency are called *constant-Q filters*. For a second-order constant-*Q* filter, the noise bandwidth is given by $B_n = \pi f_0/2Q$, where f_0 is the center frequency and Q is the quality factor. The latter is given by $Q = f_0/B_3$, where B_3 is the –3 dB bandwidth. These equations are often used to estimate the noise bandwidth of band-pass filters that are not second order.

A second type of spectrum analyzer is called a *signal analyzer*. Such an instrument uses digital signal processing techniques to calculate the spectrum of the input signal as a discrete Fourier transform. The noise bandwidth of these instruments is a constant so that the display exhibits a slope of zero when white noise is the input signal.

Fairly accurate rms noise measurements can be made with an oscilloscope. A filter should be used to limit the noise bandwidth at its input. Although the procedure is subjective, the rms voltage can be estimated by dividing the observed peak-to-peak voltage by 6 [4]. One of the advantages of using the oscilloscope is that nonrandom noise that can affect the measurements can be identified, for example, a 60 Hz hum signal.

Another oscilloscope method is to display the noise simultaneously on both inputs of a dual-channel oscilloscope that is set in the dual-sweep mode. The two channels must be identically calibrated, and the sweep rate must be set low enough so that the displayed traces appear as bands. The vertical offset between the two bands is adjusted until the dark area between them just disappears. The rms noise voltage is then measured by grounding the two inputs and reading the vertical offset between the traces.

Defining Terms

Burst noise: Noise caused by a metallic impurity in a *pn* junction that sounds like corn popping when amplified and reproduced by a loudspeaker, also called *popcorn noise*.

Crest factor: The ratio of the peak value to the rms value.

Equivalent noise input current: The noise current in parallel with an amplifier input that generates the same noise voltage at its output as all noise sources in the amplifier.

Equivalent noise input voltage: The noise voltage in series with an amplifier input that generates the same noise voltage at its output as all noise sources in the amplifier.

Equivalent noise resistance (conductance): The value of a resistor (conductance) at the standard temperature T_0 = 290 K that generates the same mean-square noise voltage (current) as a source.

Excess noise: Flicker noise in resistors.

Flicker noise: Noise generated by the imperfect contact between two conducting materials causing the conductivity to fluctuate in the presence of a dc current, also called *contact noise*, *1/f noise*, and *pink noise*.

Generation–recombination noise: Noise generated in a semiconductor by the random fluctuation of free-carrier densities caused by spontaneous fluctuations in the generation, recombination, and trapping rates.

Noise bandwidth: The bandwidth of an ideal filter having a constant passband gain that passes the same rms noise voltage as a filter, where the input signal is white noise.

Noise factor: The ratio of the mean-square noise voltage at an amplifier output to the mean-square noise voltage at the amplifier output considering the thermal noise of the input termination to be the only source of noise.

Noise figure: The noise factor expressed in dB.

Noise index: The number of μA of excess noise current in each decade of frequency per A of dc current through a resistor. Also, the amount of μV of excess noise voltage in each decade of frequency per V of dc voltage across a resistor.

Noise temperature: The internal noise generated by an amplifier expressed as an equivalent input-termination noise temperature.

Nyquist formula: Expression for the mean-square thermal noise voltage generated by a resistor.

Optimum source impedance (admittance): The complex source impedance (admittance) that minimizes the noise factor.

Parametric amplifier: An amplifier that uses a time-varying reactance to produce amplification.

Partition noise: Noise generated by the statistical process of partition when the charge carriers in a current have the possibility of dividing between two or more paths.

Shot noise: Noise caused by the random emission of electrons and by the random passage of charge carriers across potential barriers.

Schottky formula: Expression for the mean-square shot noise current.

Signal-to-noise ratio: The ratio of the mean-square signal voltage to the mean-square noise voltage at an amplifier output.

Spectral density: The mean-square value per unit bandwidth of a noise signal.

Spot noise: The rms noise in a band, divided by the square root of the noise bandwidth.

Thermal noise: Noise generated by the random collision of charge carriers with a lattice under conditions of thermal equilibrium, also called *Johnson noise*.

Varactor diode: A diode used as a variable capacitance.

Voltmeter crest factor: The ratio of the peak input voltage to the full-scale rms meter reading at which the internal meter circuits overload.

White noise: Noise that has a spectral density that is flat, that is, not a function of frequency.

References

1. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, New York: John Wiley & Sons, 1993.
2. P. Horowitz and W. Hill, *The Art of Electronics*, 2nd edn., New York: Cambridge Press, 1983.
3. W. M. Leach Jr., Fundamentals of low-noise analog circuit design, *Proc. IEEE*, 82(10), 1515–1538, 1994.

4. C. D. Motchenbacher and J. A. Connelly, *Low Noise Electronic System Design*, New York: Wiley, 1993.
5. F. Bruccoleri, E. A. M. Klumperink, and B. Nauta, *Wideband Low Noise Amplifiers Exploiting Thermal Noise Cancellation*, Springer, Dordrecht, Netherland, 2005.
6. H. W. Ott, *Noise Reduction Techniques in Electronic Systems*, 2nd edn., New York: Wiley, 1988.
7. A. Van der Ziel, Noise in solid-state devices and lasers, *Proc. IEEE*, 58, 1178–1206, 1970.
8. H. A. Haus et al., Representation of noise in linear two-ports, *Proc. IRE*, 48, 69–74, 1960.
9. R. E. Burgess, Electronic fluctuations in semiconductors, *Br. J. Appl. Phys*., 6, 185–190, 1955.
10. H. Fukui, The noise performance of microwave transistors, *IEEE Trans. Electron Devices*, ED-13, 329–341, 1966.
11. H. Rothe and W. Dahlke, Theory of noisy four-poles, *Proc. IRE*, 44, 811–818, 1956.
12. H. A. Haus et al., IRE standards on methods of measuring noise in linear two-ports, 1959, *Proc. IRE*, 48, 60–68, 1960.
13. A. L. Lance, *Introduction to Microwave Theory and Measurements*, New York: McGraw-Hill, 1964.

Further Information

Aldert Van der Ziel, *Noise: Sources, Characterization, Measurements*, Englewood Cliffs, NJ: Prentice Hall, 1970.

Bertails J. C., Low frequency noise considerations for MOS amplifier design, *IEEE J. Solid-State Circuits*, SC-14, 773–776, 1979.

Fukui H., *Low-Noise Microwave Transistors & Amplifiers*, New York: IEEE Press, 1981.

Gupta M. S., ed., *Electrical Noise: Fundamentals & Sources*, New York: IEEE Press, 1977.

Liechti C. A., Microwave field-effect transistors—1976, *IEEE Trans. Microwave Theory Tech*., MTT-24, 279–300, 1976.

Steyaert M., Z. Y. Chang, and W. Sansen, Low-noise monolithic amplifier design: Bipolar vs. CMOS, *Analog Integr. Circuits Signal Process.*, 1(1), 9–19, 1991.

III

Electromagnetic Variables

34

Microwave Measurement

Alfons Dehé
Technische Universität Darmstadt

K. Beilenhoff
Technische Universität Darmstadt

K. Fricke
Technische Universität Darmstadt

H. Klingbeil
Technische Universität Darmstadt

V. Krozer
Technische Universität Darmstadt

H.L. Hartnagel
Technische Universität Darmstadt

Microwave measurements cover the frequency range from 0.5 to about 20 GHz. Frequencies from 30 to 300 GHz are often referred to as mm waves. In the following, the most important measurements are described.

34.1 Power Measurement

The exact microwave power measurement is in demand for development, fabrication, and installation of modern telecommunication networks. It is essential for attenuation, insertion, and return loss measurements, as well as for noise measurement and six-port network analysis. This chapter gives a brief overview about the basics of microwave power measurement. Detailed information about this subject can be found in [1–4]. Power detectors usually consist of a sensor that transfers the microwave signal into a dc or low-frequency signal and a power meter to read out the measured power levels. The sensor includes load impedance for the microwave source (Figure 34.1).

First, several power definitions need to be clarified:

- *Conjugate available power* (P_{CA}): The maximum available power a signal generator can transmit. This power is delivered by the generator if the load impedance is equal to the complex conjugate of the generator's source impedance. Since measurement techniques often require different lengths of waveguides or coaxial lines, the conjugate available power can be achieved only by tuning.

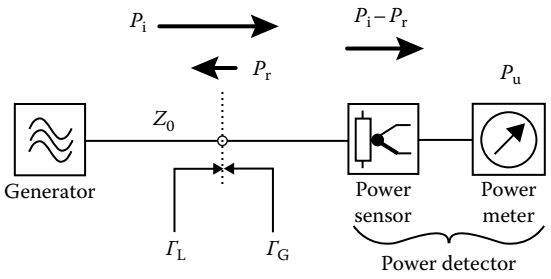


FIGURE 34.1 Setup of a power measurement in load configuration.

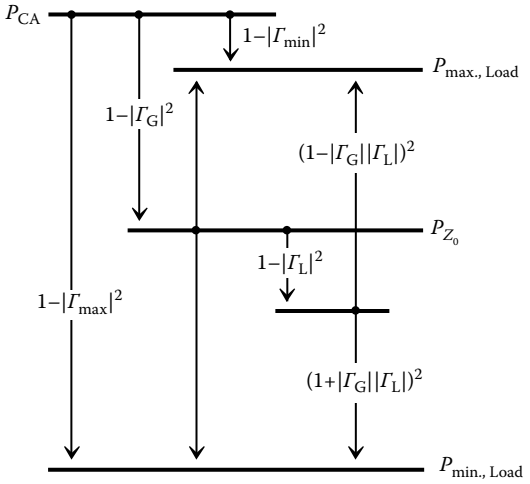


FIGURE 34.2 Relationships between conjugate power P_{CA}, Z_0 available power P_{Z0}, and the power levels available in the load of the power sensor.

- Z_0 *available power* (P_{Z0}): The power that is transferred via a coaxial line of characteristic impedance Z_0 into a load impedance Z_L equal to Z_0, while the generator has the impedance Z_G. Consequently, the available power level is reduced due to generator mismatch: $P_{Z0} = P_{CA}(1 - |\Gamma_G|^2)$, where Γ_G is the reflection coefficient of the generator. Figure 34.2 shows the relationships between P_{CA}, P_{Z0}, and the maximum and minimum power levels that can be measured in the load of a power sensor.
- *Average power* (P_{av}): The result of an averaging over many periods of the lowest modulation frequency of a modulated signal $P(t)$ (Figure 34.3).
- *Envelope power* ($P_e(t)$): The power averaged over the period of the carrier frequency.
- *Peak envelope power* (PEP): The maximum of $P_e(t)$.
- *Pulse power* (P_p): Defined for pulsed signals. If the pulse width τ and the repetition frequency $1/T$ are known, P_p is given by the measured average power:

$$P_p = P_{av}\frac{T}{\tau}$$

34.1.1 Measurement Errors and Uncertainties

Measurement errors occur due to mismatch as well as inside the power sensor and in the power meter. After correcting for these errors, the measurement uncertainties remain. Typically 75% of the total uncertainty belongs to mismatch and the smallest part is due to the power meter. The uncertainties and

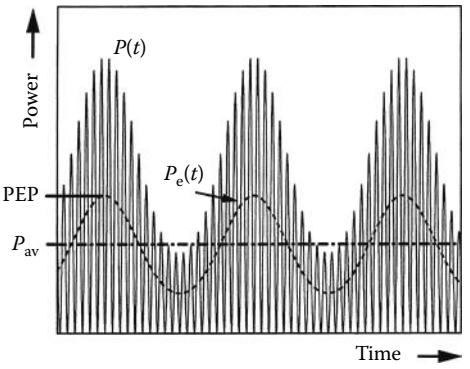


FIGURE 34.3 Power definitions for modulated signals.

errors can be power dependent, frequency dependent, or independent of both. Of course, the *total uncertainty* must be calculated from the different uncertainties u_i as the root sum of the squares: rss = $\sqrt{\sum u_i^2}$, provided that the errors are all independent. A pessimistic error definition is the *worst-case uncertainty*, which simply sums up all extreme values of the independent errors.

Mismatch errors occur due to the fact that neither the generator (G), the line, nor the load (L) exhibits exactly the characteristic impedance Z_0. Using the modulus of the reflection coefficient $|\Gamma|$, the available power of the generator P_G can be expressed as [3]

$$P_G = P_L \frac{\left|1 - \Gamma_G \Gamma_L\right|^2}{(1 - \left|\Gamma_G\right|^2)(1 - \left|\Gamma_L\right|^2)} \tag{34.1}$$

The reflection coefficients can also be expressed in terms of the voltage standing wave ratio (VSWR):

$$\left|\Gamma\right| = \frac{\text{VSWR} - 1}{\text{VSWR} + 1} = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{34.2}$$

As mentioned, the knowledge of Z_0 available power P_{Z0} is sufficient in most applications. Then the ratio between the Z_0 available power and the absorbed power is given by

$$\frac{Pz_0}{P_i - P_r} = \frac{\left|1 - \Gamma_G \Gamma_L\right|^2}{1 - \left|\Gamma_L\right|^2} \tag{34.3}$$

where the subscripts "i" and "r" denote the incident and reflected power, respectively. Expressing Equation 34.3 in dB gives the Z_0 *mismatch loss*, while the term –10 log(1 – $|\Gamma_L|^2$) is called the *mismatch loss*, which accounts for reflected power from the load only.

Since the reflection coefficients are seldom known completely but only their magnitudes are known, a *mismatch uncertainty* is defined as

$$M_u = \left\{(1 \pm \left|\Gamma_G\right\|\Gamma_L\right|)^2 - 1\right\} 100\% \tag{34.4}$$

Conversion errors are due to the specific characteristics of the individual power sensor. The conversion efficiency is frequently dependent and a *calibration factor* (CF) is defined for each power detector:

$$\text{CF} = \frac{P_u}{P_i} \tag{34.5}$$

where

P_u is the uncorrected power

P_i is the actual incident power (Figure 34.1)

Sometimes, also the *effective efficiency* η_e is used:

$$\eta_e = \frac{P_u}{P_i - P_r} \tag{34.6}$$

Both quantities are related via the reflection coefficient of the load of the power sensor:

$$CF = \eta_e(1 - |\Gamma_L|^2) \tag{34.7}$$

The CF is used to correct for efficiency loss and it also accounts for the mismatch loss. Still a remaining CF *uncertainty* has to be taken into account. It is specified for each sensor. The calibration data are usually traceable to a national bureau of standards. Power sensors under test can be compared to the standards using high directivity couplers or power splitters [2].

The next described errors are due to the electronics inside the power meter.

Some power meters exhibit an internal reference oscillator to verify and adjust for the sensitivity of the diode or thermocouple sensor. The *reference power uncertainty* is specified by the manufacturer. Since this reference has its own reflection coefficient, it is related to a *reference oscillator mismatch uncertainty*.

Instrumentation uncertainty depends on the circuit limitations of the power meter and is specified by the manufacturer.

The ±1 *count error* is for digital output and can be expressed by the relative power reading of the last significant digit.

Autozero can be used on all measurement ranges. Zero setting immediately prior to the measurement can reduce *drift errors* when measuring in the lower ranges. Still *zero set errors* remain due to noise during the zero-setting operation. This error can be very serious for measurement of low power levels [5]. *Zero carryover* is caused by ADC quantization errors in the zero readings for all measurement ranges except the most sensitive one. ADC quantization also causes a *power quantization error*. If very low power levels have to be measured, averaging can reduce random noise at the expense of measurement speed.

34.1.2 Power Sensors

A large variety of power sensors is available for the frequency range from dc up to 110 GHz and for minimum detectable power levels as low as 100 pW. The sensors differ in measurement principle, and hence, the correct choice depends on the application. A detailed description of power sensors is given in [2,3]. Thermal sensors transform the absorbed microwave power into heat that is measured with temperature-sensitive elements:

- *Calorimeters* measure the heat absorbed in a fluid (e.g., water) of well-known specific heat. Applying the substitution principle, their precision can be enhanced. Because of their high stability, they are used in the National Institute of Standards. The manufacturer usually references the sensors to these standards.
- *Bolometers* and *thermistors* [3] make use of the temperature-dependent resistivity change of a resistive load, which is highly nonlinear. Hence, dc power substitution is used to keep the load at constant temperature. The substituted dc power is the measurement. Self-balancing bridges are used for this purpose but need careful ambient temperature compensation. The effective efficiency of bolometers is known very exactly. However, they have only relatively small dynamic range (typical 10 μW–10 mW). Currently, liquid nitrogen-cooled, high-temperature superconducting bolometers with extremely high sensitivity of several thousand volts per watt are used [6]. A comprehensive overview of bolometers is given in [7].

- *Thermocouple sensors* are based on microwave power conversion into heat via matched load impedance. Its temperature is controlled by thermocouples utilizing the thermoelectric effect in thin films or semiconductors [8]. One has to distinguish between realizations where the thermocouple itself represents the load impedance [3] and galvanically decoupled thermocouples [9]. The main advantages of the latter sensors are the better flatness of the response from dc to microwave and a lower voltage standing wave ratio (VSWR). The design of such sensors is simple and a silicon micromachining technique can be applied to enhance the sensitivity [10]. The lowest measurable power levels are 1 μW.

Thermal sensors are well suited to absolute power measurement, especially with respect to their high linearity. Their drawback is the relatively long response time (>10 ms), which limits these sensors to the measurement of average power. The high speed of Schottky diodes predestines them for the measurement of peak power, envelope power, and peak envelope power (PEP):

- *Diode sensors* consist of Schottky barrier diodes that are zero biased and work in the square-law part of their *I–V* characteristics. For low power levels (<–20 dBm), these devices are very linear in response and measure the average power correctly. Still, diode sensors exhibit non-quadratic contributions [11], which can be important to account for when accurate measurement is required. Minimum detectable power is –70 dBm where the signal level is of the order of 50 nV and requires sophisticated amplification. The diode sensor is part of the sensors described later.
- *Peak power sensors* are specially designed for peak power measurements and account for measurement errors due to waveform and power level, although any diode detector can be used for this purpose if the peak voltages are ≤1 V.
- *Peak envelope analyzers* are designed to detect the envelope power. This is not possible with a simple diode sensor because the electronic setup of the diode must be different.
- *Feedthrough power sensors* are used to measure microwave power in transmission configuration. They have minor losses of approximately 0.5 dB and a limited bandwidth of typically 0.1–1 GHz. The limiting device in these systems is a directional coupler with power splitters. Such a measurement can also be implemented with discrete elements. The characteristic figure of merit of these devices is the *directivity* (a_D), relating the read incident power P_i to the read reflected power P_r in case of reflection free load ($\Gamma_L = 0$):

$$a_D = 10 \log \frac{P_i}{P_r} \text{ dB} \tag{34.8}$$

The directivity should be as high as possible.

The previously mentioned power sensors are discrete devices. The maturity in microwave monolithic integrated circuit (MMIC) design and fabrication allows the integrated realization of diode power sensors, for example, in an integrated six-port reflectometer [12]. Activities to fabricate thermal power sensors integrable to MMIC typical processes [13] can presently be implemented on commercial processes, such as the Philips Lemeill high electron mobility transistor (HEMT) processes with additional postmicromachining [14].

34.1.3 Commercially Available Power Measurement Systems

A collection of different power sensors and corresponding measurement units is shown in Table 34.1. All power meters have general purpose interface bus (GPIB) interfaces for easy use in automated measurement.

TABLE 34.1 Available Commercial Power Sensors and Power Meters

Supplier	Power Sensor* Power Meter	Frequency Range	Dynamic Range	VSWR	Remark
Power sensors in matched load configuration and related power meter					
Hewlett-Packard	HP8478B	10 MHz–18 GHz	—	1.1–1.75	Thermistor
Rohde & Schwarz	NRV-Z52	dc–26.5 GHz	1 μW–100 mW	1.11–1.22	Thermocouple, up to 30 W available
Marconi	6913/6914S	10 MHz–26.5/46 GHz	1 μW–100 mW	1.1–1.4/3.6	Thermocouple, up to 3 W available
Boonton	51100(9E)	10 MHz–18 GHz	10–100 μW	1.18–1.28	Thermocouple
Hewlett-Packard	HP8485A	50 MHz–26.5 GHz	1 μW–100 mW	1.10–1.25	Thermocouple
	HP8487A	50 MHz–50 GHz	1 μW–100 mW	1.10–1.50	Thermocouple
	HPR/Q/W8486A	26.5–40/33–50/ 75–110 GHz	1 μW–100 mW	1.4/1.5/1.08	Rectangular waveguide, thermocouple
Rohde & Schwarz	NRV-Z6	50 MHz–26.5 GHz	1 nW–20 mW	1.2–1.4	Diode
Marconi	6923/6924S	10 MHz–26.5/ 46 GHz	0.1 nW/ 0.1–10 μW	1.12–1.5/3.6	Diode
Hewlett-Packard	HP8487D	50 MHz–50 GHz	0.1 nW–10 μW	1.15–1.89	Diode
Rohde & Schwarz	*NRVS/D	dc–26.5 GHz	0.4 nW–30 W	—	One/two channel
Boonton	*4230A	l0 kHz–l00 GHz	0.1 nW–25 W	—	—
Marconi	*6960B	30 kHz–46 GHz	0.1 nW–30 W	—	—
	*6970	30 kHz–46 GHz	0.1 nW–30 W	—	Hand portable
Hewlett-Packard	*HP437B	100 kHz–110 GHz	0.07 nW–25 W	—	—
Power analyzer					
Hewlett-Packard	HP84812/13/14A	500 MHz–18/26.5/40 GHz	0.6 μW–100 mW	1.25–1.60	Resolution 100 ps
	*HP8991A	500 MHz–40 GHz	0.5 μW–100 mW	—	Rise/fall time 5 ns
Peak power sensors and related power meter					
Rohde & Schwarz	NRV-Z31/33	0.03–6 GHz	1 μW–20 mW/1 mW–20 W	1.05–1.33	With NRVS
Hewlett-Packard	HP84812/3/4A	500 MHz–18/26.5/40 GHz	1 μW–100 mW	1.25/1.35/1.50	—
Boonton	56340	500 MHz–40 GHz	—	1.25–2.00	Dual diode risetime <15 ns
	*HP8990A	500 MHz–40 GHz	—	—	—
Boonton	4500A	1 MHz–40 GHz	0.1 μW–100 mW	—	—
Feedthrough power sensor and related power meter					
Rohde & Schwarz	NAS-Z7	1.71–1.99 GHz	0.01–30 W	<1.15	a_D > 26 dB, GSM, DCS1800/1900
	*NAS	0.001–1.99 GHz	10 mW–1200 W	—	—

34.2 Frequency Measurement

Frequency measurement in the microwave regime is usually part of a more complex measurement procedure, for example, determining the scattering parameters and filter characteristics of a device under test (DUT). If one is only interested in frequency or a higher accuracy is required, one must use direct frequency measurement systems.

Two different techniques can be distinguished. The rather old-fashioned way is to use mechanically tunable resonators, the so-called *wave meters*. These are not explained in detail here. *Digital frequency counters* are an alternative and are now the state of the art (Table 34.2).

The digital frequency counter measurement system is based on the principle of counting the zero crossovers of a continuous sinusoidal signal. At low frequencies, this method can be used directly, whereas in the microwave region, direct digital counters are not available because of their limited bandwidth. Thus, a modified measurement system must be used.

The standard digital frequency counter usually consists of a mixer, a local oscillator (frequency f_0) in the lower-frequency regime, several multipliers, and the digital counter. The principal measurement technique is shown in Figure 34.4.

An extremely stable local oscillator (quartz oscillator) is used to provide the reference signal used in the measurement system. This signal is multiplied by a factor of N and mixed with the RF signal of the DUT. The IF in the low-frequency range can be easily counted, and, thus, the frequency of the signal f_s can be calculated according to the following equation:

$$f_s = f_{IF} + Nf_0 \tag{34.9}$$

For this method, an extremely stable low-frequency oscillator (often temperature controlled) must be provided, and in order to allow a sufficient bandwidth, a high number of multipliers must be implemented in this system.

TABLE 34.2 Digital Microwave Frequency Counter

Supplier	Counter	Frequency Range	Resolution (Hz)	Sensitivity	Remark	–Price $US
HewlettPackard	HP5351B	26.5 GHz	1	–40 dBm	—	7,500
HewlettPackard	HP5352B	40(46) GHz	1	–30 dBm	—	11,800

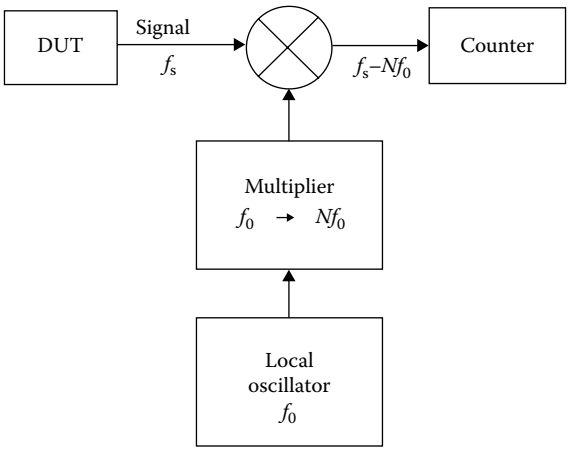


FIGURE 34.4 Block diagram of a digital frequency counter.

A pulsed oscillator with extremely short risetime can circumvent this problem. In the frequency domain, this signal is given by spectral lines at $f = if_0$, where f_0 denotes the fundamental frequency of the pulses. Using a band-pass filter, a single frequency can be separated and transferred to the mixer.

34.3 Spectrum Analysis

The expression spectrum analysis subsumes the measurements that are performed to obtain the Fourier transformation $S(f)$ of a given signal $s(t)$. The Fourier transformation of $s(t)$ in the frequency domain is defined by the following equation:

$$S(f) = \int_{-\infty}^{+\infty} s(t) e^{-j2\pi ft} dt \tag{34.10}$$

In practice, the lower and upper bounds of the integral are limited by a finite measurement time that must be fixed by the user:

For the analysis of an unknown spectrum, different methods can be distinguished:

- *Wave analyzers and selective voltmeters*: These devices utilize a tunable filter for frequency-selective measurements.
- *Spectrum analyzers* rely on the principle of heterodyne mixing with subsequent band-pass filtering.
- Calculation of the spectrum using a *fast Fourier transformation* (FET). This method can be employed only for lower frequencies, since a digital-to-analog converter is needed. For microwave frequencies, the calculation of the spectrum using the fast Fourier transform (FFT) is, therefore, difficult to realize.

Since the spectrum analyzer is most suitable for microwave frequencies, it will be described in detail in the following. Brief introductions into the spectrum analyzer measurement techniques are given in [15,16].

34.3.1 Spectrum Analyzer

The spectrum analyzer is most suitable for the analysis of microwave signals. It is a general-purpose instrument for measurements in the frequency domain and provides the user with the amplitude, power, or noise density of a signal depicted versus the frequency. The frequency scale is in most cases linear; the vertical axes can be either linear or logarithmic. Spectrum analyzers are available from a few hertz up to more than 100 GHz. They give a quick overview of the spectral power distribution of a signal. Spectrum analyzers have a large dynamic range, a resolution bandwidth of a few hertz, and a reasonable frequency resolution.

The spectrum analyzer is suitable for the following measurements:

- *Measurement of absolute and relative frequency:* Frequency drift (unstabilized oscillators), spectral purity, and frequency of harmonics.
- *Absolute and relative amplitude:* Gain of frequency multipliers, harmonics of periodic signals, intermodulation (IM) distortion, and harmonic distortion.
- Scalar network analysis (if equipped with a tracking generator): Frequency response of amplifiers and filters.
- Electromagnetic interference (EMI) measurements: Broadband spectra.
- Measurements of modulated signals: AM, FM, or PM.
- *Noise:* Many spectrum analyzers can be used for noise measurements of active devices.
- *Phase noise:* Phase noise of oscillators can be analyzed with spectrum analyzers [17].

34.3.2 Spectrum Analyzer Setup

The spectrum analyzer is basically an electronically tunable filter that allows the measurement of the amplitude, power, or noise at a certain frequency. Using the example shown in Figure 34.5, the principle of operation can be explained as follows.

The tunable filter used to separate the frequencies to be measured is realized using a chain of mixers and intermediate frequency (IF) amplifiers. In this case, three mixers convert a given input signal frequency f_s to the IF passband of the last IF amplifier. At least one of the oscillators is tunable (voltage-controlled oscillator, VCO) in order to scan the input frequency f_s. Sometimes, more than one tunable oscillator is used.

The first mixer and the following IF amplifier with a band-pass center frequency of 1.95 GHz in the given example selects the input frequency to be analyzed according to

$$f_s = f_{ol} - 1.95\ \text{GHz} \tag{34.11}$$

The input frequency for the spectrum analyzer shown in Figure 34.4 can be due to the scan of the frequency of the tunable oscillator f_{ol}, between 50 MHz and 2.05 GHz. However, the image frequency f_{is} will also be mixed to the IF:

$$f_{is} = f_{ol} + 1.95\ \text{GHz} \tag{34.12}$$

Since f_{is} is in the range of 3.95–5.95 GHz, a band-pass filter with a cutoff frequency of 3 GHz is used at the input to reject the image frequency.

Because it is difficult to realize a narrow band-pass at 1.95 GHz, the signal is converted to a lower frequency in the megahertz range. At these frequencies, stable quartz filters with high-quality factors can be employed. In principle, the RF frequency could be mixed down to the last IF section in one step. However, in such an arrangement, it would be difficult to suppress the image frequencies. In any case, the image frequency of each stage should be rejected by the preceding IF amplifier as shown in Figure 34.6 for the second stage of the given example.

The last IF amplifier is very important for the performance of the complete system. In nearly all spectrum analyzers, its bandwidth—the so-called resolution bandwidth—can be adjusted in steps. For separation of closely spaced spectral components, the bandwidth should be very small. Most spectrum analyzers offer a minimum bandwidth of a few hertz. On the other hand, larger bandwidths are needed, since for a narrow IF filter, only a slow scan speed can be allowed (see succeeding text), and therefore, the measurement over several frequency decades would result in a large sweep time. The shape of the IF filter is important for the capability of the spectrum analyzer to separate close spectral components. The performance of the IF filter is described by the shape factor. It is defined by the ratio of the 60 dB

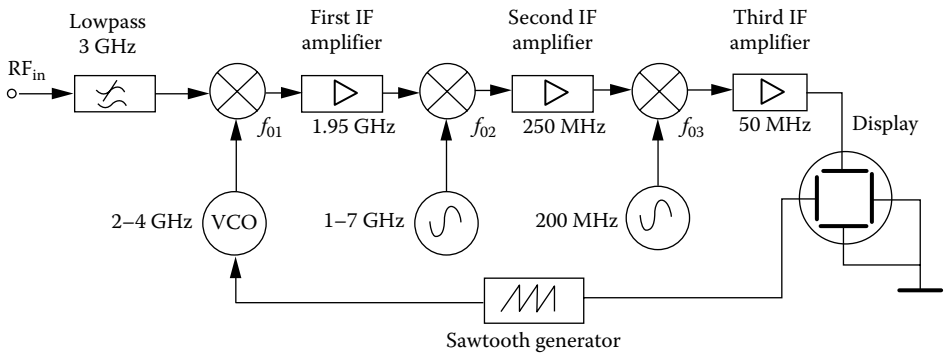


FIGURE 34.5 Simplified setup of a spectrum analyzer.

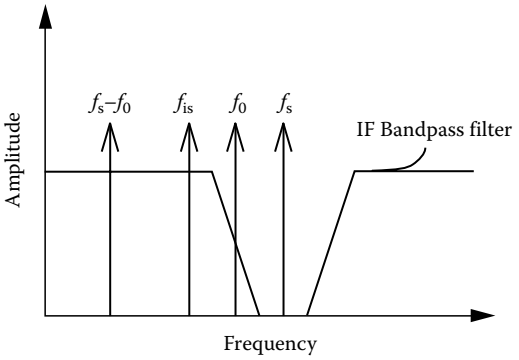


FIGURE 34.6 Example for the blocking of the image frequency by the IF filter in an arrangement according to Figure 34.5.

to the 3 dB bandwidth of the IF filter. The IF filter can be switched between the linear and logarithmic amplifications. This is performed numerically in most cases.

A sawtooth generator produces the control voltage for the voltage controlled oscillator (VCO) and the *x*-deflection voltage of the screen. A significant error is introduced in the frequency scale by a not-ideal voltage-to-frequency characteristic of the VCO. Therefore, in many cases, synthesizers with a quartz-stabilized phase-locked loop [18] are used.

The detector has to be sensitive to either the amplitude, the power, or the noise $(\text{mV}\sqrt{\text{Hz}^{-1}})$. In modern spectrum analyzer, digital signal processing is used for this purpose.

It is important to note that there are restrictions on the minimum sweep time T_s. In order to avoid settling errors of the narrow band-pass IF amplifier with bandwidth B, the scan time should be for a frequency span S larger than [19,20]

$$T_s > 20\frac{S}{B^2} \tag{34.13}$$

Most of the spectrum analyzers control the scan time automatically according to this equation.

34.3.3 Harmonic Mixing

For higher frequencies, the harmonic mixing technique is widely used. If the first 3 GHz low-pass filter in Figure 34.5 is omitted and the VCO produces harmonics, a larger number of input frequencies are converted to the passband of the spectrum analyzer. For the example shown, the possible input signals are depicted in Figure 34.7 versus the VCO frequency according to

$$f_s = nf_{01} \pm 1.95 \text{ GHz} \tag{34.14}$$

The notation of the numbers in the figure is the harmonic number n and the plus or the minus sign in the earlier equation. For a frequency $f_{ol} = 3$ GHz of the VCO, the following frequencies will appear at the same frequency location on the screen: 1.05 GHz (1–), 4.95 GHz (1+), 4.05 GHz (2–), 7.95 GHz (2+), 7.05 GHz (3–), and 10.95 GHz (3+). A tracking preselection filter (see succeeding text), which is scanned with the VCO, can select one of these harmonics.

For further extension of the frequency in the upper mm-wave range, external mixers are used. With such an arrangement, frequencies higher than 500 GHz can be measured. Additionally, the equipment for mixing of signals in the optical range is offered by some companies.

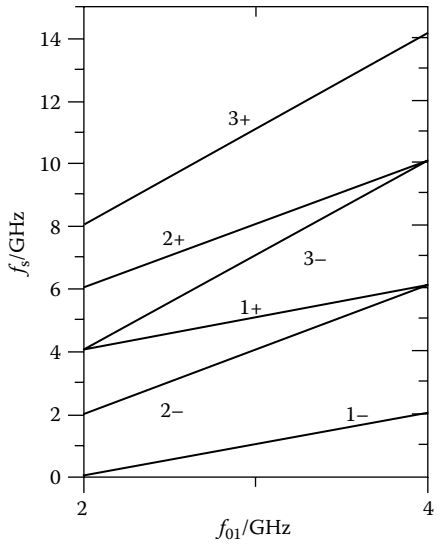


FIGURE 34.7 Measured RF frequency f_s versus the frequency of the VCO f_{01}. Parameters are n and the plus or the minus sign in Equation 34.4.

34.3.4 Tracking Preselection

For small input signals, the spectrum analyzer can be considered a linear device. However, if the input level increases, harmonics and IM products are generated due to the nonlinearities of a mixer. These products result in spurious signals on the screen of the spectrum analyzer. In addition, image frequencies will appear on the screen, as demonstrated earlier. To avoid these spurious responses, a tracking preselection filter is employed at the input of the spectrum analyzer. A tracking preselection filter is an electronically tuned band-pass filter usually realized using a yttrium iron garnet (YIG) filter.

34.3.5 Tracking Generator

Spectrum analyzers are often equipped with a tracking generator. The principle is shown in Figure 34.8. A local oscillator, with a frequency exactly on the center frequency of the IF amplifier, is mixed by an

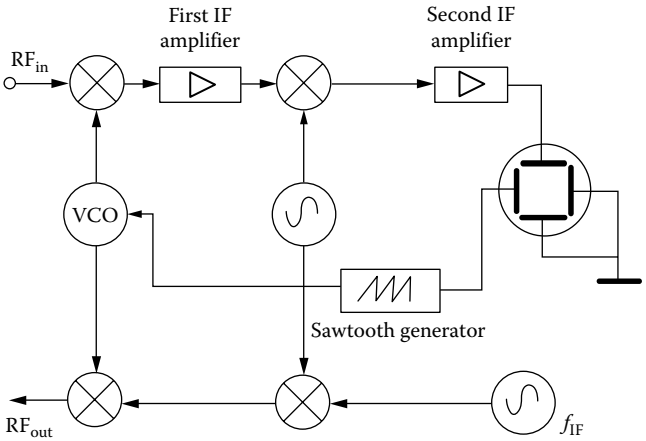


FIGURE 34.8 Principle of a tracking generator, which delivers at the port RF_{out} a signal that is precisely in the passband of the spectrum analyzer.

identical setup as in the analyzer path. Using the same local oscillator as in the analyzer path ensures that the frequency of the tracking generator follows precisely the center frequency of the swept window of the analyzed frequency band.

The tracking generator can be used for network analysis. If the tracking frequency is used as an input signal of a two-port, the amplitude of the output can be measured with the spectrum analyzer. Such a network analyzer has the advantage of being sensitive only in a very narrow band. Thus, third-order IM products and noise are suppressed. However, only scalar measurements can be performed.

34.3.6 Commercially Available Spectrum Analyzers

A number of commercially available, general-purpose spectrum analyzers for the microwave and mm-wave range are listed in Table 34.3. Only a small number of spectrum analyzers available on the market are presented. Most of the companies offer special equipment for production quality control. These spectrum analyzers can be computer controlled for fixed measurement routines. On request, spectrum analyzers with special options like fixed frequency operation, multichannel operation, support of external mixers, integrated frequency counters, and digital storage are available.

Typical specifications of microwave spectrum analyzers include

- The frequency span is several gigahertz.
- With external mixers, the upper frequency limit can be extended to more than 100 GHz.
- The frequency accuracy is between 10^{-5} and 10^{-7}.

TABLE 34.3 Commercially Available Spectrum Analyzers

Company/Model	Frequency Range	Min. Res. Bandw.	Amplitude Accuracy	Remarks
Anritsu				
MS2602A	100 Hz–8.5 GHz	—	1.1 dB	—
Avantek				
3365	100 Hz–8 GHz	10 Hz	—	Portable, tracking
3371	100 Hz–26.5 GHz	10 Hz	—	Portable, tracking
R3272	9 kHz–26.5 GHz	300 Hz	1 dB	External mixer 325 GHz
Hewlett-Packard				
HP4196A	2Hz–1.8 GHz	1 Hz	1 dB	—
HP8590L	9 kHz–1.8 GHz	1 kHz	1.7 dB	—
HP8560E	30 Hz–2.9 GHz	1 Hz	1.85 dB	Portable
HP8596E	9 kHz–12.8 GHz	30 Hz	2.7 dB	—
HP8593E	9 kHz–22 GHz	30 Hz	2.7 dB	—
HP8564E	9 kHz–40 GHz	1 Hz	3 dB	—
HP8565E	9 kHz–50 GHz	1 Hz	3 dB	—
Marconi				
2370	30 Hz–1.25 GHz	5 Hz	5 Hz	With frequency extender
2383	30 Hz–4.2 GHz	3 Hz	1 dB	Tracking
Rohde & Schwarz				
FSEA30	20 Hz–3.5 GHz	1 Hz	1 dB	—
FSEB30	20 Hz–7 GHz	1 Hz	1 dB	—
FSEM30	20 Hz–26.5 GHz	1 Hz	1 dB	External mixer 110 GHz
Tektronix				
2714	9 kHz–1.8 GHz	300 Hz	2 dB	AM/FM demodulation 50 Ω/75 Ω
2784	100 Hz–40 GHz	3 Hz	1.5 dB	Counter 1.2 THz, external mixer 325 GHz

- The resolution bandwidth (i.e., the effective bandwidth of the narrow IF filter) can be adjusted in steps between 1 Hz and a few megahertz.
- The resolution bandwidth shape factor is typically 10:1.
- The amplitude accuracy is about 1 dB.
- The noise floor is at about –140 dBm.

34.4 Cavity Modes and Cavity *Q*

Cavities are used in a variety of applications. For example, they can be used to construct filters, and they serve as those elements in microwave generators (e.g., klystrons) that determine the operating frequency. Cavities can also be applied in order to measure the frequency or the wavelength of microwaves (wavemeter). The most important parameters of a cavity are its resonant frequency and its *Q*-factor. The latter determines the sharpness of the resonance or, in case of filters, the bandwidth of the passband.

A cavity can be defined as a volume that is almost completely surrounded by a metallic surface. At one or two positions, coupling elements are applied to the metal surface in order to connect the cavity to other circuit elements. In the case of one coupling element, one speaks of a single-ended cavity, whereas a transmission cavity has two coupling elements. The coupling elements are usually small compared to the dimensions of the cavity. The inside of the cavity is filled with dielectric material (e.g., air: $\varepsilon_r = 1$).

Cavities are used as resonators for microwave applications. Therefore, they are comparable to low-frequency resonant circuits consisting of an inductance *L* and a capacitance *C*. In the low-frequency range, lumped elements (inductors and capacitors) are used, which are small in comparison with the wavelength. In contrast, cavities are distributed elements with dimensions comparable to the wavelength. This results in comparatively small losses. Since cavities are distributed elements, it is in general no longer possible to determine *L* and *C* directly. Instead of *L* and *C*, the resonant frequency $f_0 = \omega_0/2\pi$ is the most important property of a cavity.

Neither lumped elements nor cavities are completely lossless. One must take into account the finite conductivities of the materials, resulting in a resistance *R* when analyzing resonant circuits. For cavities, however, *R* cannot be determined directly due to the same reasons that hold for *L* and *C*. Therefore, a different parameter, the *Q*-factor *Q*, plays a similar role for cavities. *Q* is proportional to the stored electric and magnetic energy *W*, divided by the power loss *P*:

$$Q = \frac{\omega_0 W}{P} \tag{34.15}$$

Although cavities are distributed elements, one is able to show that their behavior near resonance can be described by a simple parallel resonant circuit consisting of *R*, *L*, and *C* if the reference plane is chosen appropriately. Therefore, the basic properties of such a parallel resonant circuit are analyzed in the following.

Assume that the losses are small $Q << 1$, which is desirable in practice. Therefore, the resistance *R* of the parallel resonant circuit is comparatively large. In this case, the resonant frequency does not depend on *R* in a first-order approximation:

$$\omega_0 \approx \frac{1}{\sqrt{LC}} \tag{34.16}$$

In Equation 34.15, which is applied to the analyzed parallel resonant circuit, one obtains

$$Q \approx \frac{R}{\omega_0 L} \approx R\sqrt{\frac{C}{L}} \tag{34.17}$$

One can characterize the width of a resonance curve by those angular frequencies $\omega_1 = 2\pi f_1$ and $\omega_2 = 2\pi f_2$ where the power has decreased to one-half of its maximum value (–3 dB). The impedance

has then decreased to $1/\sqrt{2} \approx 70.7\%$ of its maximum value. This enables calculation of the angular frequencies ω_1 and ω_2:

$$\omega_1 \approx \omega_0 - \frac{1}{2RC}, \quad \omega_2 \approx \omega_0 + \frac{1}{2RC} \tag{34.18}$$

Using these relations, one can easily derive the following equation, which is equivalent to Equation 34.17:

$$Q \approx \frac{\omega_0}{\omega_2 - \omega_1} \approx \frac{f_0}{f_2 - f_1} \tag{34.19}$$

This expression shows that Q is a symbol for the sharpness of resonance. Furthermore, it leads to the first principle as to how the Q-factor can be measured. This principle, which is often referred to as the band-pass method, is based on the measurement of a resonance curve. For example, one can measure the reflection coefficient S_{11} with a network analyzer. From this curve, the 3 dB bandwidth $\Delta f = f_2 - f_1$ and the resonant frequency $f_0 = (f_1 + f_2)/2$ can be easily determined. The application of Equation 34.19 yields the desired Q.

A second principle used to measure the Q-factor is based on the transient behavior of the cavity after excitation with a pulse. In this case, the energy decays exponentially, and the time constant of the decay is proportional to Q^{-1}. This can again be shown by analyzing the equivalent parallel resonant circuit. Up until now, all signals were time harmonic, which enabled a solution by complex quantities. This time, the corresponding differential equation must be solved. One can easily show that the amplitude of the voltage is proportional to $e^{-\omega_0 t/2Q}$, which means a $e^{-\omega_0 t/2Q}$ dependency of the energy exists. Measuring the decay time constant, therefore, enables one to determine the Q-factor.

The Q, defined here, is the unloaded Q of the cavity; that is, no further resistance is connected to the cavity. Sometimes it is desirable to determine the loaded Q (Q_L), which takes into account such resistances. Detailed information about the loaded Q can be found in [21–23]. Although the Q-factor is not influenced by the coupling structure if it is lossless, the coupling structure is of great importance. Further information about coupling parameters is presented in [21–23]. Detailed descriptions of measurement methods based on the previously mentioned principles can also be found in [21,22].

Up to now, all analyses were based on the equivalent resonant circuit. There are many other effects that can be explained by this analogy. For example, the energy in both the cavity and in its equivalent circuit oscillates between the magnetic and electric fields. Some properties of cavities, however, can only be seen by examining the electric and magnetic fields themselves. These are governed by Maxwell's equations.

A solution of Maxwell's equations (which can be very complicated for a given cavity) shows that all cavities have an infinite number of resonant frequencies. This could not be seen by analyzing the equivalent parallel resonant circuit. A description of the cavity by a discrete parallel resonant circuit is only valid in the vicinity of *one* of these resonant frequencies.

Furthermore, different modes can exist that have the same resonant frequency. This phenomenon is called *degeneration*. If the resonator will operate at such a frequency, where degenerate modes exist, one has to take care that the companion mode is suppressed. This can be accomplished by an appropriate choice of the positions of the coupling elements.

Even if the desired mode does not have any companion mode, one should take care that no other mode exists in the operating range of the cavity (this can be accomplished with a mode chart [22,23]); otherwise, these modes must be damped sufficiently in order to be sure that an observed resonance corresponds to the desired mode.

34.5 Scattering Parameter Measurements

Scattering parameters describe multiple-port structures in terms of wave variables. The introduction of scattering parameters (*S*-parameters) arises naturally in microwave circuits and systems, due to the lack of a unique definition for currents and voltages at these frequencies. Most circuits and systems at high frequencies are efficiently described in terms of *S*-parameters.

This section describes the fundamentals and properties of *S*-parameters, together with network analysis based on *S*-parameter calculations and measurements. Measurement procedures are outlined and the most frequently used systems for *S*-parameter measurement are described. Finally, information on hardware required for the experimental determination of *S*-parameters is provided, together with the corresponding suppliers.

34.5.1 Introduction and Fundamentals

At high frequencies, the wave variables are a natural extension of the voltages and currents at port terminals. In electric systems where the voltages and currents cannot be uniquely defined, the power flow in a waveguide can be described via wave variables. Whenever a transverse electromagnetic (TEM) mode of wave propagation cannot be assumed, the currents and voltages are dependent on the integration path. This situation is encountered in all rectangular, circular, and passive waveguide structures, even in the case of lossless wave propagation. It is also true for all guiding structures if losses are to be considered along the path of wave propagation [24–27]. For the case of wave propagation along a transmission line, the wave variables a and b are defined as follows:

$$\begin{aligned} a(z) &= \frac{1}{2}\left(\frac{U(z)}{\sqrt{Z_0}} + I(z)\sqrt{Z_0}\right) \\ &= \frac{U_+}{\sqrt{Z_0}} = I_+\sqrt{Z_0} \\ b(z) &= \frac{1}{2}\left(\frac{U(z)}{\sqrt{Z_0}} - I(z)\sqrt{Z_0}\right) \\ &= \frac{U_-}{\sqrt{Z_0}} = I_-\sqrt{Z_0} \end{aligned} \tag{34.20}$$

The propagation is along the z-direction. The characteristic impedance of the transmission line is Z_0, and $U(z)$ and $I(z)$ are the voltage and current, respectively, at location z along the line. The variables $a(z)$ and $b(z)$ are the complex amplitudes of the modes on the line. The voltage U_+ and U_- and the currents I_+ and I_- denote the voltage and current amplitudes, respectively, in forward and reverse direction. The wave variables are related to the power in the following form:

$$\begin{aligned} P_+ &= \frac{1}{2}\frac{|U_+|^2}{Z_0} = \frac{1}{2}|I_+|^2 Z_0 = \frac{1}{2}|a(z)|^2 \\ P_- &= \frac{1}{2}\frac{|U_-|^2}{Z_0} = \frac{1}{2}|I_-|^2 Z_0 = \frac{1}{2}|b(z)|^2 \end{aligned} \tag{34.21}$$

In Equation 34.21, it is assumed that the system is excited by a pure sinusoid and that the characteristic impedance is purely real. The wave variables have the dimensions of $\sqrt{W}$.

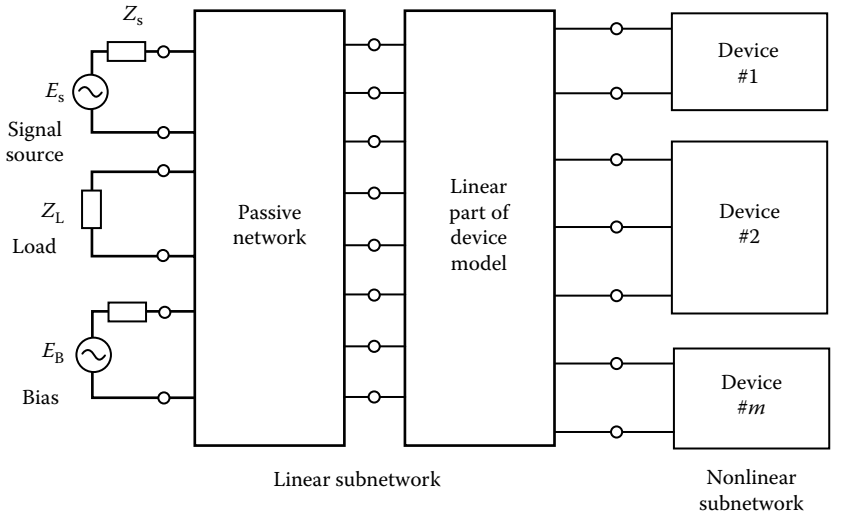


FIGURE 34.9 Schematic illustration of application of wave variables to nonlinear circuits.

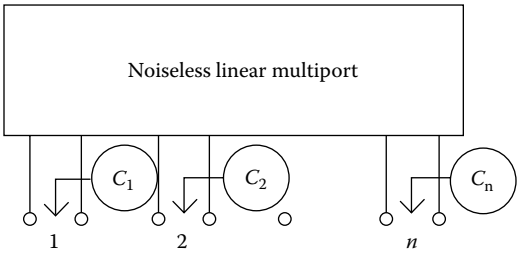


FIGURE 34.10 Schematic illustration of application of wave variables to noisy circuits.

Strictly speaking, wave variables and *S*-parameters can only be applied to linear networks. This is important because many publications are devoted to the so-called large-signal *S*-parameter measurements. The interpretation of such results is not simple, and great care must be employed in the correct determination of the characteristic impedance of the system [26]. In the case of analysis of large-signal or nonlinear circuits, two methods exist to introduce the wave variables:

- Volterra series representation [28]
- Harmonic-balance method [26]

Both methods transform the nonlinear circuit into a number of linear circuits at different frequencies and then change the terminal voltages and currents into wave variables. This situation is sketched in Figure 34.9. Particular attention must be paid to the definition of the characteristic impedance, which can vary between different frequencies. An in-depth treatment of wave variables can be found in [26].

A further utilization of wave variables can be found in the noise analysis of microwave circuits. According to Figure 34.10, a noisy multiport can be analyzed by an associated noiseless two-port with the according noise sources c_i at the corresponding ports of the circuit.

34.5.2 Calculations and Analysis with *S*-Parameters

The characterization of multiports with *S*-parameters requires embedding of the multiport into a system consisting of a signal source with characteristic impedance and appropriate terminations of all ports.

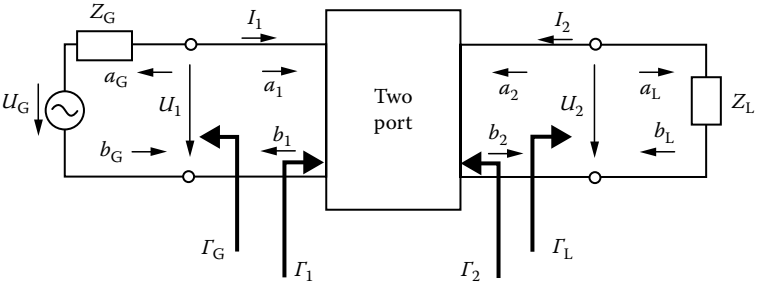


FIGURE 34.11 Two-port network indicating the wave variables and the scattering parameters. The subscripts G and L indicate the generator and the load, respectively.

This situation is shown in Figure 34.11. The outgoing wave parameters b are reflections at the corresponding ports. The wave variables are related to the scattering parameters of a two-port in the following manner:

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = (S)\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{34.22}$$

For the determination of the individual scattering matrix elements, all ports of the network must be terminated in their characteristic impedance. The impedances at the corresponding ports need not be equal for all ports. The S-parameters are, in general, complex and are defined with respect to reference planes. These reference planes can be the network terminals but could also been shifted to other locations in the circuit if this is desirable. The scattering matrix can be transformed into all circuit representations. Table 34.4 indicates the conversion formulae between the S-parameters and Z Y h parameters for arbitrary characteristic impedances. Additional conversion methods can be found in [29].

The scattering parameters in the case of a noisy two-port as indicated schematically in Figure 34.10 are defined as [26]

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{34.23}$$

The noise wave sources c_1 and c_2 represent the noise generated in the circuit and are therefore complex variables varying with time. They are characterized by a correlation matrix C_s as follows:

$$C_s = \overline{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}^H} = \begin{pmatrix} \overline{|c_1|^2} & \overline{c_1 c_2^*} \\ \overline{c_2 c_1^*} & \overline{|c_2|^2} \end{pmatrix} \tag{34.24}$$

where

bar ($\bar{}$) indicates time averaging
$(\cdot)^H$ denotes the Hermitian conjugate
* stands for the complex conjugate

TABLE 34.4 Equations for the Conversion between Different Two-Port Parameters

$S_{11} = \dfrac{(X_{11} - Z_{01}^{\star})(X_{22} + Z_{02}) - X_{12}X_{21}}{(X_{11} + Z_{01})(X_{22} + Z_{02}) - X_{12}X_{21}}$	$X_{11} = \dfrac{(Z_{01}^{\star} + S_{11}Z_{01})(1 - S_{22}) + S_{12}S_{21}Z_{01}}{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}$
$S_{12} = \dfrac{2X_{12}\sqrt{R_{01}R_{02}}}{(X_{11} + Z_{01})(X_{22} + Z_{02}) - X_{12}X_{21}}$	$X_{12} = \dfrac{2S_{12}\sqrt{R_{01}R_{02}}}{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}$
$S_{21} = \dfrac{2X_{21}\sqrt{R_{01}R_{02}}}{(X_{11} + Z_{01})(X_{22} + Z_{02}) - X_{12}X_{21}}$	$X_{21} = \dfrac{2S_{21}\sqrt{R_{01}R_{02}}}{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}$
$S_{22} = \dfrac{(X_{11} + Z_{01})(X_{22} - Z_{02}^{\star}) - X_{12}X_{21}}{(X_{11} + Z_{01})(X_{22} + Z_{02}) - X_{12}X_{21}}$	$X_{22} = \dfrac{(1 - S_{11})(Z_{02}^{\star} + S_{22}Z_{02}) - S_{12}S_{21}Z_{02}}{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}}$
$S_{11} = \dfrac{(1 - Y_{11}Z_{01}^{\star})(1 + Y_{22}Z_{02}) + Y_{12}Y_{21}Z_{01}^{\star}Z_{02}}{(1 + Y_{11}Z_{01})(1 + Y_{22}Z_{02}) - Y_{12}Y_{21}Z_{01}Z_{02}}$	$Y_{11} = \dfrac{(1 - S_{11})(Z_{02}^{\star} + S_{22}Z_{02}) + S_{12}S_{21}Z_{02}}{(Z_{01}^{\star} + S_{11}Z_{01})(Z_{02}^{\star} + S_{22}Z_{02}) - S_{12}S_{21}Z_{01}Z_{02}}$
$S_{12} = \dfrac{-2Y_{12}\sqrt{R_{01}R_{02}}}{(1 + Y_{11}Z_{01})(1 + Y_{22}Z_{02}) - Y_{12}Y_{21}Z_{01}Z_{02}}$	$Y_{12} = \dfrac{-2S_{12}\sqrt{R_{01}R_{02}}}{(Z_{01}^{\star} + S_{11}Z_{01})(Z_{02}^{\star} + S_{22}Z_{02}) - S_{12}S_{21}Z_{01}Z_{02}}$
$S_{21} = \dfrac{-2Y_{21}\sqrt{R_{01}R_{02}}}{(1 + Y_{11}Z_{01})(1 + Y_{22}Z_{02}) - Y_{12}Y_{21}Z_{01}Z_{02}}$	$Y_{21} = \dfrac{-2S_{21}\sqrt{R_{01}R_{02}}}{(Z_{01}^{\star} + S_{11}Z_{01})(Z_{02}^{\star} + S_{22}Z_{02}) - S_{12}S_{21}Z_{01}Z_{02}}$
$S_{22} = \dfrac{(1 + Y_{11}Y_{11})(1 - Y_{22}Z_{02}^{\star}) + Y_{12}Y_{21}Z_{01}Z_{02}^{\star}}{(1 + Y_{11}Z_{01})(1 + Y_{22}Z_{02}) - Y_{12}Y_{21}Z_{01}Z_{02}}$	$Y_{22} = \dfrac{(Z_{01}^{\star} + S_{11}Z_{01})(1 - S_{22}) - S_{12}S_{21}Z_{01}}{(Z_{01}^{\star} + S_{11}Z_{01})(Z_{02}^{\star} + S_{22}Z_{02}) - S_{12}S_{21}Z_{01}Z_{02}}$
$S_{11} + \dfrac{(h_{11} - Z_{01}^{\star})(1 + h_{22}Z_{02}) - h_{12}h_{21}Z_{02}}{(h_{11} + Z_{01})(1 + h_{22}Z_{02}) - h_{12}h_{21}Z_{02}}$	$h_{11} = \dfrac{(Z_{01}^{\star} + S_{11}Z_{01})(Z_{02}^{\star} + S_{22}Z_{02}) + S_{12}S_{21}Z_{01}Z_{02}}{(1 - S_{11})(Z_{02}^{\star} + S_{22}Z_{02}) - S_{12}S_{21}Z_{02}}$
$S_{12} = \dfrac{2h_{12}\sqrt{R_{01}R_{02}}}{(h_{11} + Z_{01})(1 + h_{22}Z_{02}) - h_{12}h_{21}Z_{02}}$	$h_{12} = \dfrac{2S_{12}\sqrt{R_{01}R_{02}}}{(1 - S_{11})(Z_{02}^{\star} + S_{22}Z_{02}) - S_{12}S_{21}Z_{02}}$
$S_{21} = \dfrac{-2h_{21}\sqrt{R_{01}R_{02}}}{(h_{11} + Z_{01})(1 + h_{22}Z_{02}) - h_{12}h_{21}Z_{02}}$	$h_{21} = \dfrac{-2S_{21}\sqrt{R_{01}R_{02}}}{(1 - S_{11})(Z_{02}^{\star} + S_{22}Z_{02}) - S_{12}S_{21}Z_{02}}$
$S_{22} = \dfrac{(h_{11} + Z_{01})(h_{22} - Z_{02}^{\star}) - h_{12}h_{21}Z_{02}^{\star}}{(h_{11} + Z_{01})(h_{22}Z_{02}) - h_{12}h_{21}Z_{02}}$	$h_{22} = \dfrac{(1 - S_{11})(1 - S_{22})S - S_{12}S_{21}}{(1 - S_{11})(Z_{02}^{\star}) - S_{12}S_{21}Z_{02}}$

For the calculation of the cascade connection of two networks, it is desirable to convert the *S*-parameters to *T* parameters defined in the following way:

$$\begin{pmatrix} b_1 \\ a_1 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \cdot \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \tag{34.25}$$

The conversion between *S*-parameters and *T* parameters is given in the following:

$$T_{11} = S_{12} - \frac{S_{11}S_{22}}{S_{21}} \tag{34.26}$$

$$T_{12} = \frac{S_{11}}{S_{21}} \tag{34.27}$$

$$T_{21} = -\frac{S_{22}}{S_{21}} \tag{34.28}$$

$$T_{22} = \frac{1}{S_{21}} \tag{34.29}$$

It should be emphasized that different definitions of the *T* parameters exist in the literature [2,24–27,30]. Power gain, mismatch, insertion loss, etc., can be efficiently described with the help of scattering parameters [2,30].

34.5.3 Measurement of *S*-Parameters

Network analyzers are generally used for the measurement of *S*-parameters. A schematic configuration of a network analyzer is indicated in Figure 34.12. The network analyzer consists of two structures to separate the signals and a heterodyne receiver. According to the definitions, the measurement is performed in two steps as indicated in Figure 34.13. Different error models are employed for the calibration of the network analyzer. The most simple is the one-port error model, which consists of contributions due to directivity, source mismatch, and frequency response. A flow diagram is illustrated in Figure 34.14.

For the characterization of active and passive two-ports, a more sophisticated error model is required. The signal flow graphs of the full two-port model are drawn in Figure 34.15. To determine the *S*-parameters of the DUT, a de-embedding procedure is required [31–43]. The two-port

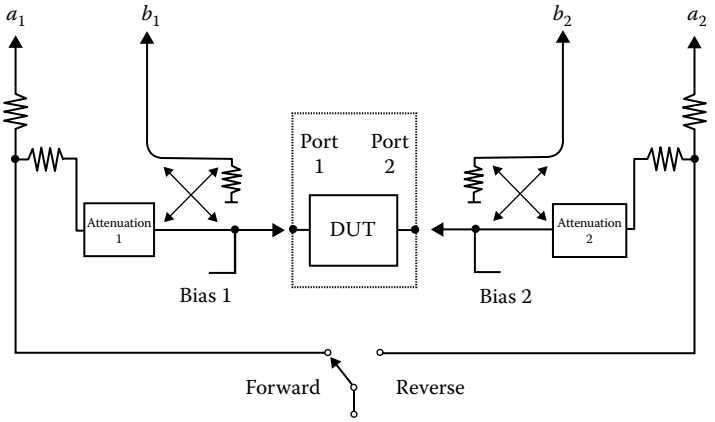


FIGURE 34.12 Schematic illustration of a network analyzer configuration.

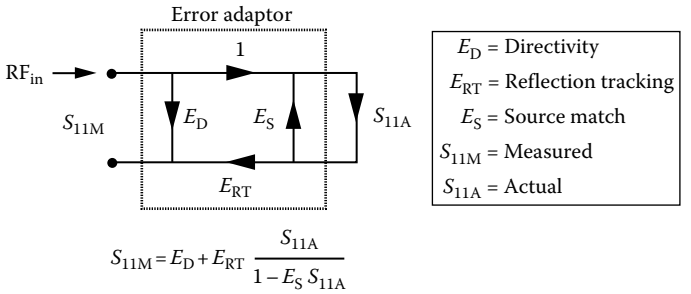


FIGURE 34.13 Flow diagram of the measurement procedure.

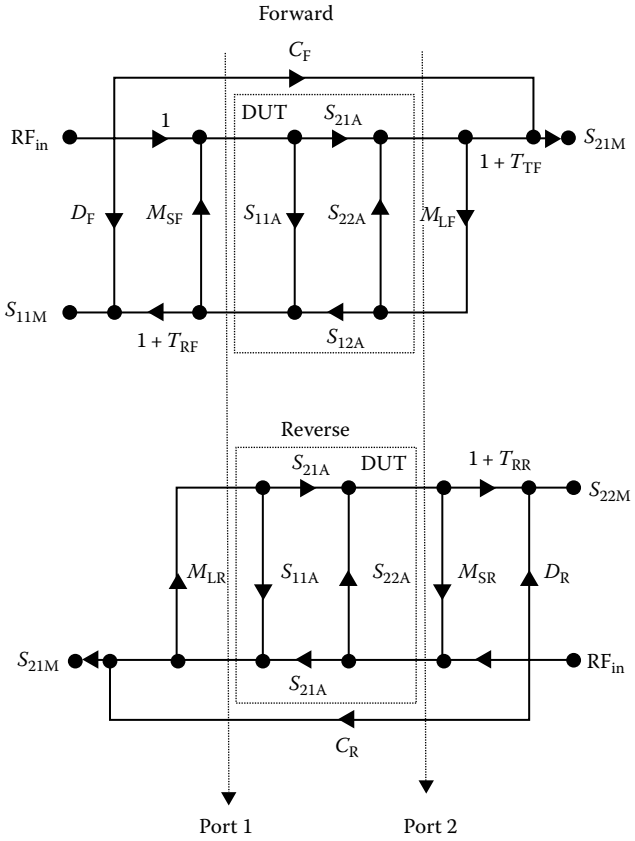
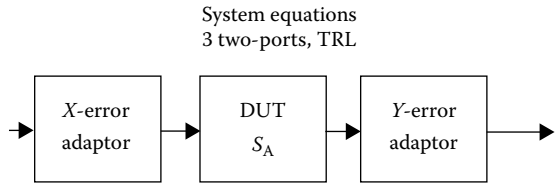


FIGURE 34.14 Flow diagram of the two-port error model.



$M = XAY$ measured DUT
$M_1 = XC_1Y$ measured two-port + cal.std1
$M_2 = XC_2Y$ measured two-port + cal.std1
$M_3 = XC_3Y$ measured two-port + cal.std1

FIGURE 34.15 De-embedding structure for the calibration of a network analyzer.

error model is then divided into two error adapters and the actual DUT, as depicted in Figure 34.16. The example shown makes reference to the so-called "through, reflect, line (TRL)" calibration procedure. This name abbreviates the three calibration standards utilized in this method: a *through* standard with zero length, a *reflecting* standard, and a *line* standard. This method cannot be applied to on-wafer measurements at low frequencies, due to the excessive line length required for a broadband measurement. Other error correction methods are summarized in Table 34.5 [44]. In addition to the measurements given in the table, a known reference impedance and port 1 to port 2 connections are required. Furthermore, at higher frequencies (above ≈ 15 GHz), a calibration measurement of the

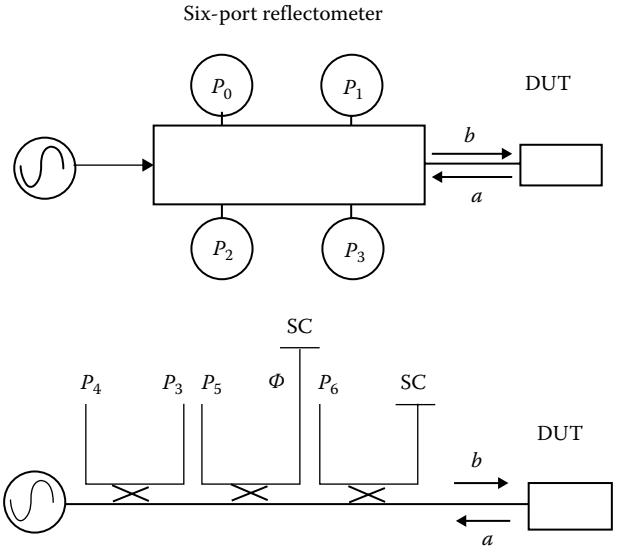


FIGURE 34.16 A network analyzer based on a six-port reflectometer and a possible realization using three couplers and four power detectors, two short circuits (SC), and one phase shifter (Φ).

TABLE 34.5 Summary of Different Calibration Methods

TOSL	Through standard (T) with known length; fulfills four conditions	Three known reflections (OSL) on port 1; fulfills three conditions	Three known reflections (OSL) on port 2; fulfills three conditions
TRL and LRL	Through or line standard (T) or (L) with known length; fulfills four conditions	Unknown equal reflection standard (R) on port 1 and port 2; fulfills one condition	Line (L) with known S_{11} and S_{22}; fulfills two conditions
TRM and LRM	Through or line standard (T) or (L) with known length; fulfills four conditions	Unknown equal reflection standard (R) on port 1 and port 2; fulfills one condition	Known match (M) on port 1 and port 2; fulfills two conditions
TXYZ and LXYZ	Through or line standard (T) or (L) with known length; fulfills four conditions	Three known reflection standards (XYZ) on port 1 or port 2; fulfills three conditions	—
UXYZ	Unknown line standard (U) with $S_{11} = S_{21}$; fulfills one condition	Three known reflection standards (XYZ) on port 1; fulfills three conditions	Three known reflection standards (XYZ) on port 2; fulfills three conditions

isolation should be performed. For this purpose, both ports are terminated by their characteristic impedances and a transmission measurement is performed. This measurement determines the values of C_F and C_R in the calibration flow diagram (see Figure 34.14).

Another possibility for performing vector network measurement is based on multiport reflectometers [45]. The advantage of such systems is the reduced complexity of the network analyzer receiver. A possible realization of a reflectometer structure is the so-called *six-port reflectometer*. The reflectometer consists of three couplers and four power sensors. No frequency conversion is required, which simplifies the test equipment.

34.5.4 Commercially Available Network Analyzers

Table 34.6 shows some of the current suppliers for network analysis. The frequency range is 10 Hz up to 800 GHz.

TABLE 34.6 Companies Supplying Network Analyzers for *S*-Parameter Measurements

	Frequency			
Company	Minimum	Maximum (GHz)	Method	Cal. Methods
Hewlett-Packard HP 8510	45 MHz	110	Heterodyne	SOLT, TLR, LRL, LRM, TRM
Wiltron	45 MHz	110	Heterodyne	SOLT, TLR, LRL, LRM, TRM
Rhode & Schwarz	10 Hz	4	Heterodyne	SOLT, TLR, LTL, LRM, TRM
AB Millimeterique	2 GHz	800	Heterodyne	TLR, LRL, proprietary

References

1. A. Fantom, *Radio Frequency and Microwave Power Measurements*, London, U.K.: Peter Peregrinus, 1990.
2. A. L. Lance, Microwave measurements, in *Handbook of Microwave Optical Components*, Vol. 1, Chapter 9, K. Chang (ed.), New York: John Wiley & Sons, 1989.
3. J. Minck, Fundamentals of RF and microwave power measurements, Hewlett-Packard Application Note 64-1A, 1997.
4. G. H. Bryant, *Principles of Microwave Measurements*, London, U.K.: Peter Peregrinus, 1988.
5. Power meter measurement errors, Tegam, http://www.microwaves101.com/encyclopedia/powermetermeasurements.cfm#errors, accessed on August 19, 2013.
6. D. Janik, H. Wolf, and R. Schneider, High-Tc edge bolometer for detecting guided millimeter waves, *IEEE Trans. Appl. Superconduct.*, 3, 2148–2151, 1993.
7. P. L. Richards, Bolometers for infrared and millimeter waves, *J. Appl. Phys.*, 76(1), 1–24, 1994.
8. D. M. Rowe, *Handbook of Thermoelectrics*, Boca Raton, FL: CRC Press, 1995.
9. P. Kopystynski, E. Obermayer, H. Delfs, W. Hohenester, and A. Löser, Silicon power sensor from dc to microwave, in *Micro Systems Technologies 90*, H. Reichl (ed.), Berlin, Germany: Springer, 1990, pp. 605–610.
10. G. C. M. Meijer and A. W. Herwaarden, *Thermal Sources*, Bristol, England: IOP Ltd., 1994.
11. T. Närhi, Nonlinearity characterisation of microwave detectors for radiometer applications, *Electron. Lett.*, 32, 224–225, 1996.
12. F. Wiedmann, B. Huyart, E. Bergeault, and L. Jallet, New structure for a six-port reflectometer in monolithic microwave integrated-circuit technology, *IEEE Trans. Instrum. Meas.*, 46(2), 527–530, 1997.
13. A. Dehé, V. Krozer, B. Chen, and H. L. Hartnagel, High-sensitivity microwave power sensor for GaAs-MMIC implementation, *Electron. Lett.*, 32(23), 2149–2150, 1996.
14. Circuits Multi Projects, *Information CMP—42*, Grenoble, France, December 1996.
15. 8 Hints for Making Better Spectrum Analyzer Measurements, Agilent Technologies, http://www.home.agilent.com/agilent/techSupport.jspx?pid=1000002155%3Aepsg%3Apro&pageMode=OV&cc=IN&lc=eng, accessed on August 19, 2013.
16. Spectrum Analyzer Basics, Agilent Technologies, http://www.home.agilent.com/agilent/techSupport.jspx?pid=1000002155%3Aepsg%3Apro&pageMode=OV&cc=IN&lc=eng, accessed on August 19, 2013.
17. A. Kiiss, Microwave instrumentation and measurements, in *Handbook of Microwave Technology*, Vol. 2, T. K. Ishii (ed.), San Diego, CA: Academic Press, 1995, p. 562.
18. H. Brand, Spectrum analyzers: Precision test instruments, *Microwave J.*, 37(3), 98, 1994.
19. Spectrum Analyzer Amplitude Accuaracy, Agilent Technologies, http://www.home.agilent.com/agilent/techSupport.jspx?pid=1000002155%3Aepsg%3Apro&pageMode=OV&cc=IN&lc=eng, accessed on August 19, 2013.
20. T. K. Ishii, Spectrum analysis: Amplitude and frequency modulation, Application Note 150-1, January 1989.
21. M. Sucher and J. Fox, *Handbook of Microwave Measurements*, Vol. II, New York: Polytechnic Press of the Polytechnic Institute of Brooklyn, 1963, Chapter VIII.

22. E. L. Ginzton, *Microwave Measurements*, New York: McGraw-Hill, 1957, Chapters 7 and 8.
23. F. E. Terman and J. M. Pettit, *Electronic Measurements*, New York: McGraw-Hill, 1952, Chapter 4–15.
24. C. A. Lee and G. C. Dalman, *Microwave Devices, Circuits and Their Interaction*, New York: John Wiley & Sons, 1994.
25. G. D. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*, New York: John Wiley & Sons, 1990.
26. J. Dobrowolski, *Microwave Circuit Analysis with Wave Variables*, Norwood, MA: Artech House, 1991.
27. G. Gonzalez, *Microwave Transistor Amplifiers*, Englewood Cliffs, NJ: Prentice Hall, 1984.
28. D. Weiner and G. Naditch, A scattering variable approach to the Volterro analysis of nonlinear systems, *IEEE Trans. Microwave Theory Tech*., MTT-24(7), 422–433, 1976.
29. D. A. Frickey, Conversion between S, Z, Y, H, ABSD, and T parameters which are valid for complex source and load impedances, *IEEE Trans. Microwave Theory Tech*., MTT-42, 205–211, 1994.
30. T. K. Ishii, *Handbook of Microwave Technology*, San Diego, CA: Academic Press, 1995.
31. R. A. Hackborn, An automatic network analyzer system, *Microwave J*., 11, 45–52, 1968.
32. S. Rehnmark, On the calibration process of automatic network analyzer systems, *IEEE Trans. Microwave Theory Tech*., MTT-22(4), 457–458, 1974.
33. J. Fitzpatrick, Error models of vector systems measurements, *Microwave J*., 21(5), 63–66, 1978.
34. Maximizing Accuracy in Noise Figure Measurements, Agilent Technologies, http://www.home.agilent.com/agilent/techSupport.jspx?pid=1000002155%3Aepsg%3Apro&pageMode=OV&cc=IN&lc=eng, accessed on August 19, 2013.
35. N. R. Franzen and R. A. Speciale, A new procedure for system calibration and error removal in automated S-parameter measurements, *5th European Microwave Conference*, Hamburg, Germany, 1975, pp. 69–73.
36. R. A. Soares and C. A. Hoer, A unified mathematical approach to two-port calibration techniques and some applications, *IEEE Trans. Microwave Theory Tech*., MTT-37(11), 1669–1674, 1989.
37. R. A. Soares, *GaAs MESFET Circuit Design*, Norwood, MA: Artech House, 1988.
38. Cascade Microtech, Microwave wafer probe calibration standards: HP8510 network analyzer input, Cascade Microtech Instruction Manual, Hillsboro, OR, 1990.
39. Optimizing Spectrum Analyzer Measurement Speed, Agilent Technologies, http://www.home.agilent.com/agilent/techSupport.jspx?pid=1000002155%3Aepsg%3Apro&pageMode=OV&cc=IN&lc=eng, accessed on August 19, 2013.
40. B. Donecker, Determining the measurement accuracy of the HP8510 microwave network analyzer, *RF & Microwave Measurement Symposium and Exhibition*, Hewlett-Packard, Santa Rosa, CA, March 1985.
41. Serious Spectrum Analyzer User Guide, Agilent Technologies, http://www.home.agilent.com/agilent/techSupport.jspx?pid=1000002155%3Aepsg%3Apro&pageMode=OV&cc=IN&lc=eng, accessed on August 19, 2013.
42. Maximizing Accuracy in Noise Figure Measurement, Agilent Technologies, http://www.home.agilent.com/agilent/techSupport.jspx?pid=1000002155%3Aepsg%3Apro&pageMode=OV&cc=IN&lc=eng, accessed on August 19, 2013.
43. UsingNetwork Analyzer Time-Domain Analysis to Verify and Troubleshoot Complex Components, Agilent Technologies, http://www.home.agilent.com/agilent/facet.jspx?t=79842.g.1&cc=IN&lc=eng&sm=g, accessed on August 19, 2013.
44. H. J. Eul and B. Schiek, Thru-match-reflect: One result of a rigorous theory for deembedding and network analyzer calibration, *18th European Microwave Conference*, Stockholm, Sweden, 1988, pp. 909–914.
45. G. F. Engen and C. A. Hoer, Thru-reflect-line: An improved technique for calibrating the dual 6-port automatic network analyzer, *IEEE Trans. Microwave Theory Tech*., MTT-27(12), 993–998, 1979.

35

Q Factor Measurement

Albert D. Helfrick
Embry–Riddle Aeronautical University

Q factor is a method of characterizing the rate of dissipation of energy from an oscillating system [1]. *Q* is defined as 2π times the energy stored in a resonant system divided by the energy dissipated per cycle. The term system used in this context refers to any type of resonance: mechanical, electric, nuclear, etc. For the purposes of this *handbook*, *Q* will be that of an electric circuit. Also, for the discussion of *Q*, very low values of *Q*, typically less than 10, will not be considered as these low values of *Q* produce highly damped oscillations that are more exponential than oscillatory and the concept of *Q* does not fit.

A common interpretation of *Q* is quality factor that explains the use of the letter, *Q*, but this is misleading. A component that has a high *Q* is not always beneficial and may not be of high quality. In many circuits, a component requires a specific value of *Q* rather than "higher is better." In other cases, a high *Q* is an advantage.

The *Q* factors encountered in common circuit components range from a low of about 50 for many inductors to nearly 1 million found in high-quality quartz crystals. *Q* can be applied to a resonant circuit or to capacitors and inductors. When *Q* is applied to a component, such as an inductor, the *Q* would be that obtained if the inductor were used in a resonant circuit with a capacitor that dissipates no energy. In this case, the value of *Q* depends on the frequency.

For most *LC* resonant circuits, the losses in the inductor dominate and the *Q* of the inductor is, essentially, the *Q* of the circuit. It is easy to make very low-loss capacitors even in the ultra high frequency (UHF) region. On the other hand, varactor diodes have considerably more loss than fixed capacitors, and a varactor can play a more significant role in setting the circuit *Q*.

35.1 Basic Calculation of *Q*

Figure 35.1 shows a simple resonant circuit. The capacitor can store energy in the electric field and the inductor in the magnetic field. The circuit oscillates with the energy transferring between the two elements. For the ideal elements shown, this continues forever. Since the definition of *Q* has the energy lost per cycle in the denominator—which is zero—the result is an infinite *Q*.

In practice, circuit elements are not perfect and the energy initially contained within this circuit would be lost by the circuit and the oscillations would decrease in amplitude as the energy diminished.

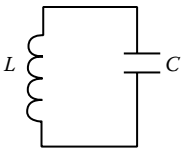


FIGURE 35.1 A simple resonant circuit with no loss elements.

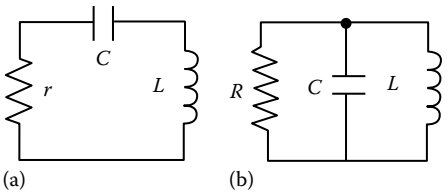


FIGURE 35.2 Resonant circuits: (a) A simple series resonant circuit with the equivalent resistance and (b) a parallel resonant circuit with a parallel equivalent resistance. For the same *Q* circuit, the values of the resistors are not the same.

Energy loss in a circuit is represented by that in a resistor, which can be included in one of two ways. The first way is shown in Figure 35.2a, where the resistor is in series with the capacitor and inductor. A second representation is a parallel resistor as shown in Figure 35.2b.

To derive the relationship between the circuit element values and the *Q* of the circuit, either the current or voltage of the circuit can be used in the equations. Current is the obvious choice for a series circuit, while voltage is the common thread for the parallel circuit. Assume that the amplitude of the current through the series circuit of Figure 35.2a is given by

$$i(t) = I(t)\cos(2\pi f_0 t) \tag{35.1}$$

where f_0 = resonant frequency

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \tag{35.2}$$

and $I(t)$ is the amplitude, which is decreasing in some unspecified fashion. The circuit's peak current occurs when the cosine function is equal to 1 and all of the energy is contained in the inductor and is equal to $(1/2)LI^2(t)$.

Assume that a relatively high *Q* is present in this circuit and that $I(t)$ changes by only a slight amount during the time of one cycle. During this cycle, the peak current is $I(t)$, and the rms value of the current is $(0.707)I(t)$. Therefore, the energy dissipated in one cycle is $(0.5)I^2(t)r/f_0$. Substituting these values in the definition of *Q* yields

$$Q = 2\pi\frac{(1/2)(LI^2(t))}{(1/2)(rI^2(t)/f_0)} = \frac{2\pi f_0 L}{r} = \frac{X_L}{r} \tag{35.3}$$

where X_L is the reactance of the inductor. The same procedure can be used with the parallel resonant circuit of Figure 35.2b using voltage equations to obtain the relationship between a parallel resistance and *Q*, which is

$$Q = \frac{R}{X_L} \tag{35.4}$$

It is very important to understand the nature of the circuit resistance in Figure 35.2a and b. This resistance represents all of the losses in the resonant circuit. These losses are from a variety of sources, such as

the resistance of the wire to make an inductor or the leakage current of a capacitor. It can also represent the deliberate addition of resistors to set the *Q* of the circuit to a specific value. The resistance shown in the circuits of Figure 35.2 represents the equivalent resistance of all of the energy losses. This resistance cannot be measured with an ohmmeter as the value of the equivalent resistor is a function of frequency and other variables such as signal level. Some of the loss in a resonant circuit is due to radiation, which is a function of frequency. The resistance of conductors is mostly due to skin effect, which increases with increasing frequency. The losses in the ferromagnetic materials used for making some inductors are nonlinear; thus, the equivalent resistance is not only a function of frequency but of signal level.

Most resonant circuits are not stand-alone circuits as shown in Figure 35.2 but are a part of other circuits where there are sources and loads. These additional resistances further remove energy from the resonant circuit. The *Q* of a resonant circuit when there are sources and loads is called the *loaded Q*. In most applications of resonant circuits, the *Q* of the resonance is set by the external loads rather than the capacitor and inductor that form the resonance.

35.2 Bandwidth and *Q*

The *bandwidth* of a resonant circuit is a measure of how well a resonant circuit responds to driving signals of a frequency near the resonant frequency and is a function of *Q*. The relationship between the 3 dB bandwidth and *Q* will be derived.

Applying a driving signal to a resonant circuit can overcome the losses of the circuit and cause the resonant circuit to oscillate indefinitely. As an example of this, consider the voltage generator in the series resonant circuit shown in Figure 35.3.

When the frequency of the voltage source is equal to the resonant frequency of the circuit, the equivalent impedance of the series resonant circuit is the resistance of the circuit and the current in the circuit, simply *E*/*r*.

At frequencies higher or lower than the resonant frequency, the impedance of the circuit is greater because the net reactance is not zero and the circuit current will be less than at resonance.

At what frequency will the circuit current be 3 dB less than at resonance? This frequency is where the circuit impedance is 1.414 that of the impedance at resonance. This is the frequency where the reactive part of the impedance is equal to the real part. This situation occurs at two frequencies. Below the resonant frequency, the net reactance is capacitive and is equal to *r*, while at a second frequency above the resonant frequency, the reactance is inductive and equal to *r*. This can be represented by two equations for the two frequencies:

$$\text{For } f_1 > f_0, \quad |X_L - X_C| = 2\pi f_1 L - \frac{1}{2\pi f_1 C} = \frac{(f_1 - f_0)^2 - 1}{2\pi f_1 C} = r$$

$$\left(\frac{f_1}{f_0}\right)^2 - \frac{1}{Q}\left(\frac{f_1}{f_0}\right) - 1 = 0 \quad \left(\frac{f_1}{f_0}\right) = \frac{1}{2Q} \pm \sqrt{\frac{1}{4Q^2}} + 1 \tag{35.5}$$

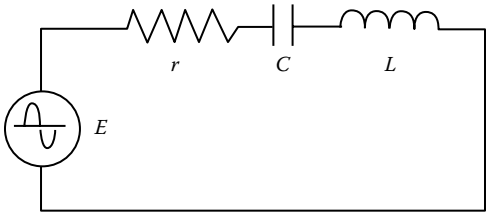


FIGURE 35.3 A series resonant circuit showing a driving source.

$$\text{For } f_2 < f_0, \quad |X_L - X_C| = \frac{1}{2\pi f_2 C} - 2\pi f_2 L = \frac{-(f_1 - f_0)^2 - 1}{2\pi f_2 C} = r$$

$$\left(\frac{f_2}{f_0}\right)^2 + \frac{1}{Q}\left(\frac{f_2}{f_0}\right) - 1 = 0 \quad \left(\frac{f_2}{f_0}\right) = \frac{1}{2Q} \pm \sqrt{\frac{1}{4Q^2}} + 1 \tag{35.6}$$

$$(f_1 - f_2) = \frac{f_0}{Q} = \text{bandwidth}$$

35.2.1 Measuring *Q*

There are a number of methods of measuring *Q* using a variety of bridges, several of which are described in [2]. One method of measuring a capacitive or inductive *Q* is to place the component in a resonant circuit. When the *Q* to be measured of a device that is, in itself, a resonant circuit such as quartz crystal, similar techniques are used except the device's own resonance is used for the measurement. Circuit voltages or currents are measured at the resonance frequency and the *Q* is determined.

35.3 *Q*-Meter

One simple and very popular method of measuring *Q* is with a device called, appropriately, the *Q*-meter. Consider the resonant circuit in Figure 35.4 for measuring the *Q* of inductors. This circuit has a very low-loss capacitor of known value and a constant voltage source.

The usual components measured by the *Q*-meter are inductors. It was previously mentioned that inductors are the weak link in resonant circuits, and the *Q* of a circuit is usually set by the inductor. The *Q*-meter can measure capacitance and capacitor *Q*. In this theoretical circuit, the circuit resistance is the equivalent series resistance (ESR) of the inductor under test. This is due to the fact the variable capacitor is assumed to be lossless, the generator has zero resistance, and the voltmeter does not appreciably load the circuit. In a real circuit, it is not possible to achieve this situation, but the instrument is designed to approach this goal.

To measure the *Q* of an inductor using the *Q*-meter, the generator is set to the desired frequency, while the variable capacitor tunes the circuit to resonance as indicated by the peak reading of the voltmeter. At resonance, the impedance of the circuit is simply the ESR of the inductor. This sets the current of the circuit as

$$I = \frac{E}{R_X} \tag{35.7}$$

where

E is the generator voltage

R_X is the equivalent resistance of the inductor

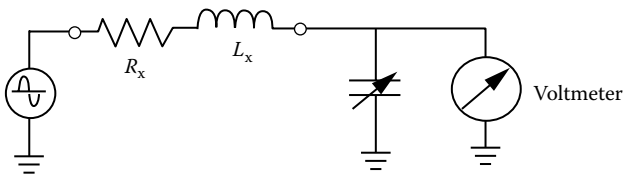


FIGURE 35.4 The basic circuit of a *Q*-meter showing the signal source, the inductor under test, and the voltmeter.

Because the circuit is at resonance, the voltages of the two reactances are equal and of opposite phase. Those voltages are

$$V = IX_C \quad \text{or} \quad V = IX_L \tag{35.8}$$

where

X_C is the capacitive reactance
X_L is the inductive reactance, which are numerically equal at resonance

Substituting the relationship of the circuit current, the result is

$$V = \frac{EX_L}{R_X} = \frac{EX_C}{R_X} = EQ \tag{35.9}$$

Therefore, the voltage across the reactances is equal to Q times the applied voltage. If, as an example, the voltage source were 1 V, the voltmeter would read Q directly. Q values of several hundred are common, and therefore, voltages of several hundred volts could be measured. Modern circuits do not typically encounter voltages of this magnitude, and many components cannot withstand this potential. Therefore, most Q-meters use a much smaller source voltage, typically 20 mV.

If the frequency of the source and the circuit capacitance are known, it is possible to calculate the inductance of the unknown.

35.4 Other *Q* Measuring Techniques

There are very few Q-meters being manufactured today, although there are thousands of old Q-meters still in use. Because the Q-meter was the only accepted method of measuring Q for so many years, it will take decades before alternative methodologies overtake the Q-meter.

Measuring Q without a Q-meter involves a variety of *RLC* measuring instruments that measure the vector impedance and calculate Q. The calculated Q value is not as valid as that determined with a Q-meter unless the *RLC* meter is capable of measuring Q at the desired frequency. Only the more sophisticated, and expensive, *RLC* meters allow the use of any test frequency. Despite the new sophisticated *RLC* measuring instruments, there is an adapter for one model *RLC* instrument that allows the classic Q-meter-style measurement to be made.

Because vector impedance measurement is covered elsewhere in this handbook, the remainder of this section will be devoted to measurements using the Q-meter.

35.5 Measuring Parameters Other than *Q*

In addition to Q, the Q-meter can be used to measure inductance, the Q or dissipation factor of a capacitor, and the distributed capacitance, C_d, of an inductor.

If an inductor with capacitance C_d is placed in the Q-meter circuit, the total circuit capacitance includes both the Q-meter's capacitance and the additional C_d. Therefore, when resonance is achieved, the actual resonating capacitance is more than what is indicated on the Q-meter capacitor's dial. If C_d is not included in the calculation of inductance, the resulting value would be too large.

In many applications, the actual inductance is not the important parameter to be measured by the Q-meter. The actual parameter being determined is "what capacitance is required to resonate the particular inductor at a specific frequency," regardless of C_d.

In other applications, such as inductors that are to be used in wide-range oscillators, where C_d will limit the tuning range, C_d is an important parameter.

The *Q*-meter can be used to determine C_d, which will also allow for an accurate calculation of inductance. Determining C_d is a matter of resonating the inductor under test at more than one frequency.

To understand how the two-frequency measurement will allow the determination of C_d, assume that an inductor is resonated at a frequency f_1. The relationship between the applied frequency, f_1, and the capacitor of the *Q*-meter to obtain resonance is

$$f_1 = \frac{1}{2\pi\sqrt{L(C_1 + C_d)}} \tag{35.10}$$

where C_1 is the capacitance set on the *Q*-meter.

Resonating the same inductor at a second, higher, frequency, f_2, requires a *Q*-meter capacitance of C_2, such that

$$f_2 = \frac{1}{2\pi\sqrt{L(C_2 + C_d)}} \tag{35.11}$$

This implies that C_2 is a smaller capacitance than C_1. Using these two equations and solving for C_d, the following result is obtained:

$$C_d = \frac{C_2 f_2^2 - C_1 f_1^2}{f_1^2 - f_2^2} \tag{35.12}$$

A convenient relationship between f_1 and f_2 is to set $f_2 = 1.414\, f_1$. With frequencies thus related, the distributed capacitance is

$$C_d = C_1 - 2C_2 \tag{35.13}$$

C_d causes errors in the measurement of *Q* because of current through the C_d. The *Q* measured by the *Q*-meter is called the "effective *Q*." Since large inductors with significant C_d are no longer in common use because of the use of active filters, the distinction between effective *Q* and real *Q* is seldom considered. For additional information about effective *Q* and distributed capacitance, see [3].

To measure capacitors on the *Q*-meter, a relatively high-*Q* inductor is connected to the inductance terminals on the *Q*-meter and resonated with *Q*-meter's internal calibrated capacitor. The capacitor to be measured is connected to the instrument's capacitor terminals, which are in parallel with the internal variable capacitor. This increases the circuit capacitance. The *Q*-meter variable capacitor is adjusted to regain resonance, which requires that the capacitance be reduced by an amount equal to the unknown capacitor.

The *Q* of the capacitor can be measured. The addition of the capacitor reduces the circuit *Q* because of the additional loss introduced by the capacitor. In the description of the *Q*-meter, the *Q*-meter's variable capacitor is assumed to have no loss.

Measuring the *Q* of a capacitor using the *Q*-meter is seldom done. This is because most capacitors have very high-*Q* values, usually significantly higher than the test inductor and the *Q*-meter's internal capacitor. Therefore, the addition of a high-*Q* unknown capacitor makes an imperceptible difference in the overall circuit *Q* making it very difficult to obtain an accurate *Q* measurement. There are special cases, such as measuring the *Q* of a capacitive transducer or a varactor diode, where low *Q* values are encountered and the *Q*-meter can be used successfully.

For best results, the *Q* of the inductor must be considerably greater than that of the unknown capacitor, and the unknown capacitance must be considerably greater than the internal capacitance. If these criteria are met, the *Q* of the unknown capacitor can be read from the *Q*-meter. If these criteria are compromised, corrections can be made but the equations become complex and the accuracy degrades.

Most *Q*-meters provide measurement ranges to about 500 or to 1000. This is sufficient for measuring inductors, which was the main purpose of the *Q*-meter. For measuring high-*Q* devices such as ceramic resonators with *Q* values greater than 1000 or for quartz resonators with *Q* values extending into the hundreds of thousands, the *Q*-meter technique is insufficient. A high-*Q* circuit implies very little energy loss, and the energy that must be removed to make a measurement must be very small if *Q* is to be measured accurately.

35.6 Measuring High-*Q* Capacitors

The original intent of the *Q*-meter was primarily for measuring the *Q* of inductors since it was the *Q* of the inductor that set the bandwidth of most tuned circuits at radio frequencies (RFs). The *Q* of a capacitor was usually much higher than the inductor in a resonant circuit and only inductors were typically measured. The *Q* of a capacitor has several other names such as dissipation factor, loss tangent, or ESR. The resistive part of a capacitor's impedance has become increasingly more important for power supplies and circuits where capacitors could experience significant currents and a capacitor with a high dissipation factor or low *Q* wastes energy and could be destroyed by internal heating.

The technique described using a *Q*-meter with a calibrated inductor for making capacitor measurements is very limited by the *Q* of the resonating inductor and particularly difficult where the capacitor to be measured is a high capacitance value and high *Q*.

Modern circuits do not rely as much on high-*Q* inductors as in the past. Signal filtering once done at intermediate and audio frequencies, IF and AF, is done routinely with digital signal processing. RF filtering is still required, but new applications extend into UHF and microwave frequencies where conventional wire-wound coils are not used. The need for wire-wound inductors and their measurements is diminishing.

As the need for high-*Q* inductors diminishes, high-*Q* capacitors are becoming in more demand as the use of switching power supplies increases. When describing capacitors for power supplies, ESR is the preferred term. But ESR = X_c/Q.

Modern switching power supplies operate at RF frequencies and therefore use components that would be suitable for RF circuits and therefore require RF style measurements.

Another factor in the increasing need for high-*Q* capacitors is the use of solid-state amplifiers for RF. High-powered vacuum tube amplifiers used resonant circuits that were high voltage and lower current when compared to the same power level for transistor amplifiers. These higher currents increase the heating of a capacitor due to the ESR. Capacitors used for filters usually required a high *Q* because of the large circulating currents, but in a transistor power amplifier, high-*Q* capacitors are also required for coupling and bypassing, applications not usually associated with high *Q*.

With the demise of the *Q*-meter, vector network analyzers were used to measure capacitor and inductor parameters. It is difficult to obtain accurate *Q* measurement for ultrahigh-*Q* capacitors using a vector network analyzer because of the large disparity between the real and imaginary part of the capacitor's impedance. This requires the network analyzer to accurately separate the real and imaginary parts and to have the ability to resolve small fractions of a dB.

Modern ultralow ESR chip capacitors can have *Q* factors well into the thousands. The classic *Q*-meter, when measuring inductors, could handle *Q* factors from less than 100 to perhaps as high as 1000 with accuracy. A very important factor when using resonance as a method of measuring *Q* of either inductors or capacitors is the *Q* of the resonating element. When measuring the *Q* of inductors in the classic *Q*-meter, the *Q* of the resonating capacitor, using an air dielectric, had sufficiently high *Q* to measure moderately high Q_s for inductors. Using resonance techniques for measuring the *Q* of ultralow ESR capacitors demands an inductor of extremely high *Q*. This need can be met with a section of coaxial transmission line.

A typical transmission line section used for measuring *Q* is on the order of a meter or two in length. One end is short-circuited, while the other is open circuited. To ensure low loss, air is used as the

dielectric, and the inner, conducting, surfaces are gold or silver plated to decrease skin effect resistance. At the open-circuited end, a sleeve extends the outer conductor of the transmission line, called a choke section, to minimize radiation from the end of the transmission line. At the shorted end of the transmission line, there is the ability to insert a leadless component usually involving a jackscrew-like clamping device.

The transmission line section is more than just an inductor but a resonant circuit. The measurement techniques call for the characterization of the transmission line in terms of resonant frequency and bandwidth. The transmission line will resonate at odd multiples of one-quarter of its wavelength. Since the dielectric is air, the resonant frequencies will be $(n/4)(3E8/d)$ where d is the length of the line and n is an odd integer.

To determine resonance and the bandwidth of the transmission line section, a signal generator is loosely coupled into the transmission line section with a loop, and the RF energy is sampled, also loosely coupled with either a capacitive probe or another loop. It is important that the coupling be very loose to prevent energy from being lost to the signal source or the sensor. Another important reason to limit the power coupled to the measurement is that very high Q_s are involved, which implies high circulating currents. The heating of the capacitor under test can skew the measurements for some capacitors. The arrangement is similar to a cavity filter but with a very high amount of insertion loss. Insertion losses of 40–60 dB are not uncommon. Typical source/detector combinations are a signal generator and a sensitive RF millivoltmeter or a spectrum analyzer. Another excellent source/detector pair is a spectrum analyzer with a tracking generator or a scalar, or vector, network analyzer. The limitations of the vector network analyzer are removed from the measurement technique as there is no need to separate the real and imaginary parts in the described technique.

The transmission line is resonated and the resonant frequency and the bandwidth of the line are determined. It is not necessary to use only the 3 dB bandwidth and the 6 dB or other values may be used. The bandwidth is used to calculate the Q of the resonator, and the calculation must reflect the measurement points used.

The capacitor under test is inserted in the shorted end of the transmission line and the resonant frequency and bandwidth are measured again. The addition of the capacitor will change the resonant frequency of the transmission line because of the capacitance being placed in series with the center conductor of the transmission line. The bandwidth of the transmission line will be wider because of the capacitor's ESR.

Using the before and after values, the value and Q of the capacitor under test can be determined. This technique is the basis behind ANSI/EIA RS 483, Effective Series Resistance (ESR) and Capacitance of Multilayer Ceramic Capacitors at High Frequencies, standard.

The Q-meter relied on causing resonance in a circuit to determine the capacitance or inductance and the attendant Q. The transmission line technique also relies on resonance. The classic Q-meter may be a thing of the past, but the techniques that defined the Q-meter live on.

Defining Terms

Bandwidth: A measurement of the amount of frequency spectrum occupied by a signal or the equivalent spectrum covered by a circuit that passes a finite band of frequencies. There are a number of methods of defining bandwidth depending on the nature of the spectrum. Relative to resonant circuits, the bandwidth is measured between the –3 dB points of the passband.

Distributed capacitance: The amount of capacitance added to an inductor typically from the capacitance due to adjacent wires in a solenoid-type inductor. The distributed capacitance is given as a single capacitance figure for a specific inductor and can be defined as the equivalent capacitance across the entire coil. This would also allow the inductor to have a self-resonant frequency where the inductor resonates with the distributed capacitance with no external capacitance.

Effective inductance: Due to distributed capacitance, less capacitance than that calculated from an inductance value is required to resonate a circuit. If the actual capacitance required to resonate a circuit is used to calculate an inductance value, the resulting inductance value will be higher than the theoretical inductance value. This higher value is called the "effective inductance." The actual inductor cannot be considered as a pure inductance of a value equal to the effective inductance because the real inductor has a resonant frequency that a pure inductance does not.

***Q*:** A measurement of the rate of energy loss in a resonant circuit.

***Q*-meter:** An instrument for measuring *Q* factor by resonating the circuit and measuring the voltage across the reactances.

RF: radio frequencies: Any frequency above about 10 kHz extending to the upper limits of the radio spectrum. Generally, radio frequencies are distinguished from power generation frequencies such as 50, 60, or 400 Hz and above the range of human hearing or 20 kHz.

References

1. D. Kajfez, *Q Factor Measurements Using MATLAB*, Artech House, Boston, MA, 2011.
2. A. Helfrick and W. Cooper, *Modern Instrumentation and Measurement Techniques*, Englewood Cliffs, NJ: Prentice Hall, 1990.
3. M. E. Van Valkenburg and W. M. Middleton, *Reference Data for Engineers: Radio Electronics, Computer and Communications*, 9th edn., Woburn, MA: Newnes, 2002.

36

EMI and EMC Test Methods

Jeffrey P. Mills
Illinois Institute of Technology

36.1 Introduction

Electric and magnetic fields must be measured for a variety of reasons. A radio or TV broadcast station is licensed to provide reliable coverage over a specified geographic area, and any properly operating receiver must pick up the signal and properly respond to it. This can be assured only if the broadcast signal is of a guaranteed minimum strength. Also, the signal must not be so strong that it interferes with a distant station sharing the same frequency. The broadcast field must be measured over its geographic area of coverage to be sure that it satisfies both criteria.

Many electric devices unintentionally radiate electromagnetic fields. Examples include

- Oscillators in superheterodyne radio or TV receivers
- Digital logic circuits
- Switching contacts, particularly if unsuppressed
- Automotive ignition systems

Stray fields (*emissions*) from these devices can interfere with other devices or even with the radiating device itself. This process is known as *electromagnetic interference*, commonly abbreviated EMI. The interference between two devices is known as *intersystem EMI*, whereas if a device interferes with itself, it is *intrasystem EMI*. Intrasystem EMI is usually easy to spot because the device itself does not operate correctly. Intersystem EMI is usually more difficult to isolate. Its result might be a simple annoyance, such as noise on a radio and TV receiver caused by an electric vacuum cleaner or a power drill. It could, however, be much more serious; a portable radio receiver might affect aircraft navigation or critical communications.

It is also possible for a device to be susceptible to fields intentionally generated by a licensed transmitter such as a broadcast or mobile-radio transmitter. Examples include

- Public-address systems
- Music (high-fidelity) systems
- Telephone lines and instruments
- Digital logic circuits

Again, the result may be only an annoyance, or it could be much more serious; aircraft control surfaces have been observed to move uncontrollably due to strong electromagnetic fields. Since the fields themselves cannot be eliminated in these cases, the devices must be made immune to electromagnetic fields.

In the previously mentioned cases, the interference is usually through electric and/or magnetic fields in space, so the process is known as *radiated coupling*. Another coupling path exists if two devices share the same power source. One device may generate undesired high-frequency voltages on its power leads, which then appear on the power leads of the other. The second device may then malfunction because of this high-frequency voltage. This is known as *conducted coupling*. So we must consider both radiated and conducted noise.

It is not practical to eliminate all interfering fields completely, so a compromise must be reached. A stray field will not cause EMI if it is very weak compared with the desired field, which might be the field of a broadcast signal. The permissible strength of the stray field depends on the strength of the desired field; the stronger the desired field, the more stray field can be tolerated. It also depends on the device that is being interfered with (the *victim*); some receivers can reject undesired signals better than others. Since there are many combinations of interference sources and victims, a worst-case scenario is sought that will protect most real-life situations. This occurs where the weakest legal radio or TV signal (in its licensed area of coverage) is received by the poorest available receiver.

The maximum stray field strength that causes no EMI for this worst-case scenario is incorporated into government regulations. The field actually radiated by every device must then be measured to be sure that it does not exceed this level at the nearest practical distance from it, usually 10 or 30 m. To specify and measure these field strengths accurately, the nature of electric and magnetic fields must be understood.

Unlike most electrical engineering topics, EMI control is not very precise because of the complexity of practical hardware. It is virtually impossible to predict interference more precisely than within a factor of three, and usually the margin of error is even worse. Measurements can vary significantly between two supposedly identical samples, due to slight variations in physical dimensions. If one measures the EMI resulting from two different designs, the design that exhibits less EMI is probably better, but not always. An engineer can often judge if an EMI problem exists, but one must never rely on the accuracy normally expected in other branches of electrical engineering.

36.2 Nature of Electric and Magnetic Fields

An electric field is generated by a distribution of electric charge. If the distribution changes with time, then so will the electric field. A magnetic field may be generated by a permanent magnet or by an electric current. If the permanent magnet or the current path moves or if the current magnitude varies with time, the magnetic field will vary with time. A time-varying electric field creates a magnetic field, and conversely.

Electric fields, designated E, are normally expressed in volts per meter (V/m). Magnetic fields are designated H and expressed in amperes per meter (A/m). More often, magnetic fields are perceived as magnetic flux densities, which are designated B and expressed in webers per square meter (W/m^2), also known as *teslas* (T). A non-SI unit, sometimes found in older literature, is the *gauss*, equal to 10^{-4} T. Of course, any unit may be preceded by a scaling prefix such as micro or pico. In free space, B is equal to $\mu_0 H$, where μ_0 is equal to 0.4π (approximately 1.257) μT m/A (equivalent to μH/m).

Near a time-varying electric field source such as a charge distribution, the magnetic field is relatively weak, but it becomes stronger when observed from farther away. At a great enough distance, the ratio of E to H approaches $\sqrt{\mu/\varepsilon}$, which in free space is equal to 120π (approximately 377) Ω. For a sinusoidal function of time with a frequency f, this occurs at any distance that is large compared with $\lambda/2\pi$ (approximately $\lambda/6$). Here, λ is the wavelength corresponding to f, equal to $3 \times 10^8/f$ m if f is specified in hertz. Distances much greater than $\lambda/2\pi$ are considered to be in the *far-field region*; nearer distances are in the *near-field region*. For a nonsinusoidal function of time, each Fourier frequency component must be considered separately, and the far-field region begins closer to the source for its higher-frequency components.

Near a time-varying magnetic field source such as a current loop, the electric field is weak, becoming stronger when observed from a greater distance. At distances that are large compared with $\lambda/2\pi$ (the far-field region), the ratio of E to H again approaches $120\pi\ \Omega$.

Since $H = \sqrt{\varepsilon/\mu}E$ and $B = \mu H = \sqrt{\mu\varepsilon}E$ in the far-field region for either type of source, only E or B must be measured, and the other can easily be calculated from it. In free space, $\sqrt{\mu/\varepsilon} \approx 10^{-8}/3\,\mathrm{T\cdot m/V}$ (equivalent to s/m), so if E is expressed in volts per meter, $B \approx 3.33E$ nT. By choice of a suitable antenna, either field can be measured. Far-field strengths are normally specified in terms of the E field, no matter whether the E or B field is measured.

Alternatively, the far-field strength may be specified in terms of *power density*, expressed in watts per square meter. This denotes the amount of radiated power passing through each square meter of a surface perpendicular to the direction away from the source. The *peak* power density P is equal to EH, and, for a sinusoidal source, the *average* power density is half this value. For a nonsinusoidal source, each frequency component must be considered separately, and the total average power is the sum of the average powers for all frequencies. Since $H = \sqrt{\varepsilon/\mu}E$, it follows that $P = E^2/377\ \Omega$.

In regions other than the far field, the ratio of E to H varies greatly, approaching infinity for an electric field source or zero for a magnetic field source. A source may generate both electric and magnetic fields; for example, a charge moving between two electrodes causes a current to flow between them. Then the ratio of E to H may be any value at all. Therefore, at distances less than $\lambda/2\pi$ from a field source, both the E and B fields must be measured separately.

In the far-field region, both the electric and magnetic fields are perpendicular to the direction that an electromagnetic wave is propagating, and they are also perpendicular to each other. This still usually allows the fields to be oriented at many different angles with respect to the surface of the Earth. The direction of the electric field is called the *polarization* of the wave, which may be vertical, horizontal, or somewhere between. Or the wave may be *elliptically polarized*, which results from two waves that are not exactly in phase, one polarized vertically and the other horizontally. If the waves are equal in magnitude and exactly 90° out of phase, the wave is *circularly* polarized. To account for all these cases, all fields must be checked separately for vertically and horizontally polarized waves.

36.3 Measurement Antennas

Most electronic components and instruments are designed to respond to voltages or currents, not fields. To measure field strength, it is necessary to convert its effect to a voltage or a current. This is achieved by an antenna. Although many antennas are simple conductor shapes, they must be analyzed carefully if accurate quantitative measurements are desired.

A straight conductor immersed in a time-varying electric field will develop a current in it. If the conductor material is linear (the usual case), the current will be proportional to the applied electric field, so their ratio will be constant. This ratio, however, depends greatly on the geometric dimensions of the conductor and the frequency of the electric field. It must be known to calibrate the antenna.

Similarly, a closed conductive loop immersed in a time-varying magnetic field will develop a current in it. Again, if the conductor is linear, the ratio of the current to the magnetic field strength is constant but depends on the dimensions of the loop and the frequency of the magnetic field.

The easiest way to calibrate an antenna is to immerse it in a known electric or magnetic field and measure the current or voltage at the antenna terminals. The principal problem is generating the known field. To find its strength, one must use a "standard" antenna for which the current-to-field ratio can be calculated.

To calculate the required ratio, Maxwell's equations must be solved subject to the boundary conditions of the antenna conductor. For most antennas, an exact closed-form solution is impossible. However, for a sinusoidally varying field encountering a straight cylindrical conductor called a *dipole antenna*, such a solution is possible, though difficult [1]. Once the solution is obtained, the required ratio becomes a simple expression if the antenna is *resonant* or *tuned*. This occurs for a precise length that

is slightly less than one-half the wavelength, λ, of the time-varying field. Obviously, the antenna will be resonant at only one frequency, so the ratio will be valid only for a field varying sinusoidally at that frequency. For nonsinusoidal fields, each Fourier frequency component must be measured separately, and the antenna length must be changed as different frequencies are measured. To simplify changing its length, two telescoping rods, mounted end to end, are normally used to make the dipole antenna. The measuring instrument is connected between these two rods via a transmission line.

For a given frequency, at any point on the antenna, there is a certain current I flowing in it, and there is also a certain voltage V on it with respect to ground. The ratio of these phasors, V/I, is known as the *driving-point impedance*. The precise resonant antenna length is that for which V and I are exactly in phase, that is, for which the driving-point impedance is purely real. As mentioned earlier, this length is slightly less than half the wavelength, λ, and it also depends on the thickness of the telescoping rods [1, pp. 547–548]. For a rod thickness of $\lambda/400$, the resonant length is 0.476λ. The driving-point impedance of a dipole antenna of these dimensions is 64 Ω. If a voltage-measuring instrument such as a radio receiver or spectrum analyzer is connected to the antenna terminals via a transmission line and is properly matched to the 64 Ω impedance, the measured voltage V_m will be equal to $0.148\lambda E$, where E is the applied field strength and λ is the wavelength at the frequency being measured. The ratio V_m/E, equal to 0.148λ, is known as the effective length (l_e) of the antenna, since it relates the field strength in volts per meter to the measured terminal voltage in volts. Obviously, it is not equal to the physical antenna length but is instead approximately one-third of that value. With this ratio known, the electric field strength E that causes a certain terminal voltage V_m can easily be calculated.

To simplify calculations, E is often expressed in decibels with respect to a reference field of 1 μV/m and is designated E_d. Similarly, V_m is expressed in decibels with respect to a reference voltage of 1 μV and is designated V_d. The antenna factor (AF) is defined as the effective length expressed in negative decibels, or $AF = -20 \log(l_e)$. Then the multiplication becomes an addition, that is, $E_d = V_d + AF$.

The aforementioned AF assumes that the antenna is perfectly matched to the receiver, which implies maximum power transfer. A mismatch would change the AF. Therefore, since the antenna driving-point impedance usually is not equal to the receiver input impedance, a matching circuit must be inserted between the antenna and receiver. Another essential consideration is antenna balance. Most receivers and spectrum analyzers have one input terminal grounded. If this grounded terminal is connected to one of the dipole antenna terminals, the impedances connected to the two antenna terminals will be unequal with respect to ground. This also will upset the AF, since one side of the antenna will not be properly matched to the receiver. To prevent this, a balanced-to-unbalanced (*balun*) network must be inserted between the antenna and the receiver. Such a circuit provides a high impedance *with respect to ground* for both input terminals while providing the correct input impedance (such as 64 Ω) *between* its input terminals. Normally, a single network provides both the matching and balancing functions.

Unfortunately, unless the dipole antenna is precisely the correct length, its AF is much more complicated. Even if the frequency being measured differs only a few percent from the antenna resonant frequency, the AF becomes unpredictable and the driving-point impedance becomes complex. Thus, the electric field cannot be easily calculated from the measured terminal voltage. To achieve the simple AF described earlier, the frequencies must be measured one at a time and the dipole antenna length properly adjusted for each frequency. It is impossible to sweep the spectrum rapidly, as when using a spectrum analyzer, unless the antenna length can somehow be varied also. This leads to mechanical difficulties and is usually impractical.

Other types of antennas, however, are less sensitive to frequency. Examples are the biconical antenna and the log-periodic antenna. A biconical antenna can perform acceptably over a range of 20–300 MHz, and a log-periodic antenna is useful from 300 to 1000 MHz. Their AFs are relatively constant, usually varying by no more than 20 dB, over their useful frequency ranges. The AFs are usually too difficult to calculate, but they may easily be measured simply by observing the terminal voltage resulting from a sinusoidally varying field of known strength. The known field is first measured using a tuned dipole

antenna, for which the AF can be calculated. The AF is measured in this manner at several frequencies throughout its useful range, and the results are plotted for use with the antenna.

Unlike the tuned dipole, the biconical and log-periodic antennas do not exhibit constant driving-point impedances over their useful frequency range. Since the receiver input impedance cannot be made to follow the variation of driving-point impedance with frequency, an exact match is impossible. This affects the AF just as it would for a mismatched tuned dipole. To compensate for this, the AF must be measured with the antenna terminated into a known impedance, which must then be used for all measurements made with that antenna. Then the mismatch is accounted for in the AF itself. The mismatch does cause the antenna to reradiate the received signal, but this effect may be minimized by performing the measurements in an open-field site, which will be discussed later.

Tuned dipole, biconical, and log-periodic antennas are *linearly polarized* antennas because they respond to only one polarization component of a propagating wave. If the antenna is oriented horizontally, only the horizontally polarized component of the wave will affect it. Similarly, only the vertically polarized component will affect a vertically oriented antenna. Thus, with two measurements, any linearly polarized antenna will detect any type of field polarization. Other types of antennas, such as the spiral antenna, are designed to detect a circularly polarized wave. They will detect vertically and horizontally polarized waves, but they could miss a wave that is circularly polarized in the reverse direction (e.g., counterclockwise instead of clockwise). Consequently, circularly polarized antennas are forbidden for many types of field measurements.

All antennas discussed earlier respond to the electric field, E. As mentioned earlier, in the far-field region, the magnetic field, B, is simply 3.33 nT times the value of E expressed in volts per meter. In any other region, however, B is not so simply related to E and must be measured separately, using an antenna that responds to magnetic fields. A circular loop or coil of wire is such an antenna. The loop is cut at one point, and the radio receiver or spectrum analyzer is connected between its two ends. For quantitative measurements, its AF must be known. The factor can be measured by immersing the antenna in a known magnetic field and measuring its terminal voltage. To find the known magnetic field strength, the electric field is first measured, in the far-field region, using a tuned dipole antenna for which the factor is known. The magnetic field is then 3.33 nT times this value expressed in volts per meter. With the magnetic field thus determined, the AF of the loop may be calculated, as required.

36.4 Measurement Environment

A major difficulty with electromagnetic field measurements is repeatability of results. Electromagnetic fields are affected by any materials in their vicinity, even by poor conductors and dielectrics. The measurement environment must therefore be carefully defined, and similar environments must be used for all comparable measurements.

The ideal environment would be one where (1) the only electromagnetic field source is the equipment under test (EUT) and (2) there is no "foreign" material at all that could affect the fields being measured. Unfortunately, the only natural location where this could be achieved is in outer space, since the Earth itself affects electromagnetic fields. Since this is impractical, attempts are made to simulate this environment on Earth.

A large outdoor open area simulates a hemisphere of free space. Such a test site is appropriately called an *open-field site*. If the conductivity, permittivity, and permeability of the Earth were constant, every open-field site would have the same effect on the electromagnetic fields radiating from the EUT. The Earth's parameters do vary, however. To compensate for this variation, a large conductive floor, or ground plane, is laid under the EUT. This causes all electromagnetic waves to be totally reflected from the ground plane, so that the Earth's properties have no effect. The ground plane must be large enough so that it appears infinite with respect to the EUT and the associated test equipment. Acceptable dimensions are $1.73d \times 2d$, where d is the distance between the measurement antenna and the EUT, normally 3 or 10 m. Radiated emissions must be measured in all directions from the EUT

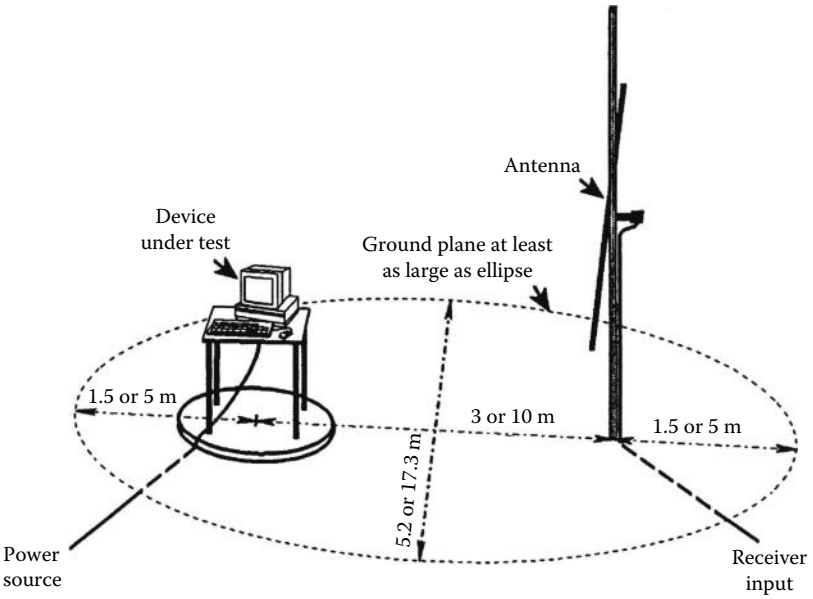


FIGURE 36.1 An open-field test site. Power and antenna cables are run under the ground plane so that they will not affect measured fields. The area outlined by the ellipse must be free of everything except the device under test, the table on which it rests, and the measuring antenna. To facilitate measuring radiation in all directions from the device, it is placed on a turntable. By rotating it during testing and simultaneously varying the height of the receiving antenna, the direction of maximum radiation is found.

and at various angles of inclination. This is most easily achieved by placing the EUT on a turntable, which is then rotated during the test. To allow measurement at various inclination angles, the receiving antenna height must be varied, and this is accomplished by mounting it on a halyard. A typical open-field site appears in Figure 36.1.

An open-field site provides repeatable data only if there are no nearby trees or structures that could cause undesirable reflections. Before it can be reliably used, it must be tested. This is done by generating a known electromagnetic field and measuring it. The field is normally generated by a radio-frequency oscillator driving a tuned dipole antenna, for which the radiation can be calculated ([2]: pp. 237–238). This radiation is then measured as though it were generated by a typical device being tested. The ratio of the voltage at the transmitting antenna terminals to that at the receiving antenna terminals is known as the *site attenuation*. If the site attenuation is within 3 dB of its calculated value, the test site is deemed acceptable.

Although an open-field site eliminates reflections, external field sources, such as licensed transmitters, still cause problems. Since electromagnetic radiation can travel thousands of miles, no open-field site will be completely free of electromagnetic fields. To eliminate the effects of these stray sources, testing must be performed inside a shielded enclosure. There, however, severe reflections occur, and measurements become inaccurate and unrepeatable.

An ideal test environment would be a shielded enclosure lined with material that does not reflect electromagnetic waves. Such an enclosure is called an *anechoic chamber*, with the understanding that the name refers to *electromagnetic* echoes. Until recently, such chambers were not practical except at very high frequencies, but improvements are constantly being made. Such an enclosure is acceptable for testing if it meets the site-attenuation requirements of a true open field. The site attenuation must be measured at several points inside the chamber, to assure that the proximity of the chamber walls has no effect. Unfortunately, such chambers are at present very expensive.

Another type of test chamber is the *transverse electromagnetic (TEM) cell*. This consists of an enlarged section of waveguide, in which the electromagnetic fields can be accurately predicted [3,4]. They are suitable only for testing small devices at relatively low frequencies. The TEM cell can be no larger than a wavelength at the frequency being tested. For example, to test at 200 MHz, the cell could be no larger than 1.5 m, or 5 ft, and the device itself must not exceed $\frac{1}{6}$ of this value, or 10 in. For small devices, however, the TEM cell is very accurate and is unaffected by stray field sources.

If a suitable anechoic chamber is not available, a device may be tested in an ordinary shielded enclosure to learn what frequencies it emits. The field strengths will be inaccurate due to the internal reflections. Then the device is tested in a true open-field site, and the suspected frequencies are measured quantitatively. Any field that exceeds the acceptable limits is then observed while the device is shut off. If it does not disappear, it is obviously not being generated by the EUT. This procedure is acceptable, although not as simple as testing inside an anechoic chamber.

Preliminary measurements may even be performed in an ordinary room. They will not be comparable with similar measurements made anywhere else, because of the effects of nearby conductors and dielectrics. Here, also, the device must be shut off to decide if any emissions are from stray external sources instead of from the EUT. This procedure provides a rough estimate of the emissions from the EUT, and it usually saves time during any later testing at a true open-field site. The various measurement methods appear in Table 36.1.

Permissible emission levels appear in the *Code of Federal Regulations* [5]. These rules assume open-field measurements, which are the most accurate possible. Even there, variations of ±6 dB are typical. Therefore, a manufacturer should allow a safety factor when performing measurements intended to assure compliance with government regulations. Otherwise, a device may pass when tested by the manufacturer but fail if later tested by the government using a supposedly identical test procedure. Since the government's measurements then prevail, the manufacturer's integrity could be questioned.

Further details on measurement techniques are available in Refs. [2,6].

TABLE 36.1 Comparison of EMI Measurement Methods

Method	Equipment Required	Space Required	Accuracy	Outside Influence	Cost	Comments
Ordinary room	Antenna and receiver	3 or 10 m radius around EUT	Medium, affected by structure, ±20 dB	May be severe, depending on location	Minimum	Usually acceptable for preliminary tests
Shielded room	Shielded room, antenna, and receiver	4–6 m radius around EUT	Poor, ±30–40 dB due to reflections	Usually none	Moderate	Use for preliminary tests in noisy areas
TEM cell	TEM cell and receiver	1–3 m^3	Very good, ±10 dB	Usually none	Moderate	Unusable for large EUT due to high-order modes
Open field	Antenna and receiver	17 × 20 m open field with no nearby structures	Excellent, usually ±6 dB	May be severe, depending on location	High, due to logistics of site (power, weather, etc.)	Standard test method
Shielded anechoic chamber	Anechoic chamber, antenna, and receiver	6–15 m radius around EUT	Very good, ±10 dB	Usually none	Very high	Use for accurate tests in noisy areas

Defining Terms

Antenna factor: Its effective length expressed in negative decibels.

Balun: An interface device used to isolate a dipole or other balanced antenna from the effects of a receiver having one grounded terminal.

Conducted coupling: Coupling due to voltages imposed on a shared power source.

Dipole antenna: An antenna consisting of two collinear rods with the feed line connected between them.

Driving-point impedance: The ratio of voltage to current at the driving point (normally the center) of an antenna.

Effective length: The ratio of the voltage observed at the driving point of an antenna to the strength of its received electric field.

Electromagnetic compatibility (EMC): The capability of two or more electric devices to operate simultaneously without mutual interference.

Electromagnetic interference (EMI): Any undesired effect of one electric device upon another due to radiated electromagnetic fields or due to voltages imposed on a shared power source.

Elliptical polarization: Polarization of an electromagnetic wave consisting of two perpendicular electric fields of differing phase.

Emissions: Fields or conducted voltages generated by an electric device.

Far-field region: Any location that is much farther than $\lambda/2\pi$ from an electric or magnetic field source, where λ is the wavelength at the frequency of concern.

Intersystem EMI: EMI between two or more systems.

Intrasystem EMI: EMI between two or more parts of the same system.

Near-field region: Any location that is much nearer than $\lambda/2\pi$ to an electric or magnetic field source, where λ is the wavelength at the frequency of concern.

Open-field site: A test location free of any conductors that would affect electromagnetic fields and taint the results.

Polarization: The direction of the electric field, E, of an electromagnetic wave.

Power density: Radiated power per unit of cross-sectional area.

Radiated coupling: Coupling due to radiated electric, magnetic, or electromagnetic fields.

Site attenuation: A measure of the degree to which electromagnetic fields at a test site are disturbed by environmental irregularities, obtained by comparing calculations with measured experimental results.

Susceptibility: The degree to which an electric device is affected by externally generated fields or conducted voltages.

Transverse electromagnetic (TEM) cell: A relatively small test chamber in which fields can be accurately controlled by its geometric properties.

References

1. E. C. Jordan and K. G. Balmain, *Electromagnetic Waves and Radiating Systems*, 2nd edn., Englewood Cliffs, NJ: Prentice Hall, 1968, pp. 540–547.
2. J. P. Mills, *Electromagnetic Interference Reduction in Electronic Systems*, Englewood Cliffs, NJ: Prentice Hall, 1993, pp. 232–233.
3. M. L. Crawford, Generation of standard EM fields using TEM transmission cells, *IEEE Trans. Electromagn. Compat.*, EMC-16, 189–195, 1974.

4. M. L. Crawford and J. L. Workman, Predicting free-space radiated emissions from electronic equipment using TEM cell and open-field site measurements, *IEEE Int. Symp. Electromagn. Compat.*, Baltimore, MD, October 7–9, IEEE, Piscataway, NJ, 1980, pp. 80–85.
5. 47 CFR Part 15—Radio frequency devices, in *Code of Federal Regulations*, *Title 47: Telecommunications*, Washington, DC: U.S. Government Printing Office, October, 2010, subpt. A–B Current data available at http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&rgn=div5&view=text&node=47:1.0.1.1.14&idno=47, accessed on August 19, 2013.
6. American National Standards Institute (ANSI) Standards, Methods of measurement of radio-noise emissions from low-voltage electrical and electronic equipment (ANSI C63.4-2009), New York: The Institute of Electrical and Electronics Engineers, 2009.

Further Readings

American National Standards Institute (ANSI) Standards, Guide for construction of open area test sites for performing radiated emission measurements (ANSI C63.7-1988), New York: The Institute of Electrical and Electronics Engineers, 1988.

Hayt W. H., *Engineering Electromagnetics*, 5th edn., New York: McGraw-Hill, 1989.

Kraus J. D., *Electromagnetics*, 4th edn., New York: McGraw-Hill, 1992.

Ott H. W., *Noise Reduction Techniques in Electronic Systems*, 2nd edn., New York: John Wiley & Sons, 1988.

Ramo S., J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics*, 2nd edn., New York: John Wiley & Sons, 1984.

37

Permittivity Measurement

Devendra K. Misra
University of Wisconsin–Milwaukee

37.1 Introduction

Dielectric materials possess relatively few free charge carriers. Most of the charge carriers are bound and cannot participate in conduction. However, these bound charges can be displaced by applying an external electric field. In such cases, the atom or molecule forms an electric dipole that maintains an electric field. Consequently, each volume element of the material behaves as an electric dipole. The dipole field tends to oppose the applied field. Dielectric materials that exhibit nonzero distribution of such bound charge separations are said to be *polarized*. The volume density of polarization $\vec{P}$ describes the volume density of those electric dipoles. When a material is linear and isotropic in nature, the polarization density is related to applied electric field intensity, $\vec{E}$, as follows:

$$\vec{P} = \varepsilon_0 \chi_e \vec{E} \tag{37.1}$$

where

ε_0 (=8.854 × 10^{-12} F m^{-1}) is the permittivity of free space
χ_e is called the electric susceptibility of the material

The electric flux density, or displacement, $\vec{D}$ is defined as follows:

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P} = \varepsilon_0 (1 + \chi_0)\vec{E} = \varepsilon_0 \varepsilon_r \vec{E} = \varepsilon \vec{E} \tag{37.2}$$

where

ε is called the permittivity of the material
ε_r is its relative permittivity or dielectric constant

Electric flux density is expressed in coulombs per meter (C m^{-1}).

Equation 37.2 represents a relation between the electric flux density and electric field intensity in frequency domain. It will hold well in time domain only if the permittivity is independent of frequency. A material is called *dispersive* if its characteristics are frequency dependent. The product of Equation 37.2 in frequency domain will be replaced by a convolution integral for the time-domain fields.

Assuming that the fields are time harmonic as $e^{j\omega t}$, the generalized Ampere's law can be expressed in phasor form as follows:

$$\nabla \times \vec{H} = \vec{J}^{\,e} + \vec{J} + j\omega\vec{D} \tag{37.3}$$

where

H is the magnetic field intensity in A m^{-1}
J^e is the current-source density in A m^{-2}
J is the conduction current density in A m^{-2}
the last term represents the displacement current density
J^e will be zero for a source-free region

The conduction current density is related to the electric field intensity through Ohm's law as follows:

$$\vec{J} = \sigma\vec{E} \tag{37.4}$$

where σ is the conductivity of material in S m^{-1}.

From Equations 37.2 through 37.4, one obtains

$$\nabla \times \vec{H} = \vec{J}^{\,e} + \sigma\vec{E} + j\omega\varepsilon\vec{E} \tag{37.5}$$

Conduction current represents the loss of power. There is another source of loss in dielectric materials. When a time-harmonic electric field is applied, the dipoles flip back and forth constantly. Because the charge carriers have finite mass, the field must do work to move them and they might not respond instantaneously. This means that the polarization vector will lag behind the applied electric field. This factor shows up at high frequencies. Therefore, Equation 37.5 is modified as follows:

$$\nabla \times \vec{H} = \vec{J}^{\,e} + \sigma\vec{E} + \omega k''\vec{E} + j\omega\varepsilon\vec{E} = \vec{J}^{\,e} + j\omega\left(\varepsilon - j\frac{\sigma + \omega k''}{\omega}\right)\vec{E} = \vec{J}^{\,e} + j\omega\varepsilon^{*}\vec{E} \tag{37.6}$$

The complex relative permittivity of a material is defined as follows:

$$\varepsilon_r^{*} = \frac{\varepsilon^{*}}{\varepsilon_0} = \frac{1}{\varepsilon_0}\left(\varepsilon - j\frac{\sigma + \omega k''}{\omega}\right) = \varepsilon_r' - j\varepsilon_r'' = \varepsilon_r(1 - j\tan\delta) \tag{37.7}$$

where ε_r' and ε_r'' represent real and imaginary parts of the complex relative permittivity. The imaginary part is zero for a lossless material. The term tanδ is called the *loss tangent*. It represents the tangent of angle between the displacement phasor and total current, as shown in Figure 37.1. Thus, it will be close to zero for a low-loss material.

Dispersion characteristics of a large class of materials can be represented by the following empirical equation of Cole–Cole:

$$\varepsilon_r^{*} = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + (j\omega\tau)^{1-\alpha}} \tag{37.8}$$

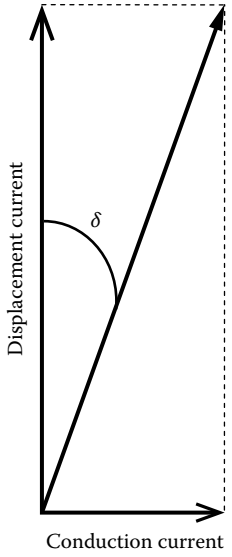


FIGURE 37.1 A phasor diagram representing displacement and loss currents.

where

ε_∞ and ε_s are the relative permittivities of material at infinite and zero frequencies, respectively
ω is the signal frequency in radians per second
τ is the characteristic relaxation time in seconds

For α equal to zero, Equation 37.8 reduces to the Debye equation. Dispersion parameters for a few liquids are given in Table 37.1.

Complex permittivity of a material is determined using lumped circuits at low frequencies, and distributed circuits or free-space reflection and transmission of waves at high frequencies. Capacitance and dissipation factor of a lumped capacitor are measured using a bridge or a resonant circuit. The complex permittivity is calculated from these data. Complex permittivities for some substances are presented in Table 37.2.

At high frequencies, the sample is placed inside a transmission line or a resonant cavity. Propagation constants of the transmission line or resonant frequency and quality factor of the cavity resonator are used to calculate the complex permittivity. Propagation characteristics of electromagnetic waves are influenced by the complex permittivity of that medium. Therefore, a material can be characterized by monitoring the reflected and transmitted wave characteristics as well.

TABLE 37.1 Dielectric Dispersion Parameters for Some Liquids at Room Temperature

Substance	ε_∞	ε_δ	α	τ (ps)
Water	5	78	0	8.0789
Methanol	5.7	33.1	0	53.0516
Ethanol	4.2	24	0	127.8545
Acetone	1.9	21.2	0	3.3423
Ethylene glycol	3	37	0.23	79.5775
Propanol	3.2	19	0	291.7841
Butanol	2.95	17.1	0.08	477.4648
Chlorobenzene	2.35	5.63	0.04	10.2920

TABLE 37.2 Complex Permittivity of Some Substances at Room Temperature

Substance	60 Hz	1 MHz	10 GHz
Nylon	3.60 – *j*0.06	3.14 – *j*0.07	2.80 – *j*0.03
Plexiglas	3.45 – *j*0.22	2.76 – *j*0.04	2.5 – *j*0.02
Polyethylene	2.26 – *j*0.0005	2.26 – *j*0.0005	2.26 – *j*0.0011
Polystyrene	2.55 – *j*0.0077	2.55 – *j*0.0077	2.54 – *j*0.0008
Styrofoam	1.03 – *j*0.0002	1.03 – *j*0.0002	1.03 – *j*0.0001
Teflon	2.1 – *j*0.01	2.1 – *j*0.01	2.1 – *j*0.0008
Glass (lead barium)	6.78 – *j*0.11	6.73 – *j*0.06	6.64 – *j*0.31

37.2 Measurement of Complex Permittivity at Low Frequencies

A parallel-plate capacitor is used to determine the complex permittivity of dielectric sheets [1,2]. For a separation d between the plates of area A in vacuum, the capacitance is given by

$$C_0 = 8.854\frac{A}{d}\ \text{pF} \tag{37.9}$$

where all dimensions are measured in meters. If the two plates have different areas, then the smaller one is used to determine C_0. Further, it is assumed that the field distribution is uniform and perpendicular to the plates. Obviously, the fringing fields along the edges do not satisfy this condition. The guard electrodes, as shown in Figure 37.2, are used to ensure that the field distribution is close to the assumed condition. For best results, the width of the guard electrode must be at least $2d$, and the unguarded plate must extend to outer edge of the guard electrode. Further, the gap between the guarded and guard electrodes must be as small as possible.

The radius of guarded electrode is r_1, and the inner radius of guard electrode is r_2. It is assumed that $R - r_2 \geq 2d$. The area A for this parallel-plate capacitor is πr^2, where r is defined as follows:

$$r = r_1 + \Delta \tag{37.10}$$

$$\Delta = \frac{1}{2}(r_2 - r_1) - \frac{2d}{\pi}\ln\left(\cosh\frac{\pi(r_2 - r_1)}{4d}\right) = \frac{1}{2}(r_2 - r_1) - 1.4659d\ln\left(\cosh 0.7854\frac{r_2 - r_1}{d}\right) \tag{37.11}$$

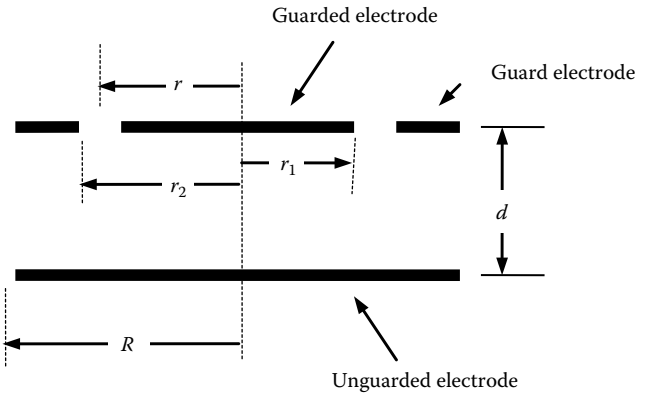


FIGURE 37.2 Geometry of a guarded capacitor.

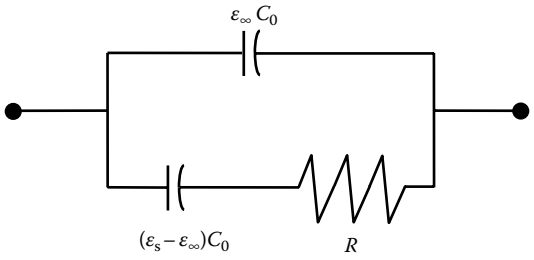


FIGURE 37.3 Equivalent circuit of a parallel-plate capacitor based on the Debye model.

Using the Debye model (i.e., $\alpha = 0$ in Equation 37.8), an equivalent circuit for a dielectric-filled parallel-plate capacitor can be drawn as shown in Figure 37.3. If a step voltage V is applied to it, then the current I can be found as follows [2]:

$$I = \varepsilon_\infty C_0 V\delta(t) + \frac{VC_0(\varepsilon_0 - \varepsilon_\infty)}{\tau}\exp\left(-\frac{t}{\tau}\right) \tag{37.12}$$

where $\tau = RC_0(\varepsilon_0 - \varepsilon_\infty)$.

The first term in Equation 37.12 represents the charging current of capacitor $\varepsilon_\infty C_0$ in the upper branch. This current is not measured because it disappears instantaneously. In practice, it needs to be bypassed at short times to protect the detector from overloading or burning. The second term of Equation 37.12 represents the current charging the lower branch of the equivalent circuit. The time constant, τ, is determined following the decay characteristics of this current. Further, the resistance R can be found after extrapolating this current–time curve to $t = 0$. The discharging current characteristics are used to remove V at $t = 0$.

A typical circuit arrangement for the characterization of dielectric materials using a step voltage is shown in Figure 37.4. A standard resistor R_1 of either 10 GΩ or 1 TΩ is connected between the guarded electrode and the load resistor R_2. A feedback circuit is used that forces the voltage drop across R_1 equal in magnitude but opposite in polarity to that of across R_2. It works as follows. Suppose that the node between capacitor and R_1 has a voltage V_1 with respect to ground. It is amplified but reversed in polarity by the amplifier. Therefore, the current through R_1 will change. This process continues until the input to the amplifier is zero. The junction between R_1 and the capacitor will then be at the ground potential. Thus, the meter M measures voltage across R_2 that is negative of the voltage across R_1. Since R_1 is known, the current through it can be calculated. This current also flows through the sample. S_1 is used to switch from charging to discharging mode, while S_2 is used to provide a path for surge currents.

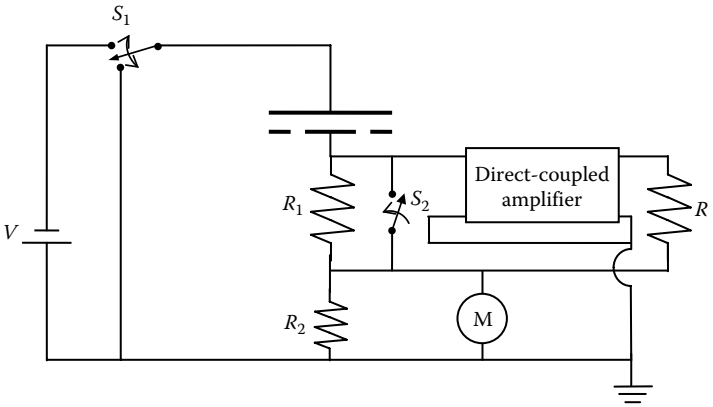


FIGURE 37.4 Circuit arrangement for the characterization of dielectric materials using a step voltage.

Capacitance and dissipation factor of the dielectric-loaded parallel-plate capacitor are used in the medium frequency range to determine the complex permittivity of materials. A substitution method is generally employed in a Schering bridge circuit for this measurement.

In the Schering bridge shown in Figure 37.5, assume that the capacitor C_v is disconnected for the time being and the capacitor C_s contains the dielectric sample. In the case of a lossy dielectric sample, it can be modeled as an ideal capacitor C_x in series with a resistor R_x. The bridge is balanced by adjusting C_d and R_c. An analysis of this circuit under the balanced condition produces the following relations:

$$R_x = \frac{C_d R_c}{C_T} \tag{37.13}$$

and

$$C_x = \frac{C_T R_d}{R_c} \tag{37.14}$$

Quality factor Q of a series RC circuit is defined as the tangent of its phase angle, while the inverse of Q is known as the dissipation factor D. Hence,

$$Q = \frac{X_x}{R_x} = \frac{1}{\omega C_x R_x} = \frac{1}{D} \tag{37.15}$$

For a fixed R_d, the capacitor C_d can be calibrated directly in terms of dissipation factor. Similarly, the resistor R_c can be used to determine C_x. However, an adjustable resistor limits the frequency range. A substitution method is preferred for precision measurement of C_x at higher frequencies. In this technique, a calibrated precision capacitor C_v is connected in parallel with C_s as shown in Figure 37.5, and the bridge is balanced. Assume that the settings of two capacitors at this condition are C_{d1} and C_{v1}. The capacitor C_s is then removed, and the bridge is balanced again. Let the new settings of these capacitors be C_{d2} and C_{v2}, respectively. Equivalent circuit parameters of the dielectric-loaded capacitor C_s are then found as follows:

$$C_x = C_{v2} - C_{v1} \tag{37.16}$$

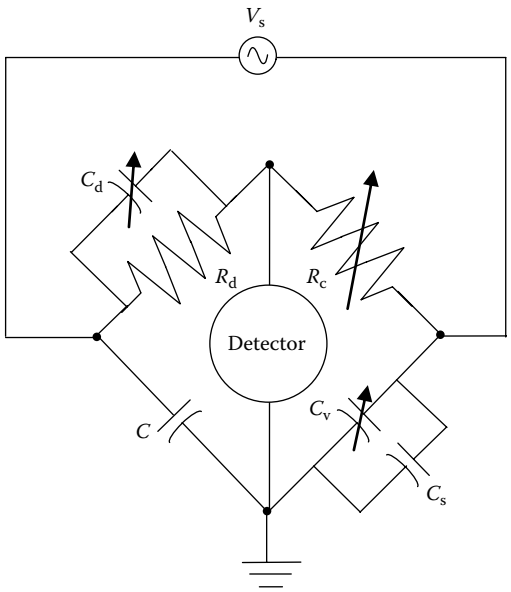


FIGURE 37.5 Schering bridge.

$$D_x = \frac{C_{v2}}{C_x}\delta D \tag{37.17}$$

where

$$\delta D = \omega R_d (C_{d1} - C_{d2})$$

Complex permittivity of the specimen is calculated from these data as follows:

$$\varepsilon_r' = \frac{C_x}{C_0} \tag{37.18}$$

and

$$\varepsilon_r'' = \frac{C_x D_x}{C_0} \tag{37.19}$$

Thus far, a series *RC* circuit equivalent model is used for the dielectric-loaded capacitor. As illustrated in Figure 37.6, an equivalent parallel *RC* model can also be obtained for it. The following equations can be used to switch back and forth between these two equivalent models:

$$G_p = \frac{R_s}{R_s^2 + \left(1/\omega^2 C_s^2\right)} = \frac{1}{R_s}\left(\frac{1}{1+Q^2}\right) \tag{37.20}$$

$$C_p = \frac{C_s}{1+(\omega R_s C_s)^2} = \frac{C_s}{1+D^2} \tag{37.21}$$

$$R_s = \frac{G_p}{G_p^2 + \omega^2 C_p^2} = \frac{1}{G_p}\left(\frac{1}{1+Q^2}\right) \tag{37.22}$$

$$C_s = \frac{G_p^2 + \omega^2 C_p^2}{\omega^2 C_p} = C_p(1+D^2) \tag{37.23}$$

and

$$Q = \frac{1}{D} = \frac{\omega C_p}{G_p} = \frac{1}{\omega R_s C_s} \tag{37.24}$$

Proper shielding and grounding arrangements are needed for a reliable measurement, especially at higher frequencies. Grounding and edge capacitances of the sample holder need to be taken into account for improved accuracy. Further, a guard point needs to be obtained that may require balancing in some cases. An inductive-ratio-arm capacitance bridge can be another alternative to consider for such application [1].

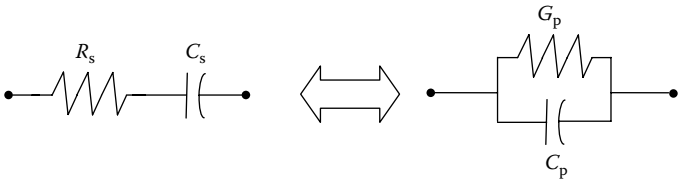


FIGURE 37.6 Series and parallel equivalent circuits of a dielectric-loaded capacitor.

37.3 Measurement of Complex Permittivity Using Distributed Circuits

Measurement techniques based on the lumped circuits are limited up to the lower end of the very high frequency (VHF) band. Characterization of materials at microwave frequencies requires the distributed circuits. A number of techniques have been developed on the basis of wave reflection and transmission characteristics inside a transmission line or in free space. Some other methods employ a resonant cavity that is loaded with the sample. Cavity parameters are measured, and the material characteristics are deduced from that. A number of these techniques, described in [3,4], can be used for a sheet material. These techniques require cutting a piece of sample to be placed inside a transmission line or a cavity. In case of liquid or powder samples, a so-called modified infinite sample method can be used. In this technique, a waveguide termination is filled completely with the sample, as shown in Figure 37.7. Since a tapered termination is embedded in the sample, the wave incident on it will be dissipated with negligible reflection, and it will look like the sample is extending to infinity. The impedance at its input port will depend on the electric properties of filling sample. Its voltage standing wave ratio (VSWR) S and location of first minimum d from the load plane are measured using a slotted line. The complex permittivity of sample is then calculated as follows [5]:

$$\varepsilon_r' = \left(\frac{\lambda}{\lambda_c}\right)^2 + \frac{[1-(\lambda/\lambda_c)^2]\times[S^2\sec^4(\beta d)-(1-S^2)^2\tan^2(\beta d)]}{[1+S^2\tan^2(\beta d)]^2} \tag{37.25}$$

and

$$\varepsilon_r'' = \frac{[1-(\lambda/\lambda_c)^2]\times[2S(1-S^2)^2\sec^4(\beta d)\tan(\beta d)]}{[1+S^2\tan^2(\beta d)]^2} \tag{37.26}$$

where

λ is the free-space wavelength
λ_c is the cutoff wavelength for the mode of propagation in empty guide
β is the propagation constant in the feeding guide

It is assumed that the waveguide supports TE_{10} mode only.

37.3.1 Resonant Cavity Method

A cavity resonator can be used to determine the complex permittivity of materials at microwave frequencies. If a cavity can be filled completely with the sample, then the following procedure can be used.

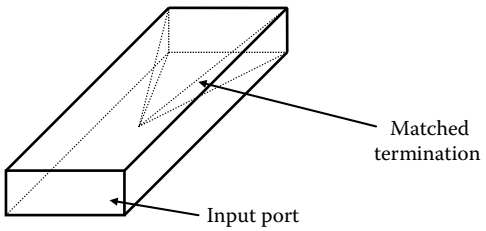


FIGURE 37.7 A waveguide termination filled with liquid or powder sample.

Measure the resonant frequency f_1 and the quality factor Q_1 of an empty cavity. Next, fill that cavity with the sample material and measure its new resonant frequency f_2 and quality factor Q_2. The dielectric parameters of the sample are then calculated from the following formulae [3]:

$$\varepsilon_r = \left(\frac{f_1}{f_2}\right)^2 \tag{37.27}$$

and

$$\tan\delta = \frac{1}{Q_2} - \frac{1}{Q_1}\sqrt{\frac{f_1}{f_2}} \tag{37.28}$$

On the other hand, a cavity perturbation technique will be useful for smaller samples [4]. If the sample is available in a circular cylindrical form, then it may be placed inside a TE_{101} rectangular cavity through the center of its broad face where the electric field is the maximum. Its resonant frequency and quality factor with and without the sample are then measured. Complex permittivity of the sample is calculated as follows:

$$\varepsilon_r' = 1 + \frac{1}{2}\frac{f_1 - f_2}{f_2}\frac{V}{v} \tag{37.29}$$

and

$$\varepsilon_r'' = \frac{V}{4v}\frac{Q_2 - Q_1}{Q_1 Q_2} \tag{37.30}$$

where V and v are cavity and sample volumes, respectively.

Similarly, for a small spherical sample of radius r that is placed in a uniform field at the center of the rectangular cavity, the dielectric parameters are as follows:

$$\varepsilon_r' = \frac{abd}{8\pi r^3}\frac{f_1 - f_2}{f_2} \tag{37.31}$$

and

$$\varepsilon_r'' = \frac{abd}{16\pi r^3}\left(\frac{Q_2 - Q_1}{Q_1 Q_2}\right) \tag{37.32}$$

where a, b, and d are the width, height, and length of the rectangular cavity, respectively. For best accuracy in cavity perturbation method, the shift in frequency $(f_1 - f_2)$ must be very small.

37.3.2 Free-Space Method for Measurement of Complex Permittivity

When a plane electromagnetic wave is incident on a dielectric interface, its reflection and transmission depend on the contrast in the dielectric parameters. Many researchers have used it for determining the complex permittivity of dielectric materials placed in free space. An automatic network analyzer and phase-corrected horn antennas can be used for such measurements [6]. The system is calibrated using the TRL (through, reflect, and line) technique. A time-domain gating is used to minimize the error due to multiple reflections. The sample of thickness d is placed in front of a conducting

plane, and its reflection coefficient S_{11} is measured. A theoretical expression for this reflection coefficient is found as follows:

$$S_{11} = \frac{jZ_{\mathrm{d}} \tan(\beta_{\mathrm{d}} d) - 1}{jZ_{\mathrm{d}} \tan(\beta_{\mathrm{d}} d) + 1} \tag{37.33}$$

where

$$Z_{\mathrm{d}} = \frac{1}{\sqrt{\varepsilon_r^{\star}}} \tag{37.34}$$

$$\beta_{\mathrm{d}} = \frac{2\pi}{\lambda}\sqrt{\varepsilon_{\mathrm{r}}^{\star}} \tag{37.35}$$

λ is the free-space wavelength of electromagnetic signal.

Equation 37.33 is solved for $\varepsilon_{\mathrm{r}}^{\star}$ after substituting the measured S_{11}. Since it represents a nonlinear relation, an iterative numerical procedure can be used.

37.3.3 Electric Probe Method for Determining the Dielectric Properties of Materials

Coaxial line-excited short electric probes have been used for the in situ determination of dielectric properties. The procedure requires an accurate measurement of the input impedance of the in situ probe and its mathematical model that is dependent on the electric characteristics of the surrounding medium. The procedure is complex, and therefore, there are a number of models reported in the literature for simplification. In one of these techniques, it is assumed that the electric probe is immersed in the medium under test and a coaxial line of inner and outer radii a and b, respectively, feeds this probe from behind the ground plane, as illustrated in Figure 37.8 [7]. The medium has conductivity σ S/m and the dielectric constant ε_{r}. Further, the opening $(b - a)$ is much smaller than the probe radius a, and the probe length λ is smaller than a quarter wavelength. The so-called induced EMF method provides an expression for the input impedance Z_{in} of this probe as follows:

$$Z_{\mathrm{in}} = \frac{j\omega\mu_0}{4\pi\beta_{\mathrm{d}} \sin(\beta_{\mathrm{d}}\ell)} \int_0^{\ell} \sin\{\beta_{\mathrm{d}}(\ell - z)\} f(z) dz \tag{37.36}$$

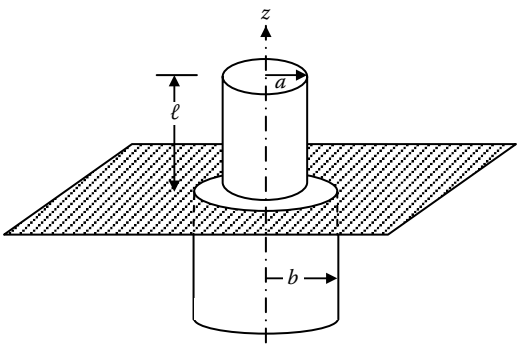


FIGURE 37.8 A coaxial line-excited short electric probe on the ground plane.

where

$$f(z) = \frac{e^{-j\beta_d r_1}}{r_1} + \frac{e^{-j\beta_d r_2}}{r_2} - 2\cos(\beta_d \ell) \cdot \frac{e^{-j\beta_d r}}{r} \tag{37.37}$$

$$r_1 = \sqrt{(\ell - z)^2 + a^2} \tag{37.38}$$

$$r_2 = \sqrt{(\ell + z)^2 + a^2} \tag{37.39}$$

$$r = \sqrt{z^2 + a^2} \tag{37.40}$$

$$\frac{\omega\mu_0}{\beta_d} = \sqrt{\frac{\mu_0}{\varepsilon^*}} = \eta = \text{Intrinsic impedance of the medium} \tag{37.41}$$

$$\varepsilon^* = \varepsilon_0\left(\varepsilon_r - j\frac{\sigma}{\omega\varepsilon_0}\right) \tag{37.42}$$

$$\beta_d^2 = \omega^2\mu_0\varepsilon^* \tag{37.43}$$

Since the input impedance Z_{in} is known precisely after measurement, Equation 37.36 can be solved numerically for complex β_d using a root-finding routine. The material characteristics are then determined from Equations 37.42 and 37.43.

37.3.4 Nondestructive Method for Measuring the Complex Permittivity of Materials

Most of the techniques described thus far require cutting and placing a part of sample in the test fixture. Sometimes, it may not be permissible to do so. Further, the dielectric parameters can change in that process. It is especially important in the case of a biological specimen to perform in vivo measurements. In one such technique, an open-ended coaxial line is placed in close contact with the sample, and its input reflection coefficient is measured using an automatic network analyzer [8,9]. As recommended by the manufacturers, the network analyzer is calibrated initially using an open-circuit, a short-circuit, and a matched load. The reference plane is then moved to the measuring end of the coaxial line using short circuit.

Assume that a and b are inner and outer radii of the coaxial line, respectively. ω is the angular frequency; μ_0 and ε_0 are the permeability and permittivity of the free space, respectively; $k = \omega\sqrt{\mu_0\varepsilon_0\varepsilon_r^*}$ is the wavenumber in material medium. Admittance of the coaxial aperture in contact with material medium is as follows:

$$Y_L = \frac{2}{\int_a^b E_\rho(\rho', 0)d\rho'} - \frac{2\pi}{\left[\sqrt{(\mu_0/\varepsilon_0\varepsilon_1)}\ln(b/a)\right]} \tag{37.44}$$

where $E_\rho(\rho', 0)$ is radial electric field intensity over the aperture. It is evaluated from the following integral equation:

$$\frac{1}{\pi\rho} + j\omega\varepsilon_1\varepsilon_o\int_a^b E_\rho(\rho', 0)K_c(\rho, \rho')\rho' d\rho' = \frac{j\omega\varepsilon_r^*\varepsilon_0}{\pi}\int_a^b E_\rho(\rho', 0)\rho' d\rho' \int_0^\pi \cos(\phi')\frac{\exp(-jkr)}{r}d\phi' \tag{37.45}$$

where

$$r = \sqrt{\rho^2 + \rho'^2 - 2\rho\rho'}\cos(\phi') \tag{37.46}$$

$$K_c(\rho, \rho') = j\sum_{n=0}^{\infty} \frac{\phi_n(\rho)\phi_n(\rho')}{A_n^2 \beta_n} \tag{37.47}$$

$$\phi_n = Y_0(\gamma_n a)J_1(\gamma_n \rho) - J_0(\gamma_n a)Y_1(\gamma_n \rho) \tag{37.48}$$

$$\beta_n = \begin{cases} \sqrt{k_1^2 - \gamma_n^2} & k_1 > \gamma_n \\ -j\sqrt{\gamma_n^2 - k_1^2} & k_1 < \gamma_n \end{cases} \tag{37.49}$$

$$A_n^2 = \frac{2}{\pi^2 \gamma_n^2}\left[\frac{J_0^2(\gamma_n a)}{J_0^2(\gamma_n b)} - 1\right] \quad n > 0; \quad A_0^2 = \ln\left(\frac{b}{a}\right) \tag{37.50}$$

The eigenvalues λ_n are solutions to the following characteristic equation:

$$J_0(\gamma_n b)Y_0(\gamma_n a) = J_0(\gamma_n a)Y_0(\gamma_n b) \tag{37.51}$$

where

J_n and Y_n are Bessel functions of the first and second kind of order n, respectively
ε_1 is the dielectric constant of the insulator
k_1 is wavenumber inside the coaxial line

Equation 37.45 is solved numerically using the method of moments. A numerical root-finding procedure, such as Muller's method, is used to solve Equation 37.44 for the complex wavenumber k. Complex permittivity, in turn, is determined from the following relation:

$$\varepsilon_r^* = \frac{k^2}{\omega^2 \mu_0 \varepsilon_0} \tag{37.52}$$

A relatively much simpler method may be used if the coaxial aperture of the probe is electrically small such that there is negligible radiation. This technique requires four standard media to calibrate the system. For example, water, open circuit, short circuit, and methanol are used to find calibration constants. Further, it does not require precise value of reflection coefficient (or input impedance), and therefore, there is no need to calibrate the automatic network analyzer to manufacturer's standards. The procedure may be summarized as follows.

With open end of the probe in free space, its reflection parameter R_1 is recorded. The probe is then dipped in distilled water and then in methanol, and the corresponding reflection parameters are recorded as R_2 and R_4. With a conductive foil tightly pressed on the aperture, the reflection parameter

R_3 is recorded. Note that the magnitudes of reflection parameters may be larger than unity sometimes because this setup is not calibrated for measuring the reflection coefficients. Now, a coefficient χ is determined to find the equivalent admittance parameters at the opening as follows:

$$\chi = \frac{(1+\Lambda)\varepsilon_{r4}^{\star} - \varepsilon_{r1}^{\star} - \Lambda\varepsilon_{r2}^{\star}}{\varepsilon_{r1}^{\star 2} + \Lambda\varepsilon_{r2}^{\star 2} - (1+\Lambda)\varepsilon_{r4}^{\star 2}} \tag{37.53}$$

where $\varepsilon_{r1}^{\star}$, $\varepsilon_{r2}^{\star}$, and $\varepsilon_{r4}^{\star}$ are the complex relative permittivities of free space (~1), water, and methanol, respectively, at the measurement frequency, and

$$\Lambda = \frac{(R_4 - R_1)\cdot(R_3 - R_2)}{(R_4 - R_2)\cdot(R_1 - R_3)} \tag{37.54}$$

A dimensionless equivalent admittance parameter y_n of the aperture with the nth medium is found as follows:

$$y_n = \varepsilon_{rn}^{\star} + \chi\varepsilon_{rn}^{\star 2} \tag{37.55}$$

Now with the test-sample terminating the coaxial aperture, reflection parameter R_s is recorded, and it is used to find the coefficient ξ as follows:

$$\xi = \frac{(R_s - R_1)\cdot(R_3 - R_2)}{(R_s - R_2)\cdot(R_1 - R_3)} \tag{37.56}$$

Using Equation 37.55, the admittance parameters y_1 and y_2 for the free space and water, respectively, are computed. The admittance parameter y_s for the sample is then found as follows:

$$y_s = \frac{y_1 - y_2\xi}{1-\xi} \tag{37.57}$$

The complex relative permittivity $\varepsilon_{rs}^{\star}$ of the sample can now be determined via Equation 37.55 as follows:

$$\varepsilon_{rs}^{\star} = \frac{-1 \pm \sqrt{1 + 4y_s\chi}}{2\chi} \tag{37.58}$$

Only one of these two values found for $\varepsilon_{rs}^{\star}$ will have positive real part and negative imaginary part. The other solution is not physically possible and, hence, can be thrown.

Note: The relevant ASTM standards and a list of manufacturers are included at the end of the text.

Defining Terms

Conduction current density: Current per unit area caused by conduction of charge carriers.

Dielectric constant: A dimensionless constant that represents the permittivity of a material relative to the permittivity of free space.

Displacement current density: It represents the time rate of change of electric flux density.

Dissipation factor: It is the inverse of the quality factor.

Electric dipole: A pair of equal and opposite electric charges separated by a small distance.

Electric flux density: A fundamental electric field quantity that is related to volume density of free charges. It is also known as the electric displacement.

Electric polarization density: The average electric dipole moment per unit volume.

Electric susceptibility: A dimensionless parameter that relates the polarization density in a material with electric field intensity.

Frequency domain field: A phasor quantity (a complex function in general) that depends on space coordinates. The time dependency is assumed to be sinusoidal.

Isotropic material: A material in which the electric polarization has the same direction as the applied electric field.

Loss tangent: A ratio of the imaginary part to the real part of the complex permittivity of a material.

Quality factor: A dimensionless quantity that represents the time average energy stored in an electric circuit relative to energy dissipated in one period.

Reflection coefficient: Defined as a ratio of reflected phasor voltage to that of incident phasor voltage at a point in the circuit.

Relaxation time: It represents the time taken by a charge placed inside a material volume to decay to about 37% of its initial value.

Time-domain field: A field expressed as a function of time. It is a real function that is dependent on time and space coordinates.

Voltage standing wave ratio (VSWR): Defined as a ratio of maximum voltage to the minimum voltage on a transmission line.

Related ASTM Standards

Standard	Material Type	Frequency Range
D5568-08	Solids	Microwaves
D1531-06	Solid insulating materials	1 kHz–1 MHz
D5493-06	Geotextiles	
D7449/D7449M-08	Solids	Microwaves
D150-11	Uses lumped impedance standards	1 Hz to several hundred megahertz
D3380-10	Polymer-based microwave circuit substrates	8–12.4 GHz
D7416-09	In-service lubricants	—
D2149-97	Solids	50 Hz–10 MHz

References

1. A. R. Von Hippel, *Dielectric Materials and Applications*, Cambridge, MA: MIT Press, 1961.
2. N. E. Hill, W. E. Vaughan, A. H. Price, and M. Davies, *Dielectric Properties and Molecular Behaviour*, London, U.K.: Van Nostrand Reinhold, 1969.
3. M. Sucher and J. Fox (eds.), *Handbook of Microwave Measurements*, Vol. II, Brooklyn, NY: Polytechnic Press, 1963.
4. R. Chatterjee, *Advanced Microwave Engineering*, Chichester, U.K.: Ellis Horwood Limited, 1988.
5. D. K. Misra, Permittivity measurement of modified infinite samples by a directional coupler and a sliding load, *IEEE Trans. Microwave Theory Technol.*, 29, 65–67, 1981.
6. D. K. Ghodgaonkar, V. V. Varadan, and V. K. Varadan, A free-space method for measurement of dielectric constants and loss tangents at microwave frequencies, *IEEE Trans. Instrum. Meas.*, 38, 789–793, 1989.

7. D. K. Misra, In-situ electric probe method for determining the dielectric properties of materials, *IEEE Trans. Instrum. Meas.*, 37, 157–160, 1988.
8. A. P. Gregory, R. N. Clarke, T. E. Hodgetts, and G. T. Symm, RF and microwave dielectric measurements upon layered materials using a reflectometric coaxial sensor, National Physical Laboratory Report, Division of Electrical Science 125, U.K., March 1993.
9. D. Misra, On the measurement of the complex permittivity of materials by an open-ended coaxial probe, *IEEE Microwave Guided Wave Lett.*, 5, 161–163, 1995.

Relevant Websites

http://www.mtt.org
http://ewh.ieee.org/soc/deis
http://iopscience.iop.org
http://www.nist.gov
http://www.npl.co.uk

Selected Test-Instrument Manufacturers

1. Agilent (www.agilent.com): 16451B dielectric test fixture for solids and 16452A dielectric test fixture for liquids. An LCR meter is needed to measure the parallel-plate capacitance. 85070E dielectric probe kit (works for 200 MHz–50 GHz frequency band with a suitable vector network analyzer).
2. QuadTech (www.quadtech.com): LD-3 dielectric cell for liquids, solids, or powdered material. Requires an LCR meter.
3. Scientifica (www.scientificainc.com): Liquid dielectric constant meter Model 871; measurement is performed at 10 kHz.
4. Brookhaven (www.brookhaveninstruments.com): BI-870 dielectric constant meter, used for measurements on liquids at 10 kHz.

38

Electric-Field Strength

David A. Hill
National Institute of Standards and Technology

Motohisa Kanda
National Institute of Standards and Technology

Electric-field strength is defined as the ratio of the force on a positive test charge at rest to the magnitude of the test charge in the limit as the magnitude of the test charge approaches zero. The units of electric-field strength are volts per meter (V m^{-1}). Electric charges and currents are sources of electric and magnetic fields, and Maxwell's equations [1] provide the mathematical relationships between electromagnetic (EM) fields and sources.

The electric field at a point in space is a vector defined by components along three orthogonal axes. For example, in a rectangular coordinate system, the electric field $\vec{E}$ can be written as

$$\vec{E} = \hat{x}E_x + \hat{y}E_y + \hat{z}E_z \tag{38.1}$$

where
$\hat{x}, \hat{y}$, and $\hat{z}$ are unit vectors
E_x, E_y, and E_z are scalar components

For electrostatic fields, the components are real scalars that are independent of time. For steady-state, time-harmonic fields, the components are complex phasors that repetition dependent, $e^{j\omega t}$, are suppressed.

38.1 Electrostatic Fields

Electrostatic fields are present throughout the atmosphere, and there are strong electrostatic fields near high-voltage dc power lines. The commonly used electrostatic field meters generate an ac signal by periodic conductor motion (either rotation or vibration). This ac signal is proportional to the electric-field strength, and field-meter calibration is performed in a known electrostatic field.

38.1.1 Field Mills

Field mills (also called generating voltmeters) determine electric-field strength by measuring modulated, capacitively induced charges or currents on metal electrodes. Two types of field mills—the shutter type and the cylindrical type—are described in the technical literature [2]. The shutter type is more common; a simplified version is shown in Figure 38.1. The sensing electrode is periodically exposed to and shielded from the electric field by a grounded, rotating shutter. The charge q_s induced on the sensing electrode and the current i_s between the sensing electrodes and grounds are both proportional to the electric-field strength E normal to the electrode:

$$q_s(t) = \varepsilon_0 E a_s(t) \quad \text{and} \quad i_s(t) = \varepsilon_0 E \frac{da_s(t)}{dt} \tag{38.2}$$

where

ε_0 is the permittivity of free space [1]
$a_s(t)$ is the effective exposed area of the sensing electrode at time t

Thus, the field strength can be determined by measuring the induced charge or current (or voltage across the impedance Z). If the induced signal is rectified by a phase-sensitive detector (relative to the shutter motion), the dc output signal will indicate both the polarity and magnitude of the electric field [3].

Shutter-type field mills are typically operated at the ground or at a ground plane, but a cylindrical field mill can be used to measure the electric field at points removed from a ground plane. A cylindrical field mill consists of two half-cylinder sensing electrodes as shown in Figure 38.2. Charges induced on the two sensing electrodes are varied periodically by rotating the sensing electrodes about the cylinder axis at a constant angular frequency ω_c. The charge q_c induced on a half cylinder of length L and the current i_c between the half cylinders are given by

$$q_c = 4\varepsilon_0 r_c L E \sin \omega_c t \quad \text{and} \quad i_c = 4\varepsilon_0 r_c L E \omega_c \cos \omega_c t \tag{38.3}$$

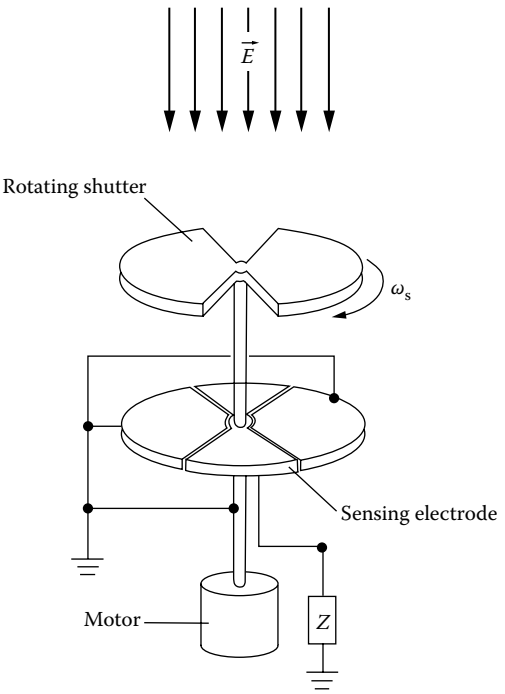


FIGURE 38.1 Shutter-type electric-field mill for measurement of the polarity and magnitude of an electrostatic field.

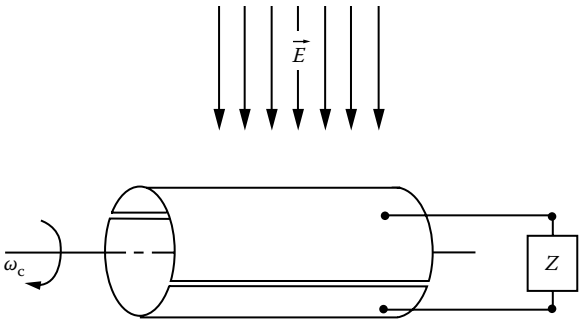


FIGURE 38.2 Cylindrical field mill for measurement of electrostatic field strength.

where r_c is the cylinder radius. Equation 38.3 is based on the 2D solution for a conducting cylinder in an electric field and neglects end effects for finite L. Equation 38.3 shows that the electric-field strength E can be determined from a measurement of the induced charge or current.

A third type of electric-field meter uses a vibrating plate [4] to generate an ac signal that is proportional to the electric-field strength. With any type of electric-field strength meter, the observer should be at a sufficient distance from the measurement location to avoid perturbing the electric field.

38.1.2 Calibration Field

A known uniform field for meter calibration can be produced between a pair of parallel plates [2]. If a potential difference V is applied between a pair of plates with a separation d_p, the field strength away from the plate edges is V/d_p. The plate dimensions should be much larger than d_p to provide an adequate region of uniform field. Also, d_p should be much larger than the field-meter dimensions so that the charge distribution on the plates is not disturbed. The parallel plates can be metal sheets or tightly stretched metal screens.

The field meter should be located in the type of environment in which it will be used. Shutter-type field mills that are intended to be located at a ground plane should be located at one of the plates. Cylindrical field mills that are not intended to be used at a ground plane should be located in the center of the region between the plates.

For simplicity, only field mills and calibration in the absence of space charge were mentioned. However, near power lines or in the upper atmosphere [5], the effects of space charge can be significant and require modifications in field mill design. A field mill for use in a space charge region and a calibration system with space charge are described in [6].

38.2 ELF and ULF Electric Fields

In this section, measurement techniques for extremely low frequency (ELF, 3 Hz to 3 kHz) and ultralow frequency (ULF, below 3 Hz) electric fields are considered. Natural ELF fields are produced by thunderstorms, and natural ULF fields are produced by micropulsations in the Earth's magnetic field [7]. Geophysicists make use of these natural fields in the magnetotelluric method for remote sensing of the Earth's crust [8]. AC power lines are dominant sources of fields at 50 or 60 Hz and their harmonics.

An ac electric-field strength meter [9] includes two essential parts: (1) an antenna and (2) a detector (receiver). Other possible features are a transmission line or optical link, frequency-selective circuits, amplifying and attenuating circuits, an indicating device, and a nonconducting handle. The antenna characteristics can be calculated for simple geometries or determined by calibration. For example,

linear antennas are often characterized by their effective length L_{eff} [10], which determines the open-circuit voltage V_{oc} induced at the antenna terminals:

$$V_{oc} = L_{eff} E_{inc} \tag{38.4}$$

where E_{inc} is the component of the incident electric field parallel to the axis of the linear antenna. The detector could respond to the terminal voltage or current or to the power delivered to the load.

38.2.1 Natural Horizontal Electric Field at the Earth's Surface

Magnetotelluric sounding of the Earth's crust requires measurement of the horizontal electric and magnetic fields at the Earth's surface [8]. The magnetic field is measured with a horizontal axis loop, and the electric field is measured with a horizontal wire antenna as shown in Figure 38.3. The antenna wire is insulated since it lies on the ground, but it is grounded at its end points. Nonpolarizing grounding electrodes should be used to avoid polarization potentials between the electrodes and the ground.

Since the natural electric-field strength to be measured is on the order of 1 μV m^{-1}, the antenna length L_h needs to be on the order of 1 km to produce a measurable voltage. Since the effective length of a grounded antenna is equal to the physical length ($L_{eff} = L_h$), the horizontal component E_h of the electric field parallel to the antenna is equal to the open-circuit voltage divided by the antenna length:

$$E_h = \frac{V_{oc}}{L_h} \tag{38.5}$$

The frequencies used in typical magnetotelluric sounding range from approximately 0.1 mHz to 10 Hz. If both horizontal components of the electric field are needed, a second orthogonal antenna is required.

38.2.2 Free-Body Electric-Field Meters

ELF electric fields in residential and industrial settings are most conveniently measured with free-body field meters [11,12], which measure the steady-state current or charge oscillating between two-halves of a conducting body in free space. (Ground reference meters [13] are also available for measuring the electric field normal to the ground or some other conducting surface.) Geometries for free-body electric-field meters are shown in Figure 38.4. Commercial field meters are usually rectangular in shape, and typical dimensions are on the order of 10 cm. A large dynamic range (1 V m^{-1} to 30 kV m^{-1}) is required to cover the various field sources (ac power lines, video display terminals, mass transportation systems, etc.) of interest. (Electro-optic field meters with less sensitivity are described in [12].) A long, nonconducting handle is normally attached perpendicular to the field-meter axis for use in measurement surveys.

The charge Q on half of the field meter is proportional to the incident electric field E along the meter axis:

$$Q = A\varepsilon_0 E \tag{38.6}$$

where

ε_0 is the permittivity of free space [1]
A is a constant proportional to the surface area

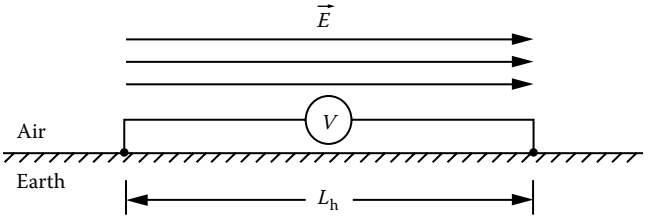


FIGURE 38.3 Grounded horizontal antenna for measurement of the horizontal component of the electric field.

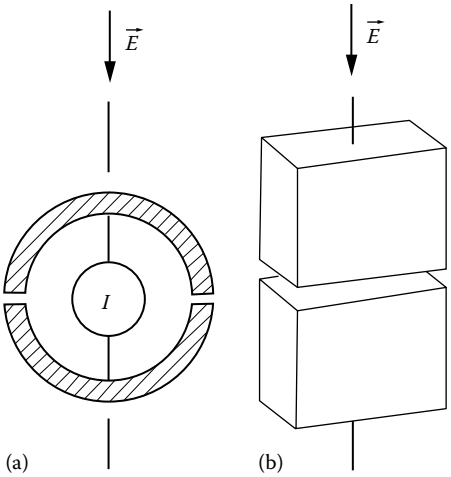


FIGURE 38.4 Electric-field meters for measurement of the axial component of the electric field: (a) spherical geometry and (b) rectangular geometry.

For the spherical geometry in Figure 38.4a, $A = 3\pi a^2$, where a is the sphere's radius. Since the current I between the two-halves is equal to the time derivative of the charge, for time-harmonic fields, it can be written as

$$I = j\omega A\varepsilon_0 E \tag{38.7}$$

This allows E to be determined from the measured current. For commercial field meters that are not spherical, the constant A needs to be determined by calibration. A known calibration field can be generated between a pair of parallel plates where the plates are sufficiently large compared to the separation to produce a uniform field with small edge effects. This technique produces a well-characterized field with an uncertainty of less than 0.5% [11]. However, the presence of harmonic frequencies can cause less accurate meter readings in field surveys.

38.3 Radio-Frequency and Microwave Techniques

38.3.1 Dipole Antennas

A thin, linear dipole antenna of length L is shown in Figure 38.5. Its effective length is approximately [10]

$$L_{\text{eff}} = \frac{\lambda}{\pi}\tan\left(\frac{\pi L}{2\lambda}\right) \tag{38.8}$$

where λ is the free-space wavelength. Resonant half-wave dipoles ($L = \lambda/2$) have an effective length of $L_{\text{eff}} = 2L/\pi$ and are of convenient length for frequencies from 30 to 1000 MHz. The physical length of a dipole at resonance is actually slightly shorter than $\lambda/2$ to account for the effect of a finite length-to-diameter ratio. Resonant dipoles are used as standard receiving antennas to establish a known standard field in the *standard antenna method* [9]. Commercial antennas and field meters are calibrated in such standard fields.

For $L < \lambda/2$, Equation 38.8 must be used to determine L_{eff}. For very short dipoles ($L/\lambda \ll 1$), the current distribution is approximately linear, and the effective length is approximately one-half the physical length ($L_{\text{eff}} \approx L/2$). Short dipoles are frequently used as electric-field probes.

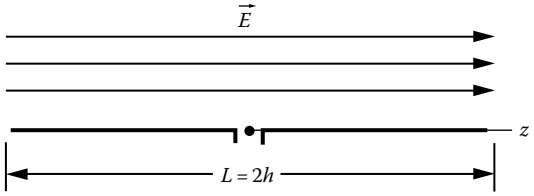


FIGURE 38.5 Dipole antenna for measurement of the axial component of the electric field.

38.3.2 Aperture Antennas

Aperture antennas are commonly used for receiving and transmitting at microwave frequencies (above 1 GHz). As receiving antennas, they are conveniently characterized by their on-axis gain g or effective area A_{eff}. Effective area is defined as the ratio of the received power P_{r} to the incident power density S_{inc} and can also be written in terms of the gain [9]:

$$A_{\text{eff}} = \frac{P_{\text{r}}}{S_{\text{inc}}} = \frac{g\lambda^2}{4\pi} \tag{38.9}$$

Equation 38.9 applies to the case where the incident field is polarization matched to the receiving antenna. The incident power density in a plane wave is $S_{\text{inc}} = E^2/\eta_0$, where E is the rms electric-field strength and η_0 is the impedance of free space (≈377 Ω). Thus, the electric-field strength can be determined from the received power:

$$E = \sqrt{\frac{P_{\text{r}}\eta_0}{A_{\text{eff}}}} = \lambda^{-1}\sqrt{\frac{4\pi\eta_0 P_{\text{r}}}{g}} \tag{38.10}$$

In general, the gain can be measured using the two-antenna method [9]. For a pyramidal horn antenna as shown in Figure 38.6, the gain can be calculated accurately from [14]

$$g = \frac{32ab}{\pi\lambda^2} R_{\text{E}} R_{\text{H}} \tag{38.11}$$

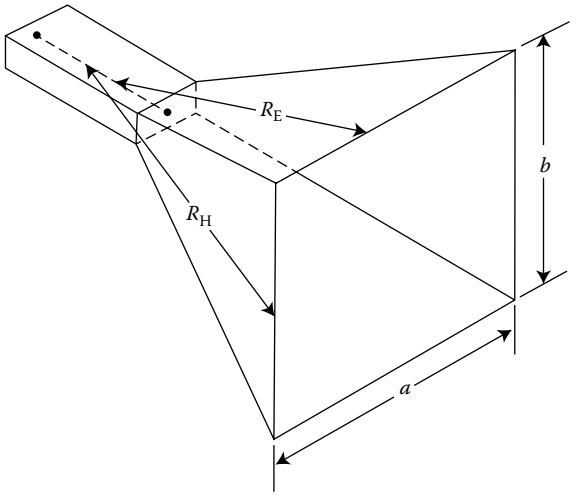


FIGURE 38.6 Pyramidal horn for measuring power density or electric-field strength.

where R_E and R_H are gain reduction factors due to the *E* and *H* plane flare of the horn. The gain reduction factors are

$$R_E = \frac{C^2(w) + S^2(w)}{w^2} \quad \text{and} \quad R_H = \frac{\pi^2\{[C(u) - C(v)]^2 + [S(u) - S(v)]^2\}}{4(u - v)^2} \tag{38.12}$$

where

$$w = \frac{b}{\sqrt{2\lambda l_E}} \quad \text{and} \quad \left.\begin{matrix} u \\ v \end{matrix}\right\} = \frac{\sqrt{\lambda l_H/2}}{a} \pm \frac{a}{\sqrt{2\lambda l_H}}$$

The Fresnel integrals *C* and *S* are expressed as [15]

$$C(w) = \int_0^w \cos\left(\frac{\pi}{2}t^2\right)dt \quad \text{and} \quad S(w) = \int_0^w \sin\left(\frac{\pi}{2}t^2\right)dt \tag{38.13}$$

Well-characterized aperture antennas are also used to generate standard fields [16] for calibrating commercial antennas and field strength meters. This method of calibration is called the *standard field method* [9]. The electric-field strength *E* at a distance *d* from the transmitting antenna is

$$E = \frac{\sqrt{\eta_0 P_{del} g/(4\pi)}}{d} \tag{38.14}$$

where P_{del} is the net power delivered to the transmitting antenna and is typically measured with a directional coupler [16]. The gain *g* for a pyramidal horn can be calculated from Equation 38.11. The National Institute of Standards and Technology (NIST)* uses rectangular open-ended waveguides from 200 to 500 MHz and a series of pyramidal horns from 500 MHz to 40 GHz to generate standard fields in an anechoic chamber [16]. The uncertainty of the field strength is less than 1 dB over the entire frequency range of 200 MHz–40 GHz.

38.4 Three-Loop Antenna System

Electronic equipment can emit unintentional EM radiation that can interfere with other electronic equipment. If the radiating source is electrically small (as is a video display terminal), then it can be characterized by equivalent electric and magnetic dipole moments. The three-loop antenna system (TLAS), shown in Figure 38.7, consists of three orthogonal loop antennas that are terminated at diametrically opposite points. The unique feature of loop antennas with double terminations [17] is that they can measure both electric and magnetic fields. For electromagnetic interference (EMI) applications, a device under test (DUT) is placed at the center of the TLAS. On the basis of six terminal measurements, the TLAS determines three equivalent electric dipole components and three equivalent magnetic dipole moments of the DUT and hence its radiation characteristics.

Here, the theory [18] is summarized for one of the three loops. The DUT in Figure 38.7 is replaced by an electric dipole moment $\vec{m}_e$ and a magnetic dipole moment $\vec{m}_m$ both located at the origin of the coordinate system. The dipole moments can be written in terms of their rectangular components:

$$\vec{m}_e = \hat{x}m_{ex} + \hat{y}m_{ey} + \hat{z}m_{ex} \quad \text{and} \quad \vec{m}_m = \hat{x}m_{mx} + \hat{y}m_{my} + \hat{z}m_{mz} \tag{38.15}$$

The loop in the *xy* plane has radius r_0 and has impedance loads Z_{xy} located at the intersections with the *x*-axis ($\phi = 0, \pi$).

* Contribution of the NIST, not subject to copyright in the United States.

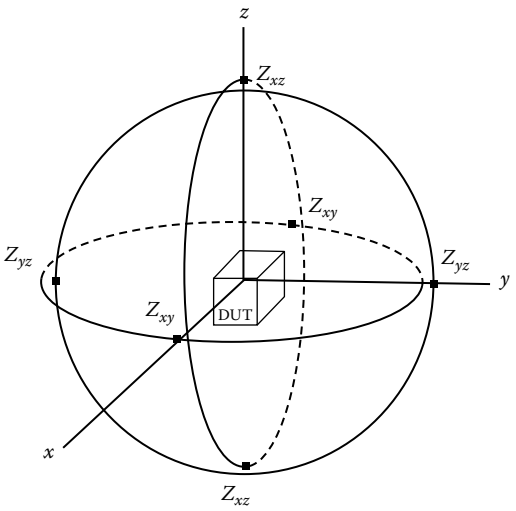


FIGURE 38.7 Geometry of the TLAS and the DUT.

The solution for the current induced in the loop is based on Fourier series analysis [19]. The incident azimuthal electric-field $E_{\phi}^{\mathrm{i}}(\phi)$ tangent to the loop is

$$E_{\phi}^{\mathrm{i}}(\phi) = A_0 + A_1 \cos\phi + B_1 \sin\phi \tag{38.16}$$

where

$A_0 = m_{\mathrm{m}x} G_{\mathrm{m}}$

$A_1 = m_{\mathrm{e}y} G_{\mathrm{e}}$

$B_1 = -m_{\mathrm{e}x} G_{\mathrm{e}}$

$$G_{\mathrm{m}} = \frac{\eta_0}{4\pi}\left(\frac{k^2}{r_0} - \frac{jk}{r_0^2}\right)e^{-jkr_0}$$

$$G_{\mathrm{e}} = \frac{-\eta_0}{4\pi}\left(\frac{jk}{r_0} + \frac{1}{r_0^2} + \frac{1}{jkr_0^2}\right)e^{-jkr_0}$$

$k = 2\pi/\lambda$ is the free-space wave number

An approximate solution [17] for the loop current $I(\phi)$ yields the following results for the load currents $I(0)$ and $i(\pi)$:

$$\begin{aligned} I(0) &= 2\pi r_0\left(\frac{m_{\mathrm{m}z} G_{\mathrm{m}} Y_0}{1 + 2Y_0 Z_{xy}} + \frac{m_{\mathrm{e}y} G_{\mathrm{e}} Y_1}{1 + 2Y_1 Z_{xy}}\right) \\ I(\pi) &= 2\pi r_0\left(\frac{m_{\mathrm{m}z} G_{\mathrm{m}} Y_0}{1 + 2Y_0 Z_{xy}} - \frac{m_{\mathrm{e}y} G_{\mathrm{e}} Y_1}{1 + Y_1 Z_{xy}}\right) \end{aligned} \tag{38.17}$$

where Y_0 and Y_1 are the admittances for the constant and cos ϕ currents [17].

One can solve Equation 38.17 for the magnetic and electric dipole components:

$$m_{mz} = \frac{I_\Sigma(1+2Y_0Z_{xy})}{2\pi r_0 G_m Y_0} \quad \text{and} \quad m_{ey} = \frac{I_\Delta(1+2Y_1Z_{xy})}{2\pi r_0 G_e Y_1} \tag{38.18}$$

where

$I_\Sigma = [I(0) + I(\pi)]/2$

$I_\Delta = [I(0) - I(\pi)]/2$

Thus, the sum current I_Σ can be used to measure the magnetic dipole moment, and the difference current I_Δ can be used to measure the electric dipole moment. The four remaining dipole components can be obtained in an analogous manner. The loop in the *xz* plane can be used to measure m_{my} and m_{ex}, and the loop in the *yz* plane can be used to measure m_{mx} and m_{ez}.

The total power P_T radiated by the source can be written in terms of the magnitudes of the six dipole components:

$$P_T = \frac{2\pi\eta_0}{3\lambda^2}\left[|m_{ex}|^2 + |m_{ey}|^2 + |m_{ez}|^2 + k^2\left(|m_{mx}|^2 + |m_{my}|^2 + |m_{mz}|^2\right)\right] \tag{38.19}$$

The expression for the power pattern is more complicated and involves the amplitudes and phases of the dipole moments [20]. The TLAS has been constructed with 1 m diameter loops and successfully tested from 3 kHz to 100 MHz [21]. It is currently being used to measure radiation from video display terminals and other inadvertent radiators.

38.5 Broadband Dipole Antennas

The EM environment continues to grow more severe and more complex as the number of radiating sources increases. Broadband antennas are used to characterize the electromagnetic (EM) environment over a wide frequency range. For electric-field measurements, electrically short dipole antennas with high capacitive input impedance are used with a capacitive load, such as a field-effect transistor (FET). The transfer function *S* of frequency *f* is defined as the ratio of the output voltage V_L of the antenna to the incident electric field E_i [22]:

$$S(f) = \frac{V_L(f)}{E_i(f)} = \frac{h\alpha/2}{1 + C/C_a} \tag{38.20}$$

where

$$C_a = \frac{4\pi h}{c\eta_0(\Omega_a - 2 - \ln 4)}$$

$$\alpha = \frac{\Omega_a - 1}{\Omega_a - 2 + \ln 4}$$

Ω_a is the antenna thickness factor and is equal to $2\ln(2h/r_a)$

r_a is the antenna radius

C is the load capacitance

C_a is the antenna capacitance

h is half the physical length of the dipole antenna, as shown in Figure 38.5

c is the free-space speed of light

Since the input impedance of an electrically short dipole is predominantly a capacitive reactance, a very broadband frequency response can be achieved with a high-impedance capacitive load. However, given the present state of the art, it is not possible to build a balanced, high-input impedance FET with high common-mode rejection above 400 MHz. For this reason, it is more common practice to use a high-frequency, beam-lead Schottky-barrier diode with a very small junction capacitance (less than 0.1 pF) and very high junction resistance (greater than several M3) for frequencies above 400 MHz.

The relationship between the time-dependent diode current $i_d(t)$ and voltage $v_d(t)$ is

$$i_d(t) = I_s[e^{\alpha_d v_d(t)} - 1] \tag{38.21}$$

where I_s and α_d are constants of the diode. For very small incident fields $E_i(t)$, the output detected dc voltage v_0 is

$$\begin{aligned} v_0 &= \frac{-b_d^2}{2\alpha_d}\left\langle \tilde{v}_i^2 \right\rangle \\ v_i &= E_i L_e \end{aligned} \tag{38.22}$$

where

$\tilde{v}_i(t) = v_i(t) - \langle v_i \rangle$

$b_d = \dfrac{C_a \alpha_d}{C_a + C_d}$

C_a is the dipole capacitance

C_d is the diode capacitance

L_e is the effective length of the dipole antenna

$\langle\rangle$ indicates time average

Thus, the dc detected voltage is frequency independent and is directly proportional to the average of $(E_i - \langle E_i \rangle)^2$.

For large incident fields $E_i(t)$, the output detected dc voltage is

$$v_0 = -\frac{b_d \tilde{V}_i}{\alpha_d} \tag{38.23}$$

where $\tilde{V}_i$ is the peak value of $v_i(t)$. Consequently, for a large incident field, v_0 is also frequency independent and is directly proportional to the peak field.

Conventional dipole antennas support a standing-wave current distribution; thus, the useful frequency range of this kind of dipole is usually limited by its natural resonant frequency. In order to suppress this resonance, a resistively loaded dipole (traveling-wave dipole) has been developed. If the internal impedance per unit length $Z_l(z)$ as function of the axial coordinate z (measured from the center of the dipole) has the form

$$Z_l(z) = \frac{60\psi}{h - |z|} \tag{38.24}$$

then the current distribution $I_z(z)$ along the dipole is that of a traveling wave. Its form is

$$I_z(z) = \frac{V_0}{60\psi(1 - j/kh)}\left[1 - \frac{|z|}{h}\right]e^{-jk|z|} \tag{38.25}$$

where

$2h$ is the total physical length of the dipole

V_0 is the driving voltage

$$\psi = 2\left[\sinh^{-1}\frac{h}{a_d} - C(2ka_d, 2kh) - jS(2ka_d, 2kh)\right] + \frac{j}{kh}(1 - e^{-j2kh}) \tag{38.26}$$

$C(x, y)$ and $S(x, y)$ are generalized cosine and sine integrals

a_d is the dipole radius

This type of resistively tapered dipole has a fairly flat frequency response from 100 kHz to 18 GHz [23].

Defining Terms

Antenna: A device designed to radiate or to receive time-varying EM waves.

Aperture antenna: An antenna that radiates or receives EM waves through an open area.

Dipole antenna: A straight wire antenna with a center feed used for reception or radiation of EM waves.

Electric-field strength: The ratio of the force on a positive test charge to the magnitude of the test charge in the limit as the magnitude of the test charge approaches zero.

Electrostatic field: An electric field that does not vary with time.

Field mill: A device used to measure an electrostatic field.

Microwaves: EM waves at frequencies above 1 GHz.

Power density: The time average of the Poynting vector.

References

1. J.A. Stratton, *Electromagnetic Theory*, New York: McGraw-Hill, 1941.
2. ANSI/IEEE Std. 1227–1990 (R2010), *IEEE Guide for the Measurement of DC Electric-Field Strength and Ion Related Quantities* New York, Current version 2002.
3. P.E. Secker and J.N. Chubb, Instrumentation for electrostatic measurements, *J. Electrostat*., 16, 1–19, 1984.
4. R.E. Vosteen, DC electrostatic voltmeters and fieldmeters, *Conference Record of Ninth Annual Meeting of the IEEE Industrial Applications Society*, Pittsburgh, PA, October 1974.
5. P.J.L. Wildman, A device for measuring electric field in the presence of ionisation, *J. Atmos. Terr. Phys*., 27, 416–423, 1965.
6. M. Misakian, Generation and measurement of dc electric fields with space charge, *J. Appl. Phys*., 52, 3135–3144, 1981.
7. G.V. Keller and F.C. Frischknecht, *Electrical Methods in Geophysical Prospecting*, Oxford, U.K.: Pergamon Press, 1966.
8. A.A. Kaufman and G.V. Keller, *The Magnetotelluric Sounding Method*, Amsterdam, the Netherlands: Elsevier, 1981.

9. IEEE Std. 291–1991, *IEEE Standard Methods for Measuring Electromagnetic Field Strength of Sinusoidal Continuous Waves*, *30 Hz to 30 GHz,* New York.
10. E.C. Jordan and K.G. Balmain, *Electromagnetic Waves and Radiating Systems*, 2nd edn., Englewood Cliffs, NJ: Prentice Hall, 1968.
11. ANSI/IEEE Std. 644–1994, *IEEE Standard Procedures for Measurement of Power Frequency Electric and Magnetic Fields from AC Power Lines,* New York.
12. IEEE Std. 1308–1994, *IEEE Recommended Practice for Instrumentation: Specifications for Magnetic Flux Density and Electric Field Strength Meters—10 Hz to 3 kHz*, reaffirmed 2010.
13. C.J. Miller, The measurements of electric fields in live line working, *IEEE Trans. Power App. Syst*., PAS-16, 493–498, 1967.
14. E.V. Jull, *Aperture Antennas and Diffraction Theory*, Stevenage, U.K.: Peter Peregrinus, 1981.
15. M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, Boulder, CO: U.S. Department of Commerce, National Bureau of Standards, Special Publication AMS 55, 1968.
16. D.A. Hill, M. Kanda, E.B. Larsen, G.H. Koepke, and R.D. Orr, Generating standard reference electromagnetic fields in the NIST anechoic chamber, 0.2 to 40 GHz, National Institute of Standards and Technology Technical Note 1335, 1990.
17. M. Kanda, An electromagnetic near-field sensor for simultaneous electric and magnetic-field measurements, *IEEE Trans. Electromagn. Compat*., EMC-26, 102–110, 1984.
18. M. Kanda and D.A. Hill, A three-loop method for determining the radiation characteristics of an electrically small source, *IEEE Trans. Electromagn. Compat*., 34, 1–3, 1992.
19. T.T. Wu, Theory of the thin circular antenna, *J. Math. Phys*., 3, 1301–1304, 1962.
20. I. Sreenivasiah, D.C. Chang, and M.T. Ma, Emission characteristics of electrically small radiating sources from tests inside a TEM cell, *IEEE Trans. Electromagn. Compat*., EMC-23, 113–121, 1981.
21. D.R. Novotny, K.D. Masterson, and M. Kanda, An optically linked three-loop antenna system for determining the radiation characteristics of an electrically small source, *IEEE Int. EMC Symp*., Dallas, TX, 1993, pp. 300–305.
22. M. Kanda, Standard probes for electromagnetic field measurements, *IEEE Trans. Antennas Propag*., 41, 1349–1364, 1993.
23. M. Kanda and L.D. Driver, An isotropic electric-field probe with tapered resistive dipoles for broadband use, 100 kHz to 18 GHz, *IEEE Trans. Microwave Theory Tech*., MTT-35, 124–130, 1987.

39

Magnetic Field Measurement

Robert L. Fagaly
Quasar Federal Systems

Steven A. Macintyre
Macintyre Electronic Design

Magnetic field strength is measured using a variety of different technologies. Each technique has unique properties that make it more suitable for particular applications. These applications can range from simply sensing the presence or change in the field to the precise measurements of a magnetic field's scalar and vector properties. A very good and exhaustive fundamental description of both mechanical and electrical means for sensing magnetic fields can be found in Lion [1]; less detailed but more up-to-date surveys of magnetic sensor technologies can be found in [2,3]. It is not possible to adequately describe all of these technologies in the space available in a handbook. This chapter concentrates on sensors that are commonly used in magnetic field measuring instruments.

As shown in Figure 39.1, magnetic field sensors can be divided into vector component and scalar magnitude types. The vector types can be further divided into sensors that are used to measure low fields (<1 mT) and high fields (>1 mT). Instruments that measure low fields are commonly called *magnetometers*. High-field instruments are usually called *gaussmeters*.

The induction coil and fluxgate magnetometers are the most widely used vector measuring instruments. They are rugged, reliable, and relatively less expensive than the other low-field vector measuring instruments. The fiber-optic magnetometer is the most recently developed low-field instrument. Although it currently has about the same sensitivity as a fluxgate magnetometer, its potential for better performance is large. The optical fiber magnetometer has not yet left the laboratory, but work on making it more rugged and field worthy is under way. The superconducting quantum interference device (SQUID) magnetometers are the most sensitive of all magnetic field measuring instruments. These sensors operate at temperatures near absolute zero and require special thermal control systems. This makes the SQUID-based magnetometer more expensive, less rugged, and less reliable.

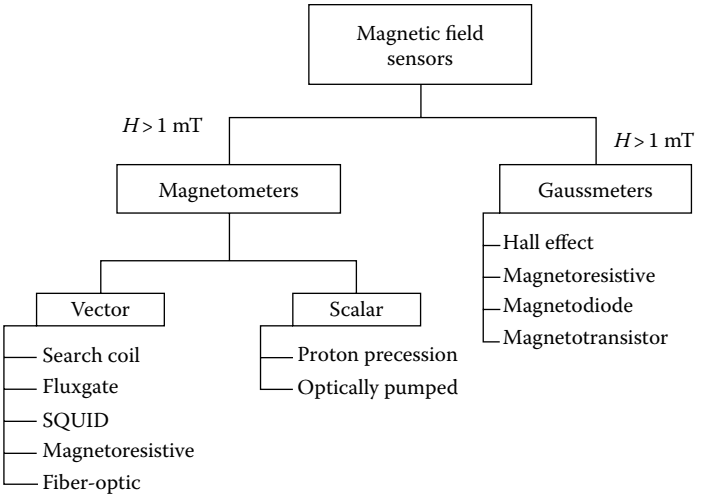


FIGURE 39.1 Magnetic field sensors are divided into two categories based on their field strengths and measurement range: magnetometers measure low fields and gaussmeters measure high fields.

TABLE 39.1 Field Strength Instrument Characteristics

Instrument	Range (mT)	Resolution (nT)	Bandwidth (Hz)	Comment
Induction coil	10^{-10}–10^{6}	Variable	10^{-1}–10^{6}	Cannot measure static fields
Fluxgate	10^{-4}–0.5	0.1	dc to 2×10^{3}	General-purpose vector magnetometer
SQUID	10^{-9}–0.1	10^{-4}	dc to 5	Highest sensitivity magnetometer
Hall effect	$0.1–3 \times 10^{4}$	100	dc to 10^{8}	Best for fields above 1 T
Magnetoresistance	10^{-3}–5	10	dc to 10^{7}	Good for midrange applications
Proton precession	0.02–0.1	0.05	dc to 2	General-purpose scalar magnetometer
Optically pumped	0.01–0.1	0.005	dc to 5	Highest-resolution scalar magnetometer

The Hall effect device is the oldest and most common high-field vector sensor used in gaussmeters. It is especially useful for measuring extremely high fields (>1 T). The magnetoresistive sensors cover the middle ground between the low- and high-field sensors. Anisotropic magnetoresistors (AMR) are currently being used in many applications, including magnetometers. The recent discovery of the giant magnetoresistive (GMR) effect, with its 10-fold improvement in sensitivity, promises to be a good competitor for the traditional fluxgate magnetometer in medium-sensitivity applications.

The proton (nuclear) precession magnetometer is the most popular instrument for measuring the scalar magnetic field strength. Its major applications are in geological exploration and aerial mapping of the geomagnetic field. Since its operating principle is based on fundamental atomic constants, it is also used as the primary standard for calibrating magnetometers. The proton precession magnetometer has a very low sampling rate, on the order of one to three samples per second, so it cannot measure fast changes in the magnetic field. The optically pumped magnetometer operates at a higher sampling rate and is capable of higher sensitivities than the proton precession magnetometer, but it is more expensive and not as rugged and reliable.

Table 39.1 lists various magnetic field strength instruments and their characteristics.

39.1 Magnetic Field Fundamentals

An understanding of the nature of magnetic fields is necessary in order to understand the techniques used for measuring magnetic field strength. The most familiar source of a magnetic field is the bar magnet. The field it produces is shown in Figure 39.2. Magnetic field is a vector quantity; that is, it has

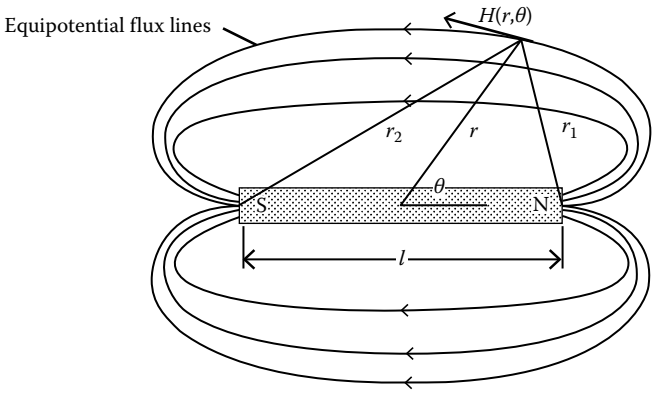


FIGURE 39.2 Magnets produce magnetic fields. A magnetic field is a vector quantity with both magnitude and direction properties.

both a magnitude and a direction. The field of a bar magnet or any other magnetized object, when measured at a distance much greater than its longest dimension, is described by Equation 39.1:

$$\vec{H} = \frac{3(\vec{m} \times \hat{a}_r)\hat{a}_r - \vec{m}}{r^3} \tag{39.1}$$

where

$\hat{a}_r$ is a unit vector along r
r is the distance between the magnetic field source and the measurement point
$\vec{m}$ is called the magnetic dipole moment

The derivation of this equation can be found in many textbooks on electromagnetics. This is a very convenient equation for estimating the field produced by many magnetized objects.

The strength or intensity of a magnetized object depends on the density of its volume-distributed moments. This intensity is called its magnetization M, which is defined as the moments per unit volume:

$$\vec{M} = \frac{\vec{m}}{\text{Volume}} \tag{39.2}$$

Like magnetic field, magnetization is a vector quantity. Magnetization is a material property that can arise from internal magnetic sources as well as be induced by an external magnetic field.

There is a third magnetic vector $\vec{B}$ called magnetic induction or flux density. In free space, magnetic field and magnetic induction are proportional to one another by a constant factor μ_0:

$$\vec{B} = \mu_0 \vec{H} \tag{39.3}$$

Things are different in matter. Equation 39.4 describes the relationship among the magnetic field, magnetic induction, and magnetization vectors in matter:

$$\vec{B} = \mu_0(\vec{H} + \vec{M}) \tag{39.4}$$

In this case, the magnetic induction and the magnetic field vectors do not necessarily have the same direction. Some materials have anisotropic magnetic properties that make these two vectors point in

different directions. The magnetization vector can consist of permanent and induced magnetization components. The permanent magnetization vector does not depend on the presence of an external field. The induced magnetization vector does depend on an external magnetic field and only exists while the inducing field is present.

Magnetic materials can be loosely classified as magnetically "soft" or magnetically "hard." In a magnetically hard material, the permanent magnetization component dominates (a magnet is an example). Magnetization in a soft magnetic material is largely induced and is described by the following equation:

$$\vec{M} = \chi \vec{H} \tag{39.5}$$

where χ is called the material's magnetic susceptibility. In an isotropic material (magnetic properties are not direction dependent), χ is a scalar quantity, and the magnetization and field vectors are proportional and aligned. In anisotropic material (magnetic properties depend on direction), χ is a tensor represented by a 3×3 matrix; therefore, the magnetization vector magnitude and direction depend on the direction and strength of the inducing field. As a result, the magnetization vector will not always align with the magnetization inducing field vectors. Equation 39.5 can be modified for magnetically "soft" material to the following:

$$\vec{B} = \mu_0(1 + \chi)\vec{H} = \mu_0 \mu \vec{H} \tag{39.6}$$

where μ is called the relative permeability of the material.

A magnetized object with a magnetic moment $\vec{m}$ will experience torque $\vec{T}$ in the presence of a uniform magnetic field $\vec{H}$. Equation 39.7 expresses this relationship:

$$\vec{T} = \vec{m} \times \vec{H} \tag{39.7}$$

Torque is the cross-product of the magnetic moment and field vectors. The magnitude equation is

$$T = mH\sin\theta \tag{39.8}$$

where θ is the angle between the direction of $\vec{m}$ and $\vec{H}$.

There is an intimate relationship between electric and magnetic fields. Oersted discovered that passing a current through a wire near a compass causes the compass needle to rotate. The compass was the first magnetic field strength sensor. Faraday found that he could produce an electric voltage at the terminals of a loop of wire if he moved a magnet near it. This led to the induction or search coil sensor.

Magnetic fields are produced by the flow of electric charge (i.e., electric currents). In effect, a magnetic field is a velocity-transformed electric field (through a Lorentz transformation). Current flowing through a straight wire, a loop of wire, or a solenoid will also produce a magnetic field as illustrated in Figure 39.3.

Units are always a problem when dealing with magnetic fields. The Gaussian centimeter, gram, and second (cgs) system of units was favored for many years. Since $\mu_0 = 1$ in the cgs system, magnetic field and flux density have the same numeric value in air, and their units (oersted for field and gauss for flux density) are often indiscriminately interchanged. This has led to great confusion. The cgs system has now been replaced by the International System of Units (SI). The SI system uses, among others, the meter (m), kilogram (kg), second (s), and ampere (A) as the fundamental units. Payne [4] gives a very good explanation of the differences between these systems of units as they relate to magnetic fields. Table 39.2 summarizes the relationships between the two systems of units.

The SI system of units is used throughout this chapter.

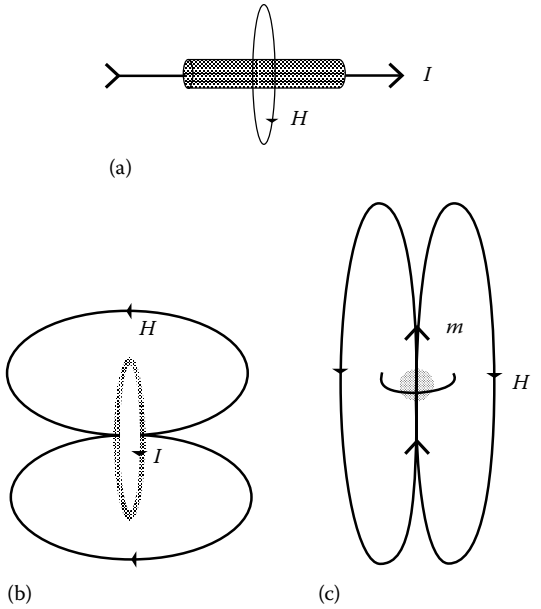


FIGURE 39.3 Magnetic fields are also produced by electric currents: (a) Straight wire, (b) loop of wire, and (c) spinning proton.

TABLE 39.2 Factors for Converting from cgs to SI Magnetic Field Units

Description	Symbol	SI Unit	Gaussian cgs Unit	Multiply By
Magnetic induction	B	Tesla	gauss	10^4
Magnetic field strength	H	A m^{-1}	oersted (oe)	$4\pi \times 10^{-3}$
Magnetization	M	A m^{-1}	emu m^3	10^{-3}
Magnetic dipole moment	m	A m^2	emu	10^3
Magnetic flux	ϕ	Weber (Wb)	maxwell	10^8
Magnetic pole strength	p	A m	emu	
Permeability of free space	μ_0	H m^{-1}	$4\pi \times 10^{-7}$	1

39.2 Low-Field Vector Magnetometers

39.2.1 Induction Coil Magnetometer

The induction or search coil, which is one of the simplest magnetic field sensing devices, is based on Faraday's law. This law states that if a loop of wire is subjected to a changing magnetic flux, ϕ, through the area enclosed by the loop, then a voltage will be induced in the loop that is proportional to the rate of change of the flux:

$$e(t) = -\frac{d\phi}{dt} \tag{39.9}$$

Since magnetic induction $\vec{B}$ is flux density, then a loop with cross-sectional area $\vec{A}$ will have a terminal voltage:

$$e(t) = -\frac{d(\vec{B} \cdot \vec{A})}{dt} \tag{39.10}$$

for spatially uniform magnetic induction fields.

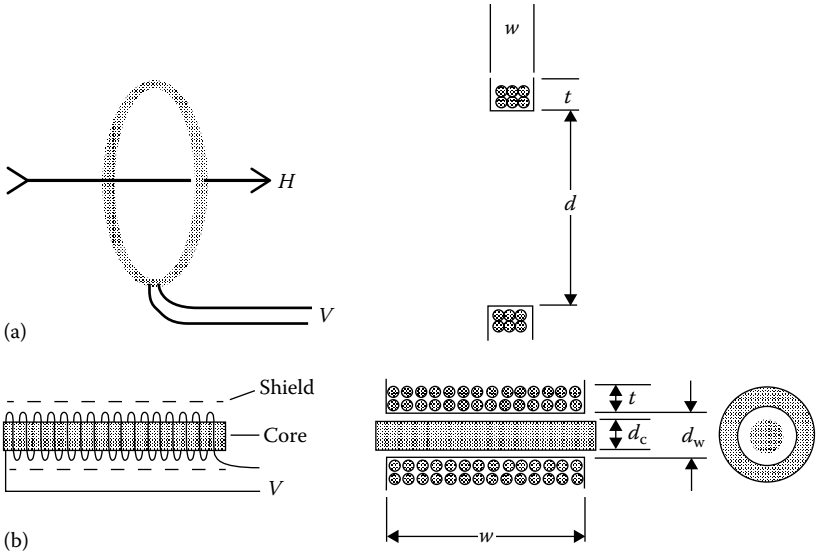


FIGURE 39.4 Induction or search coil sensors consist of a loop of wire (or a solenoid), which may or may not surround a ferromagnetic core: (a) Air-core loop antenna and (b) solenoid induction coil antenna with ferromagnetic core.

Equation 39.10 states that a temporal change in $\vec{B}$ or the mechanical orientation of $\vec{A}$ relative to $\vec{B}$ will produce a terminal voltage. If the coil remains fixed with respect to $\vec{B}$, then static fields cannot be detected, but if the loop is rotated or the magnitude of $\vec{A}$ is changed, then it is possible to measure a static field. The relationship described by Equation 39.10 is exploited in many magnetic field measuring instruments (see [1]).

Figure 39.4 shows the two most common induction coil configurations for measuring field strength: the air-core loop antenna and the rod antenna. The operating principle is the same for both configurations. Substituting $\mu_0\mu_e H(t)$ for B in Equation 39.10 and assuming the loop is stationary with respect to the field vector, the terminal voltage becomes

$$e(t) = -\mu_0\mu_e nA\frac{dH(t)}{dt} \tag{39.11}$$

where

n is the number of turns in the coil

μ_e is the effective relative permeability of the core

The core of a rod antenna is normally constructed of magnetically "soft" material so one can assume that the flux density in the core is induced by an external magnetic field and, therefore, the substitution earlier is valid. With an air (no) core, the effective relative permeability is one. The effective permeability of an induction coil that contains a core is usually much greater than one and is strongly dependent on the shape of the core and, to some extent, on the configuration of the winding.

Taking the Laplace transform of Equation 39.11 and dividing both sides by H, one obtains the following transfer function $T(s)$ for an induction coil antenna:

$$T(s) = -\mu_0\mu_e nAs = -Ks \text{ VmA}^{-1} \tag{39.12}$$

where $E(s) = T(s)\,H(s)$, $E(s)$ and $H(s)$ are the Laplace transforms of e(t) and $H(t)$, and s is the Laplace transform operator. Inspection of Equation 39.12 reveals that the magnitude of the coil voltage is proportional

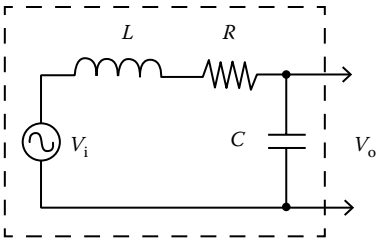


FIGURE 39.5 The induction coil equivalent circuit is a frequency-dependent voltage source in series with an inductor, resistor, and lumped capacitor.

to both the magnitude and frequency of the magnetic field being measured. The coil constant or sensitivity of the loop antenna is

$$K = \mu_0 \mu_e nA \text{ VsmA}^{-1} \tag{39.13}$$

Figure 39.5 is the equivalent circuit for an induction coil antenna. The actual voltage measured at the terminals of the loop is modified by the inductance L, resistance R, and the distributed stray and shield capacitances represented by the lumped capacitor C. These circuit parameters depend on the geometry of the core, coil, and winding.

The electrostatic shield made of nonmagnetic material shown in Figure 39.4 is an important element in the design of an induction coil. It prevents coupling of electric fields into the coil, thereby assuring that the signal seen at the coil terminals is only that due to a magnetic field. The shield should not be placed too close to the winding since it contributes to coil capacitance and noise.

39.2.2 Air-Core Loop Antenna

The air-core loop antenna consists of a circular or rectangular loop containing one or more turns of wire and no magnetic core. The diameter of the loop is usually much greater than the dimensions of the winding cross section. The sensitivity of a circular loop antenna with a winding inside diameter d and rectangular cross section is approximately

$$K = \mu_0 n\pi \frac{d^2}{4}\left[1 + 2\left(\frac{t}{d}\right) + \frac{3}{4}\left(\frac{t}{d}\right)^2\right] \tag{39.14}$$

where

t is the thickness of the winding
n is the number of turns

The resistance of the coil is

$$R = 4n\frac{d}{d_w^2}\left(1 + \frac{t}{d}\right)\rho\,\Omega \tag{39.15}$$

where

d_w is the diameter of the bare wire
ρ is its resistivity in Ω m (1.7×10^{-8} Ω m for copper)

The inductance of the coil is more difficult to compute since it depends heavily on the geometry of the coil. Those who are interested in computing very accurate inductance values for a wide variety of coil

shapes should consult [5]. Equation 39.16 is a general expression that gives a good approximation for the inductance of a circular air-core coil:

$$L = \mu_0 n^2 \pi \left(\frac{\bar{d}}{2}\right)^2 \frac{k}{w} \text{ H} \tag{39.16}$$

where
- w is the width of the winding
- $\bar{d}$ is the average diameter
- k is Nagaoka's constant

$$k = \frac{1}{1 + 0.45(\bar{d}/w) + 0.64(t/\bar{d}) + 0.84(t/w)} \tag{39.17}$$

The distributed capacitance of the coil contributes the most to the overall antenna capacitance. The parasitic capacitances can usually be ignored. Equation 39.18 can be used to estimate the distributed capacitance of a coil:

$$C_\text{d} = \left[\frac{\varepsilon_\text{w}\varepsilon_1}{\varepsilon_\text{w} t_1 + \varepsilon_1 t_\text{w}}\right]\left[\frac{0.018544\,\bar{d}w\,(n_1 - 1)}{n_1^2}\right] \tag{39.18}$$

where
- ε_w is the dielectric constant of the wire insulation
- ε_1 is the dielectric constant of the interlayer insulation if any
- t_w is the thickness of the wire insulation
- t_1 is the thickness of the interlayer insulation
- n_1 is the number of layers

Adding a second layer to a single-layer coil significantly increases the capacitance, but as the number of layers increases, the capacitance decreases.

The air-core loop antenna is particularly useful for measuring magnetic fields with frequencies from 100 Hz to several megahertz. Because it has a linear response to magnetic field strength, it has virtually no intermodulation distortion. On the negative side, the size of the sensor can get quite large for applications that require high sensitivities at low frequencies.

39.2.3 Rod Antenna

The rod antenna is a good alternative to an air-core loop antenna. It is smaller in size than a loop antenna with the same sensitivity, and it can be designed to operate at lower frequencies. Unfortunately, its response to magnetic field strength can be nonlinear and the core adds noise.

Figure 39.4b is a typical configuration for a rod antenna. It is basically a solenoid with a magnetic core. The core can have a circular or rectangular cross section and can be made from a ferrite, a nickel–iron alloy, an amorphous metal glass alloy, or some other material with high relative permeability. The winding can be wound directly on the core or on a form through which the core is inserted. Insulation is sometimes placed between layers of the winding to reduce distributed capacitance. An electrostatic shield is placed around the winding to attenuate any electric field coupling into the signal. The shield has a gap that runs the length of the winding. This prevents circulating currents in the shield from attenuating the magnetic field within the coil.

The most common rod antenna configuration is a core with a circular cross section and a tightly coupled winding that runs most of the length of the core. The sensitivity of the rod antenna is computed by substituting μ_e in Equation 39.13 with the following:

$$\mu_e = 1 + \left(\frac{d_c}{d+t}\right)^2 (\bar{\mu} - 1) \tag{39.19}$$

where

d_c is the core diameter

$\bar{\mu}$ is the core average effective permeability

The core effective or apparent permeability depends on its geometry and initial permeability, as well as the winding length relative to the core length. A rod becomes magnetized when a magnetic field is applied to it. In response, a magnetic field is created within the rod that opposes the externally applied field and reduces the flux density. The demagnetizing field is proportional to the magnetization and the net field H in the core is

$$H = H' - NM \tag{39.20}$$

where

H' is the applied external field

N is the demagnetizing factor

M is the magnetization

The apparent relative permeability of a core is the ratio of the flux density B in the middle of the core to the flux density in air:

$$\frac{B}{\mu_0 H'} = \mu_a = \frac{\mu_i}{1 + N(\mu_i - 1)} \tag{39.21}$$

where μ_i is the initial relative permeability of the core material. Initial relative permeability is the slope of the B–H magnetization curve near-zero applied field for a closed magnetic path.

The value of N is shape dependent. As the length-to-diameter ratio m of a rod increases, N decreases and the apparent relative permeability approaches the initial permeability. Table 39.3, which is

TABLE 39.3 Demagnetizing Factors N for Rods and Ellipsoids Magnetized Parallel to Long Axis

Dimensional Ratio (Length/Diameter)	Rod	Prolate Ellipsoid	Oblate Ellipsoid
0	1.0	1.0	1.0
1	0.27	0.3333	0.3333
2	0.14	0.1735	0.2364
5	0.040	0.0558	0.1248
10	0.0172	0.0203	0.0696
20	0.00617	0.00675	0.0369
50	0.00129	0.00144	0.01472
100	0.00036	0.000430	0.00776
200	0.000090	0.000125	0.00390
500	0.000014	0.0000236	0.001576
1000	0.0000036	0.0000066	0.000784
2000	0.0000009	0.0000019	0.000392

TABLE 39.4 Magnetic Properties of Typical Core Material

Name	Composition	Manufacturer	μ_i	μ_{max}
Mild steel	0.2 C, 99 Fe		120	2,000
Silicon iron	4.0 Si, 96 Fe		500	7,000
CN20	Ni–Zn ferrite	Ceramic magnetics	800	4,500
MN60	Mn–Zn ferrite	Ceramic magnetics	5,000	10,500
"49" Alloy	48 Ni, 52 Fe	Carpenter	6,500	75,000
2605S-2	Fe-based amorphous alloy	Allied signal	10,000	600,000
4–79 Permalloy	4 Mn, 79 Ni, 17 Fe	Magnetics	20,000	100,000
Mumetal	5 Cu, 2 Cr, 77 Ni, 16 Fe	Magnetics	20,000	100,000
HyMu "80"	4.2 Mo, 80 Ni, 15 Fe	Carpenter	50,000	200,000
2826MB	NiFe-based amorphous alloy	Allied signal	100,000	800,000

Note: μ_i is the slope of the magnetization curve at the origin; μ_{max} is the maximum incremental slope of the magnetization curve.

reproduced from [6], lists demagnetizing factors for a rod, prolate ellipsoid (cigar shape), and oblate ellipsoid (disk shape).

Equation 39.22 can be used to approximate the value of N for cylindrical rods with $m > 10$ and $\mu_i > 1000$:

$$N = \frac{2.01 \times \log_{10} m - 0.46}{m^2} \tag{39.22}$$

The apparent permeability of a rod with a small m and large μ_i is almost exclusively determined by m alone. Table 39.4 lists the magnetic properties of several ferromagnetic materials that can be used to construct a core.

Bozorth [7] found that the apparent permeability of a rod is not constant throughout the length of the rod. It reaches a maximum at the center of the rod and continuously drops in value until the ends of the rod are reached. The variation in permeability can be approximated by

$$\mu(l) = \mu_a \left[1 - F\left(\frac{l}{l_0} \right)^2 \right] \tag{39.23}$$

where

l is the distance from the center of the rod to the measurement point
l_0 is the half length of the rod
F is a constant that varies from 0.72 to 0.96

The average permeability seen by the coil is the integral of Equation 39.23 over the length of the coil:

$$\bar{\mu} = \mu_a \left[1 - F\left(\frac{l_w}{l_c} \right)^2 \right] \tag{39.24}$$

where

l_w is the length of the winding
l_c is the length of the core

Equation 39.24 is substituted into Equation 39.19 to compute the rod's effective relative permeability that is used in Equation 39.13 to compute sensitivity.

The inductance of the rod antenna can be computed using the following equations:

$$L = \frac{\mu_0 \mu_e n^2 \pi (d+t)^2 l_w}{4 l_c} \tag{39.25}$$

$$\mu_e = 1 + \left(\frac{d_c}{d+t}\right)^2 \left[\mu_a f\left(\frac{l_w}{l_c}\right) - 1\right] \tag{39.26}$$

$$f\left(\frac{l_w}{l_c}\right) = 1.9088 - 0.8672\left(\frac{l_w}{l_c}\right) - 1.1217\left(\frac{l_w}{l_c}\right)^2 + 0.8263\left(\frac{l_w}{l_c}\right)^3 \tag{39.27}$$

The function $f(l_w/l_c)$ accounts for the variation in flux density from the middle of the winding to its ends and assumes the winding is centered about the middle of the core.

Equations 39.15 and 39.16 can be used to compute the resistance and capacitance of a rod antenna.

39.2.4 Signal Conditioning

To be useful, the induction coil signal must be conditioned using either a voltage or a current amplifier. Figure 39.6 illustrates the circuit configurations for both of these signal conditioning methods. The voltage amplifier can have either a single-ended or differential input and it can be tuned or untuned. The signal output of the voltage amplifier is proportional to the magnitude and frequency of the field for frequencies well below resonance. Its output will peak at the resonant frequency of the coil or at the tuning frequency. Because its output signal depends on both the frequency and strength of the magnetic field, the voltage amplifier is more suited to narrow band or tuned frequency applications.

In the current amplifier configuration, the induction coil terminals are connected to a virtual ground. As long as the product of the amplifier forward gain and the coil ohmic resistance is much greater than

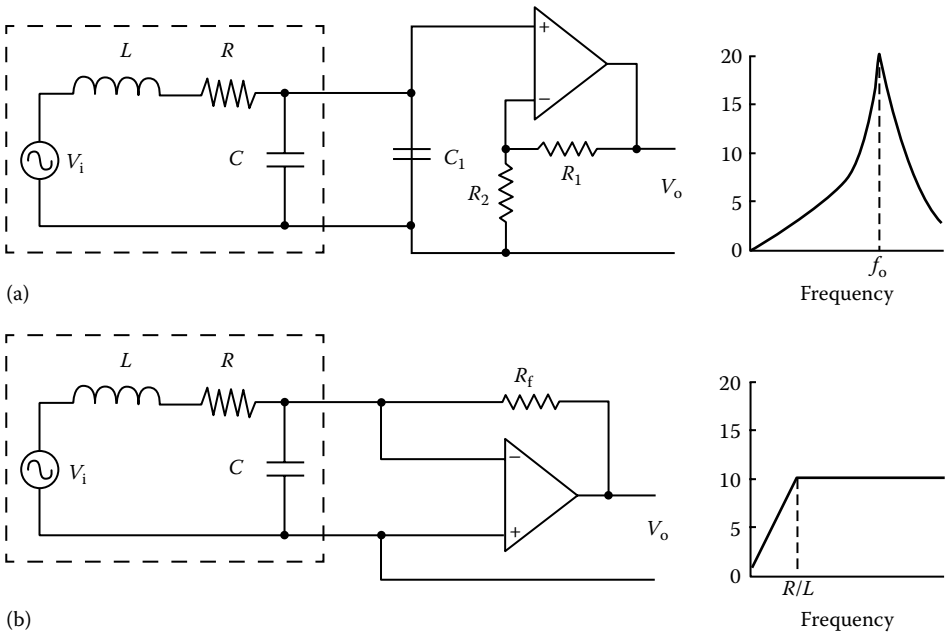


FIGURE 39.6 (a) The amplitude of a voltage-amplified induction coil signal is proportional to the frequency and strength of the field, (b) the amplitude of a current-amplified induction coil signal is only proportional to field strength beyond its L/R corner frequency.

the feedback resistor, the output signal magnitude is independent of the frequency of the magnetic field beyond the R/L (rad s^{-1}) corner of the coil. This remains true up to the coil's resonant frequency. For this reason, the current amplifier configuration is particularly suited to broadband magnetic field strength measurements. The current amplifier configuration minimizes intermodulation distortion in induction coils with magnetic cores. The current flowing through the coil produces a magnetic field that opposes the ambient field. This keeps the net field in the core near zero and in a linear region of the B–H curve.

Current amplifier–based induction coil magnetometers have been built that have a flat frequency response from 10 Hz to over 200 kHz. Some magnetometers designed for geophysical exploration applications have low-frequency corners that extend down to 0.1 Hz. For further information on this subject, refer to [8,9].

39.2.5 Fluxgate Magnetometer

The fluxgate magnetometer has been and is the workhorse of magnetic field strength instruments both on Earth and in space. It is rugged, reliable, physically small, and requires very little power to operate. These characteristics, along with its ability to measure the vector components of magnetic fields over a 0.1 nT^{-1} mT range from dc to several kHz, make it a very versatile instrument. Geologists use them for exploration and geophysicists use them to study the geomagnetic field (about 20–75 μT on Earth's surface). Satellite engineers use them to determine and control the attitude of spacecraft, scientists use them in their research, and the military uses them in many applications, including mine detection, vehicle detection, and target recognition. Some airport security systems use them to detect weapons.

39.2.6 Fluxgate

The heart of the magnetometer is the *fluxgate*. It is the transducer that converts a magnetic field into an electric voltage. There are many different fluxgate configurations. Two of the more popular ones are shown in Figure 39.7. A very comprehensive explanation of the fluxgate principle and the different fluxgate configurations is given in [10].

The ring-core fluxgate is constructed from a thin ribbon of easily saturable ferromagnetic material, such as 4–79 permalloy wrapped around a bobbin to form a ring or toroid. As shown in Figure 39.8, an alternating current is applied through a coil that is wound about the toroid. This creates a magnetic field that circulates around the magnetic core. This magnetic field causes the flux in the ferrous material to periodically saturate first clockwise and then counterclockwise. A pickup (signal) winding is wrapped around the outside of the toroid. While the ferrous material is between saturation extremes, it maintains an average permeability much greater than that of air. When the core is in saturation, the core permeability becomes equal to that of air. If there is no component of magnetic field along the axis of the signal winding, the flux change seen by the winding is zero. If, on the other hand, a field component is present along the signal winding axis, then each time the ferrous material goes from one saturation extreme to the other, the flux within the core will change from a low level to a high level. According to Faraday's law, a changing flux will produce a voltage at the terminals of the signal winding that is proportional to the rate of change of flux. For dc and low-frequency magnetic fields, the signal winding voltage is

$$e(t) = nA\frac{d(\mu_0\mu_e H)}{dt} = nA\mu_0 H\frac{d\mu_e(t)}{dt} \tag{39.28}$$

where

H is the component of the magnetic field being measured
n is the number of turns on the signal winding
A is the cross-sectional area of the signal winding
$\mu_e(t)$ is the effective relative permeability of the core

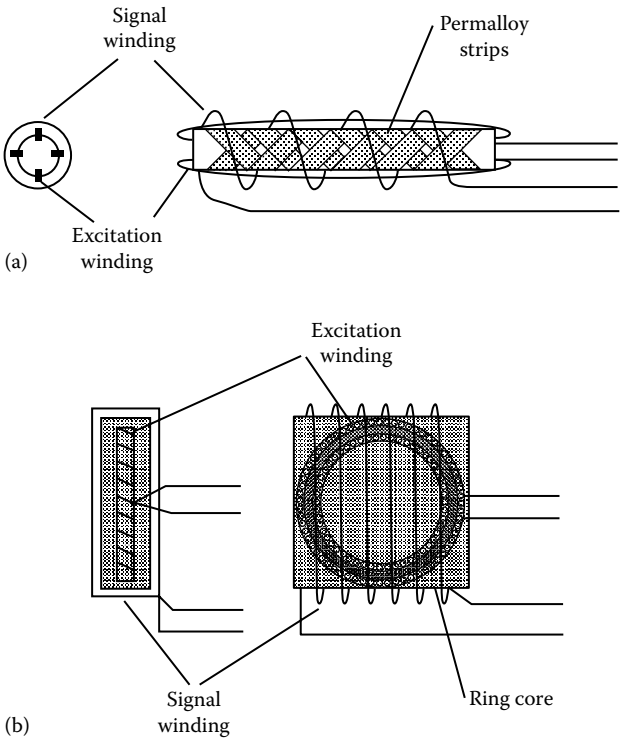


FIGURE 39.7 In Schonstedt (a) and ring-core and (b) fluxgate sensors, the excitation field is at right angles to the signal winding axis. This configuration minimizes coupling between the excitation field and the signal winding.

As the core permeability alternates from a low value to a high value, it produces a voltage pulse at the signal winding output that has amplitude proportional to the magnitude of the external magnetic field and a phase indicating the direction of the field. The frequency of the signal is twice the excitation frequency since the saturation-to-saturation transition occurs twice each excitation period.

The discussion about effective permeability in the induction coil section applies here as well. Readers can consult [10,13,14] for comprehensive discussions about fluxgate effective permeability and signal characteristics as they relate to excitation field levels, excitation waveforms, and winding geometry.

39.2.7 Signal Conditioning

The signal from the fluxgate is an amplitude-modulated suppressed carrier signal that is synchronous with the second harmonic of the excitation signal. In a simple low-power magnetometer, this signal is converted to the base band using a synchronous demodulator, filtered, and presented as the final output. Example circuits are given in [11,12]. The accuracy of magnetometers that use this open-loop architecture is limited by the linearity of the core's magnetization curve and is about 5% for Earth's field (60 μT) applications.

More precise and stable magnetometers use magnetic field feedback rather than the open-loop structure described earlier. A simplified schematic of a typical second harmonic field feedback fluxgate magnetometer is shown in Figure 39.9. The circuitry to the left of the fluxgate is called the excitation circuit. It consists of an oscillator tuned to twice the excitation frequency, a flip-flop that divides the oscillator frequency by two, and a power amplifier driven by the flip-flop and, in turn, provides the excitation current to the excitation winding.

The circuitry to the right of the fluxgate is called the signal channel circuit. It amplifies the output from the fluxgate signal winding, synchronously demodulates the ac signal using the oscillator signal

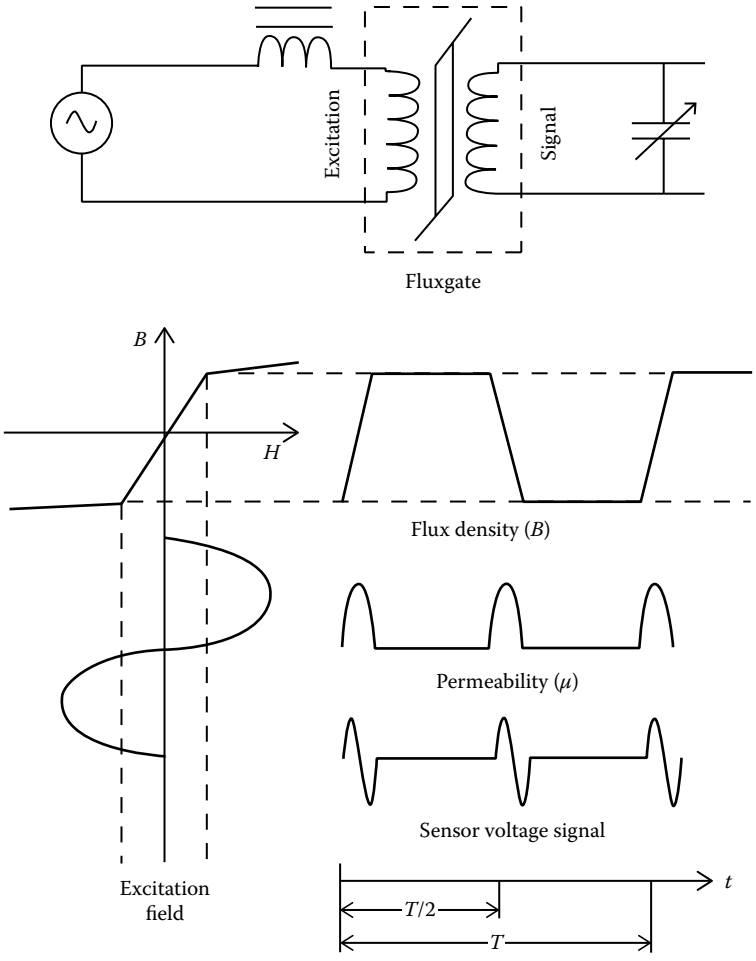


FIGURE 39.8 The excitation field of a fluxgate magnetometer alternately drives the core into positive or negative saturation, causing the core's effective permeability to switch between 1 and a large value twice each cycle.

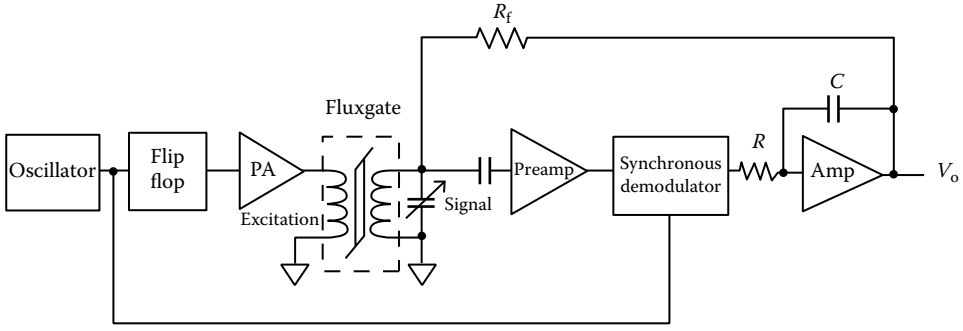


FIGURE 39.9 Typical circuit configurations for a field feedback fluxgate magnetometer. The sensor output is ac amplified, synchronously demodulated, and filtered. A magnetic field that nulls the ambient field at the sensor is produced by connecting the resistor R_f, between the output and the signal winding.

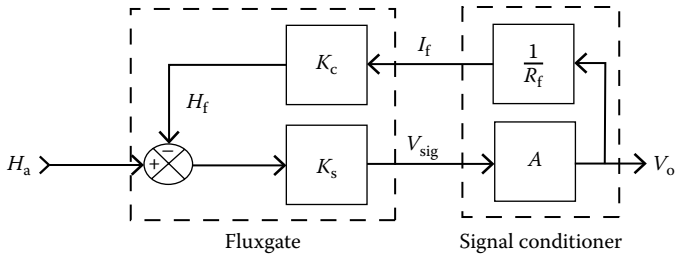


FIGURE 39.10 Block diagram of a field feedback fluxgate magnetometer. K_c is the current-to-field constant for the coil. K_s is the field-to-voltage transduction constant for the sensor. The feedback field H_f opposes the ambient field H_a, thus keeping the net sensor field very small.

as a reference, integrates and amplifies the base band output, and then feeds back the output through a resistor to the signal winding. The feed-back signal produces a magnetic field inside the sensor that opposes the external field. This keeps the field inside the sensor near zero and in a linear portion of the magnetization curve of the ferromagnetic core.

The flow diagram for the magnetometer is given in Figure 39.10. The external field H_a is opposed by the feedback field H_f, and the difference is converted into a voltage signal (K_s represents the transfer function from field to voltage). This signal is amplified (A), and the amplified signal is converted into a current I_f and then into the feedback field (K_c represents the transfer function from current to field). The overall transfer function for the magnetometer is

$$\frac{V_0}{H_a} = \frac{AK_s}{1 + (K_c AK_s / R_f)} \tag{39.29}$$

The amplifier gain is normally very high such that the second term in the denominator is much larger than one, and Equation 39.29 reduces to

$$\frac{V_0}{H_a} = \frac{R_f}{K_c} \tag{39.30}$$

Under these circumstances, the transfer function becomes almost completely determined by the ratio of R_f (the feedback resistor) to K_c (the current-to-field coil constant of the sensor winding). Both of these constants can be very well controlled. The consequence of this circuit topology is a highly stable and accurate magnetometer that is insensitive to circuit component variations with temperature or time. An accuracy of 1% over a temperature range of –80°C to 80°C is easily achievable. Accuracy and stability can be improved using a current feedback circuit, like the one described in [13] that compensates for the resistance of the signal winding or by using a separate feedback winding and a high-quality voltage-to-current converter instead of a simple feedback resistor.

39.2.8 SQUID Magnetometer

Brian D. Josephson in 1962, while a graduate student at Cambridge University, predicted that superconducting current could flow between two superconductors that are separated by a thin insulation layer. The magnitude of the superconductor (critical) current through this "Josephson junction" is affected by the presence of a magnetic field and forms the basis for the SQUID magnetometer.

Magnetometers based on the SQUID are currently the most sensitive instruments available for measuring magnetic field strength. In many cases, SQUID instrumentation offers the ability to make measurements where no other methodology is possible. Sensitivities below 1 fT/√Hz have

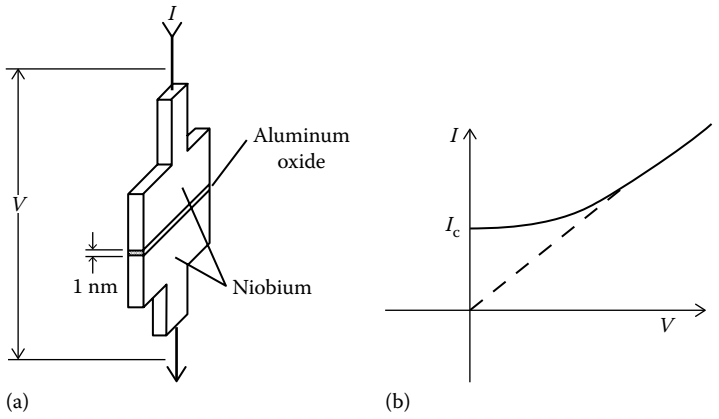


FIGURE 39.11 The Josephson junction in (a) consists of a superconductor such as niobium separated by a thin insulation layer. The voltage (V) versus current (I) curve in (b) shows that a superconducting current flows through the junction with zero volts across the junction.

been achieved. The superconducting nature of SQUID devices allows for flat frequency and phase response from dc to GHz frequencies. Chapters 5 and 14 in this book describe the principles behind SQUID devices and applications including laboratory instrumentation, nondestructive test and evaluation, geophysics, and medicine.

Figure 39.11 illustrates the general structure of a Josephson junction and the voltage–current (V–I) relationship. Two superconductors (e.g., niobium) are separated by a very thin insulating layer (e.g., aluminum oxide). The thickness of this layer is typically 1 nm. When the temperature of the junction is reduced to below 4.2 K (–269°C), a superconductor current will flow in the junction with 0 V across the junction. The magnitude of this current, called the critical current I_c, is a periodic function of the magnetic flux present in the junction. Its maximum magnitude occurs for flux values equal to $n\phi_0$ where ϕ_0 is one flux quantum (2 fW), and its minimum magnitude occurs for flux values equal to $(n + (1/2))\phi_0$. The period is one flux quantum. This phenomenon is called the "dc Josephson effect" and is only one of the "Josephson effects."

Magnetometers based on the SQUID are currently the most sensitive instruments available for measuring magnetic field strength. SQUID magnetometers measure the change in the magnetic field from some arbitrary field level; they do not intrinsically measure the absolute value of the field. Biomedical research is one of the more important applications of SQUID magnetometers. SQUID magnetometers and gradiometers (measure spatial variation in the magnetic field) have the high sensitivities needed to measure the weak magnetic fields generated by the body [15]. Other application areas include paleomagnetics (measuring remanent magnetism in rocks) and magnetotellurics (Earth resistivity measurements). Descriptions of these applications as well as the general theory of SQUIDs can be found in [16]. Clark [17], one of the pioneers in SQUID magnetometers, provides a good contemporary overview of SQUID technology and applications.

A dc SQUID magnetometer uses two Josephson junctions in the two legs of a toroid as shown in Figure 39.12a. The toroid is biased with a constant current that exceeds the maximum critical current of the junctions. When the flux through the toroid is an integral multiple of ϕ_0, the voltage across the junctions is determined by the intersection of I_b and the $n\phi_0$ $V - I$ curve (point A). As the flux increases, the critical current decreases. The V–I curve and thus the intersection point move to the right (the junction voltage increases). The critical current reaches a minimum when the flux has increased by $(1/2)\phi_0$ and the junction voltage is at its maximum (point B). As the flux continues to increase, the critical current increases back toward its maximum value and the junction voltage decreases. Thus, the period of the flux cycle is ϕ_0.

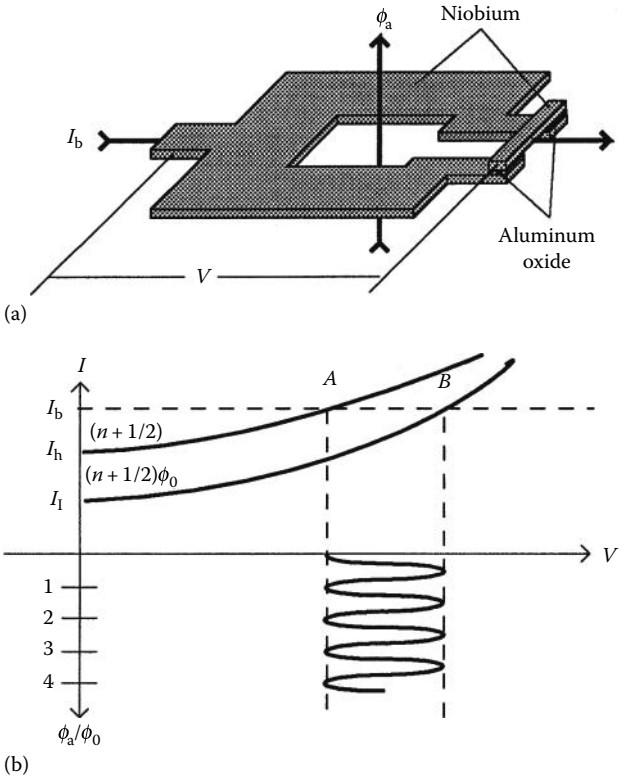


FIGURE 39.12 Use of a dc SQUID to measure magnetic flux. The dc SQUID in (a) consists of a superconductor loop and two Josephson junctions with a bias current that is greater than the maximum critical current I_h. The V–I curve in (b) illustrates how the voltage across the SQUID oscillates with a period equal to one flux quantum φ_0.

39.2.9 Signal Conditioning

Figure 39.13 is a block diagram of one implementation of a dc SQUID magnetometer that can be used for wide dynamic range field measurements. A large superconducting loop, which is exposed to the magnetic field being measured, is connected to a multiturn signal winding that is magnetically coupled directly to the SQUID. At cryogenic temperatures, the loop and signal winding effectively form a dc induction coil. External flux applied to the coil will generate a current in the loop that keeps the net flux within the loop constant, even for dc magnetic fields. The signal winding magnifies the flux that is applied to the SQUID.

The SQUID is magnetically biased at an optimal sensitivity point. A small ac magnetic field at 100–500 kHz is superimposed on the bias field. The output of the SQUID is a suppressed carrier amplitude-modulated signal where the amplitude indicates the change in magnetic field from the bias point, and the phase indicates the polarity of the change. The output signal is amplified and then synchronously demodulated down to the base band. The resulting dc signal is amplified and feed back through a resistor to a coil coupled to the SQUID. The current through the coil generates a magnetic field at the SQUID that opposes the applied field. This keeps the SQUID operating point very near the bias point. The scale factor of the magnetometer depends on the feedback resistor and the coil constant of the feedback winding in the same manner that it does for a field feedback fluxgate magnetometer.

The pickup loop, signal coil, SQUID, feedback coil, and feedback resistor are kept in a cryogenic temperature chamber and, except for the pickup coil, are magnetically shielded. The rest of the circuit is operated at room temperature.

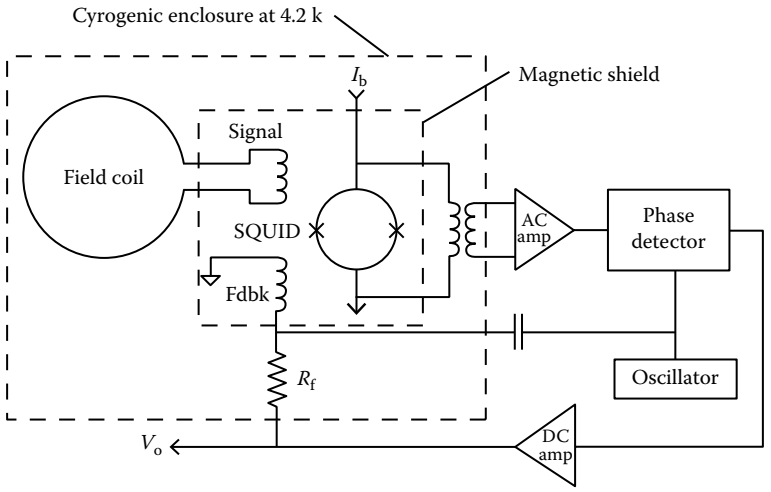


FIGURE 39.13 Wide dynamic range dc SQUID magnetometer. A magnetic field produced by connecting resistor R_f between the output and a feedback coil keeps the field in the SQUID within one flux quantum over its operating range. (Adapted from Wellstood, F. et al., *Rev. Sci. Instrum.*, 66, 952, 1984.)

39.3 High-Field Vector Gaussmeters

39.3.1 Hall Effect Gaussmeter

The Hall effect device, which is probably the most familiar and widely used sensor for measuring strong magnetic fields, is based on the discovery of the Hall effect by Edwin H. Hall in 1897. The Hall effect is a consequence of the Lorentz force law, which states that a moving charge q, when acted upon by a magnetic induction field $\vec{B}$, will experience a force $\vec{F}$ that is at right angles to the field vector and the velocity vector v of the charge as expressed by the following equation:

$$\vec{F} = -q(\vec{B} + \vec{v} \times \vec{B}) \tag{39.31}$$

The Hall effect device consists of a flat, thin rectangular conductor or semiconductor with two pairs of electrodes at right angles to one another as illustrated in Figure 39.14. An electric field E_x is applied along the x or control axis. When a magnetic field B_z is applied perpendicular to the surface of the device, the free charge, which is flowing along the x-axis as a result of E_x, will be deflected toward the y or Hall voltage axis.

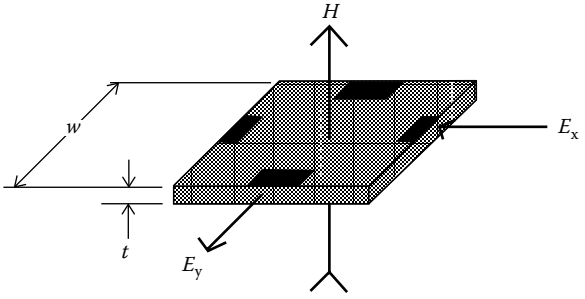


FIGURE 39.14 Hall effect sensor. A magnetic field H applied normal to the surface of the sensor, which is conducting current along the x-direction, will generate a voltage along the y-direction. E_x is the applied electric field along the x-direction, and E_y is the Hall effect electric field along the y-direction.

Since current cannot flow in the y-axis under open-loop conditions, this will cause a buildup of charge along the y-axis that will create an electric field that produces a force opposing the motion of the charge:

$$E_y = v_x B_z \tag{39.32}$$

where v_x is the average drift velocity of the electrons (or majority carriers). In a conductor that contains n free charges per unit volume having an average drift velocity of v_x, the current density is

$$J_x = qnv_x \tag{39.33}$$

and

$$E_y = \frac{J_x B_z}{qn} = R_H J_x B_z \tag{39.34}$$

where R_H is called the Hall coefficient.

A semiconductor is treated in terms of the mobility μ (drift velocity/field) of the majority carrier (electron or hole) and conductivity σ. In this case,

$$E_y = \mu E_x B_z \quad \text{and} \quad E_x = \frac{J_x}{\sigma} \tag{39.35}$$

Therefore,

$$E_y = \frac{\mu}{\sigma} J_x B_z \quad \text{and} \quad R_H = \frac{\mu}{\sigma} \tag{39.36}$$

The value of R_H varies substantially from one material to another and is both temperature and field magnitude dependent. Its characteristics can be controlled to a certain extent by doping the base material with some impurities. For example, doping germanium with arsenic can reduce the temperature dependence at the expense of magnitude.

The voltage measured across the y-axis terminals is the integral of the electric field along the y-axis. If a constant control current I is flowing along the x-axis, then

$$J_x = \frac{I}{wt} \tag{39.37}$$

and the measured output voltage is

$$e_y = \frac{R_H I B_z}{t} \tag{39.38}$$

where

t is thickness (m)

w is the distance between the y-axis terminals

Another characteristic specified by manufacturers of Hall effect devices is the magnetic sensitivity γ_b at the rated control current I_c:

$$\gamma_b = \frac{e_y}{B_z} = \frac{R_H I_c}{t} \tag{39.39}$$

Although conductors such as copper (Cu) can be used to make a Hall effect device, semiconductor materials, such as gallium arsenide (GaAs), indium antimonide (InSb), and indium arsenide (InAs), produce the highest and most stable Hall coefficients. InAs, because of its combined low temperature coefficient

of sensitivity (<0.1%/°C), low resistance, and relatively good sensitivity, is the material favored by commercial manufacturers of Hall effect devices.

The typical control current for Hall effect devices is 100 mA, but some do operate at currents as low as 1 mA. Sensitivities range from 10 mV/T to 1.4 V/T. Linearity ranges from 1/4% to 2% over their rated operating field range. The control input and the voltage output resistance are typically in the range of 1–3 Ω. The sensor element is usually tiny (on the order of 10 mm square by 0.5 mm thick), and a three-axis version can be housed in a very small package. These devices are most effective for measuring flux densities ranging from 50 μT to 30 T.

39.3.2 Signal Conditioning

A simple Hall effect gaussmeter can be constructed using the signal conditioning circuit shown in Figure 39.15. The voltage reference, operational amplifier, and sense resistor R_s form a precision constant-current source for the Hall effect device control current I_c. For best performance, the voltage reference and R_s, should be very stable with temperature and time. A general-purpose operational amplifier can be used for low control currents. A power amplifier is required for control currents above 20 mA.

The Hall voltage can be conditioned and amplified by any high input impedance (>1 kΩ) differential amplifier. A precision instrumentation amplifier is a good choice because it has adequate input impedance, its gain can be determined by a stable resistor, and the amplifier zero offset trim resistors can be used to cancel the zero offset of the Hall effect device. Some devices require a load resistor across the Hall voltage terminal to achieve optimum linearity.

The zero offset and 1/*f* noise of the Hall voltage amplifier limit the performance of a Hall effect gaussmeter for low-field strength measurements. Sometimes, these effects can be reduced by using an ac precision current source. The ac amplitude-modulated Hall voltage can then be amplified in a more favorable frequency band and synchronously detected to extract the Hall voltage signal. If the field to be measured requires this amount of signal conditioning, it probably is better to use a fluxgate magnetometer for the application.

39.3.3 Magnetoresistive Gaussmeter

The magnetoresistance effect was first reported by William Thomson (Lord Kelvin) in the middle of the nineteenth century. He found that a magnetic field applied to a ferromagnetic material caused its resistivity to change. The amount of change depends on the magnetization magnitude and the direction in which the current used to measure resistivity is flowing. Nickel–iron alloys show the greatest change

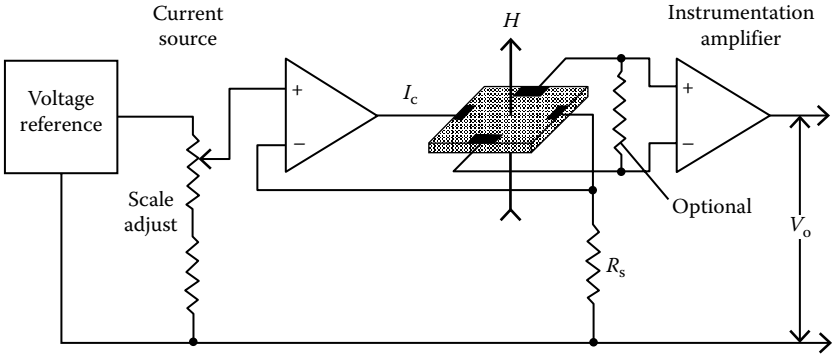


FIGURE 39.15 Example of how to construct a Hall effect gaussmeter. The operational amplifier and resistor R_s form a stable constant-current source for the Hall effect sensor. An instrumentation or differential amplifier amplifies and scales the Hall voltage. A load resistor is sometimes required across the Hall voltage output terminals.

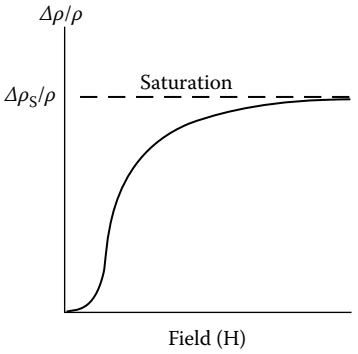


FIGURE 39.16 Change in resistivity in a ferromagnetic material. As field is applied, the resistivity changes rapidly at first. As the material approaches magnetic flux saturation, the resistivity change approaches its maximum value.

in resistivity (about 5% maximum). Figure 39.16 illustrates how the resistivity changes in permalloy (a nickel–iron alloy) for a field applied parallel to the current flow. As magnetic field is increased, the change in resistivity increases and asymptotically approaches its maximum value when the material approaches saturation. Bozorth [6] points out that the shape of the curve and the magnitude of the change depend on the composition of the alloy. Permalloy with 80% Ni and 20% Fe provides a high magnetoresistance effect with near-zero magnetostriction and is a favorite material for magnetoresistors.

The change in resistivity in permalloy film [18] is also a function of the angle θ between the magnetization direction and the current direction:

$$\rho(\theta) = \rho_0 + \Delta\rho_m \cos^2\theta \tag{39.40}$$

where

$\Delta\rho_m$ is the magnetoresistivity anisotropy change
ρ_0 is the resistivity for $\theta = \pi/2$

It was mentioned earlier that magnetic materials have anisotropic magnetic properties (their magnetic properties are direction dependent). The physical shape of an object (see the discussion on demagnetizing factor earlier) and the conditions that exist during fabrication strongly determine its anisotropic characteristics. A thin long film of permalloy can be made to have highly uniaxial anisotropic properties if it is exposed to a magnetizing field during deposition. This characteristic is exploited in the anisotropic magnetoresistance (AMR) sensor.

The basic resistor element in an AMR is a thin rectangular shaped film as shown in Figure 39.17. One axis, called the anisotropy or easy axis, has a much higher susceptibility to magnetization than

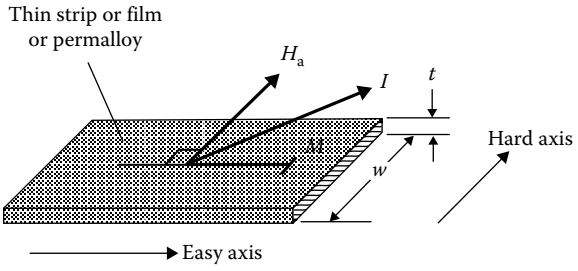


FIGURE 39.17 An AMR resistor element. During fabrication, a magnetic field is applied along the strip's length to magnetize it and establish its easy axis. Current I is passed through the film at 45° to the easy or anisotropic axis. A magnetic field H_a applied at right angles to the magnetization vector M causes the magnetization vector to rotate and the magnetoresistance to change.

the other two. The easy axis is normally along the length of the film. Because of its thinness, the axis normal to the film has virtually no magnetic susceptibility. The axis transverse to the easy axis (across the width) has very little susceptibility as well.

A bias field H_b is used to saturate the magnetization along the easy axis and establish the magnetization direction for zero external fields. For a simplified analysis, the film can be modeled as a single domain.

The effect of an external field in the plane of the film and normal to the anisotropy axis is to rotate the magnetization vector and, according to Equation 39.40, change the resistivity. Kwiatkowski and Tumanski [19] stated that the change in resistance of the film can be approximated by Equation 39.41:

$$\Delta R \approx R_s \frac{\Delta\rho_m}{\rho}\left(h_a^2\cos 2\theta + h_a\sqrt{1-h_a^2}\,\sin 2\theta - \frac{1}{2}\cos 2\theta\right) \tag{39.41}$$

where

h_a is the normalized externally applied field (i.e., $h_a = H_a/H_k$)
R_s is the nominal resistance
$\Delta\rho_m/\rho$ is the maximum resistivity change

H_k is the anisotropy field. Optimum linear performance is obtained when $\theta = \pi/4$ and Equation 39.41 reduces to

$$\Delta R \approx R_s \frac{\Delta\rho_m}{\rho}\frac{1}{H_k + H_b}H_a \tag{39.42}$$

The anisotropy field is given by

$$H_k = \sqrt{H_{k0}^2 + (NM_s)^2} \tag{39.43}$$

where

H_{k0} is the film anisotropy field
N is the demagnetizing factor (≈ thickness(t)/width(w))
M_s is the saturation magnetization

An AMR is constructed using long thin film segments of deposited permalloy. During deposition, a magnetic field is applied along the length of the film to establish its easy axis of magnetization. The shape of the film also favors the length as an easy axis. As shown in Figure 39.18, a series of these permalloy films is connected together to form the magnetoresistor. The current is forced to flow at a 45° angle to the easy axis by depositing thin strips of highly conductive material (e.g., gold) across the permalloy film. The level of magnetization of the film is controlled by a bias field that is created through the deposition of a thin layer of cobalt over the resistors, which is then magnetized parallel to the easy axis of the permalloy.

A typical AMR sensor suitable for a gaussmeter or magnetometer consists of four AMRs connected in a Wheatstone bridge as shown in Figure 39.19. The transfer function polarity of the *A* and *D* resistors is made to be opposite that of the *B* and *C* resistors by rotating the current shunt 90°. This complimentary arrangement enhances the output voltage signal for a given field by a factor of four over a single resistor. Kwiatkowski and Tumanski [19] showed that the transfer function for the bridge configuration is described by

$$v = IR_s\frac{\Delta\rho_m}{\rho}\cos 2\Delta\varepsilon h_a\sqrt{1-h_a^2} \tag{39.44}$$

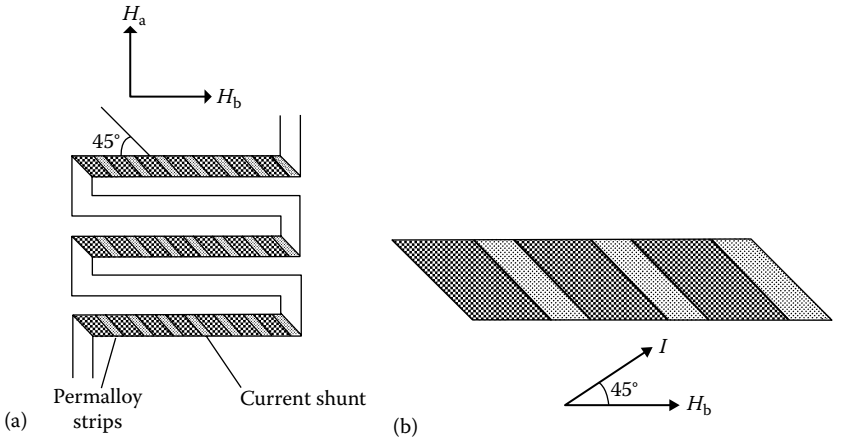


FIGURE 39.18 Magnetoresistor construction: (a) A typical AMR element consists of multiple strips of permalloy connected together in a serpentine pattern. Current shunts force the current to flow through the permalloy at 45° to the easy axis, and (b) a close-up view.

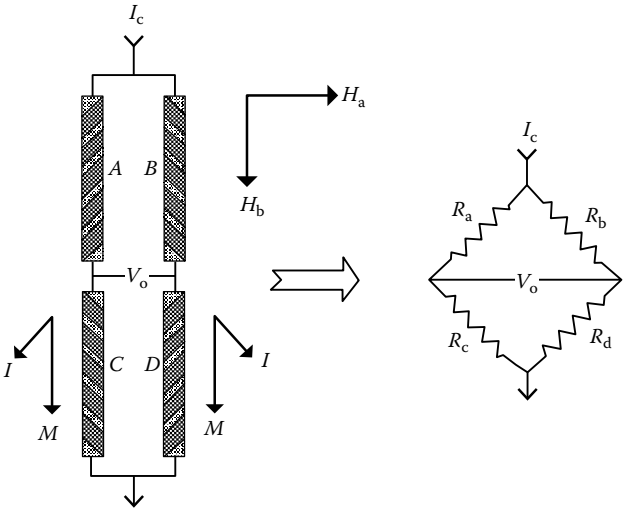


FIGURE 39.19 AMR bridge sensor. In an AMR bridge, the current shunts of resistors *A* and *D* are the same and reversed from *B* and *C*. Thus, the resistors on diagonal legs of the bridge have the same response to an applied field and opposite that of the other diagonal pair. Bridge leg resistance varies from 1 to 100 kΩ.

where

$$\cos 2\Delta\varepsilon = \frac{H_{k0}^2 + H_k^2 - (NM_s)^2}{2H_{k0}H_k} \tag{39.45}$$

$$h_a = \frac{H_a}{H_k + H_b} \tag{39.46}$$

For best linearity, $H_a < 0.1\ H_k$. The linearity of the bridge can be controlled during fabrication by adjusting the *l*/*w* ratio and H_{k0}. The bias field can also be used to optimize linearity and establish the measurement field range. Some transfer functions for a typical AMR bridge [20] are shown in Figure 39.20. A more comprehensive discussion of AMR theory can be found in [21–23].

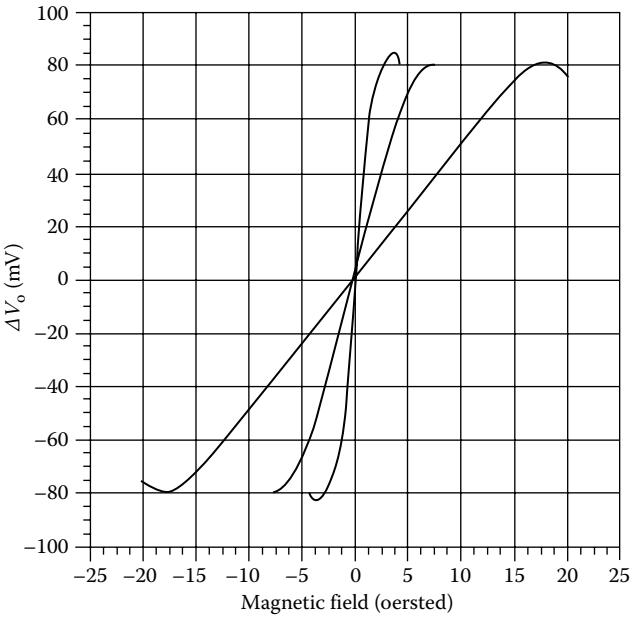


FIGURE 39.20 Typical AMR bridge sensor transfer functions. The sensitivity of an AMR bridge can be adjusted by changing its bias field. Increases in sensitivity are accompanied by corresponding decreases in range.

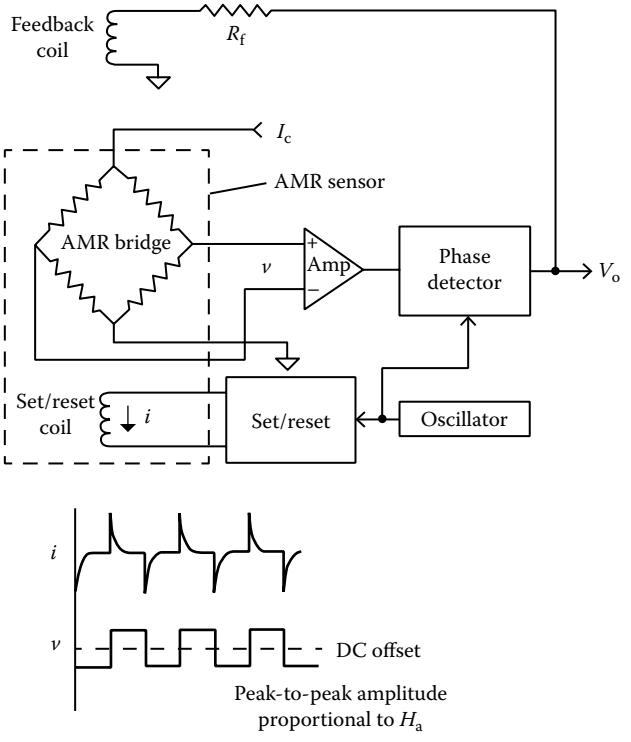


FIGURE 39.21 Example AMR gaussmeter. The magnetization direction can be alternately flipped to eliminate zero offset. The resulting ac signal can then be amplified and synchronously phase detected to recover the field-related signal. Optionally, the range and stability of the AMR gaussmeter can be increased by connecting the output voltage through a resistor to a feedback coil that produces a field that nulls the applied field.

39.3.4 Signal Conditioning

Conventional Wheatstone bridge signal conditioning circuits can be used to process the AMR bridge. The bridge sensitivity and zero offset are proportional to the bridge voltage, so it is important to use a well-regulated supply with low noise and good temperature stability.

As Equation 39.44 shows, the polarity of the transfer function is determined by the polarity of $(H_k + H_b)$. If the sensor is exposed to an external field that is strong enough to reverse this field, then the transfer function polarity will reverse. To overcome this ambiguity, the polarity should be established prior to making a measurement. This can be accomplished by momentarily applying a strong magnetic field along the easy axis of the AMR bridge. Some commercial AMR bridges come with a built-in method for performing this action.

Figure 39.21 is a block diagram for a signal conditioner that takes advantage of the bias field polarity flipping property to eliminate zero offset errors and low-frequency $1/f$ noise. A square wave oscillator is used to alternately change the direction of the bias field and thus the polarity of the transfer function. The duration of the current used to set the bias field direction should be short in order to minimize power consumption. The amplitude of the ac signal from the bridge is proportional to the field magnitude, and its phase relative to the oscillator gives the field direction. This signal can be amplified and then phase detected to extract the field-related voltage. Optionally, the output signal can be fed back through a coil that produces a magnetic field opposing the field being measured. This feedback arrangement makes the AMR bridge a null detector and minimizes the influence of changes in its transfer function on overall performance. Of course, the added circuitry increases the size, cost, and complexity of the instrument.

39.4 Scalar Magnetometers

Scalar magnetometers measure the magnitude of the magnetic field vector by exploiting the atomic and nuclear properties of matter. The two most widely used scalar magnetometers are the proton precession and the optically pumped magnetometer. When operated under the right conditions, these instruments have extremely high resolution and accuracy and are relatively insensitive to orientation. They both have several common operating limitations. The instruments require the magnetic field to be uniform throughout the sensing element volume. They have a limited magnetic field magnitude measurement range: typically 20–100 μT. And they have limitations with respect to the orientation of the magnetic field vector relative to the sensor element.

The proton precession magnetometer uses a strong magnetic field to polarize the protons in a hydrocarbon and then detects the precession frequency of the protons while they decay to the nonpolarized state after the polarizing field is turned off. The precession frequency is proportional to the magnitude of any ambient magnetic field that is present after the polarizing field is removed. This sampling of the magnetic field strength through the polarize-listen sequence makes the proton precession magnetometer response very slow. Maximum rates of only a few samples per second are typical. Because of its dependence on atomic constants, the proton precession magnetometer is the primary standard for calibrating systems used to generate magnetic fields and calibrate magnetometers.

The optically pumped magnetometer is based on the Zeeman effect. Zeeman discovered that applying a field to atoms, which are emitting or absorbing light, will cause the spectral lines of the atoms to split into a set of new spectral lines that are much closer together than the normal lines. The energy-related frequency interval between these hyperfine lines is proportional to the magnitude of the applied field. These energy levels represent the only possible energy states that an atom can possess. The optically pumped magnetometer exploits this characteristic by optically stimulating atoms to produce an overpopulated energy state in one of the hyperfine spectral lines and then causing the energy state to depopulate using a radio frequency (RF) magnetic field. The RF required to depopulate the energy state is equal to the spectral difference of the hyperfine lines produced by a magnetic field

and, therefore, is proportional to the magnetic field strength. The optically pumped magnetometer can be used to sample the magnetic field at a much higher rate than the proton precession magnetometer and generally can achieve a higher resolution. The sample rate and instrument resolution are interdependent.

39.4.1 Proton Precession Magnetometer

The proton precession magnetometer works on the principle that a spinning nucleus, which has both angular momentum $\vec{L}$ and a magnetic moment $\vec{\mu}_p$, will precess about a magnetic field like a gyroscope, as shown in Figure 39.22. The precession frequency ω_ρ is proportional to the applied field. When the magnetic field $\vec{H}_a$ is applied to the nucleus, it will produce a torque:

$$\vec{T} = \vec{\mu}_\rho \times \vec{H}_a \tag{39.47}$$

on the nucleus. Because the nucleus has angular momentum, this torque will cause the nucleus to precess about the direction of the field. At equilibrium, the relationship between the torque, precession rate, and angular momentum is

$$\mu_\rho \times \vec{H}_a = \vec{\omega}_\rho \times \vec{L} \tag{39.48}$$

Solving for the magnitude of the (Larmor) precession frequency, one finds that

$$\omega_\rho = \left(\frac{\mu_\rho}{L}\right) H_a = \gamma H_a \tag{39.49}$$

where γ is called the gyromagnetic ratio and equals $(2.6751526 \pm 0.0000008) \times 10^{-8}$ T^{-1} s^{-1}.

Figure 39.23 is a block diagram of a proton precession magnetometer. The sensor is a container of hydrocarbon rich in free hydrogen nuclei. A solenoid wrapped around the container is used to both polarize the nuclei and detect the precession caused by the ambient field. Before the polarizing field is applied, the magnetic moments of the nuclei are randomly oriented, and the net magnetization is zero. Application of the polarizing field (typically 3–10 mT) causes the nuclei to precess about the field. The precession axis can be parallel or antiparallel (nuclear magnetic moment pointing in the direction of the field) to the applied field. From a quantum mechanical standpoint, the antiparallel state is a lower energy level than the parallel state. In the absence of thermal agitation, which causes collisions between atoms, the fluid would remain unmagnetized. When a collision occurs, the parallel precession-axis nuclei lose energy and switch to the antiparallel state. After a short time, there are more nuclei with magnetic

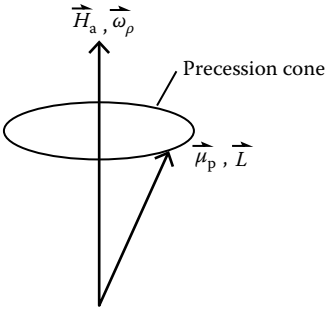


FIGURE 39.22 Nuclear precession. A spinning proton with angular momentum L and magnetic moment μ_ρ, when subjected to a magnetic field H_a, will precess about the field at an angular rate ω_ρ equal to $\mu_\rho H_a/L$.

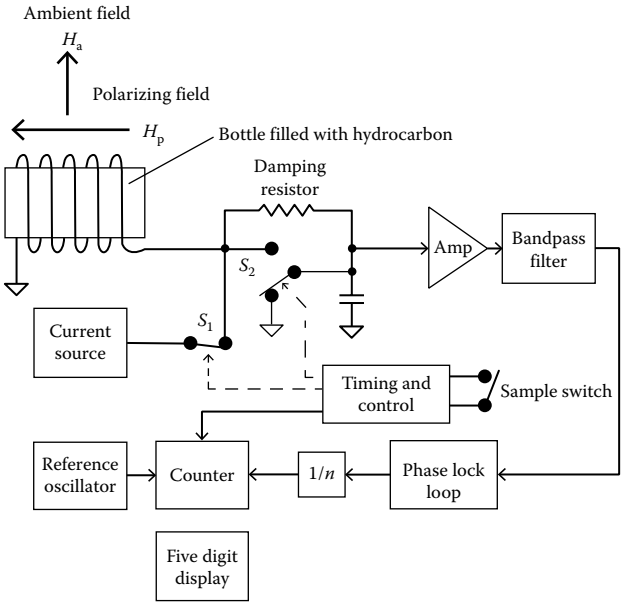


FIGURE 39.23 Typical proton precession magnetometer. A polarizing field is applied to the hydrocarbon when *S*1 is closed. The amplifier input is shorted to prevent switching transients from overdriving it. After a few seconds, *S*1 is opened and the coil is connected to the signal processor to measure the Larmor frequency.

moments pointing in the direction of the field than away from it, and the fluid reaches an equilibrium magnetization M_0. The equation that relates magnetization buildup to time is

$$M(t) = M_0(1 - e^{-t/\tau_e}) \tag{39.50}$$

where τ_e is the spin–lattice relaxation time.

The equilibrium magnetization is based on thermodynamic considerations. From Boltzmann statistics for a system with spins of 1/2,

$$\frac{n_p}{n_a} = e^{2\mu_\rho H_a/kT} \tag{39.51}$$

where

n_p is the number of precession spin axes parallel to H_a
n_a is the number of precession spin axes antiparallel to H_a
k is Boltzmann's constant
T is temperature (kelvin)

If n is the number of magnetic moments per unit volume, then

$$n = n_p + n_a = n_a(1 + e^{2\mu_\rho H_a/kT}) \tag{39.52}$$

and

$$M_0 = (n_p - n_a)\mu_\rho \approx \frac{n\mu_\rho^2 H_a}{kT} \tag{39.53}$$

Once the fluid has reached equilibrium magnetization, the field is removed and the nuclei are allowed to precess about the local ambient field until they become randomized again. This process of excitation–relaxation can take as long as several seconds.

The hydrocarbon spin–lattice relaxation time can be adjusted by mixing paramagnetic salts, such as ferrous nitrate, into the solution. The trade-off in reduced relaxation time is increased signal-to-noise and resolution. Benzene is a good general-purpose hydrocarbon that can be used in a proton precession magnetometer.

39.4.2 Signal Conditioning

The block diagram shown in Figure 39.23 is an example of the signal conditioning required for a proton precession magnetometer. The coil surrounding the bottle containing the hydrocarbon serves two purposes. At the beginning of a measurement, the current source is connected to the coil to generate the magnetic field that polarizes the fluid. This field is on the order of 10 mT. After a few seconds, the current source is disconnected and the coil, which now has a decaying nuclear precession signal at its output, is connected to the input of the amplifier. The signal is amplified and filtered and then the period of the Larmor frequency is measured, averaged, scaled, and presented to the user in magnetic field units on a digital display.

The scale factor of the proton precession magnetometer is based on the gyromagnetic ratio, which is 0.042579 Hz nT^{-1}. High resolution, up to 0.01 nT, is achieved by measuring the period of the signal rather than the frequency. The signal frequency can be divided down and used to gate a high-frequency oscillator that is driving a counter.

The sampling of the field is controlled manually in many commercially available proton precession magnetometers. Some magnetometers have an internally controlled sample rate. The sample rate and resolution are inversely related to one another. A higher sample rate produces a poorer resolution.

39.4.3 Optically Pumped Magnetometer

As explained earlier, the optically pumped magnetometer is based on the Zeeman effect. This effect is most pronounced in alkaline vapors (rubidium, lithium, cesium, sodium, and potassium). Figure 39.24 is the hyperfine spectral structure for the valence electrons of rubidium (Rb) 85, which is commonly used in these types of magnetometers. The energy-related frequency interval between these hyperfine lines is proportional to the applied field. The magnetic quantum number m is related to the angular momentum number and specifies the possible component magnitude of the magnetic moment along the applied field. The optically pumped magnetometer takes advantage of this characteristic.

Transitions occur between levels of different m values and obey the rule that the change in m can only have the values 0, 1, and –1. Table 39.5 lists the relationship between the polarization of the light stimulating the transition and the allowable change in m.

When not optically excited, the energy states of the valence electrons will be distributed according to Boltzmann statistics and will be in a state of equilibrium. If the electrons are excited with circularly polarized light at the $D1$ frequency (794.8 nm wavelength), they will absorb photons and transition from the $^2S_{1/2}$ state to the $^2P_{1/2}$ state according to the transition rules. The excited electrons will then fall back in a random fashion to the lower states, being distributed with an equal probability among all the m states.

But the rules state that the change in m can only be 1 or –1 for polarized light. If one uses right circularly polarized light, then the change in m can only be 1, and the electrons in the $m = 3$ level of the $^2S_{1/2}$ state cannot transition since there is no $m = 4$ level at the $^2P_{1/2}$ state. Therefore, these electrons remain in the $m = 3$ state. All other electrons transition to the higher state and then fall back to the lower state with equal probability of arriving at any of the m levels, including $m = 3$. Thus, the $m = 3$ level fills up,

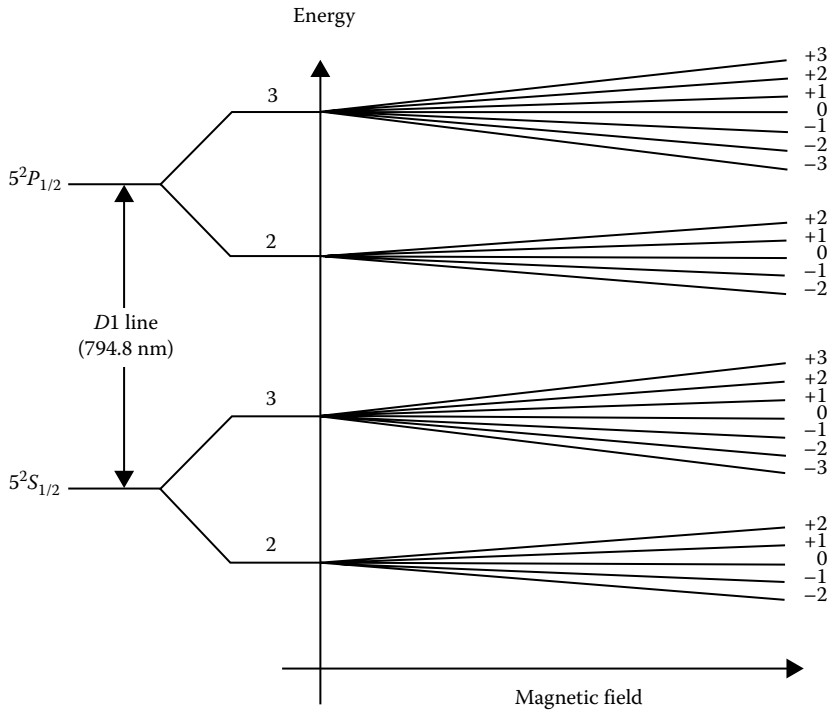


FIGURE 39.24 Rb-85 energy diagram. When a magnetic field is applied, the energy levels split into Zeeman sublevels that diverge as the field increases. Quantum mechanical factors determine the number of sublevels at each primary energy level.

TABLE 39.5 Allowable Change in *m* when Jumping from One Energy Level to Another Depends on the Polarization of the Light Causing the Transition

Polarization	*M*
Left circular	−1
Parallel	0
Right circular	1

and the other levels empty until all the electrons are in the $m = 3$ level, and no more transitions to the higher state can take place and pumping stops.

When pumping begins, the vapor is opaque. As time goes on, fewer electrons are available for absorbing photons, and the vapor becomes more transparent until, finally, pumping action stops and the vapor is completely transparent.

If a small RF magnetic field at the Larmor frequency is applied at right angles to the magnetic field being measured, the electrons in the $m = 3$ state will be depumped to the other *m* levels, making them available for further pumping. The optically pumped magnetometer exploits this situation in a positive feedback arrangement to produce an oscillator at the Larmor frequency.

The scale factors for optically pumped magnetometers are significantly higher than for the proton precession magnetometer. Table 39.6 lists these scale factors for a number of alkali vapors.

As a result, the sample rate and resolution can be much higher. A resolution of 0.005 nT is possible. Sampling rates can be as high as 15 samples per second.

TABLE 39.6 The Change in Frequency for a Change in Field Is Much Higher in Optically Pumped Magnetometers than in Proton Precession Magnetometers

Alkali	Scale Factor (Hz nT^{-1})
Rb-85	4.66737
Rb-87	~7
Cesium	3.4986

39.4.4 Signal Conditioning

Descriptions of several optically pumped magnetometers and their operating principles can be found in [24–26]. There are a number of different signal conditioning arrangements that can be used to derive a useful readout of the measured fields. Two of the more common methods are described in [26] and are shown in Figure 39.25.

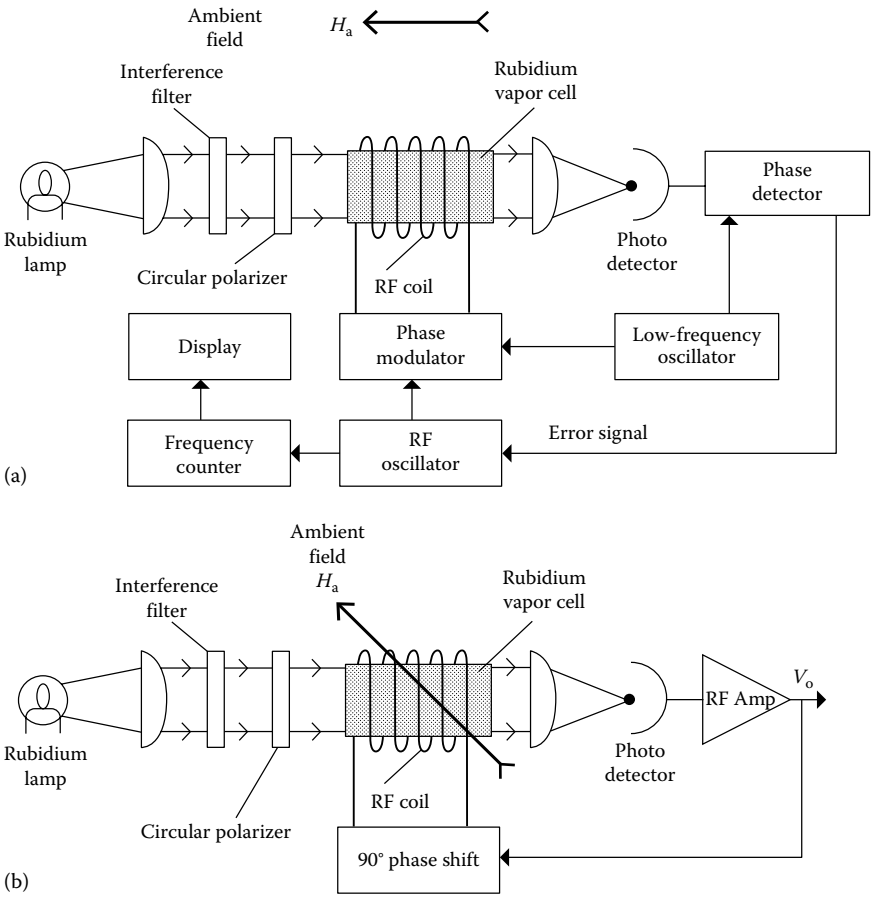


FIGURE 39.25 Two examples of optically pumped scalar magnetometers: (a) The servoed magnetometer slightly modulates the RF field at a low frequency, causing the vapor transmissivity to modulate. A phase detector provides an error signal that is used to lock the RF oscillator to the Larmor frequency; (b) A self-oscillating magnetometer: the transmissivity of the vapor, at right angles to the applied field, is made to oscillate at the Larmor frequency by phase shifting the detected light modulation and feeding it back to the RF field generator. (Adapted from Hartmann, F., *IEEE Trans. Magn.*, MAG-8, 66, 1972.)

In the servoed type shown in Figure 39.25a, the magnetic field being measured and the RF field are coaxial. The frequency of the RF oscillator is modulated with a fixed low-frequency oscillator. This causes the RF to sweep through the Larmor frequency. If the sweeped RF oscillator is not centered about the Larmor frequency, the photocell output signal will contain a fundamental component of the RF modulation frequency. The phase of the signal relative to the modulator oscillator determines whether the central RF is above or below the Larmor frequency. The photocell output is phase detected to produce an error voltage that is used to drive the RF toward the Larmor frequency. The RF can be measured to determine the magnetic field. If a linear voltage-controlled oscillator is used as the RF oscillator, its control voltage can also be used as an output since it is a measure of the Larmor frequency.

The auto-oscillating type shown in Figure 39.25b is based on the transmission of a polarized beam that is at right angles to the field being measured. The intensity of this crossbeam will be modulated at the Larmor frequency. The photocell signal will be shifted by 90° relative to the RF field. By amplifying the photocell signal, shifting it 90°, and feeding it back to drive the RF field coil, an oscillator is created at the Larmor frequency. In practice, only one light source is used, and the field being measured is set at an angle of 45°.

Defining Terms

Anisotropic: The material property depends on direction.

Gaussmeter: An instrument used to measure magnetic fields greater than 1 mT.

Induced magnetization: The object's magnetization is induced by an external magnetic field and disappears when the inducing field is removed.

Initial permeability: The slope at the origin of the magnetization curve.

Isotropic: The material property is the same in all directions.

Magnetic dipole moment: A vector quantity that describes the strength and direction of a magnetic field source, such as a small current loop or spinning atomic nucleus.

Magnetically "hard" material: The material has a significant residual (permanent) magnetization after an external magnetic field is removed.

Magnetically "soft" material: The material's magnetization is induced by an external magnetic field and the material has no significant residual (permanent) magnetization after the field is removed.

Magnetization curve: A plot of flux density B versus magnetic field H for an initially unmagnetized ferromagnetic material.

Magnetization: A vector quantity describing the average density and direction of magnetic dipole moments.

Magnetometer: An instrument used to measure magnetic fields with magnitudes up to 1 mT.

Magnetoresistance: The change in the electrical resistivity of a material in response to an applied magnetic field.

Maximum permeability: The maximum slope of the line drawn from the origin of the magnetization curve to a point on the magnetization curve.

Permanent magnetization: The source of an object's magnetization is internal and does not depend on the presence of an external field.

Permeability: A function that describes the relationship between an applied magnetic field and the resulting flux density.

Relative permeability: The permeability of a material normalized (divided) by the permeability of a vacuum.

Scalar magnetometer: A magnetometer that measures the magnitude of a magnetic field vector.

Vector magnetometer: A magnetometer that measures one or more of the individual components of a magnetic field vector.

References

1. K. S. Lion, *Instrumentation in Scientific Research*: *Electrical Input Transducers*, New York: McGraw-Hill, 1959.
2. H. R. Everett, *Sensors for Mobile Robots*: *Theory and Application*, Wellesley, MA: A. K. Peters, 1995.
3. J. E. Lenz, A review of magnetic sensors, *Proc. IEEE*, 78, 973–989, 1990.
4. M. A. Payne, SI and Gaussian cgs units, conversions and equations for use in geomagnetism, *Phys. Earth Planet. Int*., 26, P10–P16, 1981.
5. F. W. Grover, *Induction Calculations*: *Working Formulas and Tables*, New York: Dover Publications, 1973.
6. R. M. Bozorth, *Ferromagnetism*, New York: D. Van Nostrand, 1951.
7. R. M. Bozorth and D. M. Chapin, Demagnetization factors of rods, *J. Appl. Phys*., 13, 320–326, 1942.
8. S. A. Macintyre, A portable low noise low current three-axis search coil magnetometer, *IEEE Trans. Magn*., MAG-16, 761–763, 1980.
9. J. P. Hauser, A 20-Hz to 200-kHz magnetic flux probe for EMI surveys, *IEEE Trans. Electromagn. Compat*., 32, 67–69, 1990.
10. F. Primdahl, The fluxgate magnetometer, *J. Phys. E Sci. Instrum*., 1, 242–253, 1979.
11. C. J. Pellerin and M. H. Acuna, A miniature two-axis fluxgate magnetometer, *NASA Technical Note*, TN D-5325, NASA, 1970.
12. S. V. Marshall, A gamma-level portable ring-core magnetometer, *IEEE Trans. Magn*., MAG-7, 183–185, 1971.
13. W. A. Geyger, *Nonlinear-Magnetic Control Devices*, New York: McGraw-Hill, 1964.
14. M. Acuna, C. Scearce, J. Seek, and J. Schelfiele, The MAGSAT vector magnetometer—A precise fluxgate magnetometer for the measurement of the geomagnetic field, *NASA Technical Report*, NASA Tecdocs, 1978.
15. D. Cohen, Measurements of the magnetic field produced by the human heart, brain and lung, *IEEE Trans. Magn*., MAG-11, 694–700, 1975.
16. C. M. Falco and I. K. Schuller, SQUIDs and their sensitivity for geophysical applications, in *SQUID Applications to Geophysics*, Editors: Weinstock H., and W.C. Overton, Tulsa, Oklahoma: The Society of Exploration Geophysicists, pp. 13–18, 1981.
17. J. Clark, SQUIDs, *Sci. Am*., 271(2), 46–53, August 1994.
18. T. H. Casselman and S. A. Hanka, Calculation of the performance of a magnetoresistive permalloy magnetic field sensor, *IEEE Trans. Magn*., MAG-16, 461–464, 1980.
19. W. Kwaitkawski and S. Tumanski, The permalloy magnetoresistive sensors-properties and applications, *J. Phys. E*: *Sci. Instrum*., 19, 502–515, 1986.
20. Literature, Magnetic Sensors, Honeywell, http://www.magneticsensors.com/magnetic-sensor-products.php, accessed on 19 August 2013.
21. U. Dibbern and A. Petersen, The magnetoresistor sensor-a sensitive device for detecting magnetic field variations, *Elect. Comp. Appl*., 5(3), 148–153, 1983.
22. S. Tumanski and M. M. Stabrowski, Optimization of the performance of a thin film magnetoresistive sensor, *IEEE Trans. Magn*., MAG-20, 963–965, 1984.
23. L. W. Parson and Z. M. Wiatr, Rubidium vapor magnetometer, *J. Sci. Instrum*., 39, 292–299, 1962.
24. W. H. Farthing and W. C. Folz, Rubidium vapor magnetometer for near Earth orbiting spacecraft, *Rev. Sci. Instrum*., 38, 1023–1030, 1967.
25. F. Hartmann, Resonance magnetometers, *IEEE Trans. Magn*., MAG-8, 66–75, 1972.
26. F. Wellstood, C. Heiden, and J. Clark, Integrated dc SQUID magnetometer with high slew rate, *Rev. Sci. Instrum*., 66, 952–957, 1984.

40

RF/Microwave Spectrum Analysis

Roberto Ambrosini
Institute of Radioastronomy

Stelio Montebugnoli
Institute of Radioastronomy

Claudio Bortolotti
Institute of Radioastronomy

Mauro Roma
Institute of Radioastronomy

40.1 Introduction

A *signal* is usually defined by a time-varying function carrying some sort of information. Such a function most often represents a time-changing electric or magnetic field, whose propagation can be in free space or in dielectric materials constrained by conductors (waveguides, coaxial cables, etc.). A signal is said to be periodic if it repeats itself exactly after a given time T called the period. The inverse of the period T, measured in seconds, is the frequency f measured in hertz (Hz).

A periodic signal can always be represented in terms of a sum of several (possibly infinite) sinusoidal signals, with suitable amplitude and phase and having frequencies that are integer multiples of the signal frequency. Assuming an electric signal, the square of the amplitudes of such sinusoidal signals represents the power in each sinusoid and is said to be the power spectrum of the signal. These concepts can be generalized to a nonperiodic signal; in this case, its representation (spectrum) will include a continuous interval of frequencies, instead of a discrete distribution of integer multiples of the fundamental frequency. The representation of a signal in terms of its sinusoidal components is called Fourier analysis. The (complex) function describing the distribution of amplitudes and phases of the sinusoids composing a signal is called its Fourier transform (FT). The Fourier analysis can be readily generalized to functions of two or more variables; for instance, the FT of a function of two (spatial) variables is the starting point of many techniques of image processing.

A typical way to analyze a time-dependent electric signal, then in the so-called *time domain*, is through an *oscilloscope*. The flat screen of this instrument is arranged in such a way to display the signal amplitude voltage on a vertical scale, while the horizontal scale is continuously updated to account for the evolution of the time.

The *spectrum analyzer*, instead, is said to operate in the *frequency domain* because it allows one to measure the harmonic content of an electric signal, that is, the power of each of its spectral components. In this case, the vertical scales read powers available in each prefixed bandwidth, centered at the frequency reported in the horizontal scale. The two domains are mathematically well defined, and

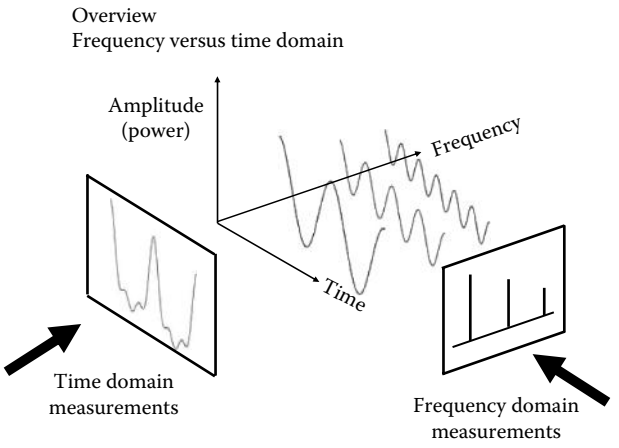


FIGURE 40.1 How the same signal can be displayed. (Courtesy of Agilent Technologies.)

through the FT algorithm, it is not too difficult to switch from one response to the other. Their graphical, easily perceivable representation is shown in Figure 40.1 where the two responses are shown lying on orthogonal planes. It is trivial to say that the easiest way to make a Fourier analysis of an electric signal is to have it displayed on a spectrum analyzer. Many physical processes, when detected by appropriate sensors, generate (electric) signals whose nature is not deterministic, but rather stochastic, or even of true random origin, like noise. Such signals can also be analyzed in terms of FT, although in a statistical sense only.

A time signal is said to be band limited if its FT is nonzero only in a finite interval of frequencies, say $(F_{max} - F_{min}) = B$, being the bandwidth. Usually, this is the case; then, an average frequency F_0 can be defined. Although the definition is somewhat arbitrary, a (band-limited) signal is referred to as radio frequency (RF) if F_0 is in the range 100 kHz–1 GHz and as a microwave signal in the range 1–1000 GHz. The distinction is not fundamental theoretically, but it has very strong practical implications in instrumentation and spectral measuring techniques. A band-limited signal can be described further as narrowband, if $B/F_0 << 1$, or wideband otherwise.

The first step in performing a spectral analysis of a narrowband signal is generally the so-called heterodyne downconversion: it consists in the mixing ("beating") of the signal with a pure sinusoidal signal of frequency F_L, called local oscillator (LO). In principle, mixing two signals of frequency F_0 and F_L in any nonlinear device will result in a signal output containing the original frequencies as well as difference $(F_0 - F_L)$ and the sum $(F_0 + F_L)$ frequencies and all their harmonic (multiple) frequencies. In the practical case, the best mixer devices are selected in order to exhibit an almost pure quadratic transfer function with the output including only the frequencies $(F_0 - F_L)$, $2F_L$, $2F_0$, and $(F_0 + F_L)$. The first term (called the intermediate frequency or IF) will be easily separated from the others, which have a much higher frequency, by a proper termination and filtering of the mixer output. The bandwidth of the IF signal will be the same as the original bandwidth B; however, to preserve the original information fully in the IF signal, stringent limits must be imposed on the LO signal, because any deviation from a pure sinusoidal law will show up in the IF signal as added phase and amplitude noise, corrupting the original spectral content. The process of downconverting a (band-limited) signal is generally necessary to perform spectral analysis in the very high-frequency (microwave) region, to convert the signal to a frequency range more easily handled by the following analyzing hardware and software.

When the heterodyne process is applied to a wideband signal (or whenever $F_L > F_{min}$), "negative" frequencies will appear in the IF signal. This process is called *double sideband* mixing, because a given IF bandwidth B (i.e., $(F_L + B/2)$ will include two separate bands of the original signal, centered at F_L + IF

("upper" sideband) and F_L – IF ("lower" sideband). This form of mixing is obviously undesirable in spectrum analysis, and input filters are generally necessary to split a wideband signal in several narrowband signals before downconversion. Alternatively, special mixers can be used that can deliver the upper and lower sidebands to separate IF output ports. A band-limited signal in the frequency interval $(F_{max} - F_{min}) = B$ is said to be converted to *baseband* when the LO is placed at $F_L = F_{min}$, so that the band is converted to the interval $(B - 0)$. Modern digital circuits are best matched to this case, being essentially limited on the maximum working frequency. Here, it is possible to employ analog-to-digital converters (ADCs) to get a discrete numerical representation of the analog signal, and the spectral analysis is then performed numerically, either by direct computation of the FT (generally via the fast Fourier transform, FFT, algorithm) or by computation of the signal autocorrelation function, which is directly related to the square modulus of the FT via the Wiener–Khinchin theorem. Considering that the ADC must sample the signal at least at the Nyquist rate (i.e., at twice the highest frequency present) and with adequate digital resolution, this process is feasible and practical only for frequencies (bandwidths) less than a few hundreds of megahertz. Also, the possibility of a real-time analysis with high spectral resolution may be limited by the availability of very fast digital electronics and special-purpose computers. The digital approach is the only one that can provide extremely high spectral resolution, up to several hundred thousand channels.

40.2 Practical Approach to Spectrum Analysis

Spectrum analysis is normally done in order to verify the harmonic content of oscillators, transmitters, frequency multipliers, etc., or the spurious components of amplifiers and mixer [1]. Other specialized applications are possible, such as the monitoring of radio-frequency interference (RFI), electromagnetic interference (EMI), and electromagnetic compatibility (EMC). These applications, as a rule, require an antenna connection and a low-noise, external amplifier. Which are then the specifications to look for in a good spectrum analyzer? First, we would suggest evaluating what are the critical parameters that are most relevant for your application. Then identify what is the instrument performance that has to drive the selection of the industrial product that best matches your needs.

Here follows a list of the typical characteristics of a spectrum analyzer:

1. It should display selectable, both wide and narrow, bands of the EM radio spectrum with amplitude power and frequency readable with good accuracy.
2. Its selectivity should range, in discrete steps, from few hertz to megahertz so that sidebands of a selected signal can be spotted in details as well as monitored in the spectral environment where that signal is immersed in.
3. It should possess a very wide dynamic range, so that signals differing in amplitude six to eight orders of magnitude can be observed at the same time on the same view of the display.
4. Its sensitivity must be compatible with the measurements to be taken. As already mentioned, specialized applications may require external wideband, low-noise amplifiers, and an antenna connection.
5. The type of power supply (mains or battery), weight, and dimensions should be compatible with the type of operation, fixed or portable, you want to have.
6. Phase noise stability of the LO shall be carefully verified before attempting to reach your ultimate target in spectral resolution measurements.
7. Consider your data logger requirements: the capability of direct driving a plotter or a printer, archive data on external memory devices, and finally the interface protocol that will allow controlling your spectrum analyzer, from a personal computer (PC).
8. Analog and/or digital averaging and other special functions can extend the measuring capabilities in dedicated measuring configurations.
9. Periodic calibration of the unit by an authorized center will guarantee that your results could be traced to international standards, along the operating life of the instrument.

A block diagram of a commercial spectrum analyzer is shown in Figure 40.2.

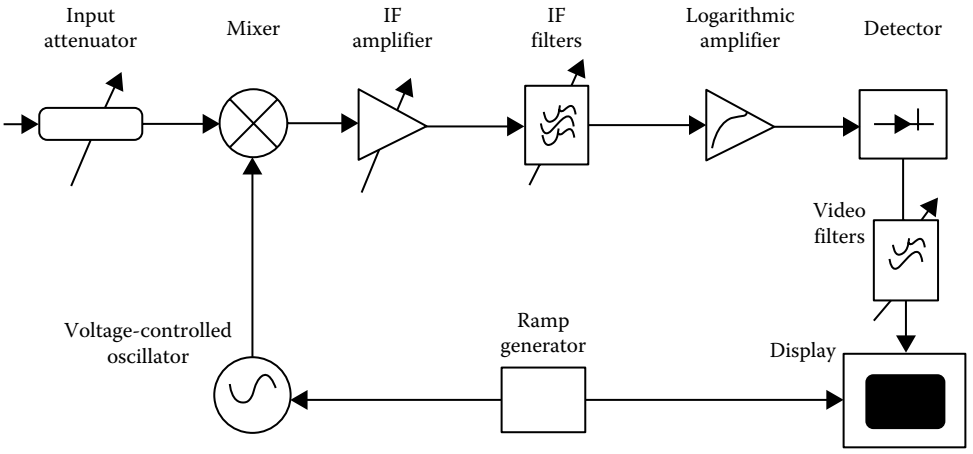


FIGURE 40.2 Block diagram of a commercial spectrum analyzer. (Courtesy of Hewlett-Packard, Rohde Schwarz, Hameg, Tektronix companies, and IEEE Microwave Measurements.)

Figure 40.2 shows that we are confronted with a radio-receiver-like superhet with a wideband input circuit. The horizontal scale of the instrument is driven by a ramp generator, which is also applied to the voltage-controlled LO [2].

A problem arises when dealing with a broadband mixing configuration like the one shown earlier, namely, avoiding receiving the image band.

The problem is successfully tackled here by upconverting the input band to a high-valued IF. An easily designed input low-pass filter, not shown in the block diagram for simplicity, will now provide the necessary rejection of the unwanted image band.

Nowadays, with the introduction of YIG band-pass filter preselectors, tunable over very wide input bands, upconversion is not always necessary. Traces of unwanted signals may, however, show up on the display although at very low level (less than −80 dBc) even on good analyzers.

A block diagram of a commercial spectrum analyzer exploiting both the mentioned principles is shown in Figure 40.3. This instrument includes a very important feature that greatly improves its performance: the LO frequency is no longer coming from a free-running source but rather from a synthesized unit referenced to a very stable quartz oscillator. The improved quality of the LO both in terms of its own noise and frequency stability optimizes several specifications of the instrument, such as frequency determining accuracy, finer resolution on display, and reduced noise in general.

Further, a stable LO generates stable harmonics that can then be used to widen the input-selected bands up to the millimeter region. As already stated, this option requires external devices, for example, a mixer-amplifier as shown in Figure 40.4a and b.

The power reference on the screen is the top horizontal line of the reticle. Due to the very wide dynamic range foreseen, the use of a log scale (e.g., 10 dB/square) seems appropriate. Conventionally, 1 mW is taken as the zero reference level: accordingly, dBm are used throughout.

The noise power level present on the display without an input signal connected (noise floor) is due to the input random noise multiplied by the IF amplifier gain. Such a noise is always present and varies with input frequency, IF selectivity, and analyzer sensitivity (in terms of noise figure).

The "on display dynamic range" of the analyzer is the difference between the maximum compression-free level of the input signal and the noise floor. As a guideline, the dynamic range of a good instrument could be of the order of 70–90 dB.

An input attenuator, always available on the front panel, allows one to apply more power to the analyzer while avoiding saturation and nonlinear readings. The only drawback is the obvious sensitivity loss.

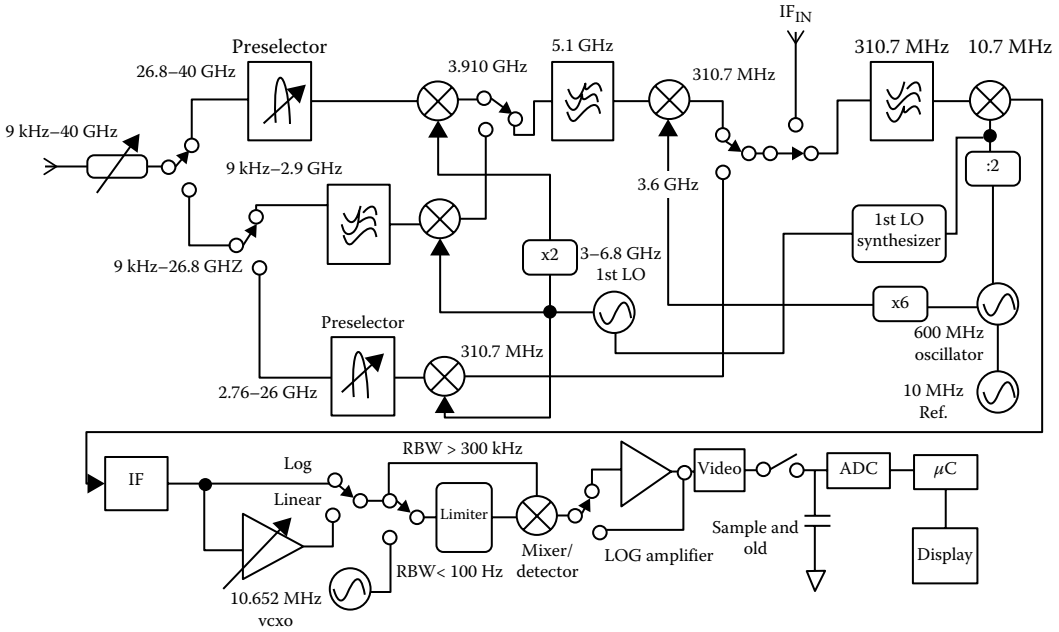


FIGURE 40.3 Standard block diagram of a modern spectrum analyzer. (Courtesy of Hewlett-Packard, Rohde Schwarz, Hameg, Tektronix companies, and IEEE Microwave Measurements.)

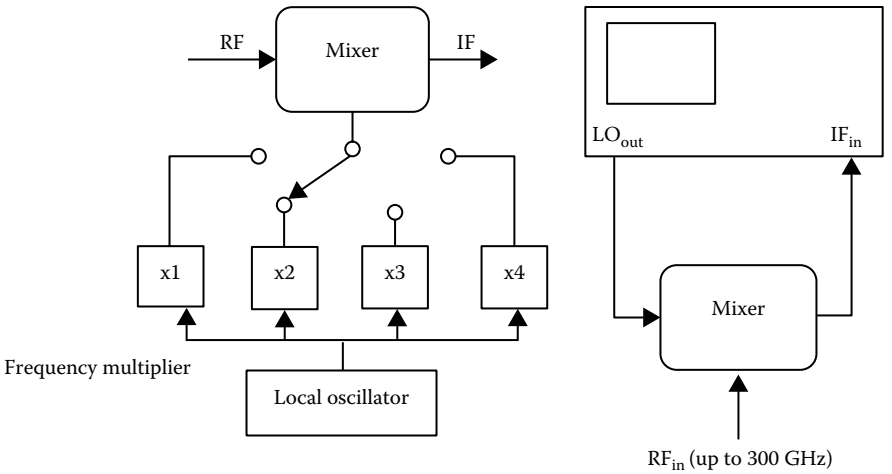


FIGURE 40.4 Increasing the input bandwidth characteristics. (Courtesy of Hewlett-Packard, Rohde Schwarz, Hameg, Tektronix companies, and IEEE Microwave Measurements.)

When kept calibrated over the specified warranty periods, modern spectrum analyzer can exhibit measurement uncertainties of a fraction of a dB and Hertz resolutions up to a few GHz.

When more accurate measurement of power sources, with slow time variations, is needed, the suggestion is to use a *bolometer* instead of a spectrum analyzer. An erratic signal pattern on display and a fancy level indication may be caused by the wrong setting of the "scan time" knob. It must be realized that high-resolution observation of a wide input band requires the proper scanning time. An incorrect parameter setting yields wrong readings, but usually an optical alarm is automatically switched on to

warn the operator. In any case, never forget that pulsed signal, in particular even when repeated with very low duty cycles, cannot be represented with fidelity by a frequency swept spectrum analyzer.

The knowledge of the noise floor level allows a good valuation of the noise temperature, T_n (and therefore of the sensitivity), of the analyzer, a useful parameter on many occasions. The relations involved are as follows.

The Nyquist relation states that

$$P = k \times T_n \times B \tag{40.1}$$

where

P is the noise floor power level read on the display (W)
k is the Boltzmann constant = 1.38×10^{-23} (J/K)
B is the passband of the selected IF (Hz)

Therefore,

$$T_n = \frac{P}{k \times B} \tag{40.2}$$

Usually, engineers prefer to quote the noise figure of receivers. By definition, we can write

$$N = \left(\frac{T_n}{T_0} \right) + 1 \tag{40.3}$$

where

N is the noise factor
$T_0 = 290$ K
F (noise figure) = 10 log N

A typical F for a good spectrum analyzer is of the order of 10–20 dB.

It must be said, however, that the "ultimate sensitivity" of the spectrum analyzer will depend not only on its noise figure but also on the setting of other parameters like the video filter, the IF bandwidth, the insertion of averaging functions, the scan speed, and the type of transfer function of the detector used.

As a rough estimate, a noise floor level of –130/–140 dBm is very frequently met by a good instrument.

Another criterion to select a spectrum analyzer is a good "IMD dynamic range," that is, the tendency to create spurious signals by intermodulation due to saturation.

This figure is generally quoted by the manufacturers, but it is also easily checked by the operator by injecting two equal amplitude sinusoidal signals at the input socket of the analyzer. The frequency separation between the two should be at least a couple of "resolution bandwidths," that is, the selected IF bandwidth. As the input levels increase, spurious lines appear at the sum and difference frequencies and spacing of the input signals.

The range in decibels between the nonoverloaded input signals on display and the barely noticeable spurious lines is known as the "spurious free dynamic range," shown graphically in Figure 40.5a, where the third-order "intercept point" is also graphically determined. If input power is increased, higher-order spurious signals appear, as shown in Figure 40.5b. The input connector of most spectrum analyzers is of the 50 Ω coaxial type. Past instruments invariably used *N*-type connectors because of their good mechanical and electric behavior up to quite a few gigahertz. Today, type SMA or type K connectors are preferred.

External millimeter wave amplifiers and converters use waveguide input terminations. As is discussed in the next section, multipurpose analyzers are available where power meter, frequency

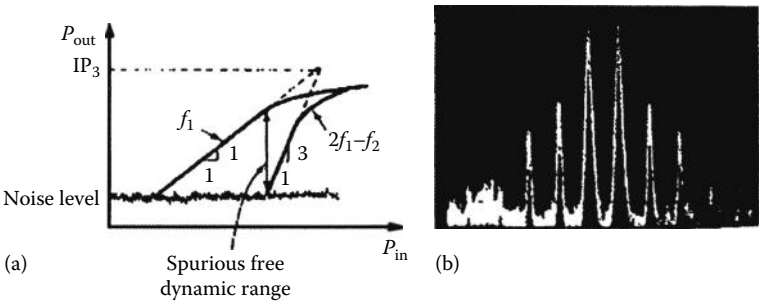


FIGURE 40.5 (a) Spurious free dynamic range and (b) higher-order spurious. (Courtesy of Hewlett-Packard, Rohde Schwarz, Hameg, Tektronix companies, and IEEE Microwave Measurements.)

counter, tracking generator, etc., can all be housed in the same cabinet. The economic and practical convenience of these units must be weighed on a case-by-case basis.

Finally, we mention that spectrum analyzers are available equipped with AM and FM detectors to facilitate their use in the RFI monitoring applications in order to identify the name of the interferer.

40.3 Selecting Correct Spectrum Analyzers for Specific Purposes

Several manufacturers offer a large number of spectrum analyzer models; the choice may be made on the basis of application field (i.e., CATV, mobile telephony, R-LAN, service, surveillance, and R&D), performance (resolution bandwidth, frequency range, accuracy, battery operation, etc.), or cost. Very often, the newest RF models exhibit also the signal analyzer function, intended as some kind of characterization in the time domain, by computing the FFT of the signal under test, that is, baseband modulation and phase noise.

In addition, it is important to know that most spectrum analyzers need some accessories generally not furnished as a standard: for example, a connectorized, coaxial, microwave cable is always required; a directional coupler, power divider, or handheld sampler antenna may be very useful to pick up the signals; and a PC is useful to collect, store, reduce, and analyze the data.

There are six main families of RF and microwave spectrum analyzers.

40.3.1 Family 1

The bench instruments are top performance but also large, heavy, and the most expensive class, intended for metrology, certification, factory reference, and radio surveillance done by government and military institutions.

Their operating frequencies span from a few tens of hertz to RF (i.e., 2.9 GHz), or up to the microwave region (i.e., 26.5 GHz), or even reach millimeter wavelengths (i.e., 40–67 GHz). This class of instruments includes lower-noise figures, approximately 20–25 dB at medium frequencies that can be decreased down to 10–15 dB with an integrated preamplifier. The synthesized LO has a good phase noise (typically 10 dB better than general-purpose spectrum analyzers) for precise, accurate, and stable measurement. Also, this class of instruments can be integrated with plug-in instruments like a power meter (for more accurate power measurements), a frequency counter (for accurate frequency measures), or a tracking generator (for network analysis and mixer testing).

An interface to a computer (and printer) such IEEE-488, USB, and in particular LAN is now standard in all units. Now, even the front panel acts as the remote control terminal for the instrument inside. Optical encoders, up and down push buttons, and the digital keyboard have substituted for all old analog or mechanical input switches. The output display is now typically a colorful flat LED monitor, with greater advantages for a clearer and much more detailed view. Having inside powerful, state-of-the-art

microprocessor controllers, RAM memories, and plug-in cards for storing data, all sort of statistical and mathematical analysis can be implemented in many different measuring configurations, each stored for later usage with a single press of recovery functions. Among the most appreciated and qualified trademarks, we mention Agilent [3] and Rohde & Schwarz [4]. Indicative prices are between $50,000 and $100,000.

40.3.2 Family 2

Less-expensive bench instruments, the workhorse class of spectrum analyzers, portable and lightweight, are associated with a synthesized LO that includes a frequency range from a few kilohertz up to the RF region (i.e., 2.9 GHz), microwave region (i.e., 26.5 GHz), or near millimeter wavelengths (i.e., 40–50 GHz). A typical noise figure of 30 dB is good enough to ensure most measurements. A large number of filters down to a few hertz of resolution are offered; digital filters are preferable to analog ones, because they give a faster refresh rate of the trace on the display. This kind of spectrum analyzer nearly always has the capability to extend the frequency range up to millimeter and submillimeter wavelengths with an external mixer. One of the most important features for a spectrum analyzer in this class is the quality of the LO; it should be synthesized phase-locked loop (PLL) to achieve stability, precision, accuracy, and low phase noise. Demodulation is also an important feature to listen to AM and FM on the loudspeaker and to display complex modulations onto the screen, which is often required by people working on surveillance, TV, and mobile telephone. The interface to a computer such as IEEE-488/USB/LAN is standard in a large number to display TV pictures or complex modulations onto the screen, which is often required by people working on surveillance, TV, and mobile telephone. The interface to a computer such as IEEE-488 or RS232 is standard in a large number of spectrum analyzers and allows the remote control and data reading, storing, and manipulation.

This kind of instrument may integrate a tracking generator, a frequency counter, and other instruments that can transform the spectrum analyzer into a compact, full-featured RF and microwave laboratory. Agilent [3], Anritsu [5], IFR [6], Rohde & Schwarz [4], and Tektronix [7] are the leader companies in the market. Prices typically span from $30,000 to $75,000.

40.3.3 Family 3

The entry level, a more economical class of spectrum analyzer, is intended for field use or for one specific application. If your need is mainly EMI/EMC, CATV, mobile telephone, or surveillance, perhaps you do not need the extreme stability of a synthesized LO, and a frequency range up to 2 GHz may be enough; however, if you need some special functions such as "quasi-peak detector" or "occupied bandwidth measurement," two functions that originate from a combination of a mathematical treatment with some legislative requirements, you can find off the shelf what you need in a dedicated instrument.

Manufacturers are the same, while costs typically run from around $10, 000 to $35,000.

40.3.4 Family 4

Handheld spectrum analyzers are very popular nowadays, being low cost, with a high level of performance for their deep usage of the newest digital processing techniques. Frequency coverage up to 43 GHz, and just a few kilograms of weight, can satisfy a large number of users in general fields and in particular when it is required to make a measurement in outdoor open spaces. Agilent [3], Anritsu [5], Rohde & Schwarz [4], Tektronix [7], and some minor companies offer such units from $5,000 to $45,000, according to their specific performance.

40.3.5 Family 5

The most economical class of spectrum analyzer, with prices around $2000–$6000, includes instruments that perform only the basic functions with a limited frequency range and filter availability.

They are intended for service, for general-purpose measurements (i.e., IP_3, harmonic distortion), or for precertification in EMI/EMC measurements. One of the most popular is HAMEG [8] and RIGOL [9].

In this class are some special spectrum analyzers that come on a PC board. Such spectrum analyzers are generally cheap (typically $2000–$4000), with a frequency range up to 2 GHz, and may include PLL LOs, tracking generators, and other advanced characteristics. The input is through a coaxial connector on the board, the output and the control are done by a virtual instrument running on the PC.

Other unusual RF spectrum analyzers working in conjunction with a PC and worth noting are the instruments for EMI/EMC measurements and reduction in power lines and power cords. For this type of instrument, the core is not the hardware but the software that performs the measurement according to international standards and may guide the engineer to meet the required compatibility.

40.3.6 Family 6

Many modern applications require "modal analysis," for example, at acoustic frequencies in mechanical systems. All devices have moving parts, like wheels, axels, and ropes so they can be applied to cars as well as civil elevators, just as the most common examples. These applications require industrial qualifications based on a deeper knowledge of their dynamical properties to avoid dangerous resonances and dissipative losses and finally improve their efficiencies.

Extremely low-frequency spectrum analyzers, with resolutions down at the millihertz level, are quite common. They are very cheap (with respect to the RF spectrum analyzers), being made of only digital hardware, acquiring the input signal in the time domain, and then computing the frequency spectrum by dedicated FFT processors. In this case, the dynamic range can be raised a few decades, ranging up to 90–110 dB. The tracking generator is generally built in, as well as many anti-aliasing input filtering choices.

40.4 Advanced Applications

Completely new families of integrated circuits now offer unprecedented functionalities. For example, off-the-shelf components now make analog-to-digital conversions up to a few giga samples per second with a large dynamic range (of 48 dB, meaning 8 bits of resolution or even more). Such a huge amount of data can only be handled by an extremely fast processing system equipped with field programmable arrays (FPGAs) or graphics processing units (GPUs).

Fundamental research like radio astronomy has found a great benefit from these technological developments opening new possibilities in advanced spectrum analysis applications. Autocorrelators, with a typical frequency resolution of ~5/25 kHz, have been extensively used in radio astronomy. Their performance is well documented; the autocorrelation function is computed online and recorded. Later, the FFT of the function is computed off-line in order to get the power spectrum.

As a matter of fact, the FPGA represents a medium-/high-cost approach to the state of the art of very high-performance silicon engines. FPGA advantages are low power consumption at high clock rates and easy reconfiguration for different very high computing demanding projects. They are limited to only linear and pipelined elaboration, under PC control for final data storage.

Fully dedicated applications allow quickly implementing even very complex algorithms on the FPGA chips. The final huge processing power is a result that comes out from combining the unreached velocity of the hardware together with the extremely high flexibility of the software, with the benefit of low power consumption per clock rate.

Current FPGA applications demonstrate the capability of processing more than 500 MHz bandwidth with up to 1,000,000 channels using polyphase filter banks (PFBs) in real time.

Here follows a simple block diagram of a high-resolution spectrum analyzer powered by FPGA using the PFB and the FFT engine blocks: our configuration is capable to synthetize in real time several hundred million frequency channels over an input band of 500 MHz [10].

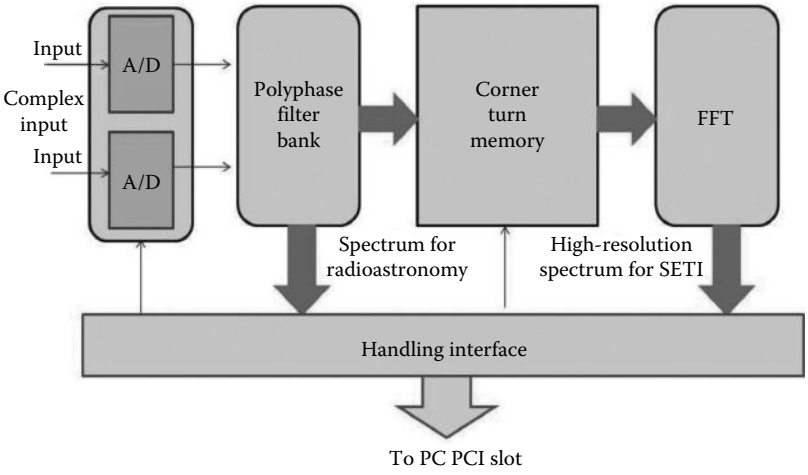


FIGURE 40.6 Block diagram of a high-resolution spectrum analyzer.

Another state-of-the-art low-cost approach for spectrum analysis is based on GPU; this is a dedicated microprocessor conceived to off-load and accelerate 3D or 2D graphics rendering, made available from its associated CPU. This kind of powerful engine is primarily used in workstations, game consoles, etc. Due to their highly parallel structure, modern GPUs are more effective than standard CPUs for implementing a wide range of complex algorithms.

Their main advantages are high computing power, suitable to drive directly data parallel computing, and easy programmability with little overhead in standard C language and with runtime configurations, scalable on different GPU clusters. In conclusion, a GPU offers an extremely powerful tool as the "number crunching" machine that is requested for the huge FFT computations needed for reaching top performance in spectrum analysis with high frequency resolution and over large bandwidths. The block diagram of a high-resolution spectrum analyzer is illustrated in Figure 40.6.

References

1. A.E. Bailey, ed., *Microwave Measurements*, 2nd edn., Peter Peregrins on behalf of IEEE, London, U.K., 1989.
2. Hewlett-Packard, *A.N 63* and *A.N 243*, Hewlett-Packard Company.
3. Agilent, www.home.agilent.com/, Accessed on July 26, 2012.
4. Rohde & Schwarz, www2.rohde-schwarz.com/, Accessed on July 25, 2012.
5. Anritsu, www.anritsu.com/, Accessed on July 22, 2012.
6. IFR, www.aeroflex.com/, Accessed on July 24, 2012.
7. Tektronix, www1.tek.com/, Accessed on July 21, 2012.
8. Hameg, www.hameg.com/, Accessed on July 27, 2012.
9. RIGOL, www.rigol.com/, Accessed on August 7, 2013.
10. S. Montebugnoli, C.B. Cosmovici et al., *SETI Back Ends Made Inexpensive*, IAF Congress, Prague, Czech Republic, 2010.

IV

Time and Frequency

41

Time Measurement

Michael A. Lombardi
National Institute of Standards and Technology

Although we have no control over its passage, we can measure time with more resolution and less uncertainty than any other physical quantity. The base unit of time, the second (s), is one of the seven base units of the International System of Units (SI) and can be measured so accurately that numerous other units, including the meter and ampere, depend upon its definition.

Time measurements can be divided into two general categories: *time interval* measurements and *time synchronization* measurements. Time interval measurements determine the duration or elapsed time between two events. Time standards typically produce 1 pulse per second (pps) signals. The period of these signals is a standard second that serves as a time interval reference. Many engineering and scientific applications require the measurement of time intervals much shorter than 1 s, such as *milliseconds* (ms, 10^{-3} s), *microseconds* (μs, 10^{-6} s), *nanoseconds* (ns, 10^{-9} s), *picoseconds* (ps, 10^{-12} s), and *femtoseconds* (fs, 10^{-15} s). Thus, the instrumentation used to measure time intervals always requires subsecond resolution.

Time synchronization measurements determine the time offset between the test signal and the Coordinated Universal Time (UTC) second. The reference for these measurements is typically a 1 pps signal synchronized to UTC. If the 1 pps signal is labeled or time tagged, it can be used to synchronize time-of-day clocks and can also be used to record when an event happened. *Time tags* typically include the year, month, day, hour, minute, second, and often the fractional part of a second. When we ask the everyday question, "What time is it?" we are asking for a current time tag. By comparing and aligning time tags, we can ensure that events are *synchronized* or scheduled to occur at the same time.

Both types of time measurements are referenced to the frequency of a *periodic event* that repeats at a constant rate. For example, this periodic event could be the swings of a pendulum. We could agree to establish a base unit of time interval by defining the second as one complete swing, or cycle,

of the pendulum. Now that we have defined the second, we can measure longer time intervals, such as minutes, hours, and days, by simply counting the swings of the pendulum. This process is how a *time scale* is formed. A time scale is simply an agreed upon way to order events and keep time. It is formed by measuring and establishing a base time unit (the second) and then counting elapsed seconds to establish longer time intervals. A device that measures and counts time intervals to mark the passage of time is called a *clock*. Let's continue our discussion by looking at the evolution of clocks and timekeeping.

41.1 Evolution of Clocks and Timekeeping

All clocks share several common features, including a device that continuously produces the periodic event mentioned previously. This device is called the *resonator*. In the case of a pendulum clock, the pendulum is the resonator. Of course, the resonator needs an energy source, such as a mainspring or motor, so it can run continuously. Taken together, the energy source and the resonator form an *oscillator* that runs at a rate called the *resonance frequency*. Another part of the clock counts the oscillations and converts them to time units. And finally, the clock must display or record the results [1].

The frequency uncertainty of a clock's resonator is directly related to the timing uncertainty of the clock, as shown in Table 41.1.

Throughout history, clockmakers have searched for better oscillators that would allow them to build more accurate clocks. As early as 3500 BC, time was kept by observing the movement of an object's shadow between sunrise and sunset. This simple clock is called a *sundial*, and the resonance frequency is the apparent motion of the sun. Later, water clocks, hourglasses, and calibrated candles allowed dividing the day into smaller units of time or hours. Mechanical clocks first appeared in the early fourteenth century. Early models used a verge and foliot mechanism for an oscillator and had an uncertainty of about 15 min/day ($\cong 1 \times 10^{-2}$). However, mechanical clocks did not display minutes until near the end of the sixteenth century.

A timekeeping breakthrough occurred with the invention of the *pendulum clock*, a technology that dominated timekeeping for several hundred years and established the second as a usable unit of time interval. In the early 1580s, Galileo Galilei observed that a given pendulum took the same amount of time to swing completely through a wide arc as it did a small arc. Galileo wanted to apply this natural periodicity to time measurement and had begun work on a mechanism to keep the pendulum in motion in 1641, a year prior to his death. In 1656, the Dutch scientist Christiaan Huygens invented an escapement that kept the pendulum swinging. The uncertainty of Huygens's clock was less than 1 min/day ($\cong 7 \times 10^{-4}$) and later reduced to about 10 s/day ($\cong 1 \times 10^{-4}$). The first pendulum clocks were weight driven, but later versions were powered by springs. Huygens is often credited with inventing the spring and balance wheel assembly still found in some of today's mechanical wristwatches.

TABLE 41.1 Relationship of Frequency Uncertainty to Timing Uncertainty

Frequency Uncertainty	Measurement Period	Timing Uncertainty
$\pm 1.00 \times 10^{-3}$	1 s	±1 ms
$\pm 1.00 \times 10^{-6}$	1 s	±1 μs
$\pm 1.00 \times 10^{-9}$	1 s	±1 ns
$\pm 2.78 \times 10^{-7}$	1 h	±1 ms
$\pm 2.78 \times 10^{-10}$	1 h	±1 μs
$\pm 2.78 \times 10^{-13}$	1 h	±1 ns
$\pm 1.16 \times 10^{-8}$	1 day	±1 ms
$\pm 1.16 \times 10^{-11}$	1 day	±1 μs
$\pm 1.16 \times 10^{-14}$	1 day	±1 ns

Major advances in timekeeping accuracy resulted from the work of John Harrison, who built and designed a series of clocks in the 1720s that kept time to within fractions of a second per day (parts in 10^6). This performance was not improved upon until the twentieth century. Harrison dedicated most of his life to solving the British navy's problem of determining longitude, by attempting to duplicate the accuracy of his land clocks at sea. He built a series of clocks (now known as H1 through H5) in the period from 1730 to about 1770. He achieved his goal with the construction of H4, a clock much smaller than its predecessors, about the size of a large pocket watch. H4 used a spring and balance wheel escapement and kept time within fractions of a second per day during several sea voyages in the 1760s.

The practical performance limit of pendulum clocks was reached in 1921, when W. H. Shortt demonstrated a clock with two pendulums, one a slave and the other a master. The slave pendulum moved the clock's hands and freed the master pendulum of tasks that would disturb its regularity. The pendulums used a battery as their power supply. The Shortt clock kept time within a few seconds per year ($\cong 1 \times 10^{-7}$) and was once used as a primary standard for time interval in the United States [1–3].

Quartz crystal oscillators, based on the phenomenon of piezoelectricity discovered by P. Curie in 1880, worked better than pendulums, resonating at a nearly constant frequency when an electric current was applied. Credit for developing the first quartz oscillator is generally given to Walter Cady, who built prototypes shortly after World War I and patented a piezoelectric resonator designed as a frequency standard in 1923. Joseph W. Horton and Warren A. Marrison of Bell Laboratories built the first clock based on a quartz crystal oscillator in 1927. By the late 1930s, quartz oscillators began to replace pendulums as the standard for time interval measurements.

Billions of quartz oscillators are now manufactured annually. They are found in nearly every type of electronic circuit, including many devices that display time such as clocks, watches, cell phones, radios, and computers. Even so, quartz oscillators are not an ideal timekeeping source. Their resonance frequency depends on the size and shape of the crystal, and no two crystals are exactly alike or produce exactly the same frequency. Their frequency is also sensitive to changes in the environment, including temperature, pressure, and vibration. These limitations make them unsuitable for high-accuracy timekeeping and led to the development of atomic oscillators [1,4–6].

41.2 Atomic Oscillators

The practice of using resonant transitions in atoms or molecules is attractive for several reasons. For example, an unperturbed atomic transition is identical from atom to atom, so that, unlike a group of quartz oscillators, a group of atomic oscillators should all generate the same frequency. Also, unlike all electrical or mechanical resonators, atoms do not wear out or change their properties over time. These features were appreciated by Lord Kelvin, who suggested using transitions in sodium and hydrogen atoms as timekeeping oscillators in 1879. However, it wasn't until the late 1930s that technology made his idea possible, when I. I. Rabi and his colleagues at Columbia University introduced the idea of using an atomic resonance as a frequency standard.

Atomic oscillators use the quantized energy levels in atoms and molecules as the source of their resonance frequency. The laws of quantum mechanics dictate that the energies of a bound system, such as an atom, have certain discrete values. An electromagnetic field can boost an atom from one energy level to a higher one. Or an atom at a high energy level can drop to a lower level by emitting electromagnetic energy. The resonance frequency (f) of an atomic oscillator is the difference between the two energy levels divided by Planck's constant (h):

$$f = \frac{E_2 - E_1}{h} \tag{41.1}$$

The basic principle of the atomic oscillator is simple: Because all atoms of a specific element are identical, they should produce exactly the same frequency when they absorb energy or release energy.

Thus, atomic oscillators easily surpassed the performance of all previous standards. In theory, the atom is a perfect pendulum whose oscillations can be used as a standard of frequency or counted to measure time interval.

The three major types of commercial atomic oscillators are based on rubidium, cesium, and hydrogen atoms, respectively. The least expensive and most common type is the *rubidium oscillator*, based on the 6.8347 GHz resonance of ^{87}Rb. Rubidium oscillators are well suited for applications that require a small, high-performance oscillator. The frequency uncertainty of a rubidium oscillator over a 1 day interval typically ranges from about $\pm 5 \times 10^{-9}$ to $\pm 5 \times 10^{-12}$.

The second type of atomic oscillator, the *cesium oscillator*, serves as a primary standard in many laboratories, because the resonance frequency of cesium (9.1926 GHz) is used to define the second (Section 41.3). Because the base unit of time is defined with respect to cesium resonance, the 1 pps output from a cesium oscillator serves as an internationally recognized time interval standard. Commercially available cesium beam oscillators typically have a frequency uncertainty over a 1 day interval near $\pm 1 \times 10^{-13}$, with the best units about one order of magnitude better. Even so, a cesium cannot recover time by itself and cannot be used as a synchronization reference unless it has been synchronized with another source.

Currently (2010), the world's best time interval standards are cesium fountains. These devices are large, sometimes several meters tall. They are not sold commercially but have been built by several national metrology laboratories. Cesium fountains work by laser cooling atoms to a temperature near absolute zero (less than 1 μK) and then lofting them vertically. The atoms make two passes (one on the way up, the second on the way down) through a microwave cavity where the atomic resonance frequency is measured. The atoms can be observed for about 1 s. This allows a cesium fountain to outperform cesium beam standards, which have a much shorter observation period, typically about 1 ms. The primary frequency and time interval standard in the United States is a cesium fountain called NIST-F1 (Figure 41.1) that can realize the second with an uncertainty of about 3×10^{-16}.

A third type of atomic oscillator, the *hydrogen maser*, is based on the 1.42 GHz resonance frequency of the hydrogen atom. Hydrogen masers typically are more stable than cesium oscillators for periods

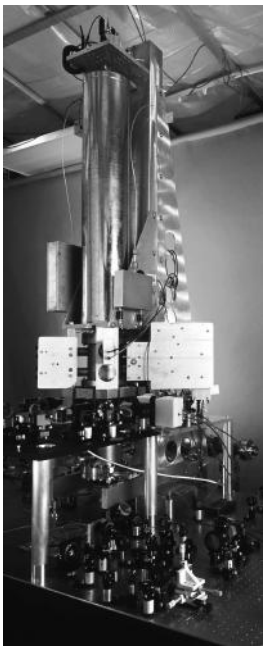


FIGURE 41.1 NIST-F1 cesium fountain.

TABLE 41.2 Evolution of Clock Design and Performance

Type of Clock	Resonator	First Appearance	Typical Time Uncertainty (1 Day)	Typical Frequency Uncertainty (1 Day)
Sundial	Apparent motion of sun	3500 BC	NA	NA
Mechanical	Verge and foliot mechanism	Fourteenth century	±15 min	$\pm 1 \times 10^{-2}$
Pendulum	Pendulum	1656	±10 s	$\pm 7 \times 10^{-4}$
Harrison chronometer	Pendulum	1761	±400 ms	$\pm 4 \times 10^{-6}$
Shortt pendulum	Two pendulums, slave and master	1921	±10 ms	$\pm 1 \times 10^{-7}$
Quartz	Quartz crystal	1927	±100 μs	$\pm 1 \times 10^{-9}$
Rubidium	Rubidium atomic resonance (6.834682610904 GHz)	1958	±1 μs	$\pm 1 \times 10^{-11}$
Cesium beam	Cesium atomic resonance (9.19263177 GHz)	1952	±10 ns	$\pm 1 \times 10^{-13}$
Cesium fountain	Cesium atomic resonance (9.19263177 GHz)	1995	±0.1 ns	$\pm 1 \times 10^{-15}$
Hydrogen maser	Hydrogen atomic resonance (1.420405751768 GHz)	1960	±10 ns	$\pm 1 \times 10^{-13}$
CSACs	Cesium or rubidium atomic resonance	2001	±10 μs	$\pm 1 \times 10^{-10}$

shorter than about 1 month. However, they lack the accuracy of cesium standards. Few are manufactured and sold due to their high cost, which often exceeds $200,000 [4–9].

Miniature atomic clocks, known as chip scale atomic clocks (CSACs), based on either rubidium or cesium resonance, were first reported on at NIST in 2001. At this writing (2010), these devices have been reduced to a volume of about 10 cm^3, have low power consumption requirements that can be satisfied by an AA battery, and are stable enough to keep time to within about 10 μs/day [10]. It seems likely that they will eventually be embedded in many types of commercial products.

Future atomic clocks will undoubtedly be based on the optical frequency of atoms, instead of the microwave frequencies. The resonance frequency of an optical atomic clock will be approximately 10^{15} Hz or about 10^5 times higher than the cesium resonance frequency. Thus, optical clocks will operate with a much smaller unit of time, a change comparable to using the second instead of the day as the base unit of time interval and making huge reductions in uncertainty possible. Many experimental optical atomic clocks have already been successfully designed and tested, but it is difficult to predict when they will replace the existing microwave standards [6,11,12].

Table 41.2 summarizes the evolution of clock design and performance. The uncertainties listed are for modern (2010) devices and not the original prototypes. Note that the performance of time and frequency standards has improved by about 14 orders of magnitude in the past 700 years and by about 10 orders of magnitude during the past 100 years.

41.3 Definition of the Second

As noted, the uncertainty of a clock depends upon the irregularity of some type of periodic motion. This periodic motion can be measured and used to define the second, which is the base unit of time interval in the SI. Because atomic time standards were clearly superior to their predecessors, they quickly became the world reference for time interval measurements. The atomic timekeeping era formally began in 1967, when the SI second was redefined based on the resonance frequency of the cesium atom:

> The duration of 9,192,631,770 periods of the radiation corresponding to the transition between two hyperfine levels of the ground state of the cesium 133 atom.

Before the invention of atomic clocks and the acceptance of atomic time, the second was defined by using astronomical time scales. The astronomical definitions of the second were very different from the atomic definition. Instead of *multiplying* the period of a short atomic event to form a longer time interval, the early definitions of the second were based on *dividing* the period of a long astronomical event to form a shorter time interval. There was no official metrological definition of the second until the SI was formed in 1960. However, prior to that date, the second was obtained by dividing the *mean solar day* or the average period of one revolution of the Earth on its axis. The *mean solar second* was 1/86,400 of the mean solar day and served as the base unit of time interval for the Universal Time (UT) family of time scales. Several variations of UT were defined:

UT0: The original mean solar time scale, based on the rotation of the Earth on its axis. UT0 was first kept by pendulum clocks. When quartz clocks became available, astronomers noticed errors in UT0 due to polar motion, which led to the UT1 time scale.

UT1: UT1 improved upon UT0 by correcting for the shift in longitude of the observing station due to polar motion. Because the Earth's rate of rotation is not uniform, the uncertainty of UT1 can be as large as a few milliseconds per day.

UT2: Mostly of historical interest, UT2 is a smoothed version of UT1 that corrects for known deviations in the Earth's rotation caused by angular momenta of the Earth's core, mantle, oceans, and atmosphere.

The *ephemeris second* was defined in 1956 and designated as the original SI second in 1960. It was defined by dividing the tropical year or the interval between the annual vernal equinoxes that occur on or about March 21. The tropical year was defined as 31,556,925.9747 ephemeris seconds. Determining the precise instant of the equinox is difficult, and this limited the uncertainty of ephemeris time (ET) to ±50 ms over a 9 year interval. It also made the ephemeris second nearly impossible to realize in a laboratory and of little or no use to metrologists or engineers. Thus, its tenure was understandably short. It remained part of the SI for just 7 years before being replaced by the much more accessible atomic definition [5,6,13].

41.4 Coordinated Universal Time

Coordinated Universal Time (UTC) is an atomic time scale that is based on the SI definition of the second and that serves as the official time reference for most of the world. UTC is maintained by the Bureau International des Poids et Mesures (BIPM) in Sevres, France. As of 2010, it is computed from a weighted average of nearly 400 atomic standards located at some 70 laboratories, including the National Institute of Standards and Technology (NIST) and the US Naval Observatory (USNO). Most of these devices are cesium beam standards, but some are hydrogen masers. In addition, about ten cesium fountain standards contribute to the accuracy of UTC.

UTC is a virtual time scale, generated from computations made on past measurements, and is distributed only through a monthly BIPM publication called the *Circular T*. No physical clock generates the official UTC. To support physical measurements of time and time interval, many laboratories maintain local UTC time scales that generate real-time signals that approximate UTC as closely as possible, often within a few nanoseconds. The *Circular T* reports the time differences between UTC and these local time scales, which are generically referred to as UTC(k), where k represents the acronym for the timing laboratory. For example, NIST maintains UTC(NIST) and the USNO maintains UTC(USNO).

The BIPM derives UTC from an internal time scale called International Atomic Time (TAI). Both UTC and TAI run at the same frequency. However, UTC differs from TAI by an integer number of seconds. This difference increases when *leap seconds* occur. When necessary, leap seconds are added to the UTC time scale on either June 30 or December 31. Their purpose is to ensure that the difference between atomic time (UTC) and astronomical time (UT1) does not exceed 0.9 s. Time codes (Section 41.5) contain a *DUT1 correction* or the current value of UT1 minus UTC. By applying this correction to UTC, users who need astronomical time can obtain UT1.

Leap seconds have been added to UTC when necessary since 1972. From 1972 to 2008, UT1 has lost (on average) more than 600 ms/year with respect to UTC; thus, 24 leap seconds were added during the 36 year interval. This means that atomic seconds are shorter than astronomical seconds and that UTC runs faster than UT1. There are two reasons for this. The first and most important reason is the definition of the atomic second. The atomic second was originally defined with respect to the ephemeris second and was shorter than the mean solar second from the beginning. The second reason is that the Earth's rotational rate is gradually slowing down and the astronomical second is gradually getting longer. When a positive leap second is added to UTC, the sequence of events is

23 h 59 m 59 s

23 h 59 m 60 s

0 h 0 m 0 s

The insertion of the leap second creates a minute that is 61 s long. This "stops" UTC for 1 s, so that UT1 can catch up [5,6,13]. Some timekeeping laboratories and other organizations have found the leap second to be cumbersome to implement and support, and there is currently an International Telecommunications Union (ITU) proposal to eliminate the leap second and instead make a larger time correction, such as a leap hour, when one is eventually needed. However, leap seconds are still in effect at this writing (2010), and the issue remains unresolved.

41.5 Time Interval Measurements

A common, but not particularly accurate, way to measure time interval is to use an oscilloscope. Oscilloscopes allow us to view the pulse waveforms, making it possible to select the start and end points that define a time interval, a major advantage when dealing with odd-shaped or noisy waveforms. Sometimes one pulse is displayed, while a different pulse is used to trigger the oscilloscope. Another method is to display both time pulses on a dual-trace oscilloscope. Triggering can be achieved from either of the two displayed pulses or from a third external pulse. The divisions along the horizontal axis of an oscilloscope are in units of time, and many oscilloscopes have a delta-time function (Δt) that will measure and displays the time interval between two cursors that can be moved along the horizontal axis. The oscilloscope display in Figure 41.2 shows a 358 ns time interval between two cursors that are aligned with the rising edge of two 1 pps signals.

Time interval measurements made with oscilloscopes are limited by a lack of resolution and range. Oscilloscope range varies, and some models can scale from 50 ps to 10 s per time division. However, the resolution of the time interval measurement is proportional to the length of the interval. If the time pulses are 10 ms apart, for instance, the best oscilloscopes can resolve the interval to only about 1 μs. However, if the pulses are only 100 μs apart, the resolution is 10 ns. In both cases, the relative uncertainty is 1×10^{-4}.

A time interval counter (TIC) is usually the best instrument for measuring all but the shortest time intervals. Universal counters can measure time interval (along with other quantities such as frequency and period), and dedicated TICs are also available. A TIC measures the time interval between a start pulse and a stop pulse. The start pulse gates the TIC (starts it counting), and a stop pulse stops the counting. While the gate is open (Figure 41.3), the counter counts zero crossings from time base oscillator cycles, and the resolution of the simplest TICs is equal to the period of these cycles (e.g., 100 ns for a 10 MHz time base). However, many TICs use interpolation or digital processing schemes to divide the period into smaller parts. For example, dividing 100 ns by 5000 allows 20 ps resolution, a feat achieved

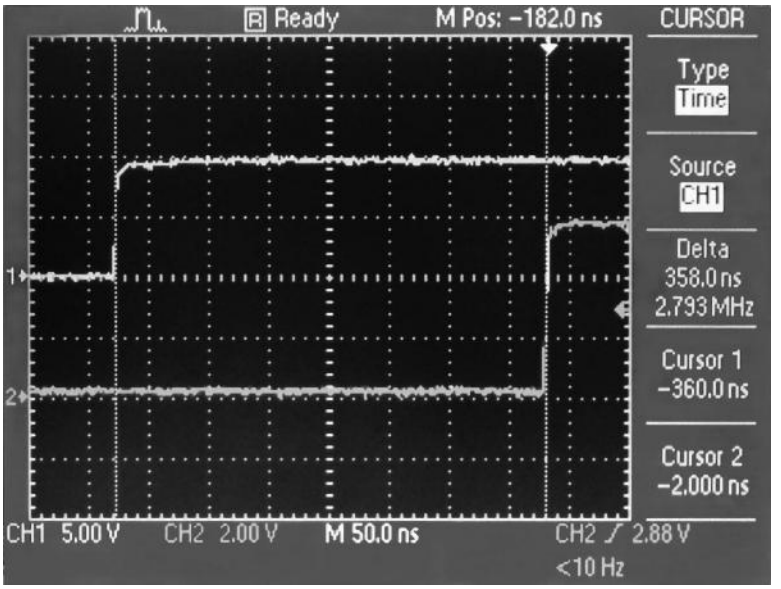


FIGURE 41.2 Measuring time interval with an oscilloscope.

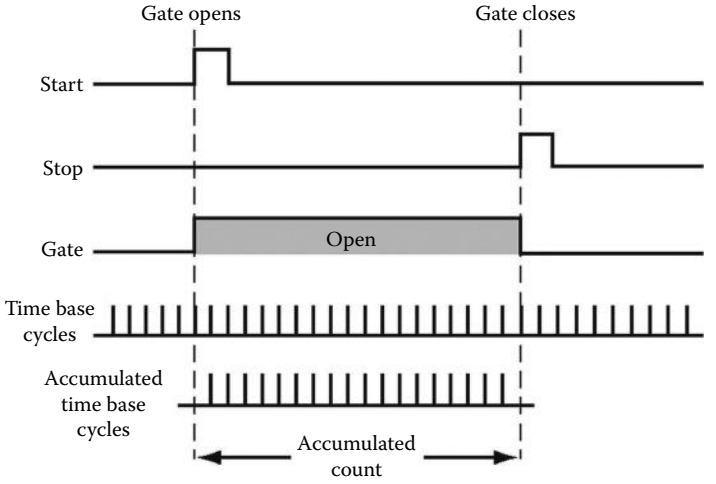


FIGURE 41.3 Gate and counting function of TIC.

by some universal counters, and some dedicated TICs achieve 1 ps resolution. These fine resolutions allow TICs to reach measurement uncertainties of 2×10^{-11} or less over a 1 s interval, often averaging down to parts in 10^{15} or 10^{16} after a few hours.

The uncertainty of TIC measurements can be limited, however, by resolution, trigger errors, or time base errors. A TIC generally has some ambiguity in its least significant digit of resolution. However, this count ambiguity often averages out when multiple readings are taken and is usually a problem only with single measurements. Trigger errors are a more significant problem. They occur when the counter does not trigger at the expected voltage level on the input signal and can be caused by incorrect trigger settings or by input signals that are noisy, too large, or too small. It is imperative to make sure that the TIC is triggering at the proper point on both the start and stop waveforms (if necessary, view the waveforms on an oscilloscope). Time base errors are generally not a problem when measuring short time intervals.

For example, consider a typical counter with 1 ns resolution and a time base stability of 1×10^{-8}. When a 1 ms interval is measured, the time base uncertainty is 10^{-3} s/$10^{8} = 10^{-11}$ s or 10 ps. This is much smaller than the uncertainty contributed by the counter's resolution. Even so, the best oscillator available should be used as the counter's external time base, because time base errors can be significant when long time intervals are measured.

A TIC can also measure medium and long time intervals if it has the necessary range. For example, a TIC that counts 100 ns periods with a 32 bit counter will overflow after $2^{32} \times 100$ ns or about 429 s. However, many TICs keep track of counter overflows, and some have the ability to measure intervals of about 1 day in time interval mode. If the TIC range is insufficient, most universal counters have a totalize mode that simply counts cycles from an external frequency source. Some commercial counters can count up to 10^{15} cycles. The period of the external gate frequency determines the resolution of this type of measurement. For example, if 1 μs resolution is desired, use a 1 MHz gate frequency. In this case, the time interval range would be $10^{15}/10^{6} = 10^{9}$ s or nearly 32 years.

The most common time interval measurements are probably made with stopwatches and similar low-cost timing devices. These devices are used to measure time intervals ranging from several seconds to many hours, and although the measurement uncertainties are high, stopwatches still must be periodically calibrated. These calibrations are typically performed by manually triggering the stopwatch while listening to an audio time signal or viewing a synchronized time display. The legally required uncertainty for stopwatch calibrations is often 1 or 2 parts in 10^{4}, a requirement limited mostly by human reaction time, which is often 10–100 times worse than the uncertainty of the oscillator inside the stopwatch.

Other stopwatch calibration methods reduce or eliminate the problem of human reaction time. One method utilizes a universal counter in the totalize mode described earlier. The counter is gated with a signal from a calibrated signal generator. The gate frequency should have a period at least one order of magnitude smaller than the resolution of the stopwatch. For example, if the stopwatch has 10 ms resolution, use a 1 kHz frequency (1 ms period). The operator then starts and stops the totalize counter by rapidly pressing the start–stop button of the stopwatch against the start–stop button on the counter. The readings from both instruments are recorded, and the equation $\Delta t/T$ is used to get the result, where Δt is the difference between the counter and stopwatch displays and T is the length of the measurement run. For example, if $\Delta t = 100$ ms and $T = 1$ h, the uncertainty is roughly 2.8×10^{-5}, which exceeds the legal requirement.

A second method involves measuring the frequency offset of the oscillator inside the stopwatch, which is usually a 32,768 Hz quartz oscillator. For example, if the stopwatch time base oscillator has a 1 Hz frequency offset, the dimensionless frequency offset ($1/32{,}768 = 3 \times 10^{-5}$) translates to a time offset of about 110 ms in 1 h or near the value for Δt in the previous example. Commercially available instruments with acoustic or inductive sensors can detect the oscillator frequency and automate this measurement [14].

41.6 Time Synchronization Measurements

Many applications require multiple clocks to be *synchronized* or set to the same time. Using the time interval measurement methods described in Section 41.5, a clock can be synchronized by comparing it to a UTC reference and adjusting the time offset until it is as near zero as possible. *Time transfer* is the practice of transferring the time from a reference clock at one location and using it to measure or synchronize a clock at another location.

The reference signals used for time transfer generally need to provide two things: an on-time marker (OTM) and a time code. The OTM is typically a 1 pps signal that is synchronized to the start of the UTC second. The time code provides time-of-day information including the UTC hour, minute, and second, as well as the month, day, and year. These reference signals usually originate from a UTC time scale maintained by a national timekeeping laboratory.

Time can be transferred through many different mediums, including coaxial cables, optical fibers, radio signals, telephone lines, and computer networks. Before discussing the available radio and Internet time transfer signals, the methods used to transfer time [15] are examined.

41.6.1 Time Transfer Methods

As noted, time transfer methods are used to synchronize or compare a local clock to a reference clock. The single largest contributor to time transfer uncertainty is *path delay* or the delay introduced as the signal travels from its source to its destination. To illustrate the path delay problem, consider a reference time signal broadcast by radio. To synchronize a remote clock at a receiving site to the reference time, we need to calculate how long it took for the signal to travel from the transmitter to our receiver. Radio signals travel at the speed of light at about 3.3 μs/km. Therefore, if the receiver is 1000 km from the transmitter, the time will be 3.3 ms late when we receive it. We can compensate for this path delay by making a 3.3 ms adjustment to the remote clock. This practice is called *calibrating the path*.

In the example earlier, we are assuming that the signal travelled the shortest path between the transmitter and receiver, which are not always true. Many factors limit how well the path can be calibrated, and our knowledge of the received time will always be limited by our knowledge of the path delay. However, many innovative ways have been developed to compensate for path delay. The various time transfer methods can be divided into five general categories:

1. *One-way method* (*user calibrates path*): This is the simplest type of time transfer, a one-way system where the user is responsible for calibrating the path (if required). The signal from the transmitter to the receiver is delayed τ_{ab} by the medium (Figure 41.4). To obtain the best results, the user must estimate τ_{ab} and calibrate the path by compensating for the delay. In many cases, however, the time is already accurate enough to meet the user's requirements, so the delay through the medium is simply ignored.
2. *One-way method* (*self-calibrating path*): This method is a variation of the simple one-way method shown in Figure 41.4. However, the time transfer system (and not the user) is responsible for estimating and removing the τ_{ab} delay.

   One of two techniques is commonly used to reduce τ_{ab}. The first technique is to simply estimate τ_{ab} and to send the time out early. For example, if we estimate that τ_{ab} will be at least 20 ms for all users, we can transmit the OTM 20 ms before the arrival of the UTC second and reduce the uncertainty for all users.

   A more sophisticated technique is to compute τ_{ab} and apply a correction. This can be done if the coordinates of both the transmitter and receiver are known. If the transmitter is stationary, a constant can be used for the transmitter position. If the transmitter is moving (e.g., a satellite), it must broadcast its coordinates in addition to broadcasting a time signal. The receiver's coordinates must either be computed by the receiver (in the case of satellite navigation systems) or input by the user. Then, the receiver firmware can compute the distance between the transmitter and receiver and compensate for the path delay by correcting for τ_{ab}. The uncertainty of this method is still limited by position errors for either the transmitter or receiver and by variations in the transmission speed of the signal along the path.

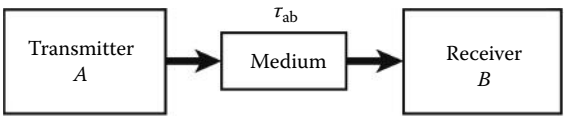


FIGURE 41.4 One-way time transfer.

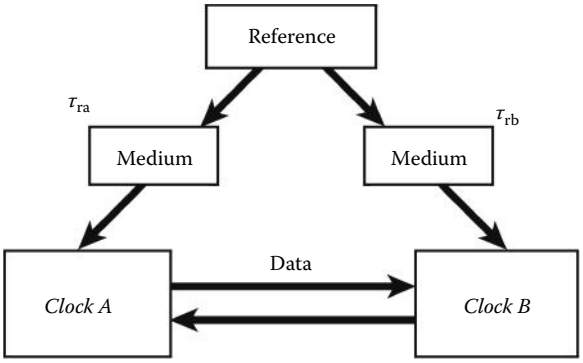


FIGURE 41.5 Common-view time transfer.

3. *Common-view method*: The common-view method involves a reference transmitter (*R*) and two receivers (*A* and *B*). The transmitter is in "common view" to both receivers. Both receivers compare the simultaneously received signal to their local clock and record the data (Figure 41.5). Receiver *A* receives the signal over the path τ_{ra} and compares the reference to its local clock (*R* – *Clock A*). Receiver *B* receives the signal over the path τ_{rb} and records (*R* – *Clock B*). The difference between the two measurements is an estimate of the difference between the two clocks. Errors from the two paths (τ_{ra} and τ_{rb}) that are common to the reference cancel out, eliminating much of the uncertainty caused by path delay. The result of the measurement is (*Clock A* – *Clock B*) – ($\tau_{ra} - \tau_{rb}$).

   Unlike the one-way methods, the common-view method cannot synchronize clocks in *real time*, because data from both receiving sites must be transferred and processed before the final measurement results are known. However, the Internet makes it possible to transfer and process the data very quickly. Therefore, common-view data can synchronize clocks in *near real time*, a method now commonly employed by NIST and other laboratories.
4. *Two-way method*: The two-way method requires two users to both transmit and receive timing signals through the same medium at the same time (Figure 41.6). This differs from the passive common-view method where timing signals are only received and not transmitted. Sites *A* and *B* simultaneously exchange time signals through the same medium and compare the received signals with their own clocks. Site *A* records $A - (B + \tau_{ba})$ and site *B* records $B - (A + \tau_{ab})$, where τ_{ba} is the path delay from *B* to *A* and τ_{ab} is the path delay from *A* to *B*. The difference between the two clocks is $(A - B)/2 + (\tau_{ba} - \tau_{ab})/2$. If the path is reciprocal ($\tau_{ab} = \tau_{ba}$), then the difference is simply $(A - B)/2$.

   When properly implemented, the two-way method outperforms all other time transfer methods and has many potential applications. It can, however, be complex to implement. When wireless mediums are used, it can require a substantial amount of equipment, as well as government licensing so that users are allowed to transmit. Like the common-view method, the two-way method requires measurement data to be transferred. However, because users have the ability to transmit, it is possible to send data through the same medium as the time signals.
5. *Loop-back method*: A variation of the two-way method, the loop-back method also requires the receiver to send information back to the timing source. For example, an OTM is sent from the transmitter (*A*) to the receiver (*B*) over the path τ_{ab}. The receiver (*B*) then sends the OTM back to

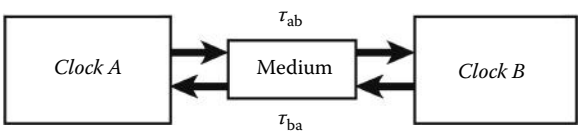


FIGURE 41.6 Two-way time transfer.

the transmitter (*A*) over the path $\tau_{ba.}$ The one-way path delay is assumed to be half of the measured round trip delay or $(\tau_{ab} + \tau_{ba})/2$. The transmitter calibrates the path by sending another OTM that has now been advanced by the estimated one-way delay. The method usually works well but is less accurate than the two-way method because the OTM transmissions are not simultaneous, and it is not known whether the path delay is the same in both directions. The loop-back method is not practical to use through a wireless medium and is typically applied to telephone or network connections where it can be implemented entirely in software.

41.6.2 Time Codes

A *time code* is a message containing time-of-day information that allows the user to synchronize a clock. ITU guidelines state that all time codes should distribute the UTC hour, minute, and second, as well as a DUT1 correction.

Time codes are broadcast in a number of different formats (including binary, binary coded decimal [BCD], and ASCII). There is very little standardization of broadcast time codes, but standards for redistributing time codes within a facility were first developed by the Inter-Range Instrumentation Group (IRIG) in 1956 and are still used by today's equipment manufacturers. IRIG defined numerous time code formats, but the most common is probably IRIG-B. These standardized time codes make it possible for manufacturers to build compatible equipment. For example, a satellite receiver with an IRIG-B output can drive a time-of-day display that accepts an IRIG-B input. Or it can provide a timing reference to a computer that can read IRIG-B.

The IRIG time code formats are serial, width-modulated codes that can be used in either dc level shift or amplitude-modulated (AM) form. For example, IRIG-B has a 1 s frame period and can be transmitted as either a dc level shift modulation envelope or a modulated 1000 Hz carrier. Time-of-day data (days, hours, minutes, and seconds) in BCD or straight binary format is included within the frame. Simple IRIG-B decoders retrieve just the encoded data and provide 1 s resolution. Other decoders count carrier cycles and provide resolution equal to the period of the 1000 Hz cycle (1 ms). More advanced decoders phase lock an oscillator to the time code and provide resolution limited only by the signal-to-noise ratio of the time code (typically ±2 μs).

41.7 Radio Time Transfer

Several types of receivers and signals are used to transfer time by radio. The cost of a time transfer receiver can vary widely, from less than $10 for a simple radio-controlled clock that synchronizes once per day and keeps time within 1 s of UTC to $20,000 or more for a receiver that adjusts a rubidium oscillator with satellite signals and keeps time within nanoseconds of UTC. When selecting a time transfer receiver, make sure that the signal can be received at your location, that the uncertainty is low enough to meet your requirements, and that the appropriate type of antenna can be mounted. The following sections summarize the various signals used for radio time transfer.

41.7.1 HF Time Signal Stations

High-frequency (HF) or *shortwave* radio signals are used for time transfer at moderate performance levels. These stations are useful because they provide worldwide coverage under optimal receiving conditions, they can be heard with any shortwave receiver, and they provide audio time announcements that can serve as a "talking clock." The audio signals are also widely used by metrologists as a time interval reference for stopwatch and timer calibrations. The practical limit of time transfer uncertainty with HF stations is about 1 ms. This is because the signals often travel by *sky wave* and the path length varies, making it difficult to accurately determine the path delay.

TABLE 41.3 HF Time Signal Station List

Call Sign	Location	Frequencies (MHz)	Controlling Agency
WWV	Fort Collins, Colorado, United States	2.5, 5, 10, 15, 20	NIST
WWVH	Kauai, Hawaii, United States	2.5, 5, 10, 15	NIST
BPM	Lintong, China	2.5, 5, 10, 15	National Time Service Center (NTSC)
CHU	Ottawa, Canada	3.33, 7.85, 14.67	National Research Council (NRC)
HLA	Taejon, Korea	5	Korea Research Institute of Standards and Science (KRISS)

Table 41.3 provides a list of HF time signal stations that are referenced to UTC as of 2010. Only a few stations remain, as many have been turned off in recent years. This is probably because far more accurate time signals can now be easily obtained from satellites and because low accuracy time (within 1 s) can now be easily obtained from the Internet or from inexpensive low-frequency (LF) radio-controlled clocks.

The best known HF time signal stations are WWV and WWVH, both operated by NIST. WWV is located near Fort Collins, CO, and WWVH is on the island of Kauai, HI. Both stations broadcast continuously on 2.5, 5, 10, and 15 MHz, with WWV also available on 20 MHz. All frequencies carry the same audio broadcast, which includes short pulses transmitted every second that sound similar to the ticking of a clock. At the start of each minute, a voice announces the current UTC hour and minute. WWV uses a male voice to announce the time, and WWVH uses a female voice. The voice announcement is followed by a long audio pulse of 800 ms in duration. In addition to the audio, a time code is sent on a 100 Hz subcarrier at a 1 bit per second (bps) rate.

41.7.2 LF Time Signal Stations

LF signals are seldom used for high-accuracy time transfer, but they have a major advantage over HF and satellite signals—they can be easily received indoors without an external antenna. This makes LF time signal stations the ideal synchronization source for radio-controlled clocks and wristwatches. Signals from NIST radio station WWVB synchronize an estimated total of more than 50 million radio-controlled clocks in the United States daily as of 2010, with many millions of new clocks being sold each year.

Several countries operate LF time signal stations at frequencies ranging from 40 to 77.5 kHz (Table 41.4). These stations do not provide voice announcements but do provide an OTM and a time code.

NIST radio station WWVB, which covers most of North America, has a format similar, but not identical, to that of the other stations listed in Table 41.4. WWVB requires a full minute to send its

TABLE 41.4 LF Time Signal Station List

Call Sign	Location	Frequencies (kHz)	Controlling Agency
WWVB	Fort Collins, Colorado, United States	60	NIST
BPC	Lintong, China	68.5	NTSC
DCF77	Mainflingen, Germany	77.5	Physikalisch-Technische Bundesanstalt (PTB)
JJY	Japan	40, 60	National Institute of Information and Communications Technology (NICT)
MSF	Rugby, United Kingdom	60	National Physical Laboratory (NPL)
RBU	Moscow, Russia	66.67	Institute of Metrology for Time and Space (IMVP)

time code in BCD format. Bits are sent a rate of 1 bps by shifting the power of the carrier. The carrier power is reduced 17 dB at the start of each second, and the first carrier cycle after the power drop serves as the OTM. If full power is restored after 200 ms, it represents a binary 0. If full power is restored after 500 ms, it represents a binary 1. Reference markers and position identifiers are sent by restoring full power after 800 ms. The WWVB time code provides year, day, hour, minute, and second information, a DUT1 correction, and information about daylight saving time, leap years, and leap seconds.

Calibrating the path is easier with LF signals than HF signals, because part of the LF signal is *ground wave* and follows the curvature of the Earth. However, due to uncertainties in the estimate of path delay and because it is difficult to determine the correct OTM to within better than a few cycles of the carrier, the practical limit for timing uncertainty with WWVB and similar stations is about 0.1 ms or roughly 1000 times larger than the timing uncertainty of GPS, as explained in the next section.

41.7.3 Global Positioning System

The Global Positioning System (GPS) is a global navigation satellite system (GNSS) developed and operated by the US Department of Defense (US DoD). GPS was designed as a positioning and navigation system but has become the main system used to distribute high-accuracy time signals worldwide. The GPS constellation includes as many as 32 satellites in semi-synchronous orbit at a height of 20,200 km. The orbital period is 11 h and 58 min, which means that each satellite passes over a given location on Earth 4 min earlier than it did on the previous day. By processing GPS signals, even a low-cost handheld receiver can determine its position on Earth with an uncertainty of a few meters.

The GPS satellites carry atomic clocks that are steered from US DoD ground stations to agree with UTC(USNO), which is normally well within 20 ns of UTC. GPS time must be accurate in order for the system to meet its specifications for positioning and navigation. To illustrate this, consider that the satellites receive clock corrections from Earth-based control stations once during each orbit or about once every 12 h. The maximum acceptable contribution from the satellite clocks to the positioning uncertainty is considered to be about 1 m. Since light travels at about 3×10^8 m/s, the 1 m requirement means that the time error between clock corrections must be less than about 3.3 ns.

All GPS satellites broadcast on the *L*1 (1.57542 GHz) and *L*2 (1.2276 GHz) carrier frequencies, with satellites launched after May 2010 also utilizing the *L*5 carrier at 1.17645 GHz. The satellites are identified by a unique spread-spectrum waveform, called a pseudorandom noise (PRN) code, which it transmits on each carrier. There are two types of PRN codes. The first type is a coarse acquisition (*C*/*A*) code with a chip rate of 1.023 megabits/s. The second type is a precision (*P*) code with a chip rate of 10.230 megabits/s. The *C*/*A* code is broadcast on *L*1, and the P code is broadcast on both *L*1 and *L*2. A 50 bit/s data message is also broadcast on both carriers [16–18]. Dual-frequency timing receivers have become more common as of 2010, but most timing devices receive only *L*1. Nearly all GPS timing receivers can simultaneously track at least eight satellites.

41.7.4 One-Way Time Transfer Using GPS

GPS receivers transfer time from the satellites to the receiver clock through a series of range measurements. The range measurements used to calculate position are derived by measuring the time required for the signals to travel from each satellite to the receiver. After the receiver position (*x, y, z*) is solved for, the solution is stored. Then, by using the travel time of the signal and the exact time when the signal left the satellite, time from the satellite clocks can be transferred to the receiver clock.

This time difference between the satellite and receiver clocks, when multiplied by the speed of light, produces not the true geometric range but rather the pseudorange. The equation for the pseudorange observable is

$$p = \rho + c \times (dt - dT) + d_{\text{ion}} + d_{\text{trop}} + r_{\text{n}} \tag{41.2}$$

where

p is the pseudorange
c is the speed of light
ρ is the geometric range to the satellite
dt and dT are the time offsets of the satellite and receiver clocks with respect to GPS time
d_{ion} is the delay through the ionosphere
d_{trop} is the delay through the troposphere
r_{n} is the effects of receiver and antenna noise, including multipath

Estimates of d_{ion} and d_{trop} are obtained from the GPS broadcast.

The receiver produces a local time estimate for each satellite by using the pseudorange data to compensate for propagation delay and by applying satellite clock corrections received from the broadcast. The receiver then combines the satellite time estimates, by simple averaging or another statistical technique, and uses this information to synchronize a 1 pps timing signal. Two additional corrections, one large and one small, are also necessary to convert GPS time to UTC(USNO). GPS time differs from UTC(USNO) by the number of leap seconds that have occurred since the origination of the GPS time scale (January 6, 1980). It also differs from UTC(USNO) by a small number of nanoseconds that continuously change. Both corrections are included in data messages broadcast by the satellites and are automatically applied by the receiver firmware. Thus, both the time-of-day and the OTM obtained from a GPS timing receiver are referenced to UTC(USNO).

Several factors limit the uncertainty of GPS time synchronization, including receiver and antenna cable delays, antenna coordinate errors, ionospheric and tropospheric delay effects noted earlier, multipath reflections, and environmental effects. Even when all of these factors are ignored, the uncertainty of the time produced by a GPS receiver will likely be less than 1 μs with respect to UTC. By calibrating the receiver and accounting for the factors listed earlier, it is usually easy to reduce this uncertainty to within ±100 ns, with ±20 ns being the practical limit.

41.7.5 Common-View Time Transfer Using GPS

Unlike the one-way method, the common-view GPS (CVGPS) method does not use GPS time as the reference source. Instead, GPS is simply a vehicle used to transfer time from one site to another. The CVGPS method compares two clocks at different locations to each other by simultaneously comparing them both to GPS signals that are in "common view." The comparison results are recorded and exchanged, and the difference between the two comparisons is the time difference between the two clocks. Because GPS is available worldwide, the CVGPS method can potentially be used to compare any two clocks on Earth to each other or to synchronize a given clock on Earth to any other clock.

The CVGPS method involves one or more GPS satellites (S) and two receiving sites (A and B), each containing a GPS receiver and a local clock (Figure 41.7). The satellites transmit signals that are received at both A and B, and both sites compare the received signals to their local clock. Thus, the measurement at site A compares GPS signals, S, received over the path d_{SA} to the local clock, *Clock A – S*. Site B receives GPS signals over the path d_{SB} and measures *Clock B – S*.

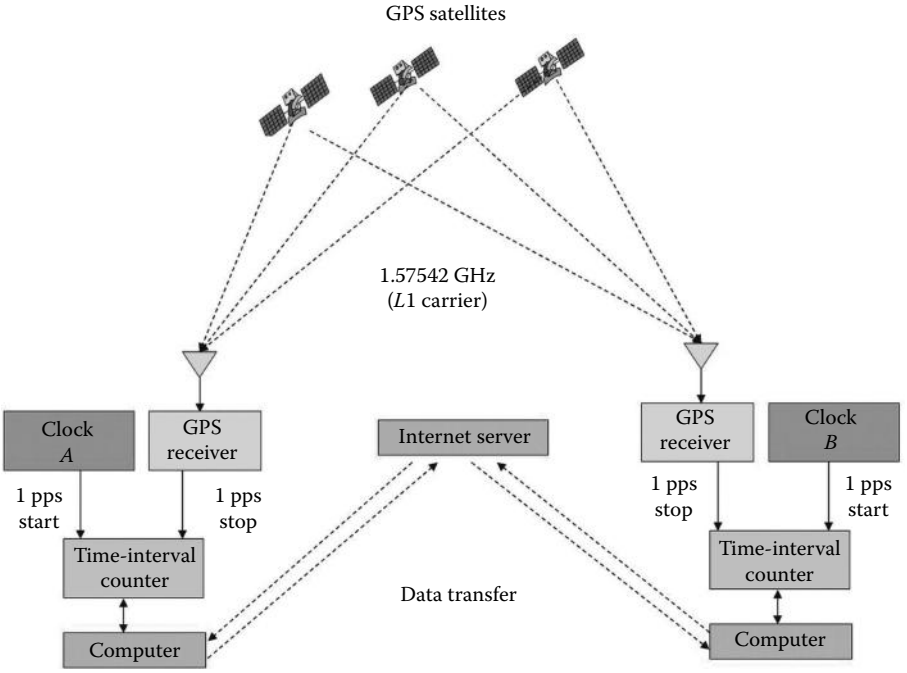


FIGURE 41.7 *C/A* code CVGPS using the *L*1 carrier.

The difference between the two measurements is an estimate of *Clock A – Clock B*. Delays that are common to both paths d_{SA} and d_{SB} cancel even if they are unknown, but uncorrected delay differences between the two paths add uncertainty to the measurement result. Thus, the basic equation for a CVGPS measurement is

$$(Clock\ A - S) - (Clock\ B - S) = (Clock\ A - Clock\ B) + (e_{SA} - e_{SB}) \tag{41.3}$$

The components that make up the $(e_{SA} - e_{SB})$ error term include delay differences between the two sites caused by ionospheric and tropospheric delays, multipath signal reflections, environmental conditions, and errors in the GPS antenna coordinates. These factors can be measured or estimated and applied as a correction to the measurement, or they can be accounted for in the uncertainty analysis. It is also necessary to calibrate the GPS receivers used at both sites and account for the local delays in the receiver, antenna, and antenna cable.

Figure 41.7 is a simplified diagram that illustrates *C/A* code common view using the *L*1 carrier, but there are several other variations of the CVGPS measurement technique that can provide lower measurement uncertainties. The uncertainty depends upon the type and quality of the GPS equipment in use and the technique used. For example, the differential ionospheric delay can be nearly eliminated by receiving both the *L*1 and *L*2 carrier frequencies. The quality of equipment also makes a difference; some GPS receivers are less sensitive to environmental changes than others, and some antennas are more effective than others at mitigating multipath. The most sophisticated techniques and equipment compare the local clock to the GPS carrier, rather than to the *C/A* code, and can reduce the time uncertainty to a few nanoseconds or less, but the incremental performance gains obtained from the additional cost and effort are relatively small. Even when inexpensive GPS hardware (*L*1 band only) and simple processing techniques are used, the measurement uncertainty of the CVGPS technique is often within ±15 ns, with ±5 ns being the practical limit for the most advanced techniques.

TABLE 41.5 Satellite Systems Used for Time Transfer

System	Controlling Region	Type of System	Completion Date	Number of Satellites
Compass	China	GNSS	2015	30
EGNOS	European Union	Augmentation, geostationary orbit	Completed	3
GPS	United States	GNSS	Completed	32
Galileo	European Union	GNSS	2015	30
GLONASS	Soviet Union	GNSS	Completed	24
QZSS	Japan	Augmentation, highly elliptical orbit	2013	3
WAAS	United States	Augmentation, geostationary orbit	Completed	2

41.7.6 Other Satellite Systems Used for Time Measurements

Several other GNSS and augmentation systems can be used a reference for time measurements. Receivers for these systems are not yet available in some cases as of 2010, but performance should eventually be comparable to GPS in both one-way and common-view modes. In many cases, the systems were designed to interoperate with each other and thus share the same or similar frequency bands as GPS, with carrier frequencies ranging from about 1.1 to 1.6 GHz. Table 41.5 summarizes the various systems, including GPS.

41.7.7 Two-Way Satellite Time Transfer via Geostationary Satellites

The two-way satellite time transfer (TWSTT) method (Figure 41.8) requires more equipment and effort than the passive GPS methods, because users are required to both transmit and receive time signals. Signals are transmitted and received from two Earth stations through the transponder on a geostationary satellite. Each Earth station contains a local clock, a spread-spectrum satellite modem, a dish antenna, a TIC, and transmitting and receiving equipment. The carrier frequency used for the radio link is typically in the K_u band.

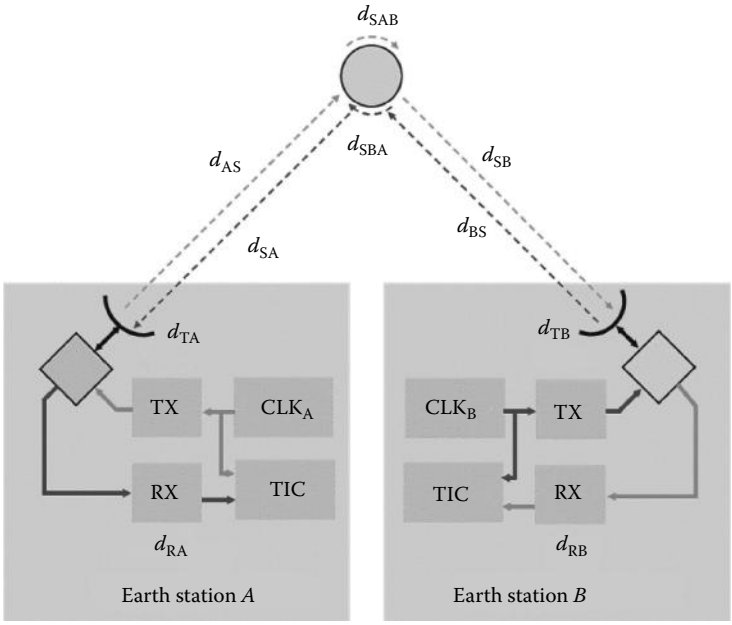


FIGURE 41.8 Two-way satellite time transfer.

TABLE 41.6 Performance of Radio Time Transfer Signals

Method	Time Uncertainty with Little or No Effort Made to Calibrate Path or Equipment	Practical Uncertainty Limit	Items Needed to Reach Uncertainty Limit
HF (WWV)	<30 ms	±1 ms	Good path delay estimate, radio propagation model
LF radio (WWVB)	<15 ms	±0.1 ms	Good path delay estimate, ability to determine the correct OTM to within a few cycles of the carrier
GPS (one way)	<1 μs	±20 ns	Calibrated receiver with antenna surveyed to within 1 m
GPS (common-view code)	±50 ns	±15 ns	Calibrated receivers at both sites with antennas surveyed to within 1 m and antennas designed to mitigate multipath
GPS (common-view carrier)	±10 ns	±5 ns	Calibrated receivers at both sites that accept the local clock input and that produce files with carrier phase observables
TWSTT	±10 ns	±1 ns	Calibrated satellite terminals at both Earth stations, good environmental control

TABLE 41.7 Suppliers of Time Standards and Measurement Equipment

Company	Website	Rubidium Oscillators	Cesium Oscillators	GPS Timing Receivers	TICs
Agilent	agilent.com				X
Accubeat	accubeat.com	X		X	
Arbiter	arbiter.com			X	
Berkeley Nucleonics	berkeleynucleonics.com				X
Brandywine	brandywinecomm.com	X		X	X
Brilliant	b-i-inc.com				X
DICOM	dicom.cz			X	
EndRun	endruntechnologies.com			X	
FEI-Zyfer	fei-zyfer.com			X	
Fluke	fluke.com	X		X	X
Frequency Electronics	freqelec.com	X	X		
FURUNO	furunogps.com			X	
Geotest	geotestinc.com				X
JAVAD	javad.com			X	
Meinberg	meinberg.de			X	
Oscilloquartz	oscilloquartz.com	X	X	X	X
Precision Test Systems	ptsyst.com	X		X	X
Spectracom	spectracomcorp.com	X		X	X
Stanford Research	thinksrs.com	X			X
Symmetricom	symmetricom.com	X	X	X	X
Timetech	timetech.de			X	X
TRAK	trak.com			X	
Trimble	trimble.com			X	

Earth stations *A* and *B* simultaneously exchange time signals by transmitting a unique PRN code and receiving the PRN code sent by the other station. Each station compares the received time signal to its local clock. Station A records $TIC_A = CLK_A - (CLK_B + \tau_{ba})$, where CLK_A is the time from the local clock, CLK_B is the time from the remote clock, and τ_{ba} includes all of the terms that make up the path delay from *B* to *A*, including delays in the transmitting modem, the satellite uplink, the satellite transponder, the satellite downlink, and the receiving modem. Station *B* records $TIC_B = CLK_B - (CLK_A + \tau_{ab})$, where τ_{ba} includes all of the terms that make up the path link from *B* to *A*.

The two stations then exchange their results and compute their time difference:

$$A - B = \frac{TIC_A - TIC_B}{2} - \frac{\tau_{ba} - \tau_{ab}}{2}. \tag{41.4}$$

If the paths were reciprocal ($\tau_{ab} = \tau_{ba}$), then the difference would simply be $(TIC_A - TIC_B)/2$. The paths will never be completely reciprocal in practice, but most of the significant differences in the path delay do cancel. When properly implemented, the two-way satellite method outperforms all other satellite time transfer methods. The uncertainty can often be reduced to ±1 ns [12].

Table 41.6 summarizes the various radio time transfer methods. Table 41.7 provides a list of companies that manufacture and sell atomic oscillators, GPS timing receivers, and TICs.

41.8 Internet and Network Time Transfer

The synchronization of computer clocks is a very common time transfer application. Computers connected to the Internet can synchronize their clocks quickly and easily because Network Time Protocol (NTP) [19] client software is often included with the operating system. NTP is a hierarchical clock system, where each level of the hierarchy is termed a stratum. Synchronization flows from primary servers at the lowest stratum to secondary servers at progressively higher strata. Servers referenced to national time standards are designated as Stratum-1. As of 2010, NIST maintains about half of the approximately 50 Stratum-1 NTP servers located in the United States that are accessible to the general public (see ntp.org for a current list). The demand for these servers is high, and the NIST servers handle several billion NTP requests per day.

Client computers (which can also be servers to computers at higher stratums) issue time requests to an NTP server via TCP/IP port 123. They receive a data packet that contains a UTC time code and other time stamps that indicate when the request was made, when it arrived, and when the packet was sent. This information is used to calibrate the path, to synchronize the client's clock, and to steer the client's clock between synchronizations. The performance of NTP depends upon the client hardware and software and network conditions and varies widely, but uncertainties of less than a millisecond have been demonstrated under optimal conditions. A typical computer, running the Microsoft Windows operating system on the public Internet, can usually keep time within 1 or 2 s of UTC with one NTP synchronization request per week.

The Precision Time Protocol (PTP) can obtain much lower uncertainties than NTP, often less than 1 μs, but is generally not suited for the public Internet. Defined by the IEEE-1588 standard, PTP is primarily intended for local area networks where path delays can be better measured and estimated. The grandmaster clock is the time reference for all other clocks in a PTP system. The other clocks are designated as ordinary clocks, which have a single PTP port, and boundary clocks, which have multiple network connections and can bridge synchronization from one network segment to another.

Time transfer methods normally associated with satellites, such as the two-way and loop-back methods, are also applicable to private or dedicated networks. In fact, experiments using optical fibers as the time transfer medium over short distances have demonstrated much lower uncertainties than any of the satellite-based methods [20]. However, there are technical and financial obstacles that make it difficult to implement these techniques over long distances. For example, it is not easy to maintain the symmetry

of the path delay in very long fibers, and it is hard to justify the cost of an optical fiber network that is dedicated to time transfer. Even so, it seems likely that optical fiber networks will play a large role in future state-of-the-art time transfer experiments.

41.9 Future Developments

Optical atomic clocks will eventually replace cesium fountain clocks as the world's most accurate time standards, and the SI second will eventually be redefined, with the new definition based on an optical atomic transition, rather than a microwave transition. When this occurs, the uncertainty of the physical realization of the SI second is likely to improve by at least two orders of magnitude.

As clocks continue to improve, time transfer systems will in turn have to lower their measurement uncertainties in order to improve the performance of UTC. The future of time transfer should involve continued enhancements to satellite systems and measurement techniques, an increased reliance on signals sent over optical fibers, and perhaps the use of technologies yet to be developed.

Time-of-day clocks that are synchronized to within 1 s of UTC will become more common and should eventually become the norm rather than the exception. Many technologies could contribute to this trend, including LF radio signals, satellite signals, mobile phone signals, Internet time codes, and CSACs.

References

1. J. Jespersen and J. Fitz-Randolph, *From Sundials to Atomic Clocks: Understanding Time and Frequency*, Dover Publications, New York, p. 308, 1999.
2. D. Howse, *Greenwich Time and the Longitude*, Philip Wilson Publishers, Ltd., London, U.K., p. 199, 1998.
3. T. Jones, *Splitting the Second*: *The Story of Atomic Time*, Institute of Physics Publishing, Bristol, England, p. 199, 2000.
4. F. G. Major, *The Quantum Beat: The Physical Principles of Atomic Clocks*, Springer-Verlag, New York, p. 475, 1998.
5. C. Audoin and B. Guinot, *The Measurement of Time*: *Time, Frequency, and the Atomic Clock*, Cambridge University Press, Cambridge, U.K., p. 335, 2001.
6. D. D. McCarthy and P. K. Seidelmann, *Time: From Earth Rotation to Atomic Physics*, Wiley-VCH, Weinheim, Germany, 2009.
7. W. M. Itano and N. F. Ramsey, Accurate measurement of time, *Sci. Am*., 269(1), 56–65, 1993.
8. M. A. Lombardi, T. P. Heavner, and S. R. Jefferts, NIST primary frequency standards and the realization of the SI second, *Measure: J. Meas. Sci*., 2(4), 74–89, December 2007.
9. M. Desaintfuscien, *Data Processing in Precise Time and Frequency Applications*, Springer-Verlag, New York, p. 217, 2007.
10. S. Knappe, MEMS atomic clocks, in *Comprehensive Microsystems*, Y. B. Giamnchondani, O. Tabata, and H. Zappe (eds.), Vol. 3, Elsevier B. V., Amsterdam, the Netherlands, pp. 571–612, 2007.
11. S. Diddams, J. Bergquist, S. Jefferts, and C. Oates, Standards of time and frequency at the outset of the 21st century, *Science,* 306, 1318–1324, November 19, 2004.
12. L. Lorini, N. Ashby, A. Brusch, S. A. Diddams, R. E. Drullinger, E. Eason, T. Fortier et al., Recent atomic clock comparisons at NIST, *Eur. Phys. J. Spec. Top*., 35, 19–35, 2008.
13. P. K. Seidelmann, Ed., *Explanatory Supplement to the Astronomical Almanac*, University Science Books, Mill Valley, CA, 1992.
14. J. C. Gust, R. M. Graham, and M. A. Lombardi, *Stopwatch and Timer Calibrations*, NIST Special Publication 960-12, Gaithersburg, MD, p. 66, January 2009.
15. J. Levine, A review of time and frequency transfer methods, *Metrologia*, 45, S162–S174, 2008.

16. E. D. Kaplan and C. Hegerty, *Understanding GPS: Principles and Applications*, 2nd edn., Artech House Publishers, Boston, MA, 2005.
17. T. E. Parker and D. Matsakis, Time and frequency dissemination: Advances in GPS transfer techniques, *GPS World*, 15, 32–38, November 2004.
18. J. Levine, Time and frequency distribution using satellites, *Rep. Prog. Phys*., 65, 1119–1164, July 2002.
19. D. L. Mills, *Computer Network Time Synchronization: The Network Time Protocol*, CRC Press, Boca Raton, FL, 2006.
20. S. M. Foreman, A. D. Ludlow, M. H. G. de Miranda, J. E. Stalnaker, S. A. Diddams, and J. Ye, Coherent optical phase transfer over a 32-km fiber with 1 s instability at 10^{-17}, *Phys. Rev. Lett*., 99, 153601, 2007.

42

Frequency Measurement

Michael A. Lombardi
National Institute of Standards and Technology

Frequency is the rate of occurrence of a repetitive event. If T is the period of a repetitive event, then the frequency is its reciprocal, $f = 1/T$. The International System of Units (SI) states that the period should be expressed in seconds (s), and the frequency should be expressed in hertz (Hz). The frequency of electrical signals is often stated in units of kilohertz (kHz), megahertz (MHz), or gigahertz (GHz), where 1 kHz equals one thousand (10^3) events per second, 1 MHz equals one million (10^6) events per second, and 1 GHz equals one billion (10^9) events per second.

Frequency metrology first became a topic of interest around 1920, when the burgeoning radio industry began erecting radio transmitters all over the world. These transmitters had to stay near their assigned frequencies to avoid interference with signals from other stations. In addition, the many millions of radio receivers that were soon manufactured had to be able to tune to a desired frequency so that the selected station could be heard. The original requirements for transmitter accuracy were low, about one part per thousand (1×10^{-3}), but at the time, they posed a challenging metrology problem. Within a few years, however, the development of the quartz oscillator soon made that type of accuracy trivial. By the 1930s, commercial quartz oscillators accurate to about one part per million (1×10^{-6}) were widely available, and by the late 1960s, quartz technology with about the same accuracy found its way into low-cost wristwatches and clocks. The production of commercial atomic oscillators also began in earnest in the 1960s, resulting in huge improvements in accuracy and resulting in many new technologies. For example, the infrastructure that we now take for granted, including telecommunication networks and the electric power grid, requires frequency accurate to about 1×10^{-11} to be simultaneously generated around the world during every hour of every day. In calibration and metrology laboratories, frequency measurements accurate to one part per ten trillion (1×10^{-13}) have now become routine, and the Global Positioning System (GPS) depends upon oscillators that are stable to parts in 10^{14} for multiple hours. This level of performance separates frequency metrology from most of the other fields of metrology, where one part per billion (1×10^{-9}) is often either unattainable or considered a major accomplishment.

This chapter provides an overview of frequency measurements. It focuses on the measurement of the electrical signals produced by oscillators. For our purposes, an *oscillator* is a device that

produces electrical signals at a specific frequency, typically in the form of either a sine or a square wave. Section 42.1 begins by discussing the concepts of accuracy and stability, which are essential to understanding oscillator specifications. Section 42.2 then describes the various types of oscillators used as frequency standards, including quartz and atomic oscillators, and oscillators disciplined to agree with an external reference signal. Section 42.3 describes the methods and techniques used to measure frequency and calibrate oscillators. Section 42.4 provides a brief look at the likely future of frequency metrology.

42.1 Frequency Accuracy and Stability

This section looks at the two main specifications used to characterize an oscillator: accuracy and stability. A good understanding of the basic concepts introduced in this section is necessary when evaluating equipment or performing measurements.

42.1.1 Frequency Accuracy

The accuracy of an oscillator is the difference between its actual frequency, as determined by measurement, and its *nominal frequency*. The nominal frequency is labeled on the oscillator output and refers to an ideal frequency with zero uncertainty. For example, an oscillator with an output labeled "10 MHz" would ideally produce perfect 10 MHz signals, but its actual signals will differ from its nominal frequency by some amount. The difference between the actual frequency and the nominal frequency is called the *frequency offset* and determines the accuracy of an oscillator at a given point in time or over a specified interval.

Frequency offset is measured by comparing a test oscillator to a more accurate reference oscillator. There are several established measurement methods (described later in Section 42.3) that can provide this comparison in either the *frequency domain* or the *time domain*. The standard equation for estimating frequency offset in the frequency domain is

$$f_{\text{off}} = \frac{f_{\text{meas}} - f_{\text{nom}}}{f_{\text{nom}}}, \tag{42.1}$$

where

f_{off} is the frequency offset
f_{meas} is the actual frequency in hertz reported by the measurement
f_{nom} is the nominal frequency in hertz that the oscillator would ideally produce

Note that in practice, f_{meas} has an associated measurement uncertainty, but f_{nom} is always an ideal value with no uncertainty. Note also that the nominal frequency is included in both the numerator and the denominator. Thus, the resulting value for f_{off} is dimensionless, and not in units of hertz.

This equation is often simplified in the literature as

$$f_{\text{off}} = \frac{\Delta f}{f}, \tag{42.2}$$

where

f_{off} is the dimensionless frequency offset
Δf is the difference between the measured and nominal frequency in hertz
f is the nominal frequency in hertz

For example, if an oscillator labeled as 10 MHz (10^7 Hz) produces a frequency that is higher than nominal by 1 Hz, the equation becomes

$$f_{\text{off}} = \frac{1}{10^7} = 1 \times 10^{-7} \tag{42.3}$$

In many cases, the frequency offset of an oscillator is obtained in the *time domain* by measuring *time interval*. This works because frequency is the reciprocal of period, which is expressed as a time interval. A mathematical definition of frequency is

$$f = \frac{1}{T} \tag{42.4}$$

where

T is the period of the signal in seconds
f is the frequency in hertz

This can also be expressed as

$$f = T^{-1} \tag{42.5}$$

If we perform mathematical differentiation on the frequency expression with respect to time and substitute in the result, we can show that the average dimensionless difference in frequency is equivalent to the average dimensionless difference in time or that $\Delta f/f$ is equivalent to $\Delta t/T$. For example,

$$\Delta f = -T^{-2}\Delta t = -\frac{\Delta t}{T^2} = -\frac{\Delta t}{T} f \tag{42.6}$$

Therefore,

$$f_{\text{off}} = \frac{\Delta f}{f} = -\frac{\Delta t}{T} \tag{42.7}$$

where

Δt is the difference between two time interval measurements
T is the elapsed time between the two measurements

To keep the sign correct, note that the first reading must be subtracted from the second, therefore,

$$f_{\text{off}} = \frac{TI_2 - TI_1}{T} = -\frac{\Delta t}{T} \tag{42.8}$$

To illustrate this, consider a simple example where a time interval (TI_1) is measured, followed by another time interval measurement (TI_2) one second (10^9 ns) later. If $TI_2 - TI_1 = 100$ ns, this produces the same value for f_{off} that was previously shown in Equation 42.3:

$$f_{\text{off}} = \frac{100}{10^9} = 1 \times 10^{-7} \tag{42.9}$$

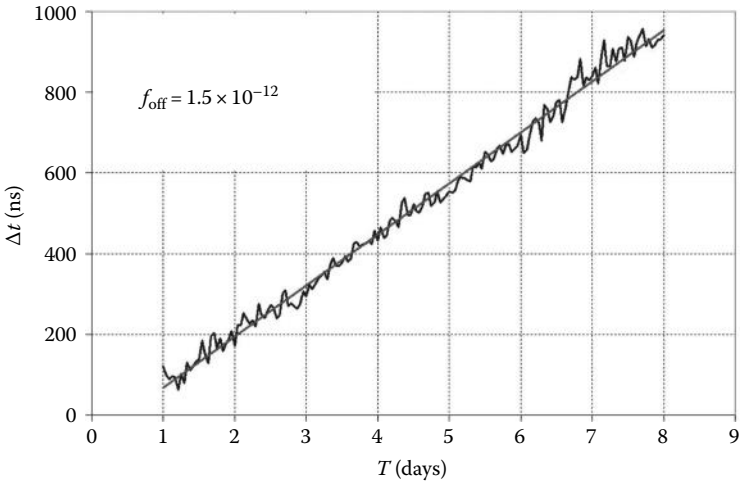


FIGURE 42.1 A sample phase graph used to estimate frequency accuracy.

In practice, more than two interval measurements are recorded, and it is common to graph the results of continuous time interval measurements recorded over multiple hours or days. These graphs are known as *phase* or time difference graphs. Phase graphs (Figure 42.1) use the standard Cartesian *x*/*y* format. The *x*-coordinate indicates elapsed time. This is the quantity T shown earlier in Equation 42.8. The values plotted as the *y*-coordinate represent the change in phase between the two electrical signals that are being compared to each other. However, since the phase changes are usually measured with instruments that express their results in units of time and not in radians or degrees, the *y*-coordinate is labeled to show the change in time or the Δt quantity in Equation 42.8.

Average frequency accuracy over a given interval can be estimated from the slope of a phase graph. In practice, a linear least squares line is often fitted to the phase data, and the slope of the least squares line is used to estimate Δt. In many cases, the slope of the least squares line is nearly identical to the slope of the actual data, because of the nearly constant frequency offset between the two oscillators being compared. In fact, to get a good estimate of frequency accuracy, the measurement period must be long enough to show this linear slope and detect a trend. A "clean" phase lot ensures that Δt is really a measure of the test oscillator's performance and indicates that neither the measurement system nor the reference oscillator has degraded the results by introducing excessive noise.

To illustrate this, Figure 42.1 shows a sample phase graph of an oscillator that was compared to a reference for a period of 7 days. During this period, the total accumulated time difference, Δt, was nearly 1000 ns, as indicated by both the actual data and the least squares line that was fitted to the data. From the slope of the least squares line, we can estimate that $f_{off} = 1.5 \times 10^{-12}$. The actual data are nosier than the least squares line, because some noise is contributed by the measurement system and the oscillators involved in the comparison. Even so, a strong linear trend is easily detected, and we can be comfortable that this is a good estimate of frequency accuracy.

If necessary, it is easy to convert a dimensionless frequency offset estimate to units of frequency (Hz) if the nominal frequency is known. To illustrate this, consider an oscillator with a nominal frequency of 10 MHz that is high in frequency by 1×10^{-11}. To find the frequency offset in hertz, multiply the nominal frequency by the dimensionless offset:

$$f_{nom} \times f_{off} = (1 \times 10^{7})\ (+1 \times 10^{-11}) = 1 \times 10^{-4} = +0.0001 \text{ Hz} = +0.1 \text{ mHz} \tag{42.10}$$

The actual frequency in this case is 10,000,000.0001 Hz, which is obtained by simply adding the offset frequency to the nominal frequency.

To summarize, frequency accuracy and frequency offset are equivalent terms that indicate how closely an oscillator produces its nominal frequency at a given point in time or over a given interval. Frequency accuracy can be estimated in either the frequency domain or the time domain. The accuracy of an oscillator can usually be at least temporarily improved by adjusting it to agree with a more accurate reference [1,2].

42.1.2 Frequency Stability

Frequency stability indicates how well an oscillator can produce the same frequency offset over a given time interval. Any frequency that "stays the same" is a stable frequency, regardless of whether the frequency is "right" or "wrong" with respect to its nominal value. To understand the difference between stability and accuracy, consider that an oscillator in need of adjustment might produce a stable frequency with a large offset. Or, an unstable oscillator that was recently adjusted might temporarily produce an accurate frequency near its nominal value. The stability of an oscillator cannot be changed by adjustment, so unlike accuracy, it tells us something about the inherent quality of an oscillator. In fact, the accuracy of an oscillator over a given interval can never be better than its stability. Figure 42.2 shows the relationship between stability and accuracy.

Frequency stability is normally estimated with statistics that quantify the frequency fluctuations of an oscillator's output over a given time interval. The fluctuations are measured with respect to a mean frequency offset, and the larger the dispersion of the fluctuations, the greater the instability of the oscillator. *Short-term stability* usually refers to fluctuations over intervals less than 100 s but is commonly used to discuss an oscillator's stability at an interval of 1 s. *Long-term stability* can refer to any measurement interval greater than 100 s, but commonly is used to discuss stability over intervals of 1 day or longer.

Normally, metrologists rely on classical statistics such as *standard deviation* (or *variance*, the square of the standard deviation) to estimate dispersion. Variance is an estimate of the numerical spread of a dataset with respect to its average or mean value. However, variance works only with stationary data, where the results must be time independent. This assumes the noise is *white*, meaning that its power is evenly distributed across the frequency band of the measurement. Oscillator data are usually nonstationary. For stationary data, the mean and variance will converge to particular values as the number of measurements increases. With nonstationary data, the mean and variance never converge to any particular values. Instead, there is a moving mean that might change each time a new measurement is added [3].

For these reasons, frequency metrologists generally rely on nonclassical statistics to estimate and specify the frequency stability of oscillators [4]. The most common statistic employed for stability estimates is often called the *Allan variance*, but because it is actually the square root of the variance, its proper name is the *Allan deviation* (ADEV). Similar to the standard deviation, ADEV is better suited for

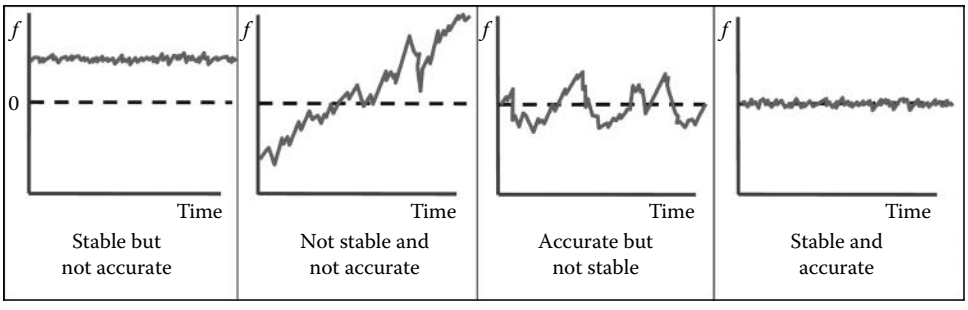


FIGURE 42.2 The relationship between frequency accuracy and stability.

frequency metrology because it has the advantage of being convergent for most types of oscillator noise. The equation for ADEV using frequency measurements and nonoverlapping samples is

$$\sigma_y(\tau) = \sqrt{\frac{1}{2(M-1)} \sum_{i=1}^{M-1} \left(\bar{y}_{i+1} - \bar{y}_i\right)^2} \tag{42.11}$$

where $\bar{y}_i$ is the ith in a series of M dimensionless frequency measurements averaged over a measurement or sampling interval designated as τ. Note that while classical deviation subtracts the mean from each measurement before squaring their summation, ADEV subtracts the previous data point. Since stability is a measure of frequency fluctuations and not of frequency offset, the differencing of successive data points is done to remove the time-dependent noise contributed by the frequency offset. Also, note that the $\bar{y}$ values in the equation do not refer to the average or mean of the entire dataset, but instead imply that the individual measurements in the dataset can be obtained by averaging.

The equation for ADEV using phase measurements and nonoverlapping samples is

$$\sigma_y(\tau) = \sqrt{\frac{1}{2(N-2)\tau^2} \sum_{i=1}^{N-2} [x_{i+2} - 2x_{i+1} + x_i]^2}, \tag{42.12}$$

where x_i is the ith in a set of N phase measurements spaced by the measurement interval τ.

To improve the confidence of a stability estimate, ADEV is normally used with overlapping samples that allow estimating stability with all possible combinations of the dataset. The equation for ADEV using phase measurements and overlapping samples is

$$\sigma_y(\tau) = \sqrt{\frac{1}{2(N-2m)\tau^2} \sum_{i=1}^{N-2m} [x_{i+2m} - 2x_{i+m} + x_i]^2}, \tag{42.13}$$

where the averaging factor, m, has been added to Equation 42.12. To understand the averaging factor, consider that τ_0 is the basic measurement interval or the shortest interval at which data are taken. To obtain stability estimates for longer intervals, τ_0 is simply multiplied by m; thus, $\tau = m\tau_0$. Even though the overlapping samples are not statistically independent, the number of degrees of freedom still increases, thus improving the confidence in the stability estimate [5].

One important advantage of ADEV over classical statistics is its ability to estimate stability over different intervals from the same dataset. Most ADEV graphs found in the literature use the *octave* method, where each successive value of τ is twice as long as the previous value. This method saved computational time, but as computers have become faster, it has become more common to estimate ADEV for all possible values of τ. A typical ADEV graph plots log τ on the x-coordinate to indicate the averaging period and log $\sigma_y(\tau)$ on the y-coordinate to indicate dimensionless frequency stability. These graphs are often referred to colloquially as "sigma–tau" graphs. ADEV graphs generally show the stability improving as the averaging period increases, until the point where the oscillator reaches its noise floor, or *flicker floor*, when no further gains will be made by averaging additional measurements. Figure 42.3 shows a sample ADEV graph that shows stability estimates using the octave method for intervals of τ ranging from 1 s to more than 2 h. This device was stable to about 3×10^{-12} at $\tau = 1$ s and reached a noise floor near 5×10^{-13} at $\tau = 512$ s. When τ exceeded 1000 s, the oscillator had begun to change frequency. Thus, further averaging degraded, rather than improved, the results.

In addition to estimating stability, ADEV can help identify the types of oscillator noise. Five noise types are commonly discussed in the time and frequency literature: *white phase*, *flicker phase*, *white frequency*, *flicker frequency*, and *random walk frequency*. A brief description of each noise type is provided

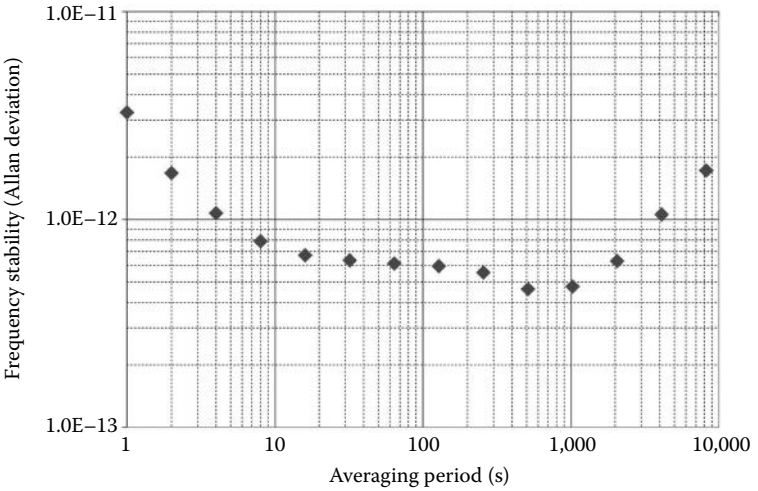


FIGURE 42.3 A sample ADEV graph used to estimate frequency stability.

in Table 42.1. The noise type can be identified from the slope of the line on an ADEV graph (Figure 42.4). Note that ADEV cannot distinguish between white phase and flicker phase noise. However, the modified ADEV, Mod $\sigma_y(\tau)$, can make this distinction, and numerous other variations of ADEV exist for specific applications, such as the improved identification of oscillator noise or improved estimates of long-term stability [4,5].

ADEV and similar statistics have proven to be very useful. However, they appear so often in the literature that confusing stability with accuracy has become a common mistake. It is important to know

TABLE 42.1 Oscillator Noise Types

Noise Type	Description	ADEV Slope
White phase	Fluctuations in the phase of a signal that have the same power at all frequencies across a given bandwidth. The stability is improving at a rate proportional to the averaging period. Mod ADEV can distinguish between white and flicker phase noise (ADEV cannot) and identifies white phase noise as having a slope of $\tau^{-3/2}$.	T^{-1}
Flicker phase	Also known as $1/f$ phase noise. As the frequency goes up, the intensity of the noise goes down. For example, if the frequency doubles, the power of the noise is cut in half. Unlike white phase noise, flicker phase noise is not evenly distributed across the frequency band.	T^{-1}
White frequency	Fluctuations in the frequency of a signal that have the same power at all frequencies across a given bandwidth. The stability is still improving, but at a rate proportional to the square root of the averaging period.	$T^{-1/2}$
Flicker frequency	Also known as $1/f$ frequency noise. The oscillator has reached a noise floor that shows its best possible stability (often called the "flicker" floor). There is nothing to be gained by more averaging.	τ^0
Random walk	Successive random steps in frequency. The difference between two steps is nearly constant, but the direction of the steps is random. Even so, it is clear that the oscillator frequency is now changing. A slope of τ^1 is sometimes used to distinguish frequency drift from random walk.	$\tau^{1/2}$

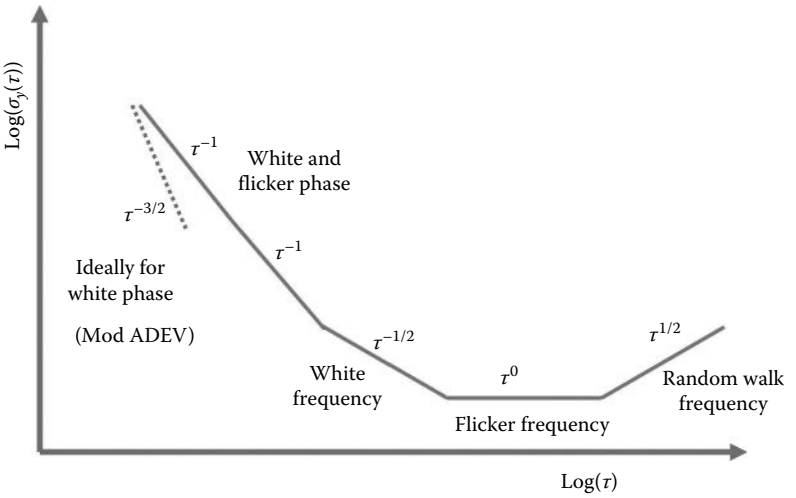


FIGURE 42.4 Identification of oscillator noise types.

that the two specifications mean different things and can have very different values for the same oscillator. For example, an oscillator accurate to only 1×10^{-8} might still be stable to 1×10^{-13} at $\tau = 1$ s. This means that even though the frequency of the oscillator is changing by only a small amount during short intervals, it is not particularly close to its nominal value.

42.2 Frequency Standards

As noted previously, stability measurements can tell us something about the inherent quality of an oscillator, and the stability of an oscillator is closely related to its *quality factor*, or *Q*. The *Q* of an oscillator is its resonance frequency divided by its resonance width. The resonance frequency is the natural frequency of the oscillator. The resonance width is the range of possible values where the oscillator will run. Obviously, a high resonance frequency and a narrow resonance width are both advantages when seeking a high *Q*. Stability and *Q* are generally correlated, because a high *Q* means that an oscillator has to stay close to its natural resonance frequency.

This section discusses the various types of oscillators used as frequency standards. It begins by discussing quartz oscillators, which achieve the highest *Q* of any mechanical-type device. It then discusses oscillators with higher *Q* factors, based on the atomic resonance of rubidium, hydrogen, and cesium. This is followed by a discussion of disciplined oscillators. These devices can be either quartz oscillators or atomic oscillators, but their frequency is automatically adjusted to agree with an external reference. The section concludes with a discussion of which type of frequency standard is best suited for use in a metrology laboratory. Table 42.2 provides a summary [2,6–8].

42.2.1 Quartz Oscillators

Billions (10^9) of quartz crystal oscillators are manufactured annually. Most are miniature, inexpensive devices that are embedded inside wristwatches, clocks, computers, cellular phones, and nearly every type of electronic circuit. However, only the larger, more expensive varieties of quartz oscillators are used as frequency standards. These devices are sometimes sold as stand-alone instruments but are more typically found inside test and measurement equipment, such as counters, signal generators, and oscilloscopes.

A quartz crystal inside the oscillator serves as resonator. The crystal strains (expands or contracts) when a voltage is applied. Reversing the polarity of the applied voltage will reverse the strain and force

TABLE 42.2 Summary of Oscillator Types

Oscillator Type	Quartz (OCXO)	Rubidium	Cesium	Active Hydrogen Maser	GPSDO
Primary standard	No	No	Yes	No	No
Resonance frequency	Mechanical (varies)	6.834682610904 GHz	9.19263177 GHz	1.420405751768 GHz	NA
Quality factor, Q	$\sim 10^6$	$\sim 10^7$	$\sim 10^8$	$\sim 10^9$	NA
Frequency accuracy (1 day average)	1×10^{-6} to 1×10^{-10}	5×10^{-9} to 5×10^{-12}	1×10^{-12} to 1×10^{-14}	$\sim 1 \times 10^{-13}$	1×10^{-12} to 5×10^{-14}
Stability, $\sigma_y(\tau)$, $\tau = 1$ s	1×10^{-11} to 1×10^{-13}	5×10^{-11} to 5×10^{-12}	1×10^{-11} to 5×10^{-12}	$\sim 2 \times 10^{-13}$	1×10^{-10} to 1×10^{-12}
Stability, $\sigma_y(\tau)$, $\tau =$ 1 day	1×10^{-10}	5×10^{-12}	8×10^{-14} to 2×10^{-14}	$\sim 2 \times 10^{-16}$	1×10^{-12} to 5×10^{-14}
Aging/year	5×10^{-9}	2×10^{-10}	None	$\sim 1 \times 10^{-13}$	None
Phase noise (dbc/Hz, 10 Hz from carrier)	−125 to −140	−90 to −130	−130 to −136	−130 to −142	−90 to −140
Life expectancy	Indefinite	>15 years	5–20 years	>15 years	>15 years
Cost (USD)	$500–$5,000	$2,000–$10,000	$30,000–$80,000	~$200,000	$1,000 to $20,000

the crystal to mechanically oscillate. This is known as the *piezoelectric effect*. The energy needed to sustain oscillation is obtained by taking a voltage signal from the resonator, amplifying it, and feeding it back to the resonator. Figure 42.5 is a simplified circuit diagram that shows the basic elements of a quartz crystal oscillator.

The rate of expansion and contraction is the resonance frequency and is determined by the cut and size of the crystal. No two crystals can be exactly alike or produce exactly the same frequency. The output frequency of a quartz oscillator is either the fundamental resonance or a multiple of the resonance, called an *overtone frequency*. Most high-stability units use either the third or fifth overtone to achieve a high Q. Overtones higher than the fifth are rarely used because they make it harder to tune the device to the desired frequency. A typical Q for a quartz oscillator ranges from 10^3 for a wristwatch-type oscillator to higher than 10^6 for the most stable devices. The maximum Q for a high-stability quartz oscillator can be roughly estimated as $Q = 16$ million/f, where f is the resonance frequency in megahertz.

Environmental changes can change the resonance frequency of a quartz crystal. Temperature changes are the largest problem, but other parameters such as humidity, pressure, and vibration can also change the frequency. There are several types of design packages that reduce these environmental problems. The most stable type of quartz oscillator is the *oven-controlled crystal oscillator* (OCXO), which encloses

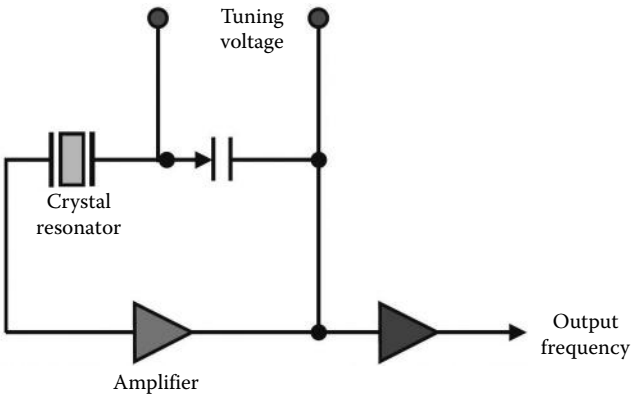


FIGURE 42.5 Block diagram of quartz oscillator.

the crystal in a temperature-controlled chamber called an oven. When an OCXO is turned on, it goes through a warm-up period while the temperatures of the crystal resonator and its oven stabilize. During this period, the performance of the oscillator continuously changes until it reaches its normal operating temperature. The temperature within the oven then remains constant, even when the outside temperature varies. An alternative solution to the temperature problem is the *temperature-compensated crystal oscillator* (TCXO). These devices include a temperature sensor that generates a correction voltage that is applied to a voltage-variable reactance, or varactor. The varactor then produces a frequency change equal and opposite to the frequency change produced by temperature. This technique does not work as well as oven control, but it generally costs less. Therefore, TCXOs are used when high stability over a wide temperature range is not required.

The best quartz oscillators have excellent short-term stability. A high-quality OCXO might be stable to 1×10^{-13} at $\tau = 1$ s. However, quartz oscillators are not stable over long intervals. Their long-term stability is limited by *aging*, which causes their frequency to change over time due to internal changes in the oscillator. Aging usually results in a nearly linear change in the resonance frequency that can be either positive or negative. A reversal in the direction of the aging occasionally occurs, and ironically, the aging rate of a quartz oscillator sometimes decreases as the device gets older. Aging has many possible causes, including a buildup of foreign material on the crystal, changes in the oscillator circuitry, or changes in the quartz material or crystal structure. A high-quality OCXO might age at a rate of less than 5×10^{-9} per year, while a TCXO might age 100 times faster.

The simple design of quartz oscillators makes them very reliable, and many devices have run continuously for decades without failing. However, their accuracy can change rapidly due to aging and environmental factors, and even a high-quality OCXO will need regular adjustments to maintain frequency accurate to within 1×10^{-9}. They are also subject to large frequency shifts when they are turned on after a power outage. For these reasons, quartz oscillators are usually a poor choice as a frequency standard, unless the measurement requirements of a laboratory are very low [8,9].

42.2.2 Atomic Oscillators

Atomic oscillators derive their resonance frequency from the quantized energy levels in atoms. The laws of quantum mechanics dictate that the energies of a bound system, such as an atom, have certain discrete values. An electromagnetic field can boost an atom from one energy level to a higher one. Or, an atom at a high energy level can drop to a lower level by emitting electromagnetic energy. The resonance frequency (f) of an atomic oscillator is the difference between the two energy levels divided by Planck's constant (h):

$$f = \frac{E_2 - E_1}{h} \tag{42.14}$$

All atomic oscillators are *intrinsic standards*, because their frequency is inherently derived from a fundamental natural phenomenon. There are currently (2011) three types of atomic oscillators sold commercially: rubidium standards, cesium standards, and hydrogen masers (discussed individually in the following sections). All three types contain an internal quartz oscillator that is locked to a resonance frequency generated by the atom of interest. This method causes the factors that degrade the long-term stability of a quartz oscillator to disappear. As a result, the long-term stability of an atomic oscillator is at least several orders of magnitude better than that of a quartz oscillator, but the short-term stability is unchanged [2,8,10].

42.2.3 Rubidium Oscillators

Rubidium oscillators are outperformed by the other types of atomic oscillators, but they have the advantage of being much smaller and less expensive. Because of their low cost and small size, rubidium

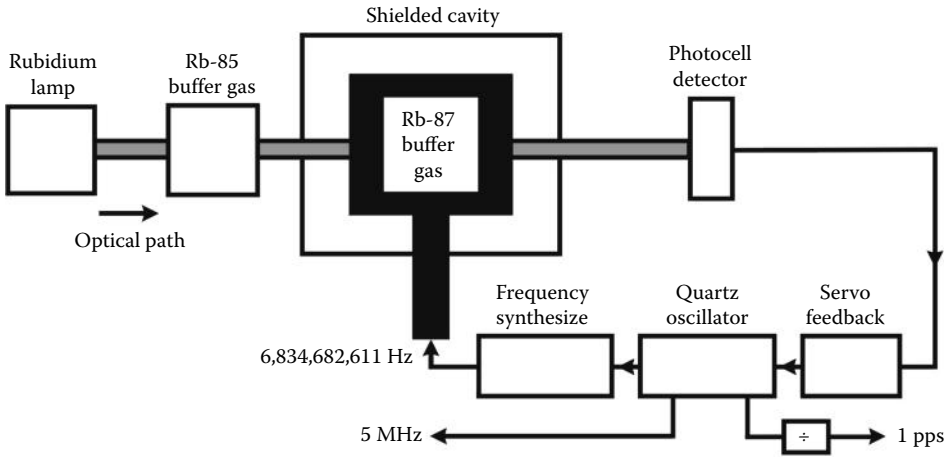


FIGURE 42.6 Block diagram of rubidium oscillator.

oscillators are often found in telecommunication networks and as time base oscillators in test and measurement equipment. They operate at 6,834,682,610.904 Hz, the resonance frequency of the rubidium atom (^{87}Rb), and use the rubidium frequency to control the frequency of a quartz oscillator. A microwave signal derived from the crystal oscillator is applied to the ^{87}Rb vapor within a cell, forcing the atoms into a particular energy state. An optical beam is then pumped into the cell and is absorbed by the atoms as it forces them into a separate energy state. A photocell detector measures how much of the beam is absorbed and tunes a quartz oscillator to a frequency that maximizes the amount of light absorption. The quartz oscillator is then locked to the resonance frequency of rubidium, and standard frequencies are derived and provided as outputs (Figure 42.6).

The Q of a rubidium oscillator is about 10^7. The shifts in the resonance frequency are mainly caused by collisions between the rubidium atoms and other gas molecules. These frequency shifts limit the long-term stability. Stability at $\tau = 1$ s is typically less than 1×10^{-11} and near 1×10^{-12} at $\tau = 1$ day. There is generally no guaranteed specification for accuracy, but after a warm-up period of a few minutes, a rubidium oscillator will typically be accurate to within parts in 10^{10} or less, and some devices might be as accurate as 5×10^{-12}. However, if an application has an accuracy requirement of parts in 10^9 or smaller, a rubidium oscillator will need to be regularly measured and adjusted, because accuracy better than about 5×10^{-9} cannot be assumed. With regular frequency adjustments, a rubidium can maintain average frequency to within a few parts in 10^{11} or 10^{12} over periods of months or years. The adjustments are made to compensate for the aging and frequency drift that changes the rubidium frequency slowly over time. Manufacturers typically specify the aging rate as less than 5×10^{-11} per month, but this is sometimes conservative, as the frequency of a well-behaved rubidium standard might change by less than 1×10^{-11} over the course of a month. Even so, the frequency change can exceed 1×10^{-10} if left unadjusted for a year, which is unacceptable for some applications [2,6–8].

42.2.4 Cesium Oscillators

Cesium oscillators are *primary frequency standards* because the SI second is defined using the resonance frequency of the cesium atom (^{133}Cs), which is 9,192,631,770 Hz. A properly working cesium oscillator should have inherent accuracy and stability and be close to its nominal frequency without adjustment.

Cesium is a complicated atom with $F = 3$ and $F = 4$ ground states (Figure 42.7). Each atomic state is characterized not only by the quantum number F but also by a second quantum number, m_F, which can have integer values between $-F$ and $+F$. There are 16 possible magnetic states of cesium, but the

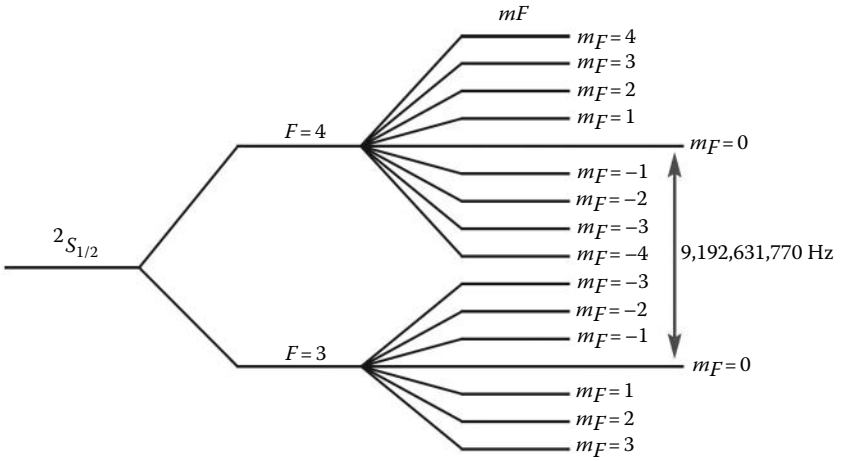


FIGURE 42.7 Cesium atomic structure.

transition between the |4,0⟩ and |3,0⟩ states is insensitive to magnetic fields. Thus, the frequency of this transition was chosen to define the SI second.

Figure 42.8 provides a simplified schematic of a cesium beam frequency standard. The design details of a cesium beam standard can vary significantly from model to model, but their basic design principles are similar. As shown on the left side of the figure, ^{133}Cs atoms are heated to a gaseous state in an oven. A beam of atoms emerges from the oven at a temperature near 100 °C and travels through a magnetic field, where the beam is split into two beams of atoms with different magnetic states. One beam is absorbed by the getter and is of no further interest. The other beam is deflected into the microwave interrogation cavity (commonly known as the Ramsey cavity).

While inside the Ramsey cavity, the cesium beam is exposed to a microwave frequency from a frequency synthesizer driven by a quartz oscillator. If this frequency is tuned to precisely match cesium resonance, some of the atoms will change their magnetic state. After leaving the Ramsey cavity, the atoms pass through a second magnetic field. These magnets direct only the atoms that changed state to the detector; the other atoms are directed to a getter and absorbed. In essence, the magnets located on

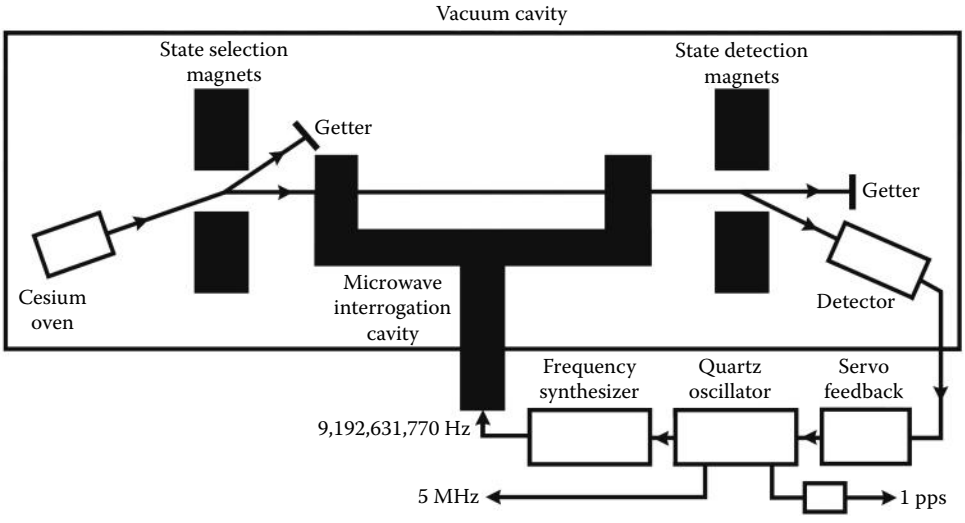


FIGURE 42.8 Block diagram of cesium oscillator.

both sides of the Ramsey cavity serve as a "gate" that allows only those atoms that undergo the desired $|4,0\rangle \leftrightarrow |3,0\rangle$ energy transition to pass through and reach the detector. The detector sends a feedback signal to a servo circuit that continually tunes the quartz oscillator so that the maximum number of atoms reaches the detector, thereby increasing the signal strength. This process is analogous to carefully tuning a radio dial until the loudest and clearest signal is heard and keeps the quartz oscillator frequency locked as tightly as possible to cesium resonance. Output frequencies, such as 1 Hz, 5 and 10 MHz, are then derived from the locked quartz oscillator.

The Q of a commercial cesium standard is about 10^8. The beam tube is typically less than 0.5 m in length, and the atoms travel at velocities of greater than 100 m/s inside the tube. This limits the observation time to a few milliseconds and the resonance width to a few hundred hertz. Stability ($\sigma_y\tau$, at $\tau = 1$ s) is typically 5×10^{-12}, normally reaching parts in 10^{14} at $\tau = 1$ day. The frequency offset is typically near 1×10^{-13} after a warm-up period of 30 min. Because the second is defined based on cesium resonance, there should be no change in frequency due to aging. However, in practice a cesium oscillator will slowly change its frequency by a very small amount, typically by parts in 10^{17} per day.

Cesium standards have a limited life expectancy and a high cost. The major component of a cesium oscillator, called the *beam tube*, typically lasts for about 5–10 years, and replacing the beam tube can cost nearly as much as replacing the entire device. When the beam tube fails, a cesium standard will no longer be locked to cesium resonance and will become a free-running quartz oscillator. For this reason, cesium frequency standards should be regularly monitored or checked to ensure that they are working properly [8,10–12].

42.2.5 Hydrogen Masers

The *hydrogen maser* is the most expensive commercially available frequency standard and is therefore found in only a small number of metrology laboratories. The word *maser* is an acronym that stands for microwave amplification by stimulated emission of radiation. The resonance frequency of the hydrogen atom is 1,420,405,751.768 Hz.

There are two types of hydrogen masers. The first type, called an *active maser*, has a microwave cavity that oscillates spontaneously, and a quartz oscillator is phase locked to this active oscillation. The second type, called a *passive maser*, frequency locks a quartz oscillator to the atomic reference in much the same fashion as a rubidium or cesium oscillator. Because active masers derive their output frequency more directly from the atomic resonance, they are more stable than passive masers in both the short and long term. Both types of maser are more stable in the short term than cesium oscillators and are well suited for applications where optimal frequency stability is required. Over long intervals, however, hydrogen masers are less accurate than cesium oscillators. This is due to several factors: their accuracy depends upon a more complex set of conditions, the resonance frequency of their microwave cavity can change over time, and also because the definition of the SI second is based on cesium resonance [2,8,10].

An active hydrogen maser works by sending hydrogen gas through a magnetic gate that only allows atoms in certain energy states to pass through. The atoms that make it through the gate enter a storage bulb surrounded by a tuned, resonant cavity. Once inside the bulb, some atoms drop to a lower energy level, releasing photons of microwave frequency. These photons stimulate other atoms to drop their energy level, and they in turn release additional photons. In this manner, a self-sustaining microwave field builds up in the bulb. The tuned cavity around the bulb helps to redirect photons back into the system to keep the oscillation going. The result is a microwave signal that is locked to the resonance frequency of the hydrogen atom and that is continually emitted as long as new atoms are fed into the system. This signal keeps a quartz crystal oscillator in step with the resonance frequency of hydrogen (Figure 42.9).

The resonance frequency of hydrogen is much lower than that of cesium, but the resonance width of a hydrogen maser is usually just a few hertz. Therefore, the Q is about 10^9 or about one order of magnitude better than that of a commercial cesium standard. As noted, a hydrogen maser is more stable than

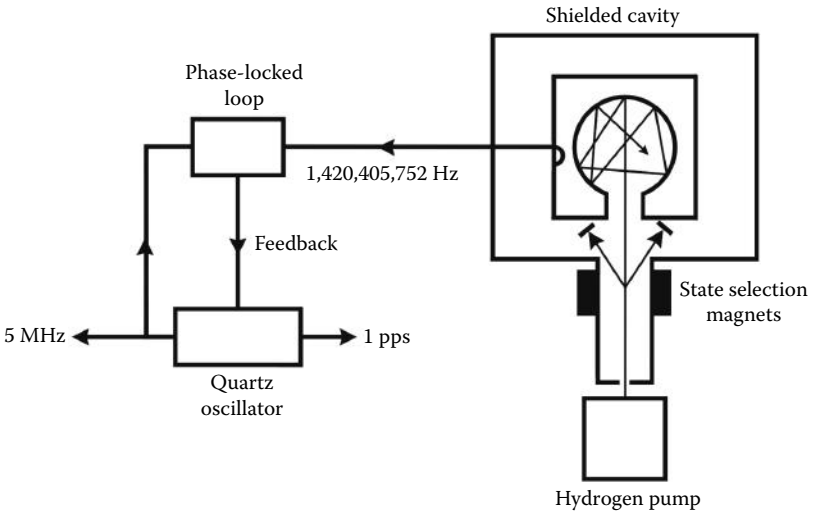


FIGURE 42.9 Block diagram of an active hydrogen maser.

a cesium oscillator for periods ranging from 1 s to perhaps weeks or months. The stability of an active maser will typically reach a few parts in 10^{13} at $\tau = 1$ s and a few parts in 10^{16} at $\tau = 1$ day.

42.2.6 Disciplined Oscillators

Oscillators whose frequency is controlled by an external reference signal are known as *disciplined oscillators*. Unlike free-running oscillators that need to be periodically adjusted to stay within specification, disciplined oscillators are locked to a reference signal and never require manual adjustment. The best disciplined oscillators can generate local signals with nearly the same accuracy and stability as the remote reference. Various types of radio signals have been used to discipline oscillators, but the vast majority of disciplined oscillators in use today (2011) employ signals from the GPS satellites as their reference source. For this reason, this section will focus entirely on GPS-disciplined oscillators (GPSDOs).

Unlike the other types of frequency standards described earlier, a GPSDO requires a small antenna to be mounted on a rooftop location with a clear view of the sky. A GPSDO will normally begin surveying its antenna position as soon as it is turned on. The survey is a one-time process that typically lasts for several hours. When the antenna survey is complete, the GPSDO is ready to use as a frequency standard and will typically produce sine wave signals of 5 and/or 10 MHz.

The basic function of a GPSDO is to receive signals from the GPS satellites and to use the information contained in these signals to control the frequency of a local quartz or rubidium oscillator. GPS signals are kept in agreement with the Coordinated Universal Time scale maintained by the United States Naval Observatory (UTC[USNO]). Nearly all GPSDOs use the coarse acquisition (C/A) code on the L1 carrier frequency (1575.42 MHz) as their incoming reference signal. The satellite signals can be trusted as a reference for two reasons: (1) they originate from atomic oscillators and (2) *they must be accurate and stable* to within parts in 10^{14} over a 12 h averaging period in order for GPS to meet its specifications as a positioning and navigation system. The best GPSDOs transfer as much of the inherent accuracy and stability of the satellite signals as possible to the signals generated by the local quartz or rubidium oscillator.

Many of the methods used to discipline oscillators are proprietary, and GPSDO manufacturers seldom disclose exactly how their products work. However, there are a few basic concepts that apply to most designs. Generally, the local oscillator is controlled with one or more servo loops, with each loop having a fixed or variable time constant. For example, one type of servo loop is a *phase-locked loop*, or PLL. In a GPSDO, the reference input signal to the PLL comes from a GPS receiver that tracks

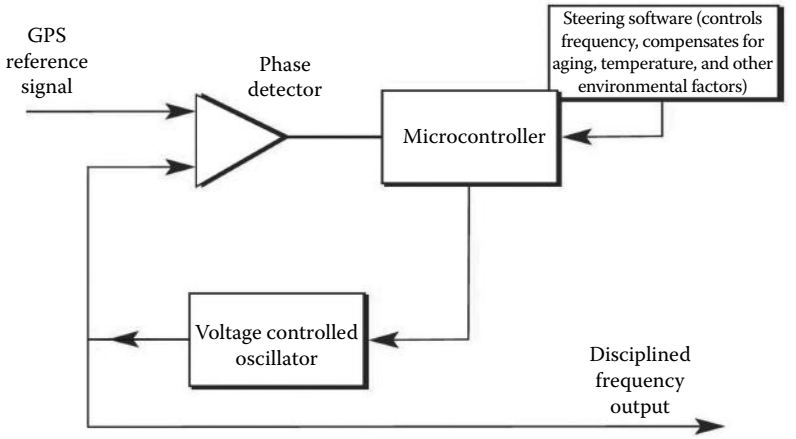


FIGURE 42.10 Block diagram of a GPSDO that steers its local oscillator.

multiple satellites and outputs a 1 pulse per second (pps) signal. A phase detector measures the difference between the 1 pps signal from the GPS receiver and a signal from a voltage-controlled oscillator (VCO). The VCO typically has a nominal frequency of 10 MHz, so its signal is divided to a lower frequency (often all the way down to 1 pps) prior to this phase comparison. A microcontroller reads the output of the phase detector and monitors the phase difference. When the phase difference changes, the software changes the control voltage sent to the VCO, so that the phase difference is held within a given range (Figure 42.10). The GPSDO is locked when the phase of the VCO has a constant offset relative to the phase of the GPS signals. Ideally, the servo loop must be loose enough to ignore the short-term fluctuations of the GPS signals, reducing the amount of phase noise and allowing the VCO to provide reasonably good short-term stability. However, the loop must be tight enough to track GPS closely and to allow the GPS signals to control the VCO frequency in the longer term. The microcontroller software often compensates not only for the phase and frequency changes of the local oscillator but also for the effects of aging, temperature, and other environmental parameters.

Another type of GPSDO design does not correct the frequency of the local oscillator. Instead, the output of a free-running local oscillator is sent to a frequency synthesizer. Steering corrections are then applied to the output of the synthesizer (Figure 42.11). Modern direct digital synthesizers (DDSs) have

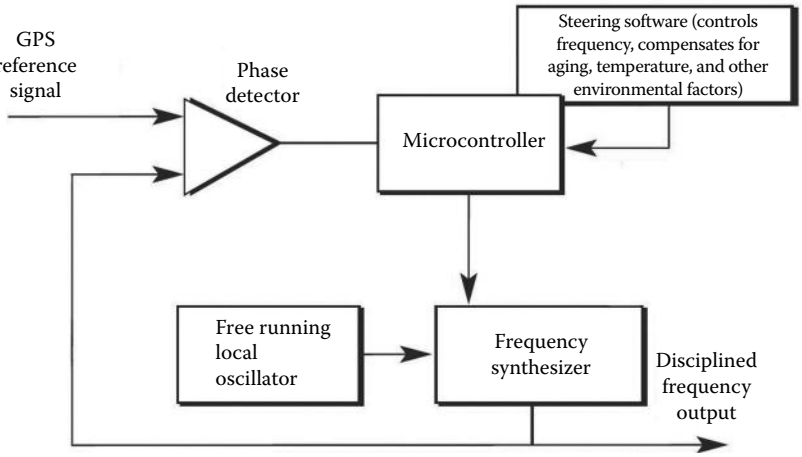


FIGURE 42.11 Block diagram of a GPSDO that adjusts a frequency synthesizer.

excellent resolution and allow very small frequency corrections to be made. For example, 1 μHz resolution at 10 MHz allows instantaneous frequency corrections of 1×10^{-13}. In addition, allowing the local oscillator to free run often results in better performance than the VCO method, where unexpected shifts in the control voltage can produce unwanted adjustments in the output frequency.

A reasonably good metric to use when evaluating GPSDO performance is its frequency stability at $\tau = 1$ day, as estimated with ADEV. Stability of 1×10^{-13} or less at $\tau = 1$ day normally indicates a device of high quality, and many (but certainly not all) GPSDOs can reach this specification. To demonstrate how the performance of GPSDOs can vary significantly, a test was conducted between two rubidium-based GPSDOs at National Institute of Standards and Technology (NIST). Both devices had the same type of rubidium local oscillator and cost approximately the same amount. During the test, both GPSDO devices were connected to the same GPS antenna with an antenna splitter. The antenna's position had previously been surveyed with an uncertainty of less than 1 m, and these precise coordinates were keyed into both units. The 10 MHz outputs of both devices were then simultaneously compared to the US national frequency standard at NIST, UTC(NIST), for a period of 80 days. The results are shown in the phase graph in Figure 42.12.

The results show that the frequency of Device A was tightly controlled. The peak-to-peak phase variation over the entire 80 day period was just 38 ns, with most of this variation due to the difference between UTC(USNO), the reference for GPS, and UTC(NIST) during the measurement interval. The frequency accuracy, as estimated from the slope of the phase, was about 1×10^{-15}. In sharp contrast, the frequency of Device B was very loosely controlled, and the phase plot shows a very large peak-to-peak phase variation of 588 ns, much larger than the dispersion of the GPS timing signals. Figure 42.13 shows the long-term frequency stability of both devices as estimated with ADEV, for values of τ ranging from 1 h to about 3 weeks. Device A is more stable than Device B at all averaging periods by roughly a factor of 10. Stability at $\tau = 1$ day, the metric discussed earlier, is about 6×10^{-14} for Device A, comparable to the performance of a cesium standard. However, the stability of Device B is worse by more than a factor of 10, about 70×10^{-14}.

In spite of the wide performance disparity between Device A and Device B, these examples still illustrate that a GPSDO that is even loosely locked to the satellite signals should be inherently accurate (parts in 10^{13} or better) and inherently stable in the long term. This is because the signals broadcast by the GPS satellites are continuously steered to agree with UTC and GPSDOs that simply "follow" the satellites will produce frequency that closely agrees with UTC over long intervals.

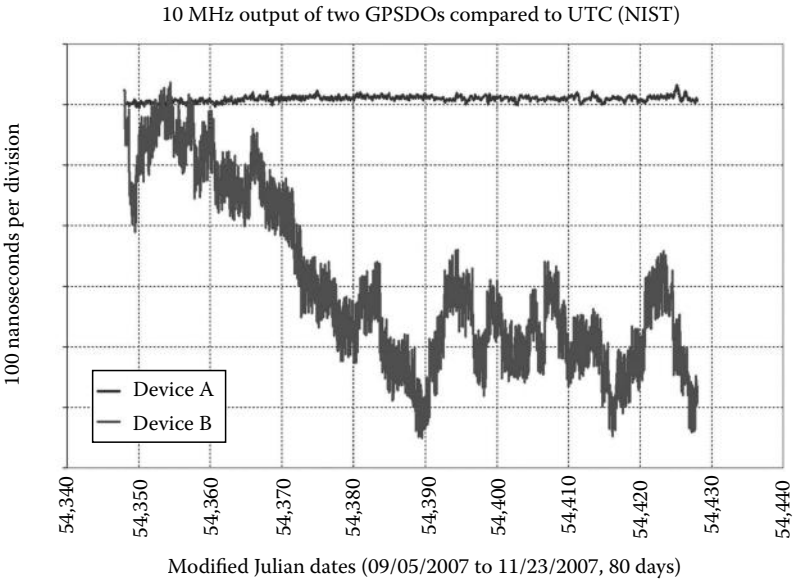


FIGURE 42.12 Phase comparison of two GPSDOs.

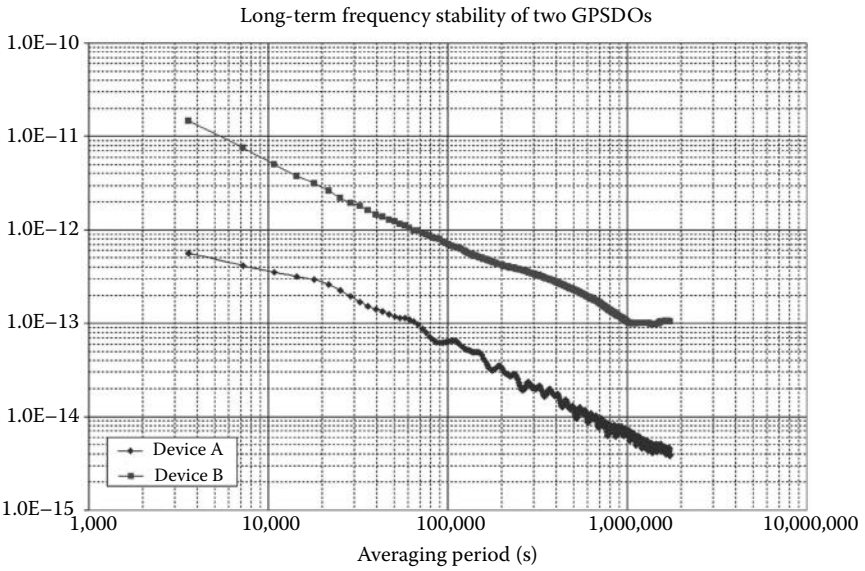


FIGURE 42.13 Stability comparison of two GPSDOs.

GPSDOs are generally reliable, but users should be aware that they will eventually fail if the GPS signal is unavailable. The most likely cause of failure is probably RF interference, because GPS signals are highly susceptible to intentional or unintentional interference due to their low power levels. A GPSDO can stop tracking satellites if there are interfering signal only a few orders of magnitude more powerful than the minimum received GPS signal strength, which is –160 dBW on Earth for the L1 carrier, equivalent to 10^{-16} W. When the GPS signal is unavailable, a GPSDO continues to produce frequency but begins relying on its *holdover* capability. In many cases, a GPSDO will simply become a free-running local oscillator while in holdover mode, in which case, its frequency accuracy will probably be several orders of magnitude worse than normal, perhaps parts in 10^9 or 10^{10} for a device with a rubidium local oscillator after a few hours without GPS and parts in 10^7 or 10^8 for a quartz based device under the same conditions. Some devices implement holdover algorithms that continue to steer the local oscillator without GPS, but the performance of these algorithms will degrade over time. As is the case with a cesium standard, it is important to verify that a GPSDO is working properly [13].

42.2.7 Choosing a Frequency Standard for a Metrology Laboratory

As noted earlier, a quartz oscillator is generally a poor choice as a frequency standard, and only those laboratories with the most demanding frequency stability requirements will be able to justify the expense of a hydrogen maser. Thus, when a metrology laboratory decides which frequency standard to buy, it will likely be choosing between a rubidium oscillator, a cesium oscillator, and a GPSDO. The specifications of the various oscillator types are summarized in Table 42.2. The specifications listed in the table were obtained from manufacturer's specification sheets and from the results of measurements performed by NIST.

As Table 42.2 indicates, a GPSDO will have better long-term frequency accuracy and stability than a stand-alone rubidium oscillator, and the GPSDO will never require adjustment. A GPSDO will normally cost more than a stand-alone rubidium standard, but in most cases, the performance and convenience of the GPSDO will easily justify the higher cost. Therefore, for most metrology laboratories, a GPSDO is probably a better choice.

Choosing between a cesium standard and a GPSDO is more difficult. As noted earlier, the SI second is defined based on energy transitions of the cesium atom; thus, cesium oscillators are often the preferred choice of frequency standard for laboratories with the best measurement capabilities and the most

demanding performance requirements. However, as also noted earlier, cesium oscillators are expensive and have a limited life expectancy, and not all laboratories can afford them. Assuming that a calibration laboratory can afford a cesium standard, should they still economize by choosing a GPSDO as their frequency standard? There are several pros and cons related to GPSDOs that should be considered before answering this question. First the pros:

- A GPSDO costs much less than a cesium standard to initially purchase, sometimes as much as 90% less. It also costs less to own, because there is no cesium beam tube to replace. This means that a calibration laboratory could buy multiple GPSDOs for less than the cost of a cesium standard and use the additional standards for crosschecks and redundancy.
- Unlike a cesium standard, a GPSDO can recover time by itself (time-of-day and an on-time pulse synchronized to UTC). This is important if a laboratory needs time synchronization capability.
- Cesium standards seldom require adjustment, but a GPSDO will never require adjustment, because its frequency is controlled by the signals from the GPS satellites.

Now, the cons:

- GPSDOs generally have poorer short-term stability and higher phase noise than cesium standards.
- GPSDOs require an outdoor antenna that must be located in an area with access to the roof. A cesium standard can be operated anywhere where electric power is available.
- Cesium standards are autonomous and independent sources of frequency, which means they can operate without input from another source. GPSDO can operate properly only where signals from the GPS satellites are available and are not suitable for applications that need an autonomous frequency source.

Based on these criteria, it seems likely that many laboratories that can afford a cesium standard will undoubtedly choose a GPSDO as a lower-cost alternative that meets all of their requirements. However, a certain percentage of laboratories do require a cesium standard, and some laboratories will operate both types of standards. Even if a laboratory already owns a cesium standard, it might be wise to acquire a GPSDO as a secondary standard so that the two devices can be compared to each other to ensure that both are working properly. Table 42.3 lists suppliers of the various types of frequency standards.

TABLE 42.3 Suppliers of Frequency Standards

Company	Website	Rubidium Oscillators	Cesium Oscillators	Hydrogen Masers	GPS Disciplined Oscillators
Accubeat	www.accubeat.com	X			X
Arbiter	www.arbiter.com				X
Brandywine	www.brandywinecomm.com	X			X
EndRun	www.endruntechnologies.com				X
FEI-Zyfer	www.fei-zyfer.com				X
Fluke	www.fluke.com	X			X
Frequency electronics	www.freqelec.com	X	X		
Meinberg	www.meinberg.de				X
Oscilloquartz	www.oscilloquartz.com	X	X		X
Precision test systems	www.ptsyst.com	X			X
Precise time and frequency	www.ptfinc.com	X		X	X
Spectracom	www.spectracomcorp.com	X			X
Stanford research	www.thinksrs.com	X			
Symmetricom	www.Symmetricom.com	X	X	X	X
Trak	www.trak.com				X
Trimble	www.trimble.com				X

42.3 Calibration and Measurement Methods

The objective of a frequency calibration is to measure the accuracy and/or stability of the *device under test* (DUT), which will be one of the oscillator types described in Section 42.2. During the calibration, the DUT is compared to a *standard* or *reference*. In most cases, both the DUT and the reference produce oscillating sine wave signals, as illustrated in Figure 42.14. The sine wave signals produce one cycle (2π radians of phase) in one period. The period is measured in units of time, and the amplitude is measured in units of voltage. The nominal frequency of the DUT is normally 1 MHz or higher, with 5 or 10 MHz being common.

In order for the calibration to be valid, the reference must outperform the DUT. The ratio by which the reference outperforms the DUT is called the *test uncertainty ratio* (TUR). A TUR of 10:1 is preferred, but not always possible. If a smaller TUR is used (4:1, for example), then the calibration will take longer to perform, because more measurements and more averaging will be required.

To further validate a measurement result, it is necessary to establish the *traceability* of a frequency measurement to the International System (SI) of units. Because frequency is the reciprocal of time interval, establishing traceability in frequency metrology means establishing traceability to the SI second through an unbroken and documented chain of calibrations. The SI second is a virtual and not a physical standard, so the chain of calibrations typically extends back to a national standard maintained by a laboratory that contributes to the UTC, such as NIST in the United States. The traceability chain will have only one link if a DUT is calibrated by NIST or an equivalent laboratory. However, there are often several links (calibrations) involved when establishing a traceability chain back to the SI.

Once a suitable reference and a plan for traceability have been chosen, the next step is to select a calibration or measurement method. The remainder of this section discusses several established techniques that are commonly used to measure frequency. With the appropriate hardware and software, measurement systems designed around these techniques can produce data that can be used to estimate either frequency accuracy or stability.

The comparison of the sine wave signal produced by the DUT to the sine wave signal produced by the reference can be made in either the time domain or in the frequency domain, but time domain

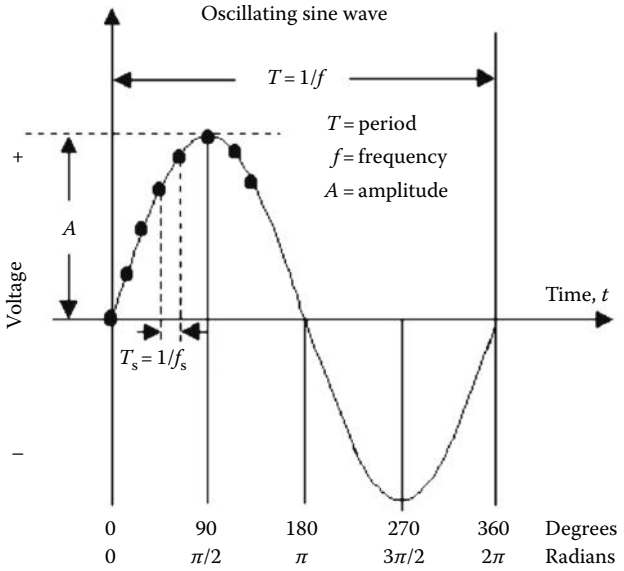


FIGURE 42.14 An oscillating sine wave.

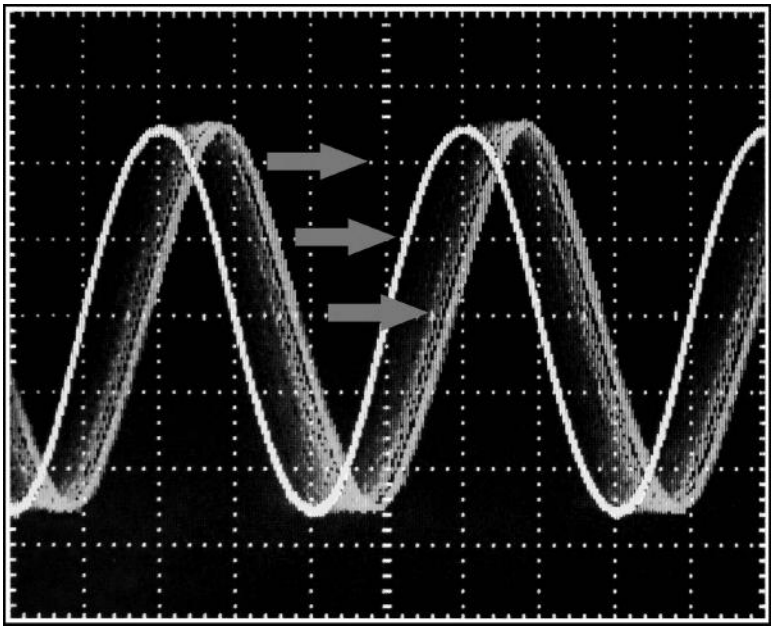


FIGURE 42.15 Phase comparison with an oscilloscope.

comparisons are more common. A simple time domain comparison that nicely illustrates the concept is the *oscilloscope pattern drift method*. This method requires a two-channel oscilloscope. In this example, both the DUT and reference signals are 10 MHz sine waves. The reference signal is produced by a cesium oscillator and the DUT signal is produced by a quartz oscillator. The scope is triggered with the reference signal on channel 2, and the DUT signal is connected to channel 1. The amplitude and position of both waveforms are adjusted on the oscilloscope display so that they overlap (Figure 42.15).

If the two frequencies were exactly the same, both sine waves would appear to be stationary on the oscilloscope display. However, because the two frequencies are not the same, their phase relationship will be continuously changing. The reference sine wave will appear to be stationary and the DUT sine wave will move. The direction of the sine wave motion will determine whether the DUT frequency is low or high with respect to the reference. A stopwatch can be started when the two signals are exactly in phase and stopped after one complete *cycle slip* has occurred and they are exactly in phase again. Frequency accuracy can be estimated as $\Delta t/T$ (see Equation 42.8), where Δt = 100 ns (the period of 10 MHz) and T is the elapsed time indicated on the stopwatch. If one cycle slip occurs per second, the frequency accuracy of the DUT is 1×10^{-7} [1].

As noted, this method is very useful for demonstrating the concept of a phase comparison, and oscilloscopes are indispensible for simple measurements and for viewing waveforms. However, the shortcomings of the pattern drift method quickly become obvious: If DUT were accurate to 1×10^{-12}, a single cycle slip would take more than a day to occur. For this reason and others, you will probably seldom use an oscilloscope to measure frequency, especially if you have a universal counter available.

Universal counters can be configured as either a *frequency counter* (FC) or a *time interval counter* (TIC). FCs are the most common instrument used for frequency measurements, and they are especially handy, because they can quickly measure a DUT's frequency and instantly display the results. The reference for this type of measurement is the FC's time base oscillator, which is usually a quartz oscillator of unknown accuracy (typically no better than 1×10^{-8}). For this reason, the laboratory's best oscillator

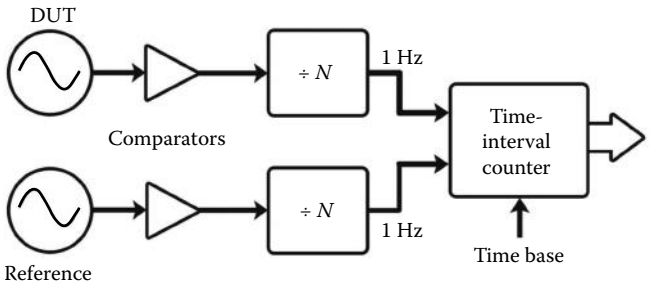


FIGURE 42.16 TIC measurement system.

should be connected to the FC's external time base input. Frequency accuracy can be estimated by using Equation 42.1, where f_{meas} is the reading taken from the FC display. The smallest frequency offset that an FC can detect with a single reading will be determined by its *resolution*, which is limited by the number of digits on the counter's display. For example, a 10-digit FC will be unable to detect a frequency change smaller than 1×10^{-9} without averaging when measuring a 10 MHz signal, but a 12-digit counter can reduce this value by two orders of magnitude to 1×10^{-11}.

TICs are commonly used to measure frequency in the time domain (Figure 42.16). When the *time interval method* is used to measure frequency, it is no longer practical to work directly with 10 MHz signals, because low-frequency input signals must be used to start and stop the counter. The solution is to use *a frequency divider* to convert standard frequency signals to a much lower frequency, typically to 1 Hz. Frequency dividers can be stand-alone instruments, integrated into the oscillator design or, in some cases, integrated into the TIC. Dividing to 1 Hz can be accomplished by blocking $f_{nom} - 1$ cycles of the frequency from passing through to the counter. The use of low-frequency signals reduces the problem of counter overflows and underflows (cycle ambiguity) and helps prevent errors that can occur when the start and stop signals are too close together. It is also important to make sure that the TIC is triggering at the correct voltage level on the input signal. This involves either carefully adjusting the trigger levels or using a TIC with fixed trigger levels and converting both signals to an identical shape and amplitude prior to the comparison.

A TIC has inputs for two signals. Typically, a signal from the DUT starts the counter and a signal from the reference stops the counter. The time interval reading will change, typically very slowly, to indicate the difference in frequency between the two signals. A single reading from an FC can immediately produce useful information, because it can be differenced from the nominal frequency. However, in the case of a TIC, at least two readings are required to produce useful frequency information. As indicated in Equation 42.8, a quick estimation of the DUT's frequency offset can be made by recording a single TIC reading, waiting for a specified period, and then recording a second TIC reading. The difference between the two readings divided by the measurement period ($\Delta t/T$) provides an estimate of the frequency offset. In practice, data are usually collected continuously, typically every second, and then averaged over longer intervals of time to get a better estimate.

Modern universal counters typically have a single-shot time interval resolution of less than 1 ns. The best devices have a resolution near 10 ps, which enables them to detect frequency changes of 1×10^{-11} in 1 s, which is equivalent to the capability of a 12-digit FC. Dedicated TICs can have a resolution near 1 ps, making it possible in theory to detect a frequency change of 1×10^{-12} in 1 s. However, the ability of a time interval system to detect small frequency changes is limited by factors other than TIC resolution, including trigger errors, time base errors, and noise from frequency dividers, and 1×10^{-11} at 1 s usually represents best case performance [1,14,15].

The time interval method is widely implemented and is an excellent way to measure long-term accuracy and stability. However, systems with higher resolution are needed to measure short-term stability.

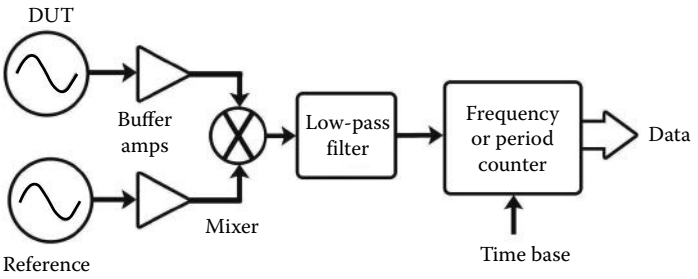


FIGURE 42.17 Simple heterodyne measurement system.

These systems continue to use counters but employ heterodyne techniques, rather than frequency dividers, to obtain the necessary low-frequency signals. The reason for this is straightforward: When dividers are used, both the frequency of the DUT and its phase fluctuations (noise) are divided by the same amount, and thus, no resolution is gained. The resolution of the system remains the same as the resolution of the counter. In contrast, heterodyne techniques convert the DUT frequency to a lower frequency, known as the intermediate frequency (IF) or *beat frequency*, without dividing the device's phase fluctuations (noise). Thus, phase information is preserved in the heterodyne process, and when the DUT changes frequency relative to the reference, the beat frequency changes by the same amount. The resolution of the counter is potentially improved by the downconversion or *heterodyne factor*, which equals f_{nom}/f_{beat}. Unlike divider systems, which typically produce 1 Hz signals, the beat frequency is usually not lower than 10 Hz, because it must be large enough to account for a wide range of possible frequency differences between the DUT and reference, including situations where a quartz oscillator is involved in the comparison.

Figure 42.17 shows a simple heterodyne system where a low-pass filter (LPF) and double-balanced mixer are used in conjunction with a universal counter in frequency or period mode. Although not shown in the figure, in most designs, the reference oscillator is intentionally offset in frequency, normally by 100 Hz or 1 kHz, by use of a frequency synthesizer. The double-balanced mixer mixes the reference frequency with the DUT signal, producing a beat frequency that is nominally the same as the intentional frequency offset of the reference. If a 10 MHz frequency is downconverted to 1 kHz, the resolution of the counter can be potentially enhanced by a heterodyne factor of 10^4 ($10^7/10^3$). The beat frequency is sent through a LPF to remove carrier frequency harmonics and then measured with either an FC or a period counter. If an FC is used, the frequency offset $f_{off} = (f_{fc} - f_{beat})/f_{nom}$, where f_{fc} is the reading on the FC, f_{beat} is the nominal beat frequency, and f_{nom} is the nominal frequency of the DUT.

The *dual-mixer time difference* (DMTD) method (Figure 42.18) combines the best features of the time interval and heterodyne methods. The DMTD method uses a transfer oscillator and two double-balanced mixers in parallel. The transfer oscillator is offset from the nominal frequencies by a small amount, typically by 10 Hz. The transfer oscillator does not have to be particularly stable or accurate because the noise it produces is common to both measurement channels and cancels when the phase difference is computed. Even so, the transfer oscillator signal is usually obtained by locking a frequency synthesizer to the reference oscillator. Typically, both the DUT and the reference must have the same nominal frequency, usually 5 or 10 MHz, and DMTD systems are often designed to measure only one nominal frequency. However, the use of frequency synthesizers allows some DMTD systems to measure a number of DUT frequencies within a specified range (e.g., 1–30 MHz) and can even allow the DUT and the reference to have different nominal frequencies.

The transfer oscillator signal is heterodyned with both the reference oscillator and the DUT to produce two beat frequencies. The two beat frequencies are out of phase by an amount proportional to the time difference between the DUT and reference, and this phase difference is measured with a TIC. Before being sent to the TIC, the beat frequencies pass through a LPF that removes harmonics

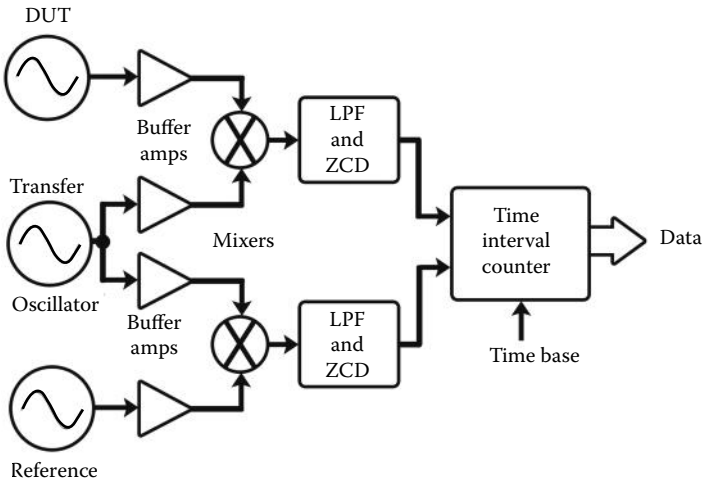


FIGURE 42.18 DMTD measurement system.

and a *zero-crossing detector* (ZCD) that determines the zero crossing of each beat frequency cycle. A DMTD system works best if the zero crossings are coincident or can be interpolated to a common epoch. Thus, a variable phase shifter is sometimes used to put the DUT signal nominally in phase with the reference before the DUT signal is mixed with the transfer oscillator. After the mixing process, some systems use an event counter to count the whole beat note cycles to eliminate the ambiguity of the zero crossings; other systems time-tag the zero crossings. The resolution of a DMTD system is determined by the resolution of the TIC or the time-tagging hardware, divided by the heterodyne factor. For example, if the TIC resolution is 100 ns and the heterodyne factor is 10^6 (based on a 10 MHz nominal frequency and a 10 Hz beat frequency), then the DMTD resolution can be estimated at 10^{-7} s/10^6 or 0.1 ps. This high resolution allows some DMTD systems to detect frequency changes of 1×10^{-13} in 1 s or two or three orders of magnitude better than a typical universal counter. It should be noted that all components of heterodyne or DMTD systems must be carefully chosen for optimum stability and not all frequency measurement systems are stable enough to support their theoretical resolution [1,5,14,16,17].

Recently developed frequency measurement systems employ digital signal processing (DSP) techniques to achieve even higher resolutions. These systems rapidly sample the DUT and reference signals with analog-to-digital converters and then fit the digitized signals with a sinusoidal waveform at an IF that is much lower than the carrier frequency. Figure 42.19 shows the relationship between the oscillator signal, the sampling clock, and the resulting IF sine wave. Again, the idea is simply to compare two

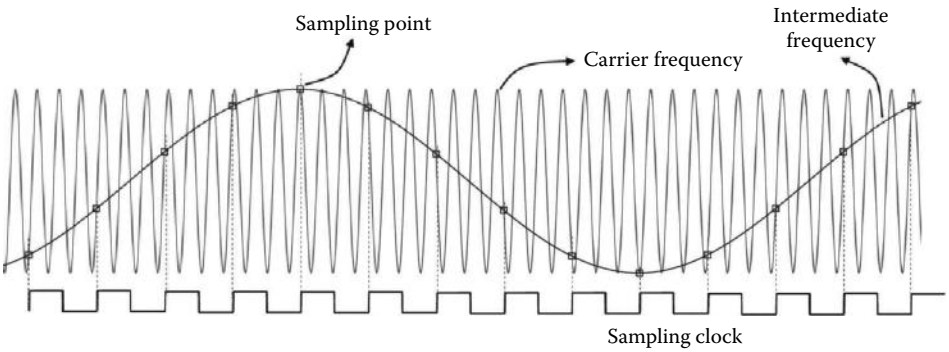


FIGURE 42.19 Digital sampling of the carrier frequency.

low-frequency signals, but note that the DSP hardware and software replaces the analog mixers shown in Figures 42.17 and 42.18 or the dividers shown in Figure 42.16. After two low-frequency signals are obtained for comparison, the subsequent filtering and phase detection is also accomplished with DSP techniques.

At this writing (2011), the available digital sampling frequency measurement systems are relatively expensive but can outperform DMTD systems by a wide margin. The best systems have a resolution of a few femtoseconds (10^{-15} s) and can detect frequency changes of less than 5×10^{-15} at 1 s. These systems can also measure phase noise, can simultaneously estimate time and frequency domain statistics with the same hardware, and can measure oscillators with a wide range of nominal frequencies [18].

To provide a practical, "real-world" comparison of the most commonly used frequency measurement systems, Figure 42.20 compares the FC, time interval, single-mixer, and DMTD methods by showing an ADEV graph of the same OCXO measured with all four systems. The same universal counter, with an external time base connection to an atomic oscillator, was used for the first three methods. The counter chosen was typical of those found in many metrology laboratories, with a 12-digit display and a time interval resolution of 150 ps. For the FC method, the counter was configured as an FC and used by itself. For the time interval method, the counter was configured as a TIC and used with a low-cost divider circuit. For the mixer method, the counter was used in conjunction with a low-cost double-balanced mixer and a frequency synthesizer that generated a signal offset by 1 kHz from the reference. A commercially available system was used for the DMTD measurement.

Only the DMTD system was able to correctly show that the OCXO reaches its noise floor at an averaging period of 10–20 s, when its stability briefly drops below 3×10^{-12}. Due to a lack of resolution and due to measurement system noise that exceeds the oscillator noise at short intervals, the other three systems incorrectly indicate that the OCXO noise floor is not reached until after several hundred seconds of averaging, when a random walk noise process had already begun. The FC and TIC methods provide similar results at all intervals, but the TIC estimates are smoother and have fewer variations. The single-mixer system was not optimized for best performance, but was somewhat quieter than the FC and TIC systems at short averaging periods. It is interesting to note that when the averaging interval exceeds several hundred seconds, all four methods produce similar results. This demonstrates that the choice of a measurement system is not critical when measuring long-term stability or accuracy, but is critical when measuring short-term stability.

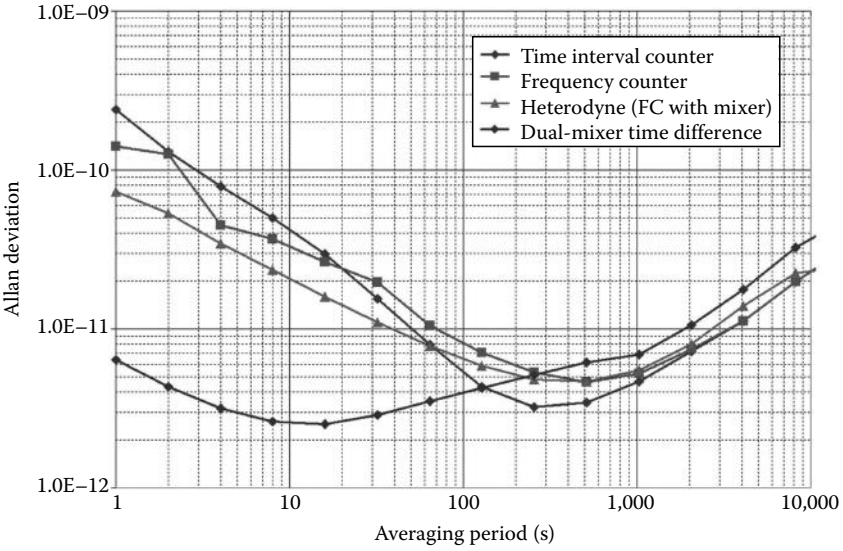


FIGURE 42.20 ADEV graph comparing different measurement methods.

42.4 Future Developments

Cesium fountain standards currently (2011) serve as the national standards of frequency at NIST and a number of other laboratories. These devices implement techniques such as a vertical microwave cavity and laser cooling to increase the interrogation period of the cesium atoms. By doing so, they reduce the resonance width and increase the oscillator Q to about 10^{10}. As a result, their accuracy and stability is about a factor of 100 times better than commercial cesium standards [19]. Perhaps cesium fountains or another type of laser-cooled microwave frequency standard will eventually be sold commercially.

It seems almost certain that optical atomic standards will have a huge impact on the future of frequency metrology. Optical standards operate at much higher resonance frequencies than microwave standards; the stabilized lasers that serve as their resonators typically operate at a frequency near 10^{15} Hz, as opposed to less than 10^{10} Hz for cesium. As a result, these standards promise accuracies and stabilities that are several orders of magnitude better than the best microwave standards. Optical frequency standards have been constructed at NIST utilizing single-ion techniques based on mercury (^{199}Hg+) and aluminum (^{27}Al+), as well as neutral atom techniques based on calcium (^{40}Ca), ytterbium (^{174}Yb), and strontium (^{87}Sr). It appears likely that the SI second will eventually be redefined, with the new definition based on one of these optical atomic transitions [20].

The accurate measurement of optical frequencies had historically been difficult, involving large and complex chains of frequency-doubled and frequency-mixed lasers. This changed in the early part of the twenty-first century, with the advent of the self-referenced femtosecond laser frequency combs. These devices, which are now sold commercially, have made the linkage of optical to microwave frequencies, or optical to optical frequencies, relatively simple. A self-referenced femtosecond laser frequency comb generates a series of discrete, equally frequency-spaced modes. The mode spacing, f_{rep}, is given by the repetition rate of a mode-locked laser, which is typically around 1 GHz. The frequency of an individual mode can be expressed as $f(m) = f_o + m \times f_{rep}$, where m is an integer. The frequency offset f_o is measured by a method called self-referencing. An optical standard can be compared to a conventional frequency standard by referencing the comb to a cesium standard or a GPSDO and then measuring the heterodyne beat frequency f_{beat} between the optical standard and the nearest tooth of the comb. This will lead to determination of the frequency of the optical standard in terms of f_o, f_{rep}, and f_{beat}, if the integer m is known. Frequency combs also allow the ratio of two optical frequencies to be compared without being limited by the accuracy of conventional standards [21]. Frequency comb comparisons of optical standards have successfully measured relative frequency differences of parts in 10^{19}, an astounding 16 orders of magnitude smaller than the requirements that existed less than a century ago for the measurements of radio frequencies.

References

1. G. Kamas and S. L. Howe, eds., *Time and Frequency User's Manual*, National Bureau of Standards Special Publication 559, Washington, DC, 256p, November 1979.
2. C. Audoin and B. Guinot, *The Measurement of Time: Time, Frequency, and the Atomic Clock*, Cambridge University Press, Cambridge, U.K., 335p, 2001.
3. J. Jesperson, Introduction to the time domain characterization of frequency standards, *Proceedings of the 25th Annual Precise Time and Time Interval* (*PTTI*) *Meeting*, Pasadena, CA, pp. 83–102, December 1991.
4. IEEE, *IEEE Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology—Random Instabilities*, IEEE Std. 1139, New York, 1999.
5. W. J. Riley, *Handbook of Frequency Stability Analysis*, National Institute of Standards and Technology Special Publication 1065, Gaithersburg, MD, 136p, July 2008.
6. H. Hellwig, *Frequency Standards and Clocks: A Tutorial Introduction*, National Bureau of Standards Technical Note 616, Washington, DC, 72p, June 1977.
7. L. Lewis, An introduction to frequency standards, *Proc. IEEE*, 79(7), 927–935, 1991.

8. F. G. Major, *The Quantum Beat: The Physical Principles of Atomic Clocks*, Springer-Verlag, New York, 475p, 1998.
9. J. R. Vig and A. Ballato, Frequency control devices, in *Ultrasonic Instruments and Devices*, Chapter 7, Academic Press, New York, pp. 637–701, 1999.
10. J. Vanier and C. Audoin, *The Quantum Physics of Atomic Frequency Standards*, 2 Vols., Adam Hilger, Bristol, England, 1989.
11. L. Cutler, Fifty years of commercial caesium clocks, *Metrologia,* 42, S90–S99, June 2005.
12. J. A. Kusters, L. S. Cutler, E. D. Powers, Long-term experience with cesium beam frequency standards, *Proceedings of the 1999 IEEE Frequency Control Symposium and European Frequency and Time Forum* (*EFTF*), Besançon, France, pp. 159–163, April 1999.
13. M. A. Lombardi, The use of GPS disciplined oscillators as primary frequency standards for calibration and metrology laboratories, *Measure: J. Meas. Sci*., 3(3), 56–65, September 2008.
14. D. B. Sullivan, D. W. Allan, D. A. Howe, and F. L. Walls, Eds., *Characterization of Clocks and Oscillators*, National Institute of Standards and Technology, Washington, DC, Technical Note 1337, 352p, March 1990.
15. Hewlett-Packard Corporation, *Fundamentals of the Electronic Counters*, Hewlett-Packard Application Note 200, 5965-7660E, Palo Alto, CA, 44p, March 1997.
16. Hewlett-Packard Corporation, *Simplify Frequency Stability Measurements with Built-in Allan Variance Analysis*, Hewlett-Packard Application Note 358-12, 5952-8006, Palo Alto, CA, 32p, January 1990.
17. J. Levine, Introduction to time and frequency metrology, *Rev. Sci. Instrum*., 70(6), 2567–2596, June 1999.
18. S. R. Stein, J. Grove, J. Hein, R. A. Martin, B. Mesander, M. Phillips, P. Schweiger, J. Shima, and W. Solbrig, Comparison of heterodyne and direct-sampling techniques for phase-difference measurements, *Proceedings of the 2005 NCSLI International Workshop and Symposium*, Washington, DC, 10p, July 2005.
19. M. A. Lombardi, T. P. Heavner, and S. R. Jefferts, NIST primary frequency standards and the realization of the SI second, *Measure: J. Meas. Sci*., 2(4), 74–89, December 2007.
20. L. Lorini, N. Ashby, A. Brusch, S. A. Diddams, R. E. Drullinger, E. Eason, T. Fortier et al., Recent atomic clock comparisons at NIST, *Eur. Phys. J. Special Topics*, 35, 19–35, 2008.
21. S. Diddams, The evolving optical frequency comb, *J. Opt. Soc. Am. B*, 27(11), B51–B62, November 2010.

V

Optical

43

Photoconductive Sensors

Fritz Schuermeyer
Wright-Patterson Air Force Base

Thad Pickenpaugh
Wright-Patterson Air Force Base

43.1 Introduction

Photoconduction has been observed, studied, and applied for more than 100 years. In the year 1873, W. Smith [1] noticed that the resistance of a selenium resistor depended on illumination by light. Since that time, photoconduction has been an important tool used to evaluate material properties, to study semiconductor device characteristics, and to convert optical into electric signals. The Radio Corporation of America (RCA) was a leader in the study and development of photoconductivity and of photoconductive devices. Richard H. Bube of RCA Laboratories wrote the classic book *Photoconductivity in Solids* [2] in 1960. Today, photoconducting devices are used to generate very fast electric pulses using laser pulses with subpicosecond rise and fall times [3]. For optoelectronic communications, photoconducting devices allow operation in the gigabit per second range.

Photoconductive devices normally have two terminals. Illumination of a photoconductive device changes its resistance. Conventional techniques are used to measure the resistance of the photoconductor. Frequently, small changes in conductivity need to be observed in the study of material or device characteristics. Also, in the measurement of light intensities of faint objects, one encounters small photoconductive signals.

Only solid photoconductors, such as Si, PbS, PbSe, and mercury–cadmium–telluride (HgCdTe), will be treated here. Photoconduction has been observed in amorphous, polycrystalline, and single-crystalline materials. During the last decade, major improvements in material growth have occurred that directly translate in better device performance such as sensitivity and stability. Growth techniques such as molecular beam epitaxy (MBE) and metal organic chemical vapor deposition (MOCVD) allow the growth of single-crystal layers with an accuracy of the lattice constant. Artificially structured materials can be fabricated with these growth techniques for use in novel photoconducting devices.

Absorption of light in semiconductors can free charge carriers that contribute to the conduction process. Figure 43.1 presents the band diagram for a direct bandgap semiconductor where the excitation

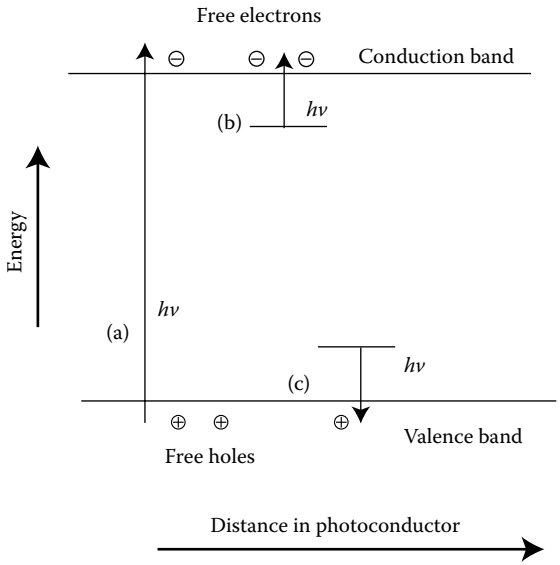


FIGURE 43.1 Example of electronic transitions in a photoconductor: (a) band-to-band excitation, (b) excitation from a trap or a donor, and (c) transition from a trap or an acceptor to the valance band; $h\nu$ is the energy of the absorbed photon.

processes are indicated. Excitation process (a) is a band-to-band transition. The photon energy for this excitation has to exceed the bandgap of the semiconductor. The absorption constant is larger for this process than for any of the other processes shown in this figure. Typical semiconductors used for electronic applications have bandgaps in excess of 1 eV, corresponding to light in the near-infrared (IR) region. Special semiconductors have been developed with narrower bandgaps to provide absorption in the mid- and long-wavelength IR regions. Indium antimonide (InSb) and HgCdTe semiconductors provide photosensitivity in the 4 and 10 μm wavelength range, respectively. The photogenerated carriers increase the electron and hole densities in the conduction and valence bands, respectively, which leads to an increase in conductivity [4]. For the simplified case with one type of carrier dominating, the conductivity σ is given by

$$\sigma = ne\mu \tag{43.1}$$

where
- n is the density of free carriers
- e is the charge
- μ is the mobility

Absorption of light results in a change in free-carrier density and a corresponding change in conductivity $\Delta\sigma$:

$$\Delta\sigma = \Delta ne\mu + \Delta\mu en \tag{43.2}$$

$\Delta\sigma$ is the definition for photoconductivity. In Equation 43.2, one assumes that due to the photon absorption, the density of carriers changes. Also, the mobility of the carriers changes due to the modified free-carrier density. The latter effect is very small except for special band transitions, as with InSb at very low temperatures.

Figure 43.1 indicates that other excitation processes exist. For example, bound electrons can be excited into the conduction band. This process can lead to persistent photoconductivity. In this example, the trapped holes have a long lifetime while the electrons move freely due to the applied electric field. Charge neutrality requires that the electrons collected at the anode be replenished by electrons supplied by the cathode. This effect leads to an amplification of the photogenerated charge (i.e., more than one electron is collected at the anode of the photoconducting device per absorbed photon). Often, the storage times are long, in the millisecond range. Hence, photoconductive devices with large amplification have a slow signal response.

Small bandgap semiconductors, such as HgCdTe and InSb, are difficult to manufacture. Thus, artificially structured layers of commonly used materials are being developed to replace these. Spatial modulation of doping has been proposed by Esaki and Tsu [5] to achieve a lattice containing a superlattice of *n*-doped and undoped, and *p*-doped and undoped layers *(n-I-p-I)*. Due to the energy configuration of this structure, the effective bandgap is less than that of the undoped material. The effective bandgap depends on the thickness of the layers and their doping concentrations. The quantum-well IR photodetector (QWIP) [6] is another approach to obtain photoconduction in the far-IR wavelength range. In this structure, energy wells exist in the conduction band of the material heterostructure due to the energy band discontinuities. Subbands form in the superlattice and electrons in these wells are confined due to the heterobarriers. IR photons can excite electrons from their confined states to the continuum, which leads to an increase in conductivity.

While it is possible to use noncontact methods to measure the conductivity in a material, electric contacts are commonly placed onto the structure during the device fabrication process. Typically, ohmic contacts are formed to fabricate metal–semiconductor–metal (MSM) structures (Figure 43.2). These contacts control the Fermi level in the material structure and provide carriers to retain charge neutrality.

43.2 Detector Performance Parameters

43.2.1 Responsivity

Variations in photon flux density incident on a photoconductor interact with the material to change the conductivity. These changes produce a signal voltage that is proportional to the input photon flux density. The detector area A collects flux contributing to the signal. J_s is the integrated power density over a spectral interval. Responsivity (R_v) is the ratio of the rms signal voltage (or current) to the rms signal power and is expressed in units of volts per watt. It is expressed as amps per watt for current responsivity:

$$R_v = \frac{V_s}{AJ_s} \tag{43.3}$$

V_s is normally linear with photon flux for low levels but can saturate under high flux conditions. One should ensure operation in the linear range for radiometric and photometric instrumentation.

43.2.2 Noise

The performance of a visible or IR instrument is ultimately limited when the signal-to-noise ratio equals one (SNR = 1). The noise from the instrument's signal processing should be less than the noise from the detector in the ideal case. This means reducing this noise within the restrictions of signal processing design limitations. These may include cost, size, and input power. The detector noise should be minimized.

Johnson noise is the limiting noise in all conductors [7]. It is frequency independent and independent of the current going through the device. Johnson noise is defined in Equation 43.4, where k is the

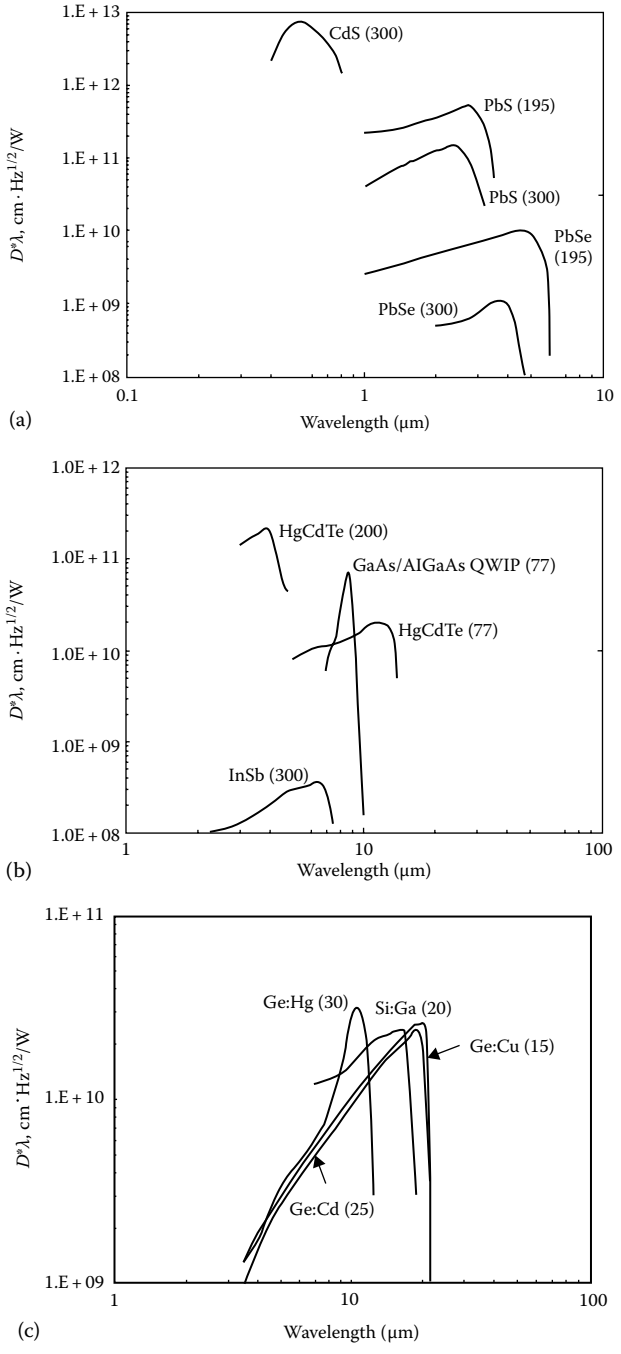


FIGURE 43.2 (a) Absolute spectral response of photoconductive detectors with the operating temperatures in K in parentheses: CdS visible and Pb salt IR detectors; (b) Absolute spectral response of photoconductive detectors with the operating temperatures in K in parentheses: III–V and II–VI intrinsic photoconductors plus a III–V QWIP detector; (c) Absolute spectral response of photoconductive detectors with the operating temperatures in K in parentheses: long-wavelength extrinsic Ge and Si photoconductors.

Boltzmann constant (1.38×10^{-23} J/K), T is the detector temperature (K), R is the resistance (Ω), and Δf is the amplifier bandwidth (Hz):

$$V_J = \sqrt{4kTR\Delta f} \tag{43.4}$$

Another type of noise known as $1/f$ noise (V_f) is present in all semiconductor detectors that carry current. The spectrum of this noise varies as $1/f^n$, with n approximately 0.5 [8].

Noise due to fluctuation in generation and recombination of charge carriers [9] varies linearly with current. This noise may be caused by the random arrival of photons from the background (photon noise), fluctuation in the density of charge carriers caused by lattice vibration (g–r noise), by interaction with traps, or between bands.

Excess noise from the amplifier or signal processing (V_{amp}) can also limit photoconductive detector performance.

These uncorrelated noises add in quadrature, giving the total noise (V_N):

$$V_N^2 = V_J^2 + V_{g-r}^2 + V_{amp}^2 \tag{43.5}$$

The total noise may be given in units of $V\sqrt{Hz}$. It may also be integrated over some frequency range to provide volts rms. Photoconductive detectors often have a g–r noise independent of frequency from dc to 100 kHz.

43.2.3 Detector Sensitivity

Minimum detectable signal power, that is, noise-equivalent power (NEP) is a convenient means to express detector sensitivity. NEP is expressed in units of watts or $W\sqrt{Hz}$:

$$\text{NEP} = \frac{V_N}{R_V} \tag{43.6}$$

The reciprocal of NEP, the detectivity D, is frequently used. In attempting to make possible comparison among detectors, detectivity can be normalized to an electronic bandwidth of 1 Hz and a detector area of 1 cm^2. This yields the highly used parameter specific detectivity or D^* (pronounced "dee-star") [10]:

$$D^* = \frac{R_V}{V_N}\sqrt{(A\Delta f)} \tag{43.7}$$

The units of D^* are $cm \cdot Hz^{1/2}/W$, sometime simplified to "Jones."

This normalization is based on evidence that noise varies as the square root of the electronic bandwidth and D varies inversely as the square root of the detector area. This relationship may not hold closely over a wide range of device sizes and bandwidths. Comparison of device performance is most meaningful among devices having similar sizes and measured under similar conditions, including operating temperature, chopping frequency/scanning rate, and detector field of view.

43.3 Preparation and Performance of Photoconductive Detectors

43.3.1 Cadmium Sulfide

CdS is normally prepared by vapor deposition or sintering a layer of high-purity CdS powder on a ceramic substrate [11]. It has the largest change in resistance with illumination of any photoconductor. The peak response of this intrinsic detector is at 0.5 μm. Its spectral response is similar to that of the human eye and operates without cooling.

43.3.2 Lead Sulfide

PbS was among the earliest IR detector material investigated. Cashman was one of the earliest researchers in the United States. [12]. This intrinsic detector material is prepared by deposition of polycrystalline thin films by vacuum sublimation or chemical deposition from a solution. The spectral response extends to approximately 3 μm. PbS operates over the temperature range from 77 K to room temperature. The frequency response slows considerably at the lowest temperatures. The spectral response extends to somewhat longer wavelengths with cooling.

43.3.3 Lead Selenide

PbSe is an intrinsic detector that operates over the temperature range from 77 K to room temperature. Its spectral response extends to longer wavelengths with cooling. Preparation of PbSe is by sublimation or chemical deposition. Noise in PbSe detectors follows a $1/f$ spectrum.

43.3.4 Indium Antimonide

InSb is prepared by melting together stoichiometric quantities of indium and antimony. It operates over the range from 77 K to room temperature. The higher performance and ease of operation with signal processing electronics lead photovoltaic InSb detectors to be much more widely used than photoconductive.

43.3.5 Mercury Cadmium Telluride

HgCdTe is a versatile intrinsic material for IR detectors. CdTe and HgTe are combined to form the alloy semiconductor $Hg_{1-x}Cd_xTe$. For the alloy with $x \approx 0.2$, the bandgap is approximately 0.1 eV, providing a long-wavelength cutoff of 12.4 μm. HgCdTe was initially grown into bulk crystals by solid-state crystallization (also called quench and anneal). Currently, thin-film growth techniques of liquid-phase epitaxy (LPE), MOCVD, and MBE are preferred to obtain larger, more uniform wafers. By appropriately choosing the alloy composition, photoconductive HgCdTe detectors are possible over the 2–20 μm range. CdZnTe wafers permit lattice-matched surfaces for HgCdTe thin-film growth. Operating temperatures can range from 77 K to room temperature, with the lower temperatures necessary for the longer wavelength devices.

43.3.6 Extrinsic Germanium and Silicon

The photoresponse of an extrinsic detector occurs when a photon interacts with an impurity added to a host semiconductor material. With an intrinsic material, the photoresponse is from the interaction with the basic material.

For the extrinsic detector, incident photons may produce free electron-bound hole pairs or bound electron-free hole pairs. The extrinsic detector's spectral response is achieved using an impurity (or doping element). Intrinsic detection occurs with a detector having the necessary bandgap width for the desired spectral response.

Extrinsic detectors require lower temperatures than do intrinsic and QWIPs, but have the advantage of longer wavelength response.

Ge and Si are zone refined to achieve high purity by making multiple passes of a narrow molten zone from one end to the other of an ingot of the material. Unwanted impurities can be reduced to levels of 10^{12}–$10^{13}/m^3$ [13]. Growth of single crystals is by the Czochralski approach of bringing an oriented seed crystal in contact with the melt and withdrawing it slowly while it is rotated or by applying the horizontal zone refining approach, whereby an oriented seed crystal is melted onto the end of a polycrystalline ingot. A molten zone is started at the meeting of the ingot and seed and moved slowly down the ingot, growing it into a single crystal. An inert atmosphere is required to prevent oxidation.

TABLE 43.1 Photoconductive Detectors

Material	Cutoff Wavelength (μm)	Temperature (K)	Responsivity (V/W)	D (cm Hz$^{1/2}$/W)
CdS	0.7	300	1×10^6	1×10^{13}
PbS	3	300	5×10^4–1×10^3	1.5×10^{11}–3.3×10^{11}
PbSe	5.8	77–300	1×10^6–1×10^3	2×10^{10}–7×10^8
InSb	7	300	5	4×10^8
HgCdTe	5	150–220	1×10^5–2×10^4	
HgCdTe	12	65–100	1×10^5	3×10^{10}
Ge/Hg	13	4–25	8×10^5	2×10^{10}
Ge/Cd	24	20–30	5×10^5	2×10^{10}
Ge/Cu	33	5	5×10^5	3×10^{10}
GaAs/AlGaAs (QWIP)	9	77	780 mA/W	7×10^{10}

Sources: Borrello, S.R. and Wadsworth, M.V.: *Encyclopedia of Chemical Technology*. pp. 862–863. 1996. Copyright Wiley-VCH Verlag GmbH & Co. KGaA.; Wolfe, W.L. and Zissis, G.J. (Eds.), *The Infrared Handbook*, Revised edition., Environmental Research Institute of Michigan, Ann Arbor, MI, 1985.

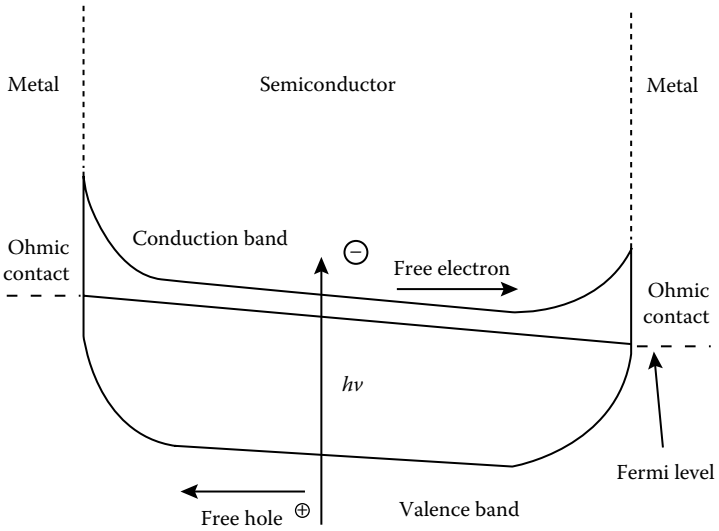


FIGURE 43.3 Energy diagram for a MSM detector.

Hg, Cd, Cu, and Zn are impurities for doping Ge detectors; Ga and As are dopants for Si detectors. See Table 43.1 and Figure 43.3.

43.3.7 Gallium Arsenide/Aluminum Gallium Arsenide QWIP

QWIP technology uses a quantum-well structure to provide intraband (intersubband) transitions to achieve an effective long-wavelength response in a wide bandgap material. Quantum wells are used to provide states within the conduction or valence bands. Since $h\nu$ of the desired spectral region is less than the bandgap of the host material, the quantum wells must be doped. Quantum-well structures are designed to permit photoexcited carriers to depart the structure and be accumulated as signal (photocurrent). The QWIP detector is generally comparable to extrinsic photoconductive detectors [16,17], in that both have lower than desirable quantum efficiency. GaAs/AlGaAs QWIPs have the advantage of higher operating temperatures than extrinsic detectors.

43.4 Instrumentation

The Stanford Research Systems SR570 low-noise current preamplifier can be used to amplify the current flowing through a photoconductive device. This preamplifier can be programmed to apply a voltage to the terminals of the photoconducting device. Its output voltage is proportional to the device current. Frequently, the IR radiation or visible light is chopped and the ac component of the device current is detected using lock-in-amplifier techniques. This approach allows the study of very small changes in device conduction. The Stanford Research Systems SR570 and the EG&G Instruments Model 651 are examples of a lock-in amplifier and a mechanical radiation/light chopper, respectively.

References

1. W. Smith, Effect of light on selenium during the passage of and electric current, *Nature*, London, 7, p. 303 (1873).
2. R.H. Bube, *Photoconductivity of Solids*, New York: John Wiley & Sons, 1960.
3. J.A. Valdmanis, G.A. Mourou, and C.W. Gabel, Pico-second electro-optic sampling system, *Appl. Phys. Lett*., **41**, 211–212, 1982.
4. R.H. Bube, Photoconductors, in *Photoelectronic Materials and Devices*, S. Larach, (Ed.), Princeton, NJ: D. Van Nostrand Company, pp. 100–139, 1965.
5. L. Esaki and R. Tsu, Superlattice and negative differential conductivity in semiconductors, *IBM J. Res. Dev*., **14**, 61, 1971.
6. B.F. Levine, Quantum-well infrared photodetectors, *J. Appl. Phys*., **74**, R1–R81, 1993.
7. P.W. Kruse, L.D. McGlauchlin, and R.B. McQuistan, *Elements of Infrared Technology*, New York: John Wiley & Sons, 1962.
8. H. Levinstein, Characterization of infrared detectors, in *Semiconductors and Semimetals*, R.K. Willardson and A.C. Beer (Eds.), Vol. 5, New York: Academic Press, p. 5, 1970.
9. K.M. Van Vliet, Noise in semiconductors and photoconductors, *Proc. IRE*, **46**, 1004, 1958.
10. R.C. Jones, Phenomenological description of the response and detecting ability of radiation detectors, *Proc. IRE*, **47**, 1495, 1959.
11. P.W. Kruse, L.D. McGlauchlin, and R.B. McQuistan, *Elements of Infrared Technology*, New York: John Wiley & Sons, pp. 417–418, 1962.
12. R.J. Cashman, Film-type infrared photoconductors, *Proc. IRE*, **47**, 1471, 1959.
13. S.R. Borrello and M.V. Wadsworth, Photodetectors, in *Encyclopedia of Chemical Technology*, R.E. Kirk and D.E. Othmer (Eds.), Vol. 18, New York: John Wiley & Sons, pp. 897–898, 1996.
14. S.R. Borrello and M.V. Wadsworth, Photodetectors, in *Encyclopedia of Chemical Technology*, R.E. Kirk and D.E. Othmer (Eds.), Vol. 18, New York: John Wiley & Sons, pp. 862–863, 1996.
15. W.L. Wolfe and G.J. Zissis (Eds.), *The Infrared Handbook*, Revised edition., Ann Arbor, MI: Environmental Research Institute of Michigan, 1985.
16. B.F. Levine, Quantum-well infrared photodetectors, *J. Appl. Phys*., **74**, R3, 1993.
17. T.R. Schimert, D.L. Barnes, A.J. Brouns, F.C. Case, P. Mitra, and L.T. Clairborne, Enhanced quantum well infrared photodetector with novel multiple quantum well grating structure, *Appl. Phys. Lett*., **68** (20), pp. 2846–2848, 1996.

44

Photojunction Sensors

Michael R. Squillante
Radiation Monitoring Devices, Inc.

Kanai S. Shah
Radiation Monitoring Devices, Inc.

James F. Christian
Radiation Monitoring Devices, Inc.

44.1 Introduction

Photojunction sensors (photodiodes and phototransistors) are semiconductor devices that convert the electrons generated by the photoelectric effect into a detectable electronic signal. The *photoelectric effect* is a phenomenon in which photons lose energy to electrons in a material. In the case of a semiconductor, when the energy of an interacting photon ($h\nu$) exceeds the energy of the semiconductor bandgap (E_g), the energy absorbed can promote an electron from the valence band to the conduction band of the material. This causes the formation of an electron–hole pair. In the presence of an electric field, these charges drift toward electrodes on the surface and produce the signal.

The junction in the photojunction device creates a diode that provides a small built-in electric field to propel the charges to the electrodes (photovoltaic mode of operation). In the photovoltaic mode, either the photocurrent or the photovoltage can be measured.

Photojunction sensors are an alternative to other light sensing technologies, such as vacuum tube detectors, for example, photomultiplier tubes. There are a variety of advantages to be gained by using photojunction sensors, including higher quantum efficiency (QE), tailored spectral response, increased ruggedness, reduced power requirements, reduced weight, compact size, elimination of a warm-up period, reduced sensitivity to temperature and voltage fluctuations, and insensitivity to magnetic fields.

In general, semiconductor photojunction sensors produce more dark current than photomultiplier tubes, especially at room temperature. Upon cooling, the noise in semiconductors can be reduced significantly due to a reduction in dark current.

44.2 Photodiodes

Photodiode devices are most often operated with a bias voltage applied opposing the junction (*reversed bias*) to provide the electric field. The presence of the junction in a diode allows for the application of a relatively large bias to be applied while maintaining a relatively low reverse leakage current and thus relatively low noise. The result of an applied bias on a junction is the increase of the "depletion region," which is the sensitive volume of the detector. Any charges generated within this volume are swept toward the electrodes by the field, adding to the reverse leakage current. The *total reverse* current is the sum of the *dark current*, which is often dominated by the thermal generation of charges in the depletion region, and the *photocurrent*, which is produced due to optical illumination. The dark, or leakage, current contains contributions from the thermal generation of charge carriers in the bulk and surface, as well as ohmic terms associated with the conduction of current through the high-field regions at the surfaces of the device. The sensitivity of the detector to optical illumination decreases as the dark current increases.

There are several fundamental types of junction photodiodes [1,2], as shown in Figure 44.1. A *Schottky barrier* diode is a device in which the junction is formed at the surface of the semiconductor by the application of a metal electrode that has a work function that is different from the work function of the semiconductor; a *heterojunction diode* is a device in which two different

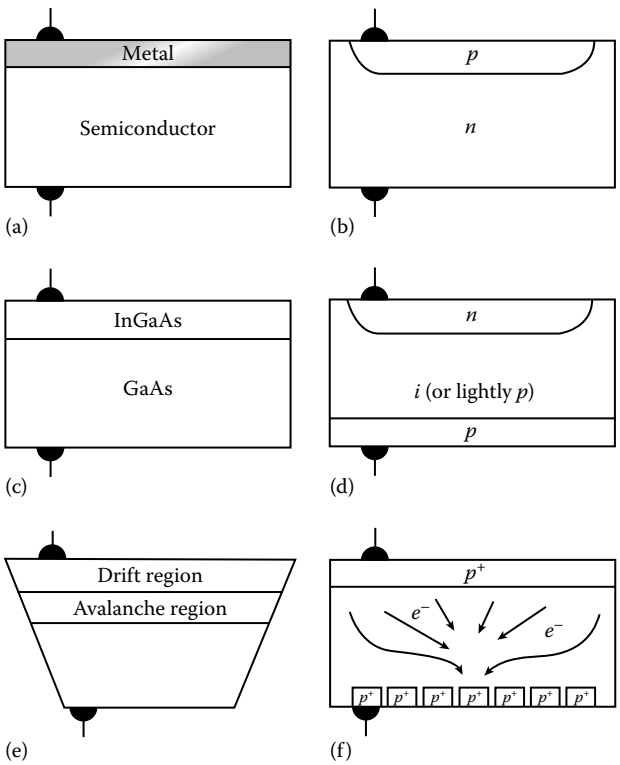


FIGURE 44.1 Schematic of photodiode device structures: (a) Schottky junction, (b) homojunction, (c) heterojunction, (d) *p-i-n*, (e) APD, and (f) drift diode.

semiconductor materials with differing work functions are joined; a *homojunction diode* is a device in which the junction is created at an interface in a single material and the difference in work function is created by doping the material *n*-type and *p*-type. Most photodiodes are homojunction devices made using silicon. Other, more complex types of photojunction devices, which are discussed later, include *p-i-n photodiodes, avalanche photodiodes (APDs), drift photodiodes, Geiger photodiodes (GPDs), and phototransistors*.

Photodiodes are typically characterized by several properties, including spectral response, operating bias, operating temperature, dark current, junction capacitance, noise equivalent power (NEP), and peak wavelength. Other specifications usually provided by manufacturers include size, packaging details, operating temperature range, capacitance, and price. The bandwidth is generally not an independent intrinsic property of the photodetector but is often related to the readout electronics and the selected integration time.

Photodiodes are used in numerous applications, including CD-ROM systems, television remote control systems, fax machines, copiers, optical scanners, fiber optic telecommunication repeaters, surveillance systems such as motion detectors, certain smoke detectors, and light meters. They are also used in a wide variety of scientific instrumentation including spectrophotometers, scintillation detectors, optical trackers, laser range finders, Laser Interferometry Detection and Ranging (LIDAR), LAser Detection And Ranging (LADAR), analytical instrumentation, optical thermometers, nephelometers, densitometers, radiometers, laser detectors, shaft encoders, and proximity sensors. Photodiode arrays are available for use as position-sensitive detectors that can either be used for imaging (such as in laser scanners, night vision equipment, spectrophotometers, and edge detection) or alignment systems. Medical imaging applications such as x-ray computed tomography (CT) scanners also use large arrays of photodiodes.

Varieties of materials are used in the fabrication of photodiodes, but most are fabricated using silicon. Other materials used include Ge; Se; II–VI compounds such as CdS, PbS, CdSe, and $Hg_xCd_{(1-x)}Te$; and III–V compounds such as GaAs, GaP, InP, InGaAs, and GaAlAs. In addition, materials with unique properties can be used to fabricate photojunction sensors that solve very specific and unusual problems, including HgMnTe, HgI_2, and InI.

Table 44.1 lists a few examples of available photodiodes. These are only a small fraction of the commercially available photodiodes and photodiode manufacturers. Lists of photodiode manufacturers are available [3,4].

TABLE 44.1 Examples of the Variety of Commercially Available Photodiodes

Commercial Source	Type	Region	Example Device	Comments
RMD	Si APD	Visible	SH8S	169 mm², $2850
Hamamatsu	Si	Visible	S2386-44K	13 mm² $13
OSI	Si-*p-i-n*, UV enhanced	To 200 nm	UV50	50 mm², $54
OSI	Si-*p-i-n*	Visible	PIH-HR040	0.8 mm², fast, $19
Hamamatsu	Si APD	Visible	S2385	20 mm², $536
OSI	InGaAs	Near IR	FCI-InGaAs-300	0.1 mm², $28
Hamamatsu	GaP	Near UV	G1961	1 mm², $50
Kolmar	HgCdTe	2–12 μm	KV104-l-a	1 mm², $4050
Kolmar	HgCdTe	To 18 μm	KMPC18-l-bl	1 mm², $2300
Kolmar	InSb	5.1 μm	KISD-l-a	1 mm², $2100

Note: Hamamatsu, Corp. Bridgewater, NJ (http://www.hamamatsu.com/); Kolmar Technologies, Newburyport, GA (http://www.kolmartech.com); Radiation Monitoring Devices, Inc. (RMD), Watertown, MA (www.rmdinc.com); Optoelectronics (OSI), Hawthorne, CA (http://www.osioptoelectronics.com).

44.2.1 Equivalent Circuit

The following parameters describe the current–voltage relationship for a diode: I_d = dark current (current present with no incident photons), I_j = reverse saturation current, I_o = output current, I_p = photocurrent current, and V_j = junction voltage. The dark current contains two contributions, one from an ideal rectifier, I_{ir}, which includes generation and recombination in only the charge neutral regions, that is, excluding contributions from the depletion region, and a second contribution from generation and recombination in the depletion region, I_{dw} [5]. The current from the diode, I_{ir}, assuming the diode is an ideal rectifier, is given by [6]

$$I_{ir} = I_j \cdot \left(\exp\left(\frac{qV_j}{kT} \right) - 1 \right) \tag{44.1}$$

where
- k is the Boltzmann's constant
- q is the electronic charge
- T is the absolute temperature

This expression neglects recombination and generation in the depletion region, which underestimates the dark current. The reverse saturation current depends on the intrinsic carrier density, n_i, acceptor and donor concentrations, diffusion constants, and the electron and hole lifetimes in the charge neutral regions, which excludes the depletion region. Neglecting the contribution from the charge neutral regions, which is often reasonable for many photodiodes, the following expression describes the current–voltage relationship for the diode, I_{dw} [7,8]:

$$I_{dw} = I_{jd} \cdot \left(\exp\left(\frac{qV_j}{2kT} \right) - 1 \right), \quad \text{where} \quad I_{jd} = \frac{qAn_i w}{2\tau} \tag{44.2}$$

where
- I_{jd} is the saturation current from the depletion region, which is dominated by midband, trap-assisted transitions
- A is the area
- w is the depletion width
- τ is the minority carrier lifetime

When operating photodiodes under reverse bias, the sum of the "ideal diode" saturation current and the saturation current from the depletion region describes the dark current. The depletion width contains a dependence on bias, V_j. In many photodiodes, the contribution from the current in the depletion region is dominant. Presumably, the sum of Equations 44.1 and 44.2 describes the current–voltage relationship of photodiodes. The following expression describes the total current under illumination:

$$I_o = I_{ir} + I_{dw} + I_p \tag{44.3}$$

For a more rigorous treatment, refer to [1, pp. 752–754] and [2,9].

44.2.2 *I–V* Characteristics of Photodiodes

With no illumination, photodiodes have *I–V* curves to equivalent standard diodes connected to an external power supply given by the sum of Equations 44.1 and 44.2. Illumination by light causes the

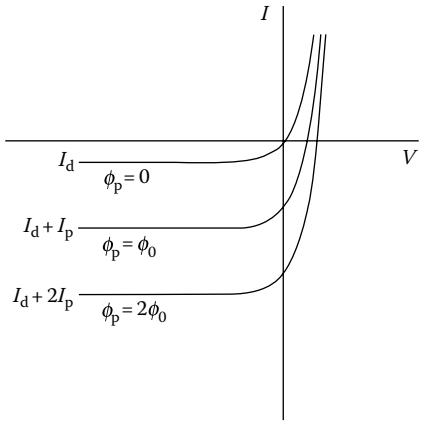


FIGURE 44.2 *I–V* characteristics of photodiode under illumination. The variable ϕ_p represents the photon flux.

current to increase. Figure 44.2 shows a family of *I–V* curves for a photodiode under illumination with equally increasing increments of incident light intensity, quantified by the photon flux φ_p. As the illumination on the device increases, the curve shifts downward by the amount of current generated by the incident light. The lower right-hand quadrant represents the photovoltaic mode of operation. When a photovoltaic device is operated in "current mode" with low or no load resistance (as with an operational amplifier, as in Figure 44.4a), the output is linear with incident light intensity. When operated in "voltage mode" with a high load resistance, there is an exponential relationship between the output and the incident illumination. The lower left-hand quadrant shows the reversed bias mode of operation. Again, in this mode, the output is nearly linear with the incident intensity.

44.2.3 Output Current under Reverse Bias

In a reverse-bias *p-n* junction under bias, the depletion width (W) increases as a function of applied bias (V_b) until the device is fully depleted. The dark leakage current (I_d) under reverse-bias conditions can arise from the generation–recombination effects (I_{jd}) and from diffusion (I_D) as well as surface effects. In most cases, the diffusion current is significantly smaller than the generation–recombination component. Thus, it is possible to assume that $I_d \cong I_{dw}$, as illustrated in Equation 44.2. This dark current is often referred to as the bulk dark current, which arises from trap-assisted generation in the depletion region.

The total diode current under illumination (I_o) is the sum of the dark leakage current (I_d) and the photocurrent (I_p):

$$I_o = I_d + I_p \tag{44.4}$$

44.2.4 Quantum Efficiency

In an ideal diode, the entire light incident on the photodiode surface is converted to electron–hole pairs, and all of the charges drift to the electrodes and are collected. In a real device, there are reflection losses at the surface, attenuation by the electrode and/or front dead layers of the device, and charge trapping so that only a fraction of the charges are collected at the electrodes. The behavior of real sensors deviates

from the ideal, and the QE characterizes the ideality of the sensor. The QE of a photodiode is the ratio of the charge pairs generated to the incident photons:

$$\eta = \frac{I_p/q}{P_i/h\upsilon} \tag{44.5}$$

where

P_i is the optical power incident on the photodiode
$h\upsilon$ is the energy of the photons

In the days of yore, before LEDs and calibrated photodiodes, light intensity was measured with power meters and bolometers. With these devices, the measured response is proportional to the energy of the photons, as well as the number of photons. Therefore, the responsivity, R, is the ratio of the photocurrent to the incident optical power (A/W):

$$R = \frac{I_p}{P_i} = \frac{\eta q}{h\upsilon} \tag{44.6}$$

Rearranging Equation 44.3 gives the following expression for the photocurrent:

$$I_p = \frac{q\eta P_i}{h\upsilon} \tag{44.7}$$

44.2.5 Simplified Equivalent Circuit

A simplified version of the equivalent circuit for a photodiode operated in photovoltaic mode is shown in Figure 44.3, where C_j = junction capacitance, I_j = reverse saturation current, I_o = output current, I_p = photocurrent current, R_j = junction shunt resistance (or parallel resistance), R_s = series resistance, V_j = junction voltage, V_o = output voltage, and R_L is the load resistance. In this mode, the diode is forward biased, and the reverse-bias saturation current, I_j, which is injected into the junction, subtracts from the photocurrent.

In the equivalent circuit for the photodiode, the junction shunt resistance is a measure of the ohmic leakage along the diode surfaces, the junction capacitance is inversely proportional to the depletion width and proportional to the device area, and the series resistance includes resistance associated with the diode material and contacts.

44.3 Photodiode Circuits

Figure 44.4a shows a simple circuit for operating a photodiode in the photovoltaic mode. In this mode, photocurrent is usually measured because the photocurrent is nearly proportional to the input signal. The output of the photodiodes is typically connected to the input of an op-amp current-to-voltage converter. Figure 44.4b shows a simple circuit for operating a photodiode under reverse bias.

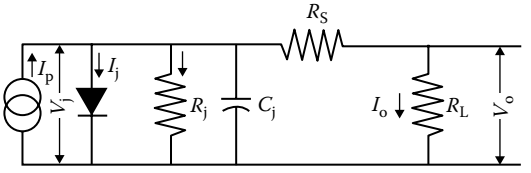


FIGURE 44.3 Equivalent circuit for a photodiode.

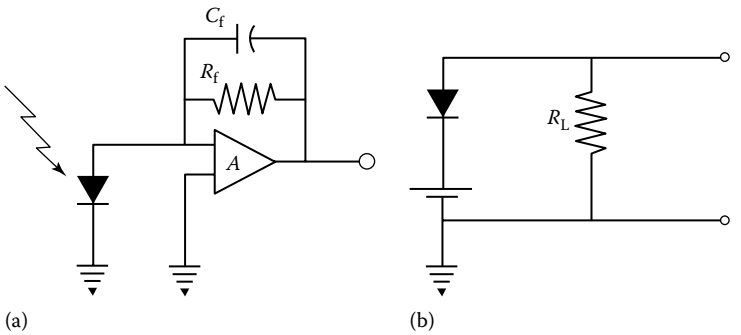


FIGURE 44.4 Typical circuits for operation of a photodiode: (a) circuit for photodiode operation in photovoltaic mode and (b) circuit for photodiode operating under reverse bias.

44.3.1 Noise

There are two main sources of noise when using a photodiode: *shot noise* in the diode and *thermal noise*. The shot noise current is related to the bulk dark current by the formula:

$$I_s = (2qI_dB)^{1/2} \tag{44.8}$$

where B is the bandwidth of the readout circuit or measurement. It should be noted that the noise currents described in this section are actually the fluctuations in the dark, or noise, current. In other words, the shot noise current is the root-mean-square (rms) variance in the average dark current.

Assuming that the diode shunt resistance and the input resistance of the measuring circuit to be used to measure the output of the photodiode are high relative to the load resistance, the thermal noise current, I_t, is given by

$$I_t = \left(\frac{4kTB}{R_L}\right)^{1/2} \tag{44.9}$$

where R_L is the load resistance. The total noise current is the sum of these currents in quadrature.

44.3.2 Quadrant Detector Arrays

Many manufacturers offer photodiodes fabricated in a *quadrant* configuration, where four photodiodes are fabricated in a square, 2 × 2 geometry. When coupled to a lens or a pinhole, they can be operated as position-sensitive detectors. In operation, the outputs of the four photodiodes are monitored, and the position of the light source can be determined by the projection of the light spot on the detector surface. Applications of quadrant detectors include position sensing, target tracking, laser ranging, fluorescence spectroscopy, laser alignment, and atomic force microscopy [10]. More recently, manufacturers are offering linear and area arrays of photodiodes that can be used as imaging devices.

44.4 Phototransistors

Phototransistors are photojunction devices similar to transistors except that the signal amplified is the charge pairs generated by the optical input. Like transistors, phototransistors can have high gain. This high gain, however, is often restricted to large signals in detectors with a macroscopic size. As the size of the phototransistor increases, the amount of charge needed to change the field in the transistor

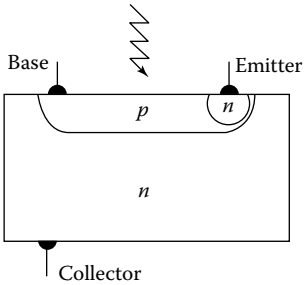


FIGURE 44.5 Schematic representations of a simple bipolar phototransistor. Note that the phototransistor has a large *p*-*n* junction region that is the photosensitive portion of the device.

to provide gain increases. In other words, phototransistors generally do not provide efficient gain for small signals. Phototransistors can be made on silicon using *p*- and *n*-type junctions or can be heterostructures. Figure 44.5 shows a sketch of the structure of a simple bipolar phototransistor, which is essentially the same as that of a simple bipolar transistor. The main difference is the larger base–collector junction, which is the light-sensitive region. This results in a larger junction capacitance, and, although the devices have gain, the capacitance gives phototransistors lower-frequency response than photodiodes.

Using thin-film transistor (TFT) technology developed for flat panel displays, large arrays of phototransistors can be fabricated on amorphous silicon to form imaging devices that can be used in place of other imaging technologies such as vidicon tubes or even film. Examples of this are the very large-area detectors (hundreds of square centimeters) being investigated for use in medical radiography by combining the TFT arrays with radiographic phosphor screens [11] or coupled to semiconductor films [12–15].

44.5 Silicon Photojunction Detector Structures

44.5.1 Silicon *p-i-n* Detectors

Silicon *p*-*i*-*n* diodes are an extension of the standard *p*-*n* junction diodes but are more attractive for low-noise applications due to reduced capacitance in these devices [1]. The reduction in capacitance is achieved by incorporating an intrinsic region between the *p* and *n* regions. This increases the depletion width of the detector and thereby lowers its capacitance. Silicon *p*-*i*-*n* detectors can be designed to have a lower junction capacitance, and therefore a higher frequency response, than *p*-*n* junctions.

In operation, *p*-*i*-*n* detectors are similar to *p*-*n* junction detectors, but the surface region (either *p* or *n*) is made thin so that the optical photons penetrate this entrance layer and are stopped in the intrinsic (*i*-region) where electron–hole pairs are produced, as shown in Figure 44.1. These electron–hole pairs are swept toward the appropriate electrodes by the applied electric field. For fabrication, *p*-*i*-*n* detectors require high-resistivity material, and typical photodiodes have a thickness ranging from 100 to 500 μm. Important applications of *p*-*i*-*n* detectors include optical sensing of scintillated light in CT scanners, general scintillation spectroscopy, charged particle spectroscopy, and high-speed sensing applications.

44.5.2 Silicon Drift Detectors

Silicon drift photodiodes are an extension of the *p*-*i*-*n* geometry. In recent extensive studies, they have been found to provide very low capacitance (<1 pF for 1 cm^2 detector with 300 μm thickness) [16,17]. This is achieved by reducing the area of ohmic electrode (anode in most cases) significantly as compared to the entrance electrode as shown in Figure 44.1f. Since the charge sensing electronics are connected to the smaller electrode, the device capacitance is proportional to its size and not to the actual

detector area. Thus, by exploiting this concept, significantly lower capacitance has been achieved than in a comparable *p-i-n* detector.

In drift detectors, it is important to ensure that charges created over the entire active volume will be collected at the anode. In order to achieve this, drift rings are provided around the anode. The outermost ring is biased at highest potential, with the inner rings biased to lower potentials in a successive manner. This arrangement creates a potential minimum at the anode and thereby enables efficient charge collection over the entire detector volume. A variety of device geometries have been explored based on this concept, and in some instances, these detectors are capable of providing position-sensitive and imaging detection as well [18,19]. While many of these detectors are in the research stage, excellent performance has been demonstrated by prototypes.

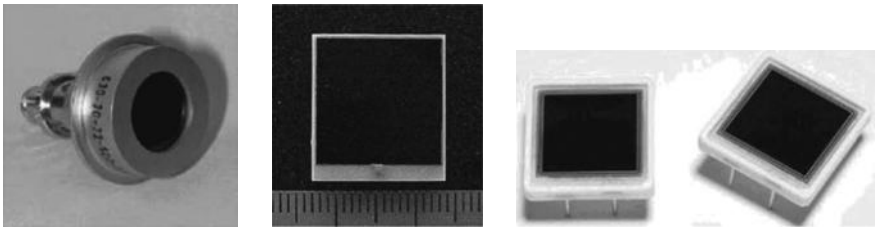
44.5.3 Silicon Avalanche Photodiodes

While the conventional silicon diodes, such as *p-n* junction diodes, *p-i-n* diodes, and drift diodes, have no gain, silicon APDs have internal gain that enables them to operate with high signal-to-noise ratios (SNRs) placing less stringent requirements on supporting electronics. Figure 44.6 shows a photograph of commercially available silicon APDs.

In its simplest form, an APD is a *p-n* junction formed in a silicon wafer structured in such a way that it may be operated near breakdown voltage under reverse bias. The absorption of either photons or charged particles in the silicon generates electron–hole pairs that are accelerated by the high electric field. These electrons gain sufficient velocity to generate free carriers by impact ionization, resulting in internal gain [20]. Small-area APDs have been in routine use in the telecommunications industry for some years [21–23] and in a variety of other applications such as optical decay measurement, time domain reflectometry, and laser ranging [24]. The standard large-area avalanche diode is a single element sensor made on a thin silicon wafer.

Initial research on large-area silicon APDs was carried out in the late 1960s [25], but large-area APDs attracted little attention until the 1980s when improved fabrication techniques led to higher gains and better reproducibility [26]. These devices have been used for a wide variety of nuclear spectroscopy applications [27–29]. In comparison to photomultiplier tubes, these sensors are smaller, more rugged, and more stable and use less power. They are also far more sensitive and have inherently better SNRs than the other semiconductor photosensors. In addition, they may be operated without cooling and are insensitive to magnetic fields and vibration.

The APD generally consists of several regions including the "drift" region and the active junction that contains the multiplication region. The drift region is typically 20 μm thick, while the active region of the device is approximately 120 μm thick, and the multiplication region is less than 10 μm thick. Ionizing radiation or light with energy greater than the bandgap that strikes the drift region will cause the generation of free charge carriers, which are then transported to the active region. No external field is present here, but a gentle gradation in dopant concentration causes a small electric field to impart



API 16 mm diameter APD (SD630-70-75-500)

RMD 14 × 14 mm^2 APD

Hamamatsu S8664-1010

FIGURE 44.6 Photographs of commercially available silicon APD devices.

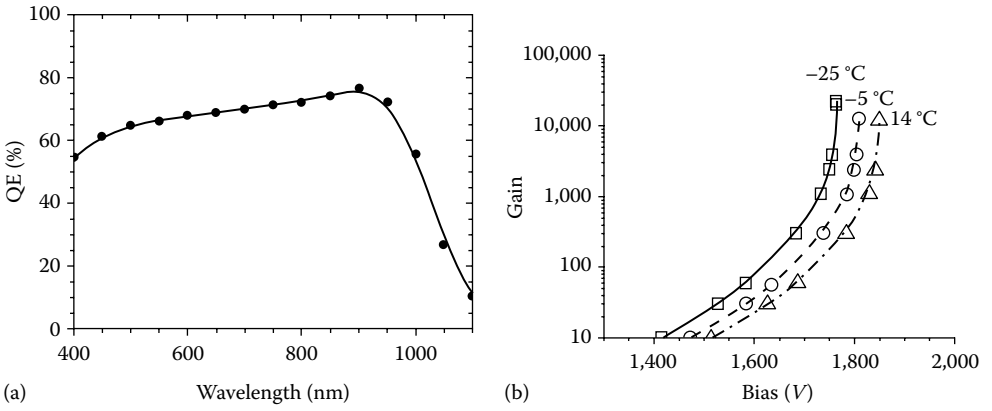


FIGURE 44.7 The diagram (a) shows a typical QE versus wavelength plot for high-gain silicon APD. The plot (b) shows the gain versus bias relationship for a high-gain APD. Higher gains are achievable at lower voltages as the temperature is decreased.

a net movement of electrons toward the multiplication region. Since the dopant concentration in this region is relatively light, the carrier lifetimes are quite large, and efficient transport is easily attained with the high-quality silicon now available. Within the multiplication region, the charge carriers are amplified in accordance with the gain of the device. Electrons entering this region quickly attain velocities large enough to cause knock-on collisions with bound electrons in the lattice. This process then frees additional electrons that can also undergo new collisions. The multiplication process occurs many times, with the result that a single electron will commonly generate hundreds of free electrons, which produces a significant net gain in the electric signal. This built-in signal multiplication allows very low-energy particles, such as beta emissions from tritium beta decay, to be amplified and detected with high efficiency.

APDs with small areas (a few mm in diameter) can be manufactured with standard planar processing and have a gain of a few hundred. These detectors are widely used in the telecommunications industry. It is difficult to fabricate high-gain detectors with large areas using the planar process; however, special detector designs with beveled edges (see Figure 44.1) have been fabricated to provide high gain (>10,000) in large areas (>1 cm^2) [30]. The APD gain versus bias behavior for such a device is shown in Figure 44.7. Recent advances in surface preparation and dead layer reduction have extended the application of these detectors to the UV region. While they are relatively expensive, these detectors are well suited to a number of commercial applications such as medical imaging, astronomy, charged particle and x-ray detection, scintillation spectroscopy, and optical communications.

44.5.3.1 Equivalent Circuit

The diagram for the equivalent circuit of an APD is similar to that of a photodiode, shown in Figure 44.3; however, there is no component that would introduce an injection current because the diode must be reverse biased to provide gain. The contributions from the shunt resistance can increase because the reverse operating bias can be large. When considering diodes with avalanche gain, it is often useful to delineate bulk dark current, I_{db}, which undergoes impact ionization, from surface leakage, I_{ds}, which is often approximated using a linear, or ohmic, dependence on the applied bias. Equation 44.10 relates the previously defined terms for dark current from the depletion region and the ideal rectifier to the bulk and surface terms:

$$I_d = (I_{dw} + I_{ir}) + I_{ds} = I_{db} + I_{ds} \tag{44.10}$$

The expression for the total current under illumination of an APD contains an additional term for the gain, *M*, as described in the following equation:

$$I_o = M \cdot I_{db} + I_{ds} + M \cdot I_p \tag{44.11}$$

where
I_{db} is the bulk dark current
I_{ds} is the surface dark current, which can include ohmic components

For simplicity, all of the dark bulk current is assumed to experience multiplication in Equation 44.11. Although there could be gain associated with the surface components, the gain should be small, or negligible, in an ideal device.

44.5.3.2 Quantum Efficiency

The fundamental expression for the QE of a photodiode [2] applies to the APD, as follows:

$$\eta = (1 - \Re) \cdot \beta \cdot (1 - Exp(-\mu \cdot \ell)) \tag{44.12}$$

where
$\Re$ is the reflectivity
η is the charge collection efficiency
β is the absorption coefficient
ℓ is the width of the active region

The QE of an APD can be different than that of the equivalent photodiode due to the collection of charge at the higher biases needed to produce gain through impact ionization.

The QE and gain are separate properties; however, most measured quantities include the product of the QE and gain. Therefore, analysis is needed [31] to determine the gain dependence of the QE and the wavelength dependence of the gain.

44.5.3.3 Noise

In addition to the sources of noise present in photodiodes, multiplied by the gain, APDs introduce an additional term due to the noise generated by the multiplication process, defined as the excess noise, *F*. For an ideal APD, the excess noise ranges from two to higher, depending on the amount of hole multiplication, which generally increases with the gain. The shot noise current in the APD from the dark current can be written as

$$I_s = (2 \cdot q \cdot (M \cdot F \cdot I_{db} + I_{ds}) \cdot B)^{1/2} \tag{44.13}$$

Assuming the APD shunt resistance and the input resistance of the measuring circuit to be used to measure the output of the photodiode are high relative to the load resistance, the thermal noise current, I_t, is given by

$$I_t = \left(\frac{4kTBM}{R_L} \right)^{1/2} \tag{44.14}$$

where R_L is the load resistance. The total noise current is the sum of these currents in quadrature. It should be noted that these expressions include terms, such as *B* and R_L, which are specific to the measurement.

44.5.4 Geiger Photodiodes

The gain of an APD increases with increasing bias voltage, and as the bias voltage approaches the breakdown voltage, the gain goes to infinity. At the breakdown voltage, both electrons and holes have sufficient velocity to cause knock-on multiplication, which leads to a self-propagating avalanche or Geiger avalanche. At this voltage, the device effectively becomes a short circuit with the current limited by the resistance of the circuit. By adding ballast, or quenching, resistor, the GPD devices function like a Geiger–Mueller tube that discharges the junction capacitance, which determines the gain that can be 10^6–10^8. A single charge generated in the device causes breakdown, whether from a photon or thermally generated.

When properly constructed, as described by Mcntyre [32,33], it is possible to raise the bias above breakdown for macroscopic periods of time between thermally generated charges. In this overbiased state, the device becomes sensitive to single visible light photons. Such devices are APDs operated in Geiger mode, called GPDs by Cova [34,35] and [36]. Figure 44.8 shows a photograph of GPD devices designed by Vasile and Christian et al. at RMD.

44.5.4.1 Equivalent Circuit

Figure 44.9 shows a SPICE-equivalent model for a GPD element with a quenching resistor, R_q. The current source is a current pulse with an integrated charge of ~200 fC, which represents the gain, G, of the GPD.

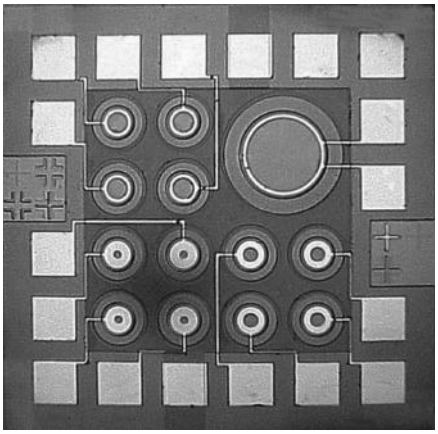
The expression for the total current under illumination of a GPD contains a term for the gain, G, as described by:

$$I_o = G \cdot I_{db} + I_{ds} + G \cdot I_p \tag{44.15}$$

where

I_{db} is the bulk dark current

I_{ds} is the surface dark current, which can include ohmic components



(a)

(b)

FIGURE 44.8 Photograph of two chips containing round GPD elements in sizes ranging from 5 μm diameter to as large as 180 μm in diameter. The chip (a) was fabricated with a custom process, *circa* 1997; the chip (b) uses a commercial CMOS process and was fabricated around 2005. The photographs are roughly to scale: the square contact pads at the perimeter of each chip are 100 μm × 100 μm.

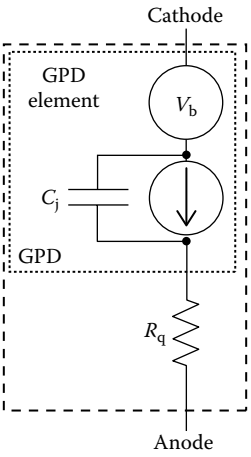


FIGURE 44.9 SPICE equivalent of a GPD element with an integrated quenching resistor, R_q. The internal voltage source subtracts the breakdown voltage from the pulse generated by the current source, which is in parallel with the junction capacitance, C_j.

Notice that like the APD, the gain minimizes the contribution of nonbulk leakage current; however, the gain associated with the GPD is appreciably larger, ~10^6, than that of an APD. The excess bias, V_x, is the bias above the breakdown voltage, V_b, as described in the following equation:

$$V_x = V_{bias} - V_b \tag{44.16}$$

where V_{bias} is the reverse bias applied to the diode. The gain, G, is the product of the junction capacitance, C_j, and the excess bias:

$$G = C_j \cdot V_x \tag{44.17}$$

The individual GPD element is effectively a binary device that detects individual optical photons. The signal associated with each Geiger pulse is only proportional to the gain of the GPD element and is not related to the number of photons that may have triggered the GPD. The current expressions assume that the photocurrent is continuously distributed in time, that is, the illumination is not pulsed.

44.5.4.2 Quantum Efficiency

The QE of a GPD is similar to that of an APD; however, the single-photon detection efficiency (DE) contains an additional term for the probability to generate a Geiger avalanche or Geiger probability, P_g:

$$\mathrm{DE} = \eta \cdot P_g \sim \eta(\lambda) \cdot P_g(V_x) \tag{44.18}$$

In Equation 44.18, the approximation refers to the dependence of the QE on the wavelength and the Geiger probability on the excess bias. GPD elements are often very thin, unlike many APDs that have a separate drift and multiplication region, and the QE and Geiger probability depend on both the wavelength and the bias. The dependence of the QE on bias arises from the bias dependence of the charge

collection and the wavelength dependence of the absorption depth. The dependence of P_g on wavelength arises from the wavelength dependence of the absorption depth and differing Geiger probabilities for electron-initiated Geiger events compared to hole-initiated Geiger events.

44.5.4.3 Noise

The GPD is essentially a digital device, and thus there is no corresponding multiplication, that is, excess, noise in an ideal GPD. Of course, there is the shot noise associated with the input referenced dark or photocurrent. The shot noise in the GPD from the dark current can be written as

$$I_s = (2 \cdot q \cdot (G \cdot I_{db} + I_{ds}) \cdot B)^{1/2} \tag{44.19}$$

Due to the very high gain provided by the GPD element, the surface leakage term, I_{ds}, can often be neglected. The expression for the thermal noise current variance, I_t, is essentially the same as that for an APD, where G replaced M for the gain variable.

44.5.5 Solid-State Photomultiplier

A solid-state photomultiplier (SSPM) is an array of APD elements operated in Geiger mode, that is, above the reverse-bias avalanche breakdown voltage, with integrated quenching components [37–39]. The elements contain a quenching component, such as a resistor, that serves the same purpose as the ballast resistor in a Geiger tube. Each element, referred to as a GPD, is a Geiger detector for optical photons, and proportional information, that is, the number of photons detected, is proportional to the number of triggered GPD elements when the array is uniformly illuminated. Important characteristics of the SSPM detectors are that they provide single photoelectron resolution, they exhibit low excess noise factors, and they can withstand illumination by room lights when biased without damage, that is, they are robust.

Single GPD elements were introduced by McIntyre [32,33] and advanced by Cova et al. [34,35] and Vasile et al. [36,40]. Buzhan et al. [37,41], who refer to SSPM as SiPMs, introduced a process where arrays of GPD elements with an integrated quenching resistor could be fabricated. Phelan et al. [42] also developed analogous sensors, which is the foundation for devices manufactured by SensL, Inc. RMD [43] and Rochas et al. [44] developed SSPMs using commercial CMOS foundry processes. We prefer to call these devices SSPMs because we are interested in generalizing the concept of an array of GPD elements to materials other than silicon. Hamamatsu has introduced the multi-pixel photon counter (MPPC), which is an equivalent device as an SSPM. Companies offering SSPM devices, or their equivalent, include Hamamatsu (Japan), Philips (Netherlands), RMD (Watertown, MA), SensL (Ireland), FBK (Italy), Photonique SA (Switzerland), and Excelitas (New York).

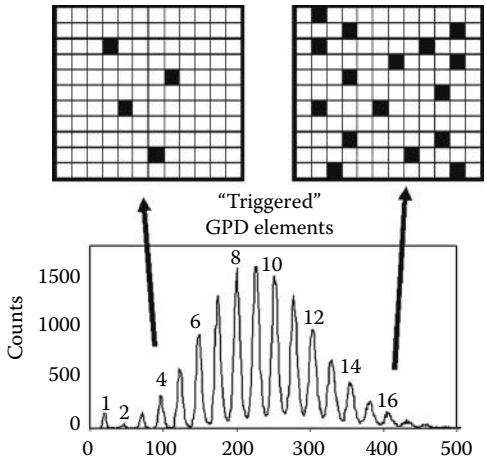
Figure 44.10a shows a photograph of 1 cm × 1 cm CMOS SSPM devices from RMD, next to a table of characteristics for the devices (Figure 44.10b), along with an illustration of the pulse-height spectra (Figure 44.10c), showing the proportionality to the number of triggered GPD elements.

SSPM detectors are particularly well suited for specialized scintillation detection applications, such as medical imaging. An example is as dual-mode imaging, where nuclear imaging is coupled to magnetic resonance imaging (MRI). In these applications, the compact size and insensitivity to magnetic field are critical performance characteristics. Images can be generated from events detected and localized in segmented scintillation detectors coupled to segmented SSPM detectors. Other specialized SSPM detector configurations, for example, position-sensitive configurations, are being developed to minimize the number of readout channels.

(a)

Suggested range of operating bias	27.5–32.5 V
DEmax	15% at 500 nm
Breakdown voltage	27.2 ± 0.2 V
Temperature sensitivity	50 mV/°C
Gain	~V_x × (10^6)
Typical dark current (output referenced)	100 μA (V_x = 2 V, 1 Quadrant)
Equivalent dark count rate (input referenced)	~350 MHz (~56 pA) (V_x = 2 V, 1 Quadrant)

(b)



(c)

FIGURE 44.10 Photograph of 1 cm × 1 cm SSPM device (a) and a table of properties (b). The illustration (c) shows the dependence of the pulse-height spectra on the number of triggered GPD elements.

44.5.5.1 Equivalent Circuit

Like the GPD description, the expression for the equivalent circuit of an SSPM contains a term for the gain, G, as described in the following equation:

$$I_o = G \cdot I_{db} + I_{ds} + G \cdot I_p \tag{44.20}$$

where

I_{db} is the bulk dark current

I_{ds} is the surface dark current, which can include ohmic components

Notice that like the APD, the gain minimizes the contribution of nonbulk leakage current. Unlike the GPD, however, the SSPM can accommodate the detection of multiple photon events when the photons are distributed over the array into distinct GPD elements.

44.5.5.2 Quantum Efficiency

The QE and DE of an SSPM is similar to that of the constituent GPD elements; however, the DE, DE_{SSPM}, contains an additional term for geometric fill factor, *ff*:

$$DE_{SSPM} = DE_{GPD} \cdot ff \tag{44.21}$$

44.5.5.3 Noise

Although the constituent GPD elements nominally do not produce gain, that is, multiplication noise, the SSPM performance can be characterized by an excess noise factor. The primary processes responsible for producing excess noise are cross talk between GPD elements and after pulsing in the GPD element. The after pulsing contribution contains a dependence on the integration time, which is specific to event detection applications. Generally, the excess noise factors of SSPM devices are equivalent to that of PMTs but depend on the excess bias. The characterization of the excess noise for SSPM devices is described in the literature [45–48]. After pulsing and cross talk trigger GPD elements in addition to an incident stimulus and can thus be described as gain or multiplication terms. The fluctuations in the amplification terms associated with cross talk and after pulsing produce gain noise, that is, excess noise.

The shot noise in the SSPM from the dark current can be written as shown in the following equation:

$$I_s = (2 \cdot q \cdot (G \cdot F \cdot I_{db} + I_{ds}) \cdot B)^{1/2} \tag{44.22}$$

In applications detecting pulses of photons, such as the detection of gamma-ray events by a scintillation detector, the noise not only broadens the peaks in the histogram of pulse heights, that is, the gamma-ray spectrum, but also introduces a noise floor that produces a lower limit to the detected energy of the gamma ray or amplitude of the scintillation pulse. In these applications, the dark current represents an uncorrelated source of single-photon events, unlike the detection of the scintillation pulse associated with the detection of a gamma-ray event.

For the detection of scintillation pulses, the dark counts will produce an amplitude spectrum, which defines the noise floor of the SSPM. In the absence of cross talk and after pulsing, a Poisson distribution (see Equation 44.23) characterizes this idealized noise floor when operating a GPD array as an SSPM [45]:

$$P(n) = \frac{(I_d \tau)^{n-1}}{(n-1)!} e^{-I_d \tau} \tag{44.23}$$

$P(n)$ describes the probability for generating an n-photon event from the thermally generated dark counts in the SSPM. The term I_d denotes the dark count rate of the SSPM, which is the sum of the dark count rate for each of the individual pixels, and τ represents the integration time of the amplifier. In this expression, the dark count rate refers to the input referenced bulk dark current, that is, the dark current divided by the gain.

In pulse detection applications, the width of the zero-amplitude peak defines the noise floor, which should be convolved with the features associated with the detected events. The relative amplitude of the noise floor peak, however, is related to the maximum detection rate, or bandwidth, of the readout electronics and multichannel analyzer that histograms the pulse amplitudes.

44.6 Photodiode Arrays

Photodiode arrays consist of multiple photodiode elements, which can be in the form of a linear, square, or rectangular array. CMOS imaging sensors and SSPMs are essentially arrays of photodiodes. Photodiode arrays are used in applications where CCDs are not suitable because of sensitivity or speed requirements and where very high spatial resolution is not as critical. They are frequently the detector of choice in medical x-ray and gamma-ray imaging applications, such as CT, single-photon emission tomography (SPECT), and positron-emission tomography (PET). Arrays made from silicon photodiodes, APDs, and GPDs are available for response to specific wavelengths. Photodiodes made in other materials are also available for response to specific wavelengths based on the bandgap of the material. Nominally, CMOS active pixel sensor cameras [49,50] can be considered as an array of individually addressable photodiodes.

44.6.1 APD Arrays

A number of researchers have investigated the use of arrays of APD for use in medical imaging. Figure 44.11 shows a monolithic, 8 × 8-element APD array. APDs have the advantage of being insensitive to magnetic fields and are being investigated for use in dual-modality PET/MRI and SPECT MRI medical imaging cameras [51–57]. The ability to operate in magnetic fields also makes these photosensors attractive for nuclear physics and high-energy physics applications.

The internal gain of the APD reduces the effect of the readout noise on the SNR. The gain provided by the monolithic array shown in Figure 44.11 exhibits a low excess noise factor compared to small, shallow CMOS APD structures. The excess noise factor, F, degrades the signal-to-noise performance as described in the following:

$$\text{SNR} \propto \frac{\text{QE}}{F} \tag{44.24}$$

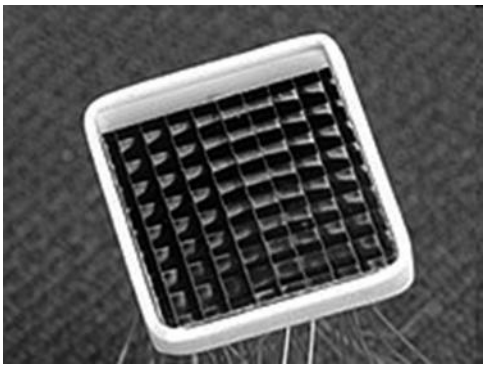
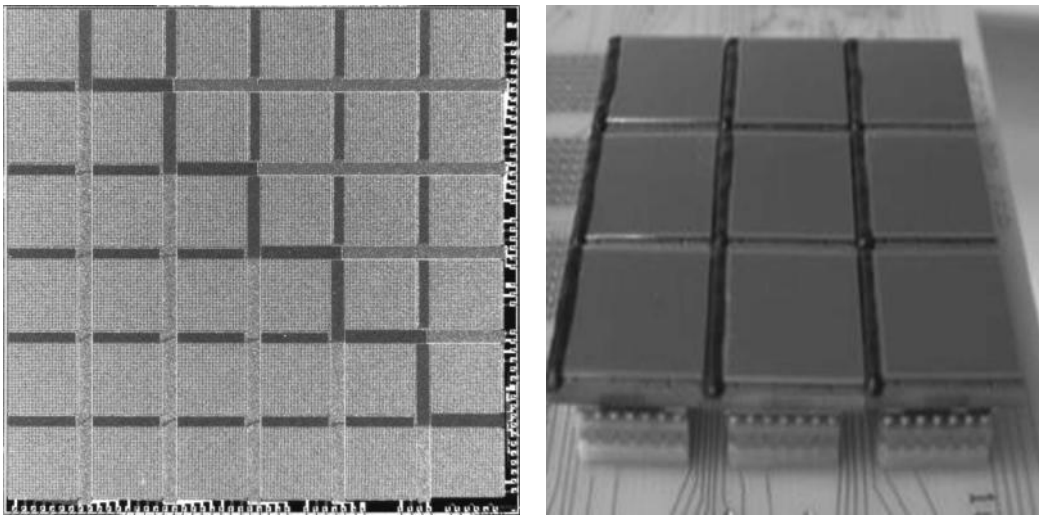


FIGURE 44.11 Monolithic 8 × 8 element APD array. The size of each individual element in the array is approximately 1 mm × 1 mm.



(a) (b)

FIGURE 44.12 (a) Monolithic 6 × 6 array of SSPM detectors. The size of each SSPM in the array is 1.5 × 1.5 mm, and the size of the detector is 11 mm × 11 mm. (b) Segmented SSPM array formed from a 3 × 3 array of 2 × 2 SSPM elements. Each chip in the 3 × 3 array is 11 mm × 11 mm and comprised of a 2 × 2 array.

44.6.2 SSPM Arrays

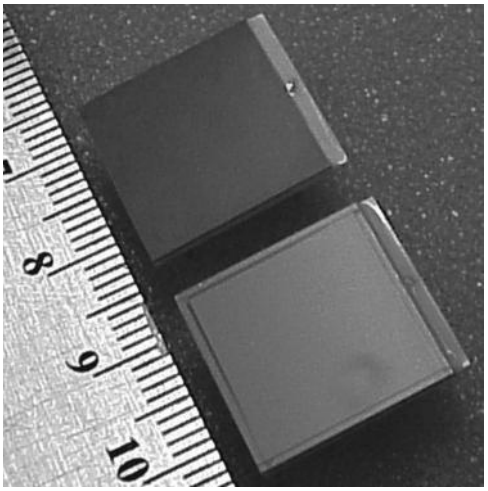
SSPM devices are inherently an array of GPD elements. Monolithic and discrete arrays of SSPM devices have been developed for nuclear imaging applications using segmented scintillation detectors. Figure 44.12a shows a 6 × 6 monolithic SSPM array chip, which is 11 mm × 11 mm; Figure 44.12b shows a 6 × 6 SSPM of discrete 3 × 3 chips, where each chip is a 2 × 2 SSPM array.

44.6.3 Position-Sensitive APDs

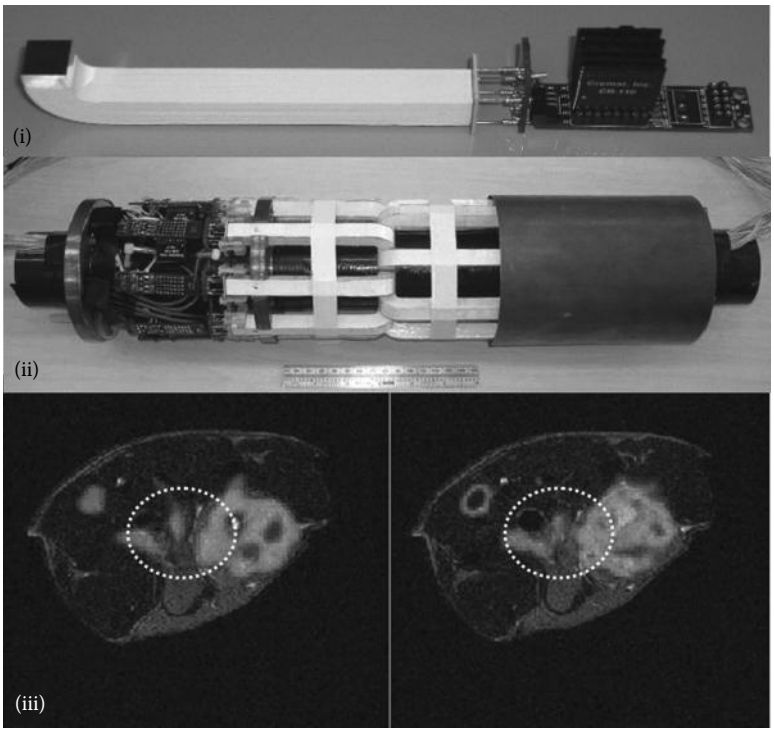
Traditional gamma cameras used in nuclear imaging applications generate images with the signals from multiple detectors connected by a resistive network using Anger logic. Position-sensitive photomultiplier tubes (PMTs) distribute the detected charge pulse into a resistor network that divides the signal into four or more readout contacts. The spatial location of the scintillation event on the photocathode is determined by the analysis of the relative signal amplitudes at the four contacts. Analogous to a position-sensitive PMT, the charge from an APD can be distributed by the sheet resistance of the silicon to divide the charge into four contacts, producing a position-sensitive APD device. These devices are well suited for nuclear imaging applications that involve strong magnetic fields, such as PET and SPECT imaging coupled with MRI. Figure 44.13a, shows two position-sensitive APD detectors; Figure 44.13b shows the detector module constructed with the devices, along with an image generated with the detector module.

44.6.4 Position-Sensitive SSPM

The fabrication of SSPM in standard CMOS processes facilitates the integration of components for specialized configurations. For example, a charge-dividing resistive network produces a position-sensitive SSPM (PS-SSPM) for nuclear imaging applications [38,45,58] analogous to the charge-division network used in position-sensitive PMTs [59,60]. In nuclear imaging applications, the location of the event, as well as the energy of the event, must be resolved. In an Anger camera, the scintillation light is distributed over a number of PMT detectors. In some incarnations, the signal from the PMT detectors is connected to a charge-dividing resistive network. The ratio of the signal at the four contacts of the resistive network yields the location of the scintillation event, while the sum yields the energy.



(a)



(b)

FIGURE 44.13 (a) Monolithic position-sensitive APD device. (b) PET nuclear imaging module based on PS-APD detectors.

In a PS-SSPM, the ratio of the signal at four readout contacts determines the location of the event [61], and the sum of the signals determines the energy of the event. Figure 44.14a shows a photograph of a segmented LYSO scintillation array, next to a photograph of a 5 mm × 5 mm PS-SSPM (Figure 44.14b). When coupled to the PS-SSPM, the segmented scintillation detector produced the image shown in Figure 44.14c when irradiated with a flood field of gamma rays from a 22-Na source.

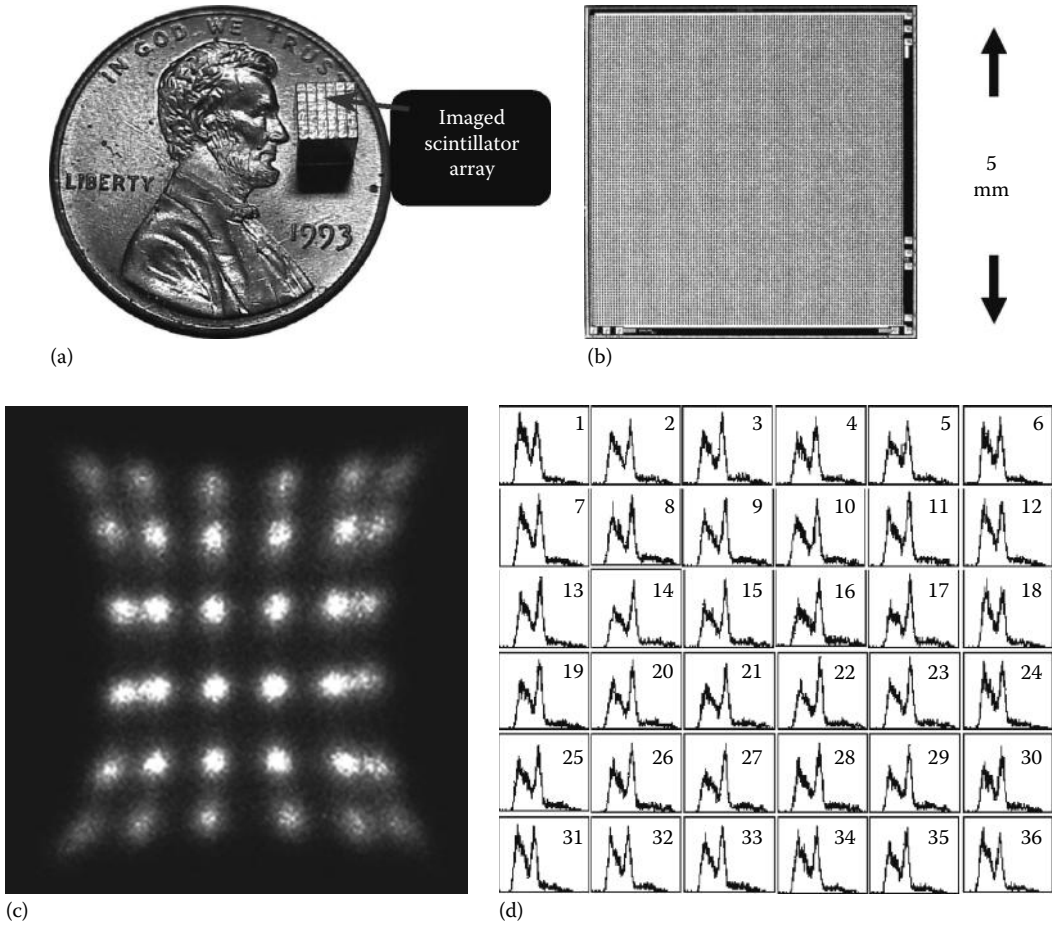


FIGURE 44.14 (a) Segmented LYSO scintillation array, (b) 5 mm × 5 mm PS-SSPM, (c) image when irradiated with a flood field of gamma rays from a 22-Na source, (d) energy spectrum for each of the scintillation segments.

As seen in the image, the location of each of the scintillation segment is cleanly resolved. Figure 44.14d shows the energy spectrum obtained for each of the scintillation segments resolved in the image.

44.6.5 Amorphous Silicon Detectors

While impressive results have been obtained with various device structures on crystalline silicon such as drift detectors, APDs, and CCDs, they are limited to active areas of only a few square centimeters. As a result, considerable attention has been devoted to the development of hydrogenated amorphous silicon (a-Si:H) [11]. This material is produced by an RF plasma technique in large areas (30 cm × 30 cm) on glass substrates. The films are typically a few micrometers in thickness, although films as thick as 200 μm have been reported. Device structures such as *p-n* junctions were developed initially for use in solar cells with lower cost than crystalline silicon devices.

Recently, more complicated devices such as *p-i-n* sensors and TFTs have been fabricated from a-Si:H and have been configured in an array format as shown in Figure 44.15. In these arrays (as large as 20 cm × 20 cm), each pixel consists (200–500 μm) of a *p-i-n* sensor connected to a TFT, and the entire array is read out in matrix fashion. These arrays are well suited for high-resolution document imaging and also for medical x-ray imaging applications with phosphors.

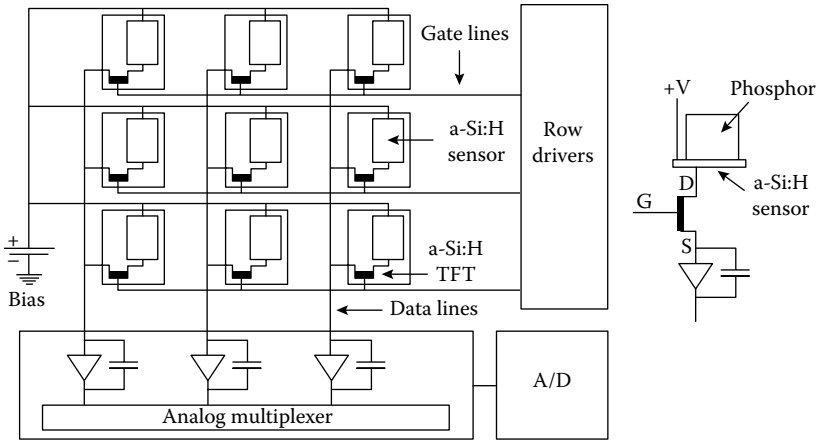


FIGURE 44.15 Schematic representation of a 2D x-ray imager consisting of a-Si:H *p-i-n* diodes connected to a-Si:H TFTs for readout. The *p-i-n* diodes are coupled to a phosphor layer to increase sensitivity.

44.7 Other Materials for Photojunction Detectors

44.7.1 Novel Materials for Photodiodes

Applications involving the detection of infrared (IR) light require the use of detector materials with small bandgaps, <1 eV, compared to silicon. Although silicon is by far the most common material used for semiconductor photodetectors, silicon cannot be used for making IR sensors because the long wavelength light is not absorbed by the silicon. To address near-IR sensing, III–V semiconductors are used for making photosensors and photosensor arrays. Table 44.2 lists materials, and their properties, used for the construction of photodiodes designed to operate at room temperature.

Applications such as scintillation detection with certain emerging UV-emitting scintillation material require the detection of UV light in addition to visible light. For these applications, a material with a wide bandgap, >1 eV, may provide lower dark current enabling an increase in the area of the device. Although the bandgap energy certainly affects the dark current, the quality of the material may be a limiting factor in the dark current performance, especially for nonsilicon devices. This is important because the sensitivity of many scintillation-based nuclear detectors scales with the size of the detector, and the ability of optics to concentrate the spatially distributed emission is limited. Table 44.3 lists materials that are ideal for detecting 380 nm scintillation light.

As stated in Table 44.3, GaP is an indirect bandgap material with bandgap energy of approximately 2.26 eV. The maximum photo absorption wavelength is around 550 nm based on the calculation from the bandgap energy. $GaAs_{1-x}P_x$ is an indirect bandgap material when the phosphorus mole fraction of the compound reaches beyond 50% ($x > 0.5$). At this condition, the bandgap is around 2.12–2.33 eV. Similar to GaP, the high impact ionization ratio and dark noise of GaAsP make it less ideal for GPD development in the project.

Important requirements for photodiodes include high QE, good charge collection efficiency, and low noise. The low-noise requirement is satisfied by reducing detector capacitance and its dark current. In order to satisfy the QE requirements over a wide range of wavelengths, new semiconductor materials are being extensively investigated [78]. Since most semiconductors show high optical response near their bandgap, special materials are developed for various applications. Furthermore, since the bandgap represents a cut-off point in the optical response of the material, by selecting an appropriate material, it is possible to obtain response in a desired band [79]. For example, materials such as GaN and SiC are being explored to obtain UV detection with no sensitivity in the visible region. Other materials are being studied to exploit their unique properties such as high QE, high-temperature operation, and high-speed response.

TABLE 44.2 Properties of Semiconductor Materials Used for Construction of Photodiodes at 25°C

Material	Bandgap	Dielectric Constant	Resistivity (Ω-cm)	Electron Mobility (cm^2/V s)	Electron Lifetime (s)	Hole Mobility (cm^2/V s)	Hole Lifetime (s)	μ_τ (e) (cm^2/V)	μ_τ (h) (cm^2/V)
HgTe	0.14	6.4		22,000		100			
InAs	0.36	12.5	30,000	240					
Ge	0.67	16	50	3,900	$>10^{-3}$	1900	1×10^{-3}	>1	>1
Si	1.12	11.7	$\leq 10^4$	1,400	$>10^{-3}$	480	2×10^{-3}	>1	=1
InP	1.35	12.5	10^7	4,600	1.5×10^{-9}	150	$<10^{-7}$	4.8×10^{-6}	$<1.5 \times 10^{-5}$
GaAs	1.43	12.8	10^7	8,000	10^{-8}	400	10^{-7}	8×10^{-5}	4×10^{-6}
CdSe	1.73	10.6	10^8	720	10^{-6}	75	10^{-6}	7.2×10^{-4}	7.5×10^{-5}
a-Si	1.8	11.7	10^{12}	1	6.8×10^{-9}	.005	4×10^{-6}	6.8×10^{-8}	2×10^{-8}
InI	2.01	26	10^{11}					7×10^{-5}	
HgI_2	2.13	8.8	10^{13}	100	10^{-6}	4	10^{-5}	10^{-4}	4×10^{-5}
SiC	2.2								
T1BrI	2.2–2.8		10^{10}					9×10^{-5}	
GaP	2.24								
a-Se	2.3	6.6	10^{12}	.005	10^{-6}	0.14	10^{-6}	5×10^{-9}	1.4×10^{-7}
Pbl_2	2.32		10^{12}	8	10^{-6}	2		8×10^{-6}	
CdS	2.5	11.6		300		50			
TIBr	2.68	29.8	10^{12}	6	2.5×10^{-6}		1.6×10^{-5}	1.5×10^{-6}	
GaN	3.4	12	$>10^{10}$	300–1,000					
Diamond	5.4	5.5		2,000	10^{-8}	1600	$<10^{-8}$	2×10^{-5}	$<1.6 \times 10^{-5}$

TABLE 44.3 List of Candidate Wide-Bandgap Materials for SSPM Development

Wide-Bandgap Material	Bandgap Type	Bandgap Energy (eV)	Max Absorption (Based on BGE) (nm)	Dark Noise	Impact Ionization Ratio (k)
GaP [62,63]	Indirect	~2.26	~550	Similar to Si	Close to 1
$GaAs_{1-x}P_x$ [64,65]	Indirect	~2.12–2.33	~532	Similar to Si	Close to 1
SiC-3C [66,67]	Indirect	~2.3	~540	Lower than Si	n/a
SiC-4H [68–70]	Indirect	~3.23	~384	Lower than Si	<0.1
GaN [71,72]	Direct	3.4	~365	Very high	<0.1
$Al_xGa_{1-x}N$ [73,74]	Direct	$3.43 + 1.44x + 1.33x^2$	<250	Very high	<0.1
$Al_xGa_{1-x}As$ [75–77]	Indirect for $x > 0.4$	~1.42–2.16	~574	Lower than Si	<0.1

44.7.2 GaN

GaN is an attractive material for *UV photodiode* fabrication due to its wide bandgap ($E = 3.4$ eV). Due to the difficulty in growing bulk crystals of GaN, much of the work is done with films of GaN prepared by chemical vapor deposition or molecular beam epitaxy [80]. A variety of optical devices (e.g., blue LEDs and lasers, field effect transistors, photoconductive detectors, and photodiodes) have been fabricated using GaN films. GaN photodiodes have the capability of solar blind UV detection and are capable of fast response time due to high electron mobility (as high as 1000 cm^2/V s), which is comparable to silicon. Even though GaN has a low impact ionization ratio, it has extremely high dark noise and suffers from microplasma-related quality issues. Direct bandgap materials, like GaN, also pose

challenges associated with the generation of light when used as detector materials. AlGaN alloy suffers similar performance and quality issues.

44.7.3 SiC

SiC is another material that has shown promise for UV detection due to its wide bandgap (E_g = 3.0 and 3.2 eV for 6H and 4H phases, respectively) [81]. Silicon carbide has long been hailed as the future replacement of silicon in the microelectronics and optoelectronics industries. Its distinct characteristics—wide bandgap, low dark noise, and low impact ionization ratio—make it a suitable candidate for the development of next-generation photodetectors. Two polytypes of SiC are of high interest. 3C-SiC has a bandgap of 2.3 eV, which is ideal for detecting photons in the 200–500 nm range. However, the present quality of starting substrate material for 3C-SiC is lacking compared to its 4H sibling, and the most recent studies in literature [66] have shown comparably lower optical performance (shown in Figure 44.16). The recent improvements in 4H-SiC material quality and processing maturity have accelerated the development of electronic and optoelectronic devices based on this polytype. 4H-SiC has an indirect bandgap around 3.23 eV. This makes the material sensitive to photons in the UV and near-UV range (220–380 nm); however, the sensitivity drops significantly beyond 380 nm, which would raise concerns for scintillation lights in the shallow UV/blue region.

44.7.4 InI

Indium iodide (InI) is a wide-bandgap semiconductor (E_g = 2.0 eV) being developed for detection in visible and near-UV regions [82]. The resistivity of the material is quite high (>10^{10} Ω-cm), and Schottky diodes are fabricated using evaporated palladium electrodes. Because of the high resistivity, it is possible to deplete relatively large thicknesses (0.5–1 mm) at low bias (<200 V). One of the unique properties of InI photodiodes is their high QE (>70%) in the 300–600 nm wavelength region, as shown in Figure 44.17, which in combination with its low dark current makes it attractive for low light level detection applications such as scintillation spectroscopy.

44.7.5 Alloys and Bandgap Engineering

In many instances, required properties are attained by bandgap engineering where two or more semiconductors are alloyed together to create a *ternary semiconductor.* The use of a ternary semiconductor

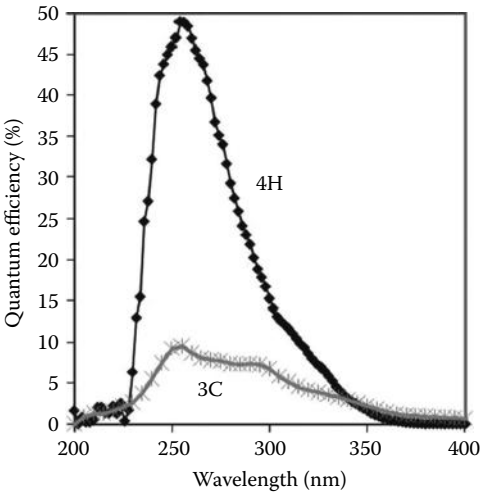


FIGURE 44.16 QE comparison between 4H-SiC and 3C-SiC avalanche diodes. (From Rowland, L.B. et al., Cubic silicon carbide avalanche photodiodes, Presented at *2009 International Semiconductor Device Research Symposium, ISDRS 2009.*)

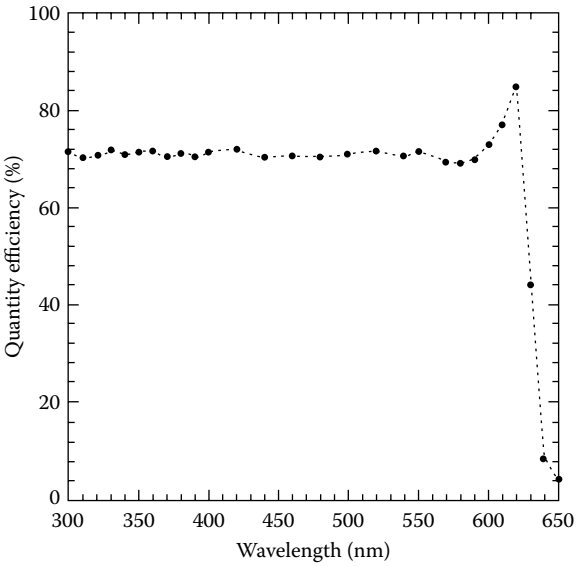


FIGURE 44.17 QE of an InI photodiode. The QE peaks at over 80% near the band edge and has a spectral sensitivity of about 70% into the near UV.

provides the ability to tune the peak wavelength of a photodiode. When two binary compounds are combined, the resulting ternary material usually has properties that lie between those of the constituent binary compounds. Examples of some alloys developed in this manner include $Hg_xCd_{1-x}Te$ for IR detection, created from HgTe (E_g = 0.14 eV) and CdTe (E_g = 1.45 eV); Si_xGe_{1-x}, which is a mixture of Si (E_g = 1.1 eV) and Ge (E_g = 0.7 eV) for IR and visible detection; and $TlBr_xI_{1-x}$, which is a mixture of TIBr (E_g = 2.7 eV) and T1I (E_g = 2.1 eV) for visible detection. A discussion of some recent novel materials being developed is presented in the following section, and a compilation of relevant properties of various semiconductor materials is presented in Table 44.2.

44.7.6 III–V Ternary Materials

Ternary alloys of GaN and AlN (E_g = 6.2 eV), $Ga_xAl_{1-x}N$, are also being investigated to create optimized UV detectors for desired wavelengths. In the ternary compound, the bandgap depends on the material composition, or x, and varies almost linearly from 3.4 to 6.2 eV. Such bandgap engineering is desirable to create material with required photoresponse. These devices are also expected to be capable of high-temperature operation due to the wide semiconducting bandgap of the material. Similar devices have also been studied, such as GaP (E_g = 2.1 eV) and A1P (E_g = 2.9 eV), for visible and near-UV detection.

Another example is indium gallium arsenide, which is a mixture of InAs (E_g = 0.36 eV) and GaAs (E_g = 1.43 eV) and has been recently commercialized as an IR detector material in the 1000–1700 nm region. InGaAs photodiodes in *p-n* diode, *p-i-n* diode, and APD configurations are available. InGaAs photodiode arrays coupled to amorphous silicon TFTs are being developed for large-area IR imaging. Other ternary III–V materials that have been investigated for similar reasons include GaAsP, GaNP, and BNP.

$Al_xGa_{1-x}As$ is an indirect bandgap material when the mole fraction of Al reaches 80%. The bandgap energy for such an arrangement is approximately 2.16 eV. This is suitable especially for photon detection in the shallow UV region. The impact ionization ratio is also low (<0.1), which indicates low excess noise associated with impact ionization. AlGaAs is epitaxially grown from GaAs wafers, and since GaAs material and processing is a mature technology, it is very suitable for developing high-quality photodetectors.

For the shallow silicon devices, the trade-off between the reflectance and the absorption depth as a function of wavelength mimics the QE curve. Specifically, at low wavelengths, that is, blue to UV light, the reflectance increases, one minus the reflectance decreases, and the collection of photocurrent decreases because of a "dead layer" at the surface, where charge is inefficiently collected when the absorption coefficient is too large. At long wavelengths, that is, red light, the DE decreases because the light passes through the active region, as predicted by the decrease in the absorption coefficient. The reflectance and absorption coefficient are important in determining the expected DE of the detector. Figure 44.18 shows plots of one minus the reflectance, on the left axis, against the absorption coefficient, in μm^{-1} on the right axis as a function of wavelength for silicon [83], $Al_xGa_{1-x}As$ [84],

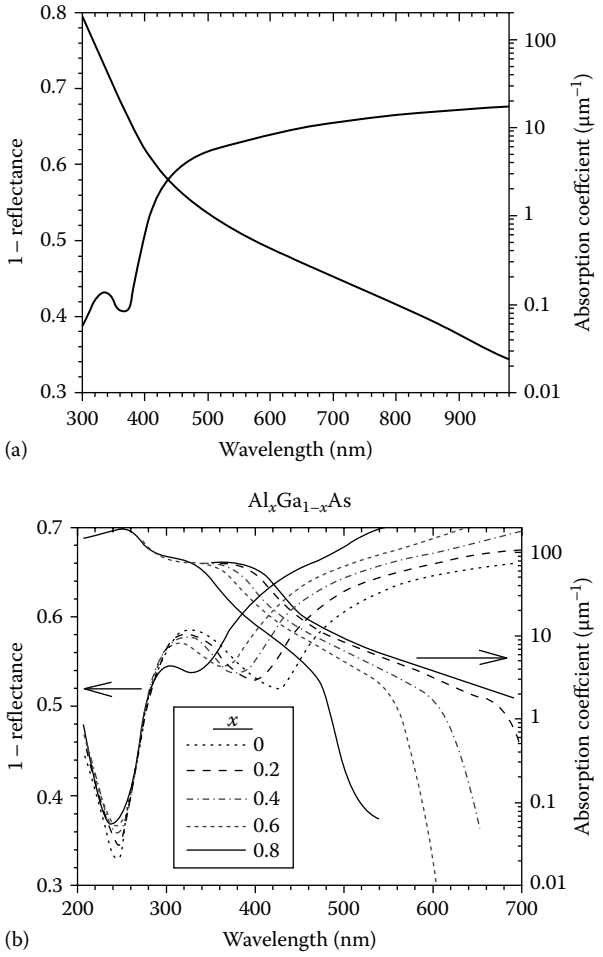


FIGURE 44.18 Plot of 1 – reflectance, which should be roughly proportional to the light that can be absorbed, against the absorption coefficient, in μm^{-1}, which should be roughly proportional to the light that can be absorbed, for (a) silicon, (b) Si, $Al_xGa_{1-x}As$, respectively. These two factors often contribute to the wavelength dependence of the QE. At short wavelengths, the QE is expected to decrease because the reflectance increases, and the collection of the photocurrent from the surface decreases, as indicated by the large absorption coefficients. At long wavelengths, the QE is expected to decrease because the absorption length decreases, so the light passes through the material. The purpose of these plots is to examine anticipated correlations to the QE and to estimate the penetration depth, which corresponds to the maximum thickness of any "dead layer" at the surface for the wavelengths of interest.

(*continued*)

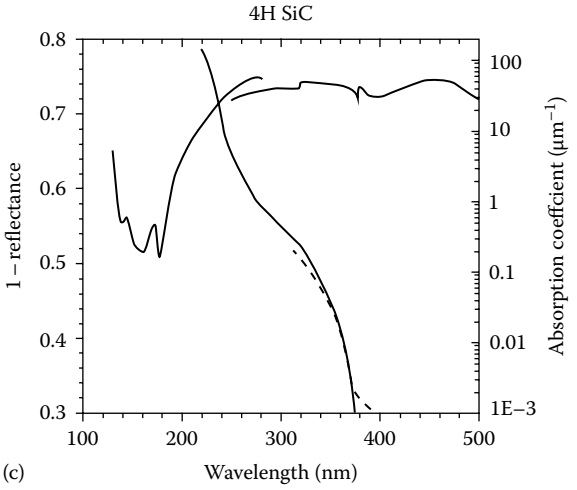


FIGURE 44.18 (continued) Plot of 1 – reflectance, which should be roughly proportional to the light that can be absorbed, against the absorption coefficient, in μm^{-1}, which should be roughly proportional to the light that can be absorbed, for (c) 4H SiC, respectively. These two factors often contribute to the wavelength dependence of the QE. At short wavelengths, the QE is expected to decrease because the reflectance increases, and the collection of the photocurrent from the surface decreases, as indicated by the large absorption coefficients. At long wavelengths, the QE is expected to decrease because the absorption length decreases, so the light passes through the material. The purpose of these plots is to examine anticipated correlations to the QE and to estimate the penetration depth, which corresponds to the maximum thickness of any "dead layer" at the surface for the wavelengths of interest.

and SiC (4H) [85–87]. Note that the reflectance is plotted on a linear scale, whereas the absorption coefficient is plotted on a log scale.

As seen in the plot for GaAs, which is $Al_xGa_{1-x}As$ for $x = 0$, the absorption depths at 320 and 380 nm are 75 and 72 μm, respectively. This suggests that a "dead layer" of more than 10 nm would have a significant impact on reducing the DE for these wavelengths. The absorption depths at 320 and 380 nm for $Al_{0.8}Ga_{0.2}As$ are 80 and 20 μm, respectively. This suggests that a "dead layer" of 10 nm would effectively block 320 nm, but some 380 nm light would pass through. The values for the absorption depths for other compositions are between those for $x = 0$ and $x = 0.8$.

The plot for the combination of 1 – reflectance and the absorption depth mimics the QE plot shown in Figure 44.16. The absorption depth is 0.2 μm at 320 nm. At 380 nm and other wavelengths near and in the visible region, the absorption depth is smaller, suggesting that the material is relatively transparent in this region. For SiC (4H), the effect of the "dead layer" at 320 nm is expected to be insignificant for layer thicknesses of <5 μm.

44.7.7 Heterojunction Photojunction Detectors

A heterojunction exists at the interface of two different semiconductors. This concept can be exploited to produce photodiodes with unique properties such as tuned optical response in the region of interest (by adjusting the composition) and reduced optical absorption at the entrance (by irradiating the wider bandgap semiconductor that is transparent to the optical signal). A number of optical sensors have been fabricated using this concept with III–V compounds that can be tuned in composition to create heterojunctions with similar lattice constants in both the semiconductors. The research in the heterojunction devices has been aided considerably by the progress in molecular beam epitaxy.

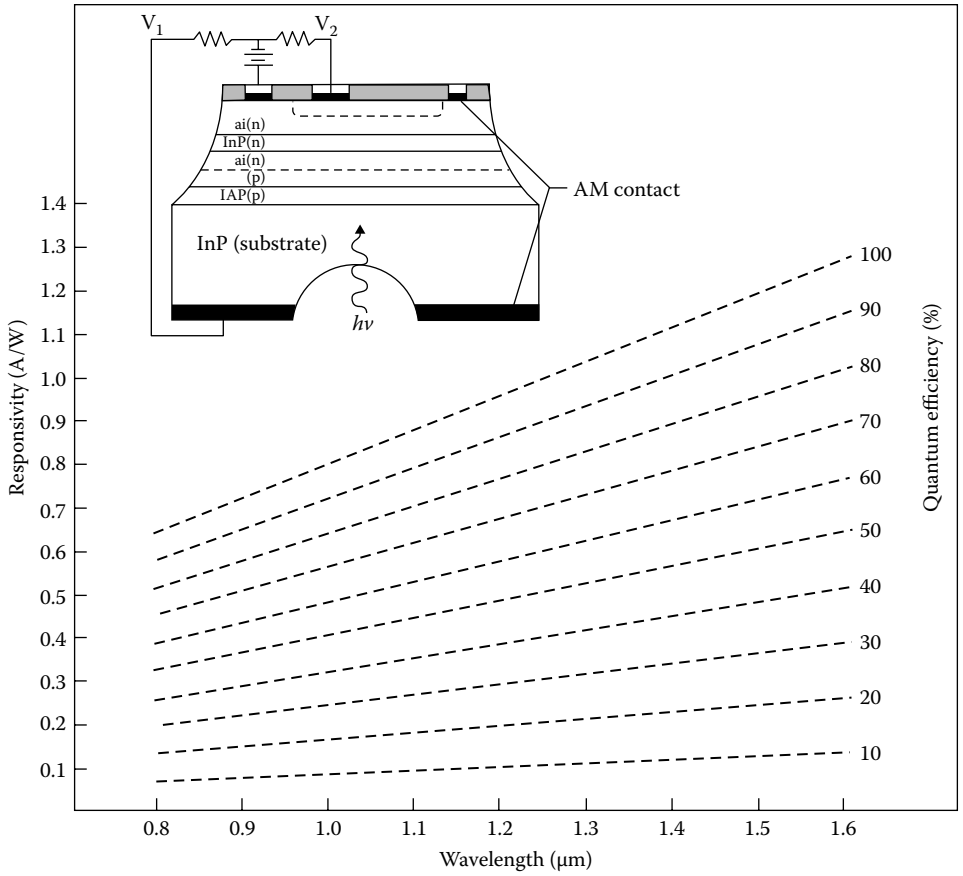


FIGURE 44.19 Responsivity and QE of a heterojunction photodiode versus wavelength. The insert shows the cross section of the photodiode. (From Sze, S.M., *Physics of Semiconductors*, 2nd edn., John Wiley & Sons, New York, 1981, p. 765.)

One unique application of the heterojunction concept is to fabricate detectors that are capable of distinguishing wavelengths above or below a certain level. This has been accomplished using a multilayer device (see Figure 44.19) that consists of two layers of $Ga_xIn_{1-x}As_yP_{1-y}$, which have different compositions and, therefore, different bandgaps. The layer Q1 has a larger bandgap than Q2, and both are grown on InP. The optical response of this device when irradiated through the InP substrate is shown in Figure 44.19 and shows minimal overlap in the desired bands indicating successful wavelength discrimination.

Defining Terms

Bandwidth, *B*: The range of frequencies over which the photodiode operates.

Breakdown voltage, V_b: The reverse-bias voltage at which the applied field overcomes the ability of the junction to block current and the device acts like a resistor. The reverse leakage current increases abruptly near this voltage.

Dark current or reverse leakage current, I_d: The leakage current through the device when at the operating voltage with no incident signal.

Depletion region thickness: The depth of the depleted portion of the diode when at the operating voltage. Photodiodes are frequently operated fully depleted.

Junction capacitance, C_j: Capacitance of the photodiode that decreases as the depletion width increases.

Noise equivalent power, NEP: The incident power that generates a signal equal to the noise, that is, S/N or SNR equals one.

Operating bias: The applied voltage at which the device operates.

Peak wavelength: The wavelength with the highest QE.

Quantum efficiency, η, or QE: The efficiency of converting photon incident on the photodiode into electrons that are detected. Reflection of light from the surface and loss of electrons in the semiconductor reduce the efficiency. Reflection losses can be minimized using an antireflection coating on the surface of the device.

Responsivity, R (A/W): A measure of the signal current produced as a function of the optical power incident on the photodiode.

Spectral response: The QE as a function of wavelength.

References

1. S. M. Sze, *Physics of Semiconductors*, 2nd edn. New York: John Wiley & Sons, 1981.
2. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. New York: John Wiley & Sons, 1991.
3. *Laser Focus World Buyer's Guide*. Nashua, NH: Pennwill Publishing Co., 2011.
4. *Photonics Buyer's Guide*. Pittsfield, MA: Laurin Publishing Co., 2010.
5. J. L. Moll, The evolution of the theory for voltage-current characteristic of *P-N* junctions, *Proc. IRE*, 46, 1076, 1958.
6. W. Shokley, The theory of *p-n* junctions in semiconductors and *p-n* junction transistors, *Bell Syst. Tech. J.*, 28, 435, 1949.
7. W. Shokley and W. T. Read, Statistics of the recombination of holes and electrons, *Phys. Rev.*, 87, 835, 1952.
8. R. N. Hall, Electron-hole recombination in germanium, *Phys. Rev.*, 87, 387, 1952.
9. B. V. Zeghbroeck, *Principles of Electronic Devices*, 2007. Available at http://ecee.colorado.edu/~bart/book/book/contents.htm, accessed on December 2011.
10. Radiation Monitoring Devices, Inc., RMD quadrant array specification sheet, Vol. 2011, 2009. http://www.rmdinc.com/, accessed on September 6, 2013.
11. R. A. Street, Amorphous silicon sensor arrays for radiation imaging, *MRS Proc.*, 192, 441, 1990.
12. R. L. Weisfield, R. B. Apte, S. Ready, A. Moore, M. Nguyen, W. B. Jackson, and P. Nylen, Amorphous silicon sensor array for X-ray and document imaging, *Thin Solid Films*, 296, 172–176, 1997.
13. R. A. Street, S. E. Ready, J. T. Rahn, M. Mulato, K. S. Shah, P. R. Bennett, P. Mei et al., High-resolution direct-detection X-ray imagers, *Proc. SPIE*, San Diego, CA, 2000.
14. R. A. Street, K. S. Shah, S. E. Ready, R. B. Apte, P. R. Bennett, M. Klugerman, and Y. N. Dmitriev, Large-area X-ray image sensing using a PbI_2 photoconductor, *Proc. SPIE*, San Diego, CA, 1998.
15. R. A. Street, J. T. Rahn, S. E. Ready, K. S. Shah, P. R. Bennett, Y. N. Dmitriev, P. Mei et al., X-ray imaging using lead iodide as a semiconductor detector, *Proc. SPIE*, San Diego, CA, 1999.
16. E. Gatti and P. Rehak, Semiconductor drift chamber—An application of novel charge transport scheme, *Nucl. Instrum. Methods Phys. Res. A*, 225, 608–614, 1984.
17. E. Gatti and P. Rehak, Review of semiconductor drift detectors, *Nucl. Instrum. Methods. A*, 541, 47–60, 2005.
18. J. Ulrici, S. Adler, P. Buchholz, P. Fischer, P. Klein, M. Locker, G. Lutz, W. Neeser, L. Struder, M. Trimpl, and N. Wermes, Spectroscopic and imaging performance of DEPFET pixel sensors, *Nucl. Instrum. Methods. Phys. Res. A*, 465, 247–252, 2001.

19. M. Porro, G. Ferrari, P. Fischer, O. Halker, M. Harter, S. Herrmann, N. Hornel et al., Spectroscopic performance of the DePMOS detector/amplifier device with respect to different filtering techniques and operating conditions, *IEEE Trans. Nucl. Sci*., 53, 401–408, 2006.
20. R. J. McIntyre, Multiplication noise in uniform avalanche diodes, *IEEE Trans. Electron Dev.*, 13, 164–168, 1966.
21. S. Barber, Photon counting with avalanche photodiodes, *Electron. Eng*., 56, 63–70, 1984.
22. R. G. W. Brown, K. D. Ridley, and J. G. Rarity, Characterization of silicon avalanche photodiodes for photon correlation measurements. 1: Passive quenching, *Appl. Opt*., 25, 4122–4126, 1986.
23. R. G. W. Brown, R. Jones, J. G. Rarity, and K. D. Ridley, Characterization of silicon avalanche photodiodes for photon correlation measurements. 2: Active quenching, *Appl. Opt*., 26, 2383–2389, 1987.
24. M. Ghioni and G. Ripamonti, Improving the performance of commercially available Geiger-mode avalanche photodiodes, *Rev. Sci. Instrum*., 62, 163, 1991.
25. G. C. Huth, Recent results obtained with high field, internally amplifying semiconductor radiation detectors, *IEEE Trans. Nucl. Sci*., 13, 36–42, 1966.
26. G. Reiff, M. R. Squillante, H. B. Serreze, G. Entine, and G. C. Huth, Large area silicon avalanche photodiodes: Photomultiplier tube alternative, *Nuclear Radiation Detector Materials Symposium*, Boston, MA, pp. 131–140, 1982.
27. M. R. Squillante, R. Farrell, J. C. Lund, F. Sinclair, G. Entine, and K. R. Keller, Avalanche diode low energy X-ray and nuclear particle detector, *IEEE Trans. Nucl. Sci*., NS-33, 336–339, 1986.
28. M. R. Squillante, J. S. Gordon, R. Farrell, S. Vasile, K. Daley, C. Oakes, and K. Vanderpuye, Recent advances in avalanche photodiode technology, *Proc. SPIE Int. Soc. Opt. Eng*., 2009, 64–71, 1993.
29. Radiation Monitoring Devices, Inc., Radiation monitoring devices application note on Si APDs, http://www.rmdinc.com/, accessed on September 6, 2013.
30. R. Farrell, K. Vanderpuye, L. Cirignano, M. R. Squillante, and G. Entine, Radiation detection performance of very high gain avalanche photodiodes, *Nucl. Instrum. Methods. A*, 353, 176–179, 1994.
31. M. McClish, R. Farrell, K. Vanderpuye, and K. S. Shah, A reexamination of deep diffused silicon avalanche photodiode gain and quantum efficiency, *IEEE Trans. Nucl. Sci*., 53, 3049–3054, 2006.
32. R. J. McIntyre, On the avalanche initiation probability of avalanche diodes above the breakdown voltage, *IEEE Trans. Electron Dev*., ED20, 637–641, 1973.
33. R. J. McIntyre, Recent developments in silicon avalanche photodiodes, *Measurement*, 3, 146–152, 1985.
34. S. Cova, A. Longoni, and A. Andreoni, Towards picosecond resolution with single-photon avalanche diodes, *Rev. Sci. Instrum*., 52, 408–412, 1981.
35. S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, Avalanche photodiodes and quenching circuits for single-photon detection, *Appl. Opt*., 35, 1956, 1996.
36. S. Vasile, P. Gothoskar, D. Sdrulla, and R. Farrell, Photon detection with high gain avalanche photodiode arrays, Presented at *IEEE Nuclear Science Symposium Conference Record*, Albuquerque, NM, Cat. No. 97CH36135, 1997.
37. P. Buzhan, B. Dolgoshein, A. Ilyin, V. Kantserov, V. Kaplin, A. Karakash, A. Pleshko et al., The advanced study of silicon photomultiplier, *Proceedings of the 7th International Conference on Advanced Technology & Particle Physics (ICATPP-7)*, Como, Italy, pp. 717–728, 2001.
38. P. Buzhan, B. Dolgoshein, L. Filatov, A. Ilyin, V. Kantzerov, V. Kaplin, A. Karakash et al., Silicon photomultiplier and its possible applications, *Nucl. Instrum. Methods. Phys. Res. A*, 504, 48–52, 2003.
39. C. J. Stapels, W. G. Lawrence, J. Christian, M. R. Squillante, G. Entine, F. L. Augustine, P. Dokhale, and M. McClish, Solid-state photomultiplier in CMOS technology for gamma-ray detection and imaging applications, *Nuclear Science Symposium Conference Record*, Fajardo, PR, Vol. 5, pp. 2775–2779, 2005.
40. S. Vasile, J. S. Gordon, R. Farrell, and M. R. Squillante, Fast avalanche photodiode detectors for the superconducting super collider, *Semiconductors Room-Temperature Radiation Detector Applications Symposium*, San Francisco, CA, pp. 537–542, 1993.

41. P. Buzhan, B. E. Dolgoshein, L. Filatov, A. Ilyin, V. Kantzerov, V. Kaplin, A. Karakash, F. Kayumov, S. Klemin, E. Popova, and S. Smirnov, Silicon photomultiplier and its possible applications, *NIMA*, 504, 48–52, 2003.
42. D. Phelan, C. Jackson, R. Redfern, A. P. Morrison, and A. Mathewson, Geiger mode avalanche photodiodes for microarray systems, *Proc. SPIE Int. Soc. Opt. Eng.*, 4626, 89–97, 2002.
43. J. F. Christian, G. Svolos, A. I. Kogan, F. L. Augustine, M. R. Squillante, and G. Entine, Characterization & modeling of APD pixels made with CMOS technology, Presented at the *Nano Materials for Defense Applications*, Maui, HI, 2004.
44. A. Rochas, M. Gani, B. Furrer, P. A. Besse, R. S. Popovic, G. Ribordy, and N. Gisin, Single photon detector fabricated in a complementary metal–oxide–semiconductor high-voltage technology, *Rev. Sci. Instrum.*, 74, 3263–3270, 2003.
45. C. J. Stapels, W. G. Lawrence, F. L. Augustine, and J. F. Christian, CMOS solid-state photomultiplier for detecting scintillation light in harsh environments, *Proceedings of the International Symposium on Development Detectors Particle, Astro Particle, Synchrotron Radiation Experiments*, Stanford, CA, 2006.
46. J. F. Christian, C. J. Stapels, E. B. Johnson, M. McClish, P. Dokhale, K. S. Shah, S. Mukhopadhyay, E. Chapman, and F. L. Augustine, Advances in CMOS solid-state photomultipliers for scintillation detector applications, *Nucl. Instrum. Methods A*, 624, 449–458, 2010.
47. E. B. Johnson, C. J. Stapels, X. J. Chen, F. L. Augustine, and J. F. Christian, Large-area CMOS solid-state photomultipliers and recent developments, *Nucl. Instrum. Methods A*, 652, 494–499, 2011.
48. E. B. Johnson, C. J. Stapels, X. J. Chen, C. Whitney, E. C. Chapman, G. Alberghini, R. Rines, F. Augustine, and J. Christian, CMOS solid state photomultipliers for ultra-low light levels, *Proc. SPIE*, Orlando, FL, 2011.
49. E. R. Fossum, Active pixel sensors: Are CCD's dinosaurs? *Proc. SPIE*, 1900, 2–14, 1993.
50. E. Fossum and A. Krymski, *High Speed CMOS Imaging*. Piscataway, NJ, 2000.
51. Y. Shao, R. W. Silverman, R. Farrell, L. Cirignano, R. Grazioso, K. S. Shah, G. Vissel, M. Clajus, T. O. Tumer, and S. R. Cherry, Design studies of a high resolution PET detector using APD arrays, *IEEE Trans. Nucl. Sci.*, 47, 1051–1057, 2000.
52. K. S. Shah, R. Grazioso, R. Farrell, J. Glodo, M. McClish, G. Entine, P. Dokhale, and S. R. Cherry, Position sensitive APDs for small animal PET imaging, *IEEE Trans. Nucl. Sci.*, 51, 91–95, 2004.
53. M. McClish, R. Farrell, R. Myers, F. Olschner, G. Entine, and K. S. Shah, Recent advances of planar silicon APD technology, *Nucl. Instrum. Methods A*, 567, 36–40, 2006.
54. Y. Wu, C. Catana, R. Farrell, P. A. Dokhale, K. S. Shah, J. Qi, and S. R. Cherry, PET performance evaluation of an MR-compatible PET insert, *IEEE Trans. Nucl. Sci.*, 56, 574–580, 2009.
55. Y. Yang, J. Qi, Y. Wu, S. St. James, R. Farrell, P. A. Dokhale, K. S. Shah, and S. R. Cherry, Depth of interaction calibration for PET detectors with dual-ended readout by PSAPDs, *Phys. Med. Biol.*, 54, 433–445, 2009.
56. W. C. Barber, T. Funk, M. McClish, K. S. Shah, and B. H. Hasegawa, PSAPD gamma camera for SPECT imaging, Presented at *IEEE Nuclear Science Symposium on Conference Record*, Rome, Italy, 2004.
57. S. Kim, M. McClish, F. Alhassen, Y. Seo, K. S. Shah, and R. G. Gould, Temperature dependant operation of PSAPD-based compact gamma camera for SPECT imaging, *IEEE Trans. Nucl. Sci.*, 58, 2169–2174, 2011.
58. M. McClish, P. Dokhale, J. Christian, C. Stapels, E. Johnson, F. Augustine, and K. S. Shah, Performance measurements from LYSO scintillators coupled to a CMOS position sensitive SSPM detector, *Nucl. Instrum. Methods A*, 652, 264–267, 2011.
59. K.-P. N. Ziock and L. F. Nakae, Large area PSPMT-based gamma ray imager with edge reclamation, *IEEE Trans. Nucl. Sci.*, 49, 1552–1559, 2002.
60. A. J. Chaudhari, Y. Yongfeng, R. Farrell, P. A. Dokhale, K. S. Shah, S. R. Cherry, and R. D. Badawi, PSPMT/APD hybrid DOI detectors for the PET component of a dedicated breast PET/CT system—A feasibility study, *IEEE Trans. Nucl. Sci.*, 55, 853–861, 2008.

61. J. Zhang, A. M. K. Foundray, P. D. Olcott, and C. S. Levin, Performance characterization of a novel thin position-sensitive avalanche photodiode-based detector for high resolution PET, *IEEE Nuclear Science Symposium Conference Record*, Fajardo, PR, Vol. 5, pp. 2478–2482, 2005.
62. A. L. Beck, B. Yang, S. Wang, C. J. Collins, J. C. Campbell, A. Yulius, A. Chen, and J. M. Woodall, Quasi-direct UV/blue GaP avalanche photodetectors, *IEEE J. Quantum Electron*., 40, 1695–1699, 2004.
63. A. N. Pikhtin, S. A. Tarasov, and B. Kloth, Ag-GaP Schottky photodiodes for UV sensors, *IEEE Trans. Electron Dev*., 50, 215–217, 2003.
64. T. B. Ramachandran, Gallium arsenide phosphide-gallium arsenide heterojunction photodetectors, Presented at *International Electron Devices Meeting*, 1966.
65. J. Blazej, New materials for photon counting avalanche photodiodes, Presented at *12th International Workshop on Laser Ranging Instrumentation*, Matera, Italy, 2000.
66. L. B. Rowland, J. L. Wyatt, and S. M. Bishop, Cubic silicon carbide avalanche photodiodes, Presented at *International Semiconductor Device Research Symposium* (*ISRDS*), College Park, MD, 2009.
67. Y. Hirabayashi, S. Karasawa, K. Kobayashi, S. Misawa, and S. Yoshida, Spectral response of a photodiode using 3C-SiC single crystalline film, *Sensors Actuat. A Phys*., 43, 164–169, 1994.
68. J. C. Campbell, H.-D. Liu, D. McIntosh, and X. Bai, SiC avalanche photodiodes, *International Semiconductor Device Research Symposium*, College Park, MD, pp. 276-277, 2007.
69. A. Owens and A. Peacock, Compound semiconductor radiation detectors, *Nucl. Instrum. Methods Phys. Res. A*, 531, 18–37, 2004.
70. X. Guo, A. Beck, B. Yang, and J. C. Campbell, Low dark current 4H-SiC avalanche photodiodes, *Electron. Lett*., 39, 1673–1674, 2003.
71. S. Verghese, K. A. McIntosh, R. J. Molnar, L. J. Mahoney, R. L. Aggarwal, M. W. Geis, K. M. Molvar, E. K. Duerr, and I. Melngailis, GaN avalanche photodiodes operating in linear-gain mode and Geiger mode, *IEEE Trans. Electron Dev*., 48, 502–511, 2001.
72. Z. Vashaei, E. Cicek, C. Bayram, R. McClintock, and M. Razeghi, GaN avalanche photodiodes grown on m-plane freestanding GaN substrate, *Appl. Phys. Lett*., 96, 201908–201908-3, 2010.
73. J. K. Sheu, M. L. Lee, and W. C. Lai, Aluminum gallium nitride ultraviolet photodiodes with buried p-layer structure, *Appl. Phys. Lett*., 87, 043501–043501-3, 2005.
74. F. Barkusky, C. Peth, A. Bayer, K. Mann, J. John, and P. E. Malinowski, Radiation damage resistance of AlGaN detectors for applications in the extreme-ultraviolet spectral range, *Rev. Sci. Instrum*., 80, 093102–093102-3, 2009.
75. S. D. McDougall, M. J. Jubber, O. P. Kowalski, J. H. Marsh, and J. S. Aitchison, GaAs/AlGaAs waveguide pin photodiodes with non-absorbing input facets fabricated by quantum well intermixing, *Electron. Lett*., 36, 749–750, 2000.
76. B. K. Ng, J. P. R. David, S. A. Plimmer, R. C. Tozer, G. J. Rees, and M. Hopkinson, $Al_{0.8}Ga_{0.2}As$: A very low excess noise multiplication medium for avalanche photodiodes, *IEE Proc. Optoelectron*., 148, 243–246, 2001.
77. J. P. Anthes, X-ray response of AlGaAs/GaAs radiation hardened double heterostructure photodiode compared to Si:p-n photodiodes, *Rev. Sci. Instrum*., 59, 1846–1848, 1988.
78. R. H. Bube, *Photoelectronic Properties of Semiconductors*. Cambridge, U.K.: Cambridge University Press, 1992.
79. J. I. Pankove, *Optical Processes in Semiconductors*. New York: Courier Dover Publications, 1971.
80. M. A. Khan, J. N. Kuznia, D. T. Olson, M. Blasingame, and A. R. Bhattarai, Schottky barrier photodetector based on Mg-doped p-type GaN films, *Appl. Phys. Lett*., 63, 2455–2456, 1993.
81. J. A. Edmond, H.-S. Kong, and C. H. Carter, Blue LEDs, UV photodiodes and high-temperature rectifiers in 6H-SiC, *Physica B*, 185, 453–460, 1993.
82. K. S. Shah, P. Bennett, L. P. Moy, M. M. Misra, and W. W. Moses, Characterization of indium iodide detectors for scintillation studies, *Nucl. Instrum. Methods A*, 380, 215–219, 1996.
83. A. S. Grove, *Physics and Technology of Semiconductor Devices*. New York: John Wiley & Sons, 1967.

84. D. E. Aspnes, S. M. Kelso, R. A. Logan, and R. Bhat, Optical properties of $Al_xGa_{1-x}As$, *J. Appl. Phys*., 60, 754–767, 1986.
85. W. R. L. Lambrecht, B. Segall, W. Suttrop, M. Yoganathan, R. P. Devaty, W. J. Choyke, J. A. Edmond, J. A. Powell, and M. Alouani, Optical reflectivity of 3C and 4H-SiC polytypes: Theory and experiment, *Appl. Phys. Lett*., 63, 2747–2749, 1993.
86. S. G. Sridhara, R. P. Devaty, and W. J. Choyke, Absorption coefficient of 4H silicon carbide from 3900 to 3250 [A-ring], *J. Appl. Phys*., 84, 2963–2964, 1998.
87. J. C. Campbell, Optical properties of SiC (4H), Private Communication, 2010.

45

Vision and Image Sensors

Stanley S. Ipson
University of Bradford

Chima Okereke
University of Bradford

Vision is the act of seeing, a human capability derived from the combination of the image-forming optical system of the eye, the array of light-sensitive receptors in the retina of the eye, and the information processing capacity of the retina and human brain. Applications for instruments with a similar ability to sense a scene include broadcast television, monitoring industrial processes, quality control during manufacture, viewing inaccessible or hazardous places, aiding medical diagnosis, and remote sensing from satellites and space probes, to name but a few. Each application often has particular requirements more critical than others, and the problem is to find the most economical solution satisfying requirements closely.

The human eye is a familiar imaging system, highly optimized to aid survival under the naturally occurring range of illumination conditions. It is useful to describe its basic characteristics [1] to provide a benchmark against which machine vision can be compared. Light intensity can be measured using radiometric quantities such as radiant power, which is the radiant energy transferred per second in watts; irradiance, which is the radiant power falling onto a surface of unit area in W/m^2; and radiance, which is the radiant power leaving unit area of a surface per unit solid angle in $W/st/m^2$. It can also be measured using photometric units that take into account the variation in response of a standard human eye with wavelength, the Commission Internationale de l'Éclairage (CIE) response curve. The photometric equivalent of radiant power is luminous flux, measured in lumens (lm); the photometric equivalent of irradiance is illuminance, measured in lm/m^2 or lux; and the photometric equivalent of radiance is luminance, measured in $lm/st/m^2$ or cd/m^2. At the wavelength of peak sensitivity (555 nm) in the CIE sensitivity curve, the conversion factor from radiometric to photometric units is 680 lm/W.

The eye has a response (observed brightness) to incident light intensity, which is roughly logarithmic and is capable of adapting, given time, to an enormous range of different illuminance levels, from full sunlight (100,000 lux) to starlight (0.001 lux). The illumination generally recommended for surfaces in an office environment (about 100 lux) represents a comfortable working level. It takes about 30 min for the eye to become fully dark adapted and a further 3 min to adapt again to increased lighting levels. Despite this adaption capability, when viewing any one point in a scene

under bright light conditions, the eye is capable of discerning only about 25 different intensity levels. The eye achieves its enormous range of brightness adaptation partly through the action of a variable size iris (about 2–8 mm in diameter) but mainly through the presence of two types of light-sensitive receptor in the retina of the eye.

At high illumination levels, corresponding to photopic vision, the eye perceives color due to the excitation of three different types of cone receptors. There is considerable overlap in the spectral sensitivity curves of these red, green, and blue receptors, which have peak responses near 600, 550, and 500 nm, respectively. The overall photopic response is greatest at a wavelength near 555 nm in the yellow-green region of the spectrum and falls to zero toward the red and the blue ends of the spectrum at about 750 and 380 nm, respectively. At low levels of illumination, corresponding to scotopic vision, only the rod receptors are excited with a peak luminous efficiency of about 1700 lm/W near 510 nm in the green region of the spectrum. The sensitivity to contrast of the dark-adapted eye is much poorer than that of the light-adapted eye. The transition between photopic and scotopic vision is gradual, with both excited over a luminance range from about 0.001 to 0.1 cd/m^2. The photopic response extends higher to the glare limit, about five orders of magnitude brighter, while the scotopic response extends lower to the threshold of vision, about three orders of magnitude lower. The photopic (CIE) and scotopic spectral responses of the eye are shown in Figure 45.1 with the response of a typical silicon-based sensor for comparison. The latter has a spectral response that is quite different from that of the human eye, extending well beyond the red end of the visible spectrum toward 1100 nm, with maximum sensitivity around 800 nm. Assuming the average spectral response of silicon, an irradiation of 1 μW/cm^2 corresponds to an illumination of about 0.22 lux.

The distribution of the cones and rods over the inner surface of the eyeball is nonuniform, with the cones most densely packed in the region of the fovea, a circular region about 1.5 mm in diameter situated toward the rear of the eyeball. To achieve sharpest vision, the eye muscles automatically rotate the eyeball so that the object of interest in the scene is imaged onto the fovea. The separation of the approximately 400,000 cones in the fovea is such that a normal young eye is able to distinguish alternating black and white bands, each 1 mm wide at a distance of about 5 m, corresponding to an angular resolution of about 0.2 mrad. The resolution of the dark-adapted eye is very much poorer because, although the maximum number of rods per square millimeter is similar to that of cones, several rods are connected to a single nerve end, whereas only a single cone is connected to a nerve end.

The overall performance of the eye is difficult to equal. However, a machine vision system can surpass it in individual respects such as sensitivity to infrared and ultraviolet (UV) radiations invisible to the eye, speed of response (which can be as short as nanoseconds), and sensitivity to low-light

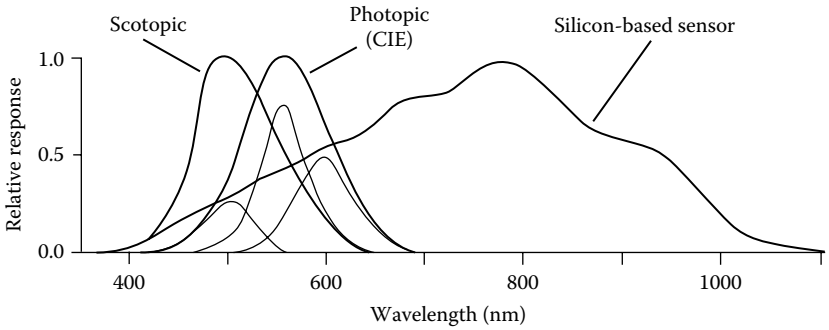


FIGURE 45.1 The linear responses, normalized to unity, of a typical bright-adapted human eye (photopic), a typical dark-adapted eye (scotopic), and a typical charge-coupled-device sensor to different wavelengths of light. The CIE response curve is an internationally agreed response curve for the photopic vision of a standard observer, which is used to convert radiometric measurements to photometric. The individual responses of the red, green, and blue cones are also shown, drawn using thinner lines, beneath the photopic curve.

levels corresponding to the detection of individual photons of radiation over periods of several hours. Machine vision characteristics depend on the type of image sensor employed and the modifying effects of additional components such as image intensifiers or optical fiberscopes. The machine vision equivalent of the eye includes a lens to project a 2-D image of the 3-D object of interest onto a sensor that transforms the light energy into an electric signal. In this form, it may then be transmitted to a remote location, subjected to computer analysis, or displayed on a television screen. A rectangular 2-D image can be sensed using three different approaches. A single, small sensor can be moved in a zigzag or raster fashion to sense the light intensity on a grid of points covering the whole image. A second approach is to use a line sensor composed of many individual sensors. If it is equal in length to one side of the image, it need only be moved in steps equal to the width of the individual sensors, a distance equal to the other side to cover the whole image. The third approach is to use an area sensor comprising a 2-D array of individual sensors. Whatever method is used, each individually sensed small region of the image is called a *picture element* or *pixel*, and in general, the resolution of the imaging system improves with increasing number of pixels. Although the first two methods require relative movement between sensor and scene, they are not limited by the sensor in the amount of resolution that can be achieved along the direction of motion. In addition, the inevitable variation in response between different sensors is easier to correct than with comparable area sensors, because there are fewer sensors. Two common applications of line sensors are monitoring objects on moving conveyer belts and scanning documents. In many cases, however, it is not practical to arrange relative movement of image and sensor, and the majority of image sensors in use are monochrome or color area sensors designed for use in television cameras. Monochrome and color cameras both use the same basic sensors, which have an inherently broad spectral response. Color sensitivity is achieved with the aid of color filters. A common approach is to deposit color filters directly on the sensor surface in a mosaic or stripe pattern. This is a cheap and compact solution but results in a lower spatial resolution in each color, compared with the resolution of the equivalent unfiltered sensor, and loses much of the incident light through absorption in the filters. Better sensitivity and resolution are obtained using three-sensor cameras that incorporate an optical arrangement (based on dichroic mirrors) that separates the incident light into three components—say red, green, and blue—which are each directed at a different sensor. However, this approach is both bulky and expensive, because it requires three sensors rather than one and precision optics to align the images correctly on the individual sensors.

Most area sensors currently sold are designed to produce electric signals compatible with either the 525-line American television standards (RS-170 for monochrome and National Television System Committee (NTSC) for color) or the 625-line European television standards (Comité Consultatif International des Radio Communications (CCIR) for monochrome and Phase Alternating Line (PAL) for color) [2]. These are interlaced television standards in which each complete image or frame is made up of two fields, each containing either odd- or even-numbered lines from the frame. Half the remaining lines, making up the odd total number in the frame, appear in each field, ensuring that the two fields interlace properly on the television display. Interlacing is used to avoid picture flicker without having to double the rate of information transmitted. According to the Ferry–Porter law, the critical frequency below which flicker is observed depends on the logarithm of the luminance of the picture highlights. At the brightness levels corresponding to normal viewing, flicker would be observable in pictures interrupted at 25 or 30 Hz, but it is reduced to an acceptable level at the repetition rates of the fields (60 Hz for NTSC and 50 Hz for PAL). A consequence of the Ferry–Porter law is that an NTSC picture can be about six times brighter than a PAL picture and still be acceptable in terms of flicker. Many companies supply frame grabber computer boards that sample the TV-compatible voltage signals (about 1 V range) at a frequency of at least 10 MHz (for at least 512 samples per line), quantize the analog voltage to 256 or more discrete levels (8 bits), and store the resulting digital data for computer analysis. Interlaced operation is not ideal for some applications, and at higher cost, cameras are available that provide higher frame rates, higher resolutions, or progressive scan (no interlacing). These usually produce a standard digital output signal for easy input into a computer. A very wide range of

sensors and cameras are currently commercially available, based on several different types of technology. Although each technology has particular advantages, the characteristics of different devices based on a specific technology can vary significantly. For example, the spectral responses of photodiode-based sensors generally extend to shorter wavelengths than those of CCD (charge-coupled device)-based sensors, but some CCD devices are available with extended blue-end responses. Improved and/or cheaper sensors and cameras are appearing on the market all the time, and when considering a new imaging system, it is wise to review the characteristics of the devices currently available in comparison with the requirements of the application in order to select the most suitable. One of the first steps in the creation of a machine vision system for a new application is the design or selection of an appropriate optical system, and this is the subject of the next section. Later sections discuss the various sensor technologies that are available and other related devices that can be used to improve sensitivity or to allow image sensing in restricted spaces such as body or machine cavities.

45.1 Image Formation

The machine vision optical system has the tasks of matching the required field of view to the dimensions of the sensor and gathering sufficient light to achieve a sensor signal that has adequate signal-to-noise ratio (SNR) while maintaining the image sharpness required by the application. Resolutions of area sensors are sometimes specified by manufacturers as the maximum number of horizontal TV lines that can be distinguished. This value may be obtained from measurements of the visibility of a test chart (Electronic Industries Alliance (IEA) test pattern), or it may simply be the number of horizontal pixels divided by 1.33, in the case of a monochrome sensor. In special circumstances, measurements to subpixel accuracy can be achieved by interpolation between pixel values, and Ref. [3] describes the measurement of a knife edge to about 0.1 pixel accuracy. As a rule of thumb, however, for size measurement applications, the sensor should have a number of pixels at least equal to twice the ratio of the largest to smallest object sizes of interest [4], and a lens is then selected to provide the required magnification and working distance. A lens operating in a single medium such as air is characterized [5] by two focal points and two principal planes as indicated in Figure 45.2. The principal planes coincide with the lens center in the case of an ideal thin lens but may be separated by +20 to –10 mm (the negative sign indicating reversed order), depending on the design, in multielement lenses. They can be determined by

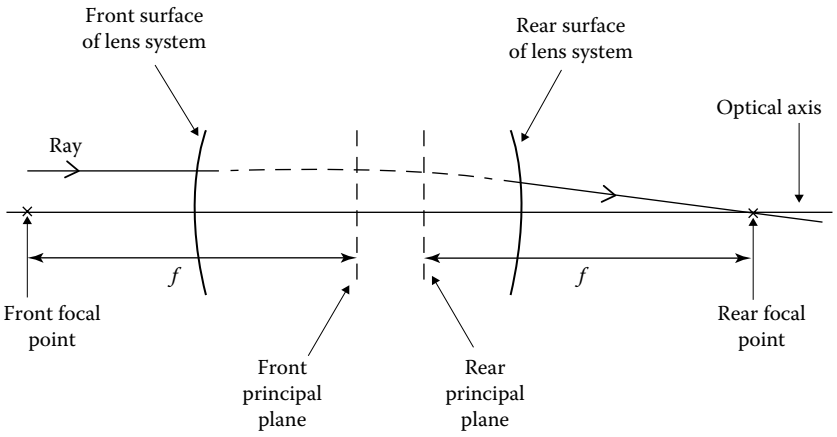


FIGURE 45.2 The trace of a ray from a distant object through a typical multielement lens system showing typical positions of focal points and principal planes. The distance from the back focal length to the lens flange mount is 17.3 mm for C-mount lenses, 12.5 mm for CS-mount lenses, and 46.5 mm for 35 mm photographic format Nikon F (bayonet) mount lenses. Adapters are available to accommodate the differences in distances and fittings of the different mounting systems.

first locating the two focal points and then measuring along the axis from these distances equal to the focal length. A lens of focal length f produces an image in best focus at a distance l_i when the object is at distance l_o where

$$\frac{1}{l_o}+\frac{1}{l_i}=\frac{1}{f} \tag{45.1}$$

and the distances are measured to the corresponding principal planes. The image magnification m, defined as the ratio of image to object size, is equal to the ratio of the image to object distances and is related to the total distance between object and image D_T by

$$D_T=\frac{F(m+1)^2}{m}+D_N \tag{45.2}$$

where D_N is the separation between the principal planes. Lenses generally have a focusing adjustment of 5%–10% of the focal length, and extension rings between the lens and sensor are required for image magnifications greater than about 0.05. The lens extension required is simply the product of the magnification and the focal length. The depth of field F_o of an imaging system [6] is the displacement of the object along the optic axis that produces no significant degradation in the sharpness of the image. The corresponding movement of the image is the depth of focus. In the case of a lens with *f*/number $f_\#$ (equal to the ratio of focal length to lens iris diameter) and an image sensor with pixel size P, it is given by

$$F_o=2f_\#\frac{P(m+1)}{m^2} \tag{45.3}$$

Television camera sensors are manufactured to standard sizes such as 1, 2/3, 1/2, or 1/3 in. The size is defined to be twice the horizontal dimension of a rectangular image with 4:3 aspect ratios so that a 1 in. sensor has a width of 12.7 mm, a height of 9.5 mm, and a diagonal length of 15.9 mm. Lens sizes are similarly specified to allow easy matching of lenses to sensors. Because image distortion and sharpness worsen toward the edges of the field of view, it is permissible, for example, to use a 2/3 in. lens with a 1/2 in. sensor, but not the converse. A 35 mm camera lens, designed for a 24 by 36 mm image size, generally performs much better, at relatively low cost, than a corresponding C-mount lens supplied for a TV camera, but a C-mount to Pentax, Canon, or Nikon mount converter will then be required.

It is frequently necessary to relate lighting of a scene to camera sensitivity. Accurate calculations are difficult, and it is usually better to make a simple estimate and then make fine adjustments to the lighting or lens aperture. Manufacturers often specify camera sensitivities by quoting an illumination level in lux at the sensor faceplate. This may be the illumination required to achieve maximum signal output, some proportion of this maximum, or simply a "usable" signal level from the camera. The illumination of the sensor L_S in lux is related to the luminance of the object B in cd/m^2 by

$$L_S=\frac{TB\pi}{[2f_\#(m+1)]^2} \tag{45.4}$$

where losses in the lens are characterized by a transmission coefficient T. If the object is a uniform diffuse reflector (Lambertian surface) illuminated by L_o lux, then the luminance of the object is given by

$$B=L_o\frac{R}{\pi} \tag{45.5}$$

where R is the reflection coefficient of the surface. Some practical examples of radiometric calculations are given in a Dalsa application note [7].

45.2 Image Sensing

The primary functions occurring within a standard image sensor are the conversion of light photons falling onto the image plane into a corresponding spatial distribution of electric charge, the accumulation and storage of this charge at the point of generation, the transfer or readout of this charge, and the conversion of charge to a usable voltage signal. Each of these functions can be accomplished by a variety of approaches, but only the principal sensor types will be considered here. Sensors can be divided into two groups: (1) vacuum tube devices in which the charge readout is accomplished by an electron beam sweeping across the charge in a raster fashion similar to that in a television picture tube and (2) solid-state devices based on CCDs or photodiodes. These three types of sensor will be described in the next three sections.

45.2.1 Television Camera Tubes

For many years, vacuum tubes [8] provided the only technology available for television applications, and they are still widely used because of the high-quality image signals they provide. Companies supplying tubes include Burle and Philips. Most modern tube cameras are based on the vidicon design whose basic components are indicated in Figure 45.3. Light from the scene is imaged by a lens onto a photoconductive target formed on the inner surface of an end window in a vacuum tube. An electron gun is placed at the opposite end of the tube to the window, and it provides a source of electrons that are focused into a beam, accelerated toward the target by a positive potential on a fine mesh placed just in front of the target, and scanned across the target by an electrostatic or magnetic deflector system. The target consists of a glass faceplate on the inner surface, upon which is placed a transparent electrically conducting coating of indium tin oxide. On top of this is deposited a thin layer of photoconductive material in a pattern of

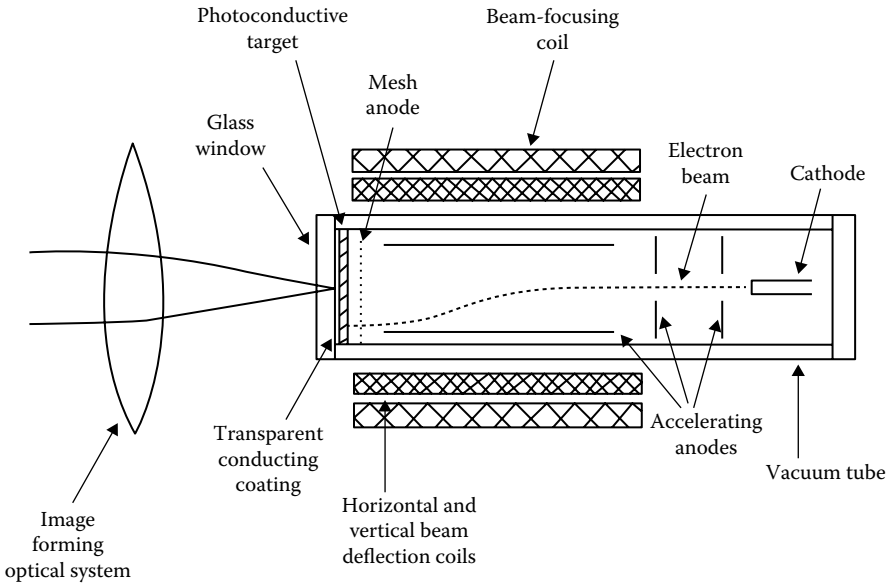


FIGURE 45.3 A schematic diagram of the major components of a vidicon camera tube. Many variants of this existing basic design utilize different target materials to improve particular characteristics. The electron beam readout of the optically generated charge in the target makes the tube bulky and less reliable than solid-state sensors. Most tubes currently sold are for replacement in existing equipment rather than in new cameras, which almost always incorporate solid-state sensors.

tiny squares, each insulated laterally from its neighbors. The transparent coating is connected to a positive voltage through an electric load resistor across which the signal voltage is developed. In the absence of light, the electron beam causes the external surface of the photoconductor to be charged to near zero potential. Light causes the resistance of the photoconductor to decrease and its surface to acquire a more positive potential, due to the accumulation of positive charge. At each point on the surface touched by the electron beam, some of the beam current is deposited to neutralize the positive charge present due to the illumination. The rest passes through the load resistor generating an output voltage that is a function of the light intensity at that point. The precise timing of the scanning process ensures that the location of the beam is known at each instant in time.

Vidicons have some unique characteristics. They can have spectral responses extending into the UV and infrared regions of the spectrum, and they allow scan rate and adjustments in scan alignment to be made electronically by adjusting the scanning pattern. This allows very simple region-of-interest readouts. On the other hand, the electron beam readout method confers a number of disadvantages, which are not found in solid-state devices, including large size, fragility, susceptibility to shock and vibration, high-power consumption, and a high degree of sensitivity to external electric and magnetic fields. The principal imaging defects, which again are not present in solid-state devices, include

- Lag when displaying changing images, due to slow response speed
- Image burn that is damage to the sensor surface caused by intense light
- Geometric distortion
- Drift in the apparent position of the image
- A nonlinear relationship between signal output and light intensity that varies with target voltage

The degree to which these effects are present depends on the type of tube and its particular construction.

The original vidicon tube design employed antimony trisulfide as the photoconductive material and a low-velocity electron beam and gave only adequate performance. Newer tubes employ the same basic design but use better electron beam readout techniques and target materials to improve characteristics such as spectral response, sensitivity, dark current, lag, and burn-in. Two tubes suitable for broadcast television use are the plumbicon developed by Philips and the saticon introduced more recently by Hitachi. The plumbicon employs a lead-oxide photoconductive target, while the saticon target comprises several layers with different combinations of selenium, arsenic, and tellurium. It can provide greater resolution than the plumbicon, but the lag performance and range of operating temperature are worse. Examples of nonbroadcast quality tubes are the newvicon, chalnicon, and si-vidicon. The newvicon target has a double layer of cadmium and zinc tellurium, which achieves exceptional sensitivity, but lag and nonuniformity are excessive. The chalnicon employs a cadmium selenide target and provides good performance in most respects but has excessive lag. The si-vidicon target effectively consists of an array of silicon photodiodes that are reverse biased and store charge released by the incident light. The resulting tube has very high sensitivity and little lag, and it is virtually burn proof. However, it exhibits excessive blooming and nonuniformity. The silicon-intensified camera, or SIT camera, is a modification of the si-vidicon camera that employs a photocathode that generates electrons when struck by light photons. These electrons are accelerated by a voltage of several hundred volts and focused on the target of the si-vidicon sensor that produces an output in the normal way. The high-speed electrons landing on the target produce large numbers of electron–hole pairs and a corresponding larger response. The resulting SIT sensor is used in very low-light-level applications such as tracking satellite debris and can produce useful video images at illumination levels down to 0.001 lux. By comparison, a solid-state CCD sensor (without an image intensifier) typically requires about 1 lux to produce a useful video signal.

45.2.2 Charge-Coupled Devices

A CCD is fabricated on a single crystal wafer of *p*-type silicon and consists of a 1-D or 2-D array of charge storage cells on centers typically about 10 μm apart. Each cell has several closely spaced

electrodes (gates) on top, separated from the silicon by an insulating layer of silicon dioxide. The charge is stored under one of the electrodes, and its location within the cell is defined by the pattern of positive voltages applied to the electrodes. By applying a coordinated sequence of clock pulses to all the electrodes in the array, packets of stored charge (of between 10 and 10^6 electrons) are transferred from one cell to the next until they finally reach a sensing amplifier (floating gate diffusion) that generates a voltage signal proportional to charge, usually assumed to be 1 μV per electron. The result is a device that has an inherently linear variation of output voltage with light from the minimum useful level set by noise to the maximum useful level set by saturation of the output amplifier or the limited capacity of the charge storage cells. The dynamic range of the device is defined as the ratio of the maximum output signal to the output resulting from noise. Manufacturers published figures that may use peak-to-peak or root-mean-square noise values (typically five times smaller) in this calculation, but the former is more relevant for imaging applications.

Any light penetrating into the underlying silicon generates electron–hole pairs. The holes are swept away to the substrate electrode, while the electrons accumulate beneath the nearest electrode in a potential well created by the applied voltage. The sensitivity of silicon is of the order of 1 μA of generated charge per microwatt of incident light. The cells intended to function as light-sensitive photosites have electrodes made of semitransparent polysilicon so that the light can penetrate into the storage region, whereas those intended to function only as part of a shift register for charge transfer are covered by an opaque surface layer. Due to manufacturing imperfections, the photosites do not have perfectly uniform characteristics, and a photoresponse nonuniformity (PRNU) of about 5%–10% is fairly typical. This is easy to measure using uniform sensor illumination, and its effects can be removed if necessary by calibration. The basic spectral response of the silicon extends from 200 to 1100 nm, but the first figure is typically reduced to 450 nm by absorption in the surface layers of the CCD photosites. Infrared radiation penetrates deeper into the silicon than shorter wavelengths, and charge created by an infrared photon may be collected by a different cell to the one entered by the photon. This reduces the resolution of the device, and if infrared operation is not required but the illumination contains infrared (e.g., from a tungsten lamp), then an infrared reflecting filter (a hot-mirror filter) is often used. CCD cells also accumulate charge linearly with time due to thermally generated electrons produced within the cells and at electrode interfaces. Like the photoresponse, this dark signal varies from cell to cell and can be compensated for by calibration. These thermally generated contributions are most significant for low-light-level applications and can be reduced by cooling the sensor using either a thermoelectric cooler, a Joule–Thomson cooler, or a liquid nitrogen Dewar. The dark signal reduces by 50% for every 7 °C reduction in temperature, and at −60 °C, produced by a Peltier cooler, the dark signal is typically reduced to about one electron per pixel per second. Another important temperature-dependent characteristic of the CCD sensor that improves with cooling is the noise floor of the output amplifier, which is proportional to $T^{1/2}$ and typically equivalent to about 300 electrons at room temperature. A particular CCD device used in astronomy, for example, and operated at about −110 °C has a readout noise of about 10 electrons, a dark current less than 0.3 electrons/min, and quantum efficiency for converting visible photons into electrons of between 70% and 80%. Light can be integrated for periods of hours, compared with the approximately 1/8–1/4 s integration period of the dark-adapted eye. Compared with photographic film previously used for low-light level imaging in astronomy, CCDs are from 10 to 100 times more sensitive, linear in response rather than nonlinear, and have a much greater dynamic range so that both faint and bright objects can be recorded in the same exposure.

The short-exposure, high-frequency performance of CCD devices is limited by another effect. The process of transporting charge in a CCD sensor is not 100% efficient, and in practice, a small amount of charge is left behind at each transfer to the next cell, contributing noise and degrading resolution. This effect limits the maximum clock frequency and the maximum number of transfers in a CCD sensor. Manufacturers' data sheets for commonly available CCD sensors quote values ranging from 0.9995 to 0.99999 for the charge-transfer efficiency (CTE) of a single transfer. There are many variations in CCD technology. For example, virtual-phase CCDs [9] have some of the electrodes replaced by ion-implanted

regions, resulting in improved blue response and higher sensitivity because of the removal of some of the blocking surface gates and simpler drive circuitry because of the reduction in number of gates per cell. A manufacturing technique known as *pinning* can be used to passivate the interface states, which are the biggest contribution to the dark signal, producing an order of magnitude improvement in dark signal as well as improved quantum efficiency and CTE. A signal-processing technique called *correlated double sampling* can also be applied to the output signal from the sense amplifier to improve the readout noise performance.

45.2.3 Linear Charge-Coupled Devices

The basic structure of a linear CCD sensor is shown in Figure 45.4. It consists of a line of up to several thousand photosites and a parallel CCD shift register terminated by a sensing amplifier. Each photosite is separated from a shift register cell by a transfer gate. During operation, a voltage is applied to each photosite gate to create empty storage wells, which then accumulate amounts of charge proportional to the integral of the light intensity over time. At the end of the desired integration period, the application of a transfer pulse causes the accumulated charge packets to be transferred simultaneously to the shift register cells through the transfer gates. The charges are clocked through the shift register to the sensing amplifier at a rate of up to 20 MHz, producing a sequence of voltage pulses with amplitudes proportional to the integrated light falling on the photosites. In practice, it is common for shift registers to be placed on both sides of the photosites, with alternate photosites connected by transfer gates to the right and left registers. This halves the number of cells in each register and the time required to clock out all the data. Another 2 × reduction of transfer time is achieved if each shift register is split in two with an output amplifier at each end. There is a limit, typically between 10^5 and 10^6 depending on photosite size and dimensions, to the number of electrons that can be stored in a particular cell, beyond which electrons start to spill over into adjacent cells. The saturation charge in electrons is about 1000–2000 times the area of the photosite in square micrometers. This spread of charge or blooming is a problem with images containing intense highlights. It is reduced by about a factor of 100 by adding antiblooming gates between adjacent photosites and transfer gates and channel stops between adjacent photosites.

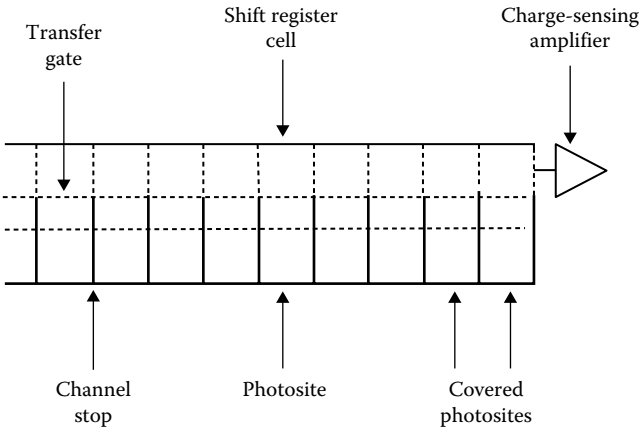


FIGURE 45.4 The architecture of a basic linear CCD showing the arrangement of photosites, transfer gates, shift register cells, and sensing amplifier. The latter produces a sequence of voltage pulses each proportional to the charge accumulated in one of the photosites. Although these pulses are often displayed on an oscilloscope during setting up of the optics and illumination, in normal use, they are digitized to 8 or 12 bits, and the resulting values are stored in memory. In practice, most CCDs have two or more covered cells at each end of a line of photosites to allow the dark current to be monitored and subtracted during signal processing. Applying uniform illumination to the CCD enables the relative response of the individual photosites to be measured and compensated for.

The voltage on the antiblooming gates is set at a value that allows surplus charge to drain away instead of entering the transfer gates and shift register. By clocking this voltage, variable integration times that are less than the frame pulse to frame pulse exposure time can also be attained.

45.2.4 Area Charge-Coupled Devices

Three basic architectures are used in area CCDs and are illustrated in Figure 45.5. The simplest is the full-frame CCD, consisting of an imaging area separated from a horizontal CCD shift register by a transfer gate. In the imaging area, each photosite is one stage of a vertical shift register separated from neighboring shift registers by channel stops and antiblooming structures. During the light integration period, the vertical clocks are stopped, creating potential wells that collect photoelectrons. At the end of this period, the charge is clocked out vertically, one row at a time, into the horizontal shift register. The charge in the horizontal shift register is then very rapidly shifted toward the output amplifier by the application of a horizontal clock signal. An example of full-frame CCD is the RA1001J from EG&G Reticon, which has pixels arranged in a 1024 × 1024 configuration and dual horizontal shift registers and outputs to achieve a 30 frame/s readout rate. To avoid image smearing during the readout period, full-frame CCD sensors must be operated with external shutters, as in digital cameras produced by Kodak and other companies, or used in low-light-level applications requiring very long integration times compared with the readout time, as in astronomy. The full-frame CCD is also the basis for time delay and integration (TDI) sensor manufactured by companies such as Dalsa and EG&G Reticon. In the application of this device, which may have 96 rows each with 1024 pixels, a moving scene illuminates a region on the sensor only one pixel high, and this is arranged to move vertically down the sensor at the same rate that the generated photocharge is shifted down the sensor. This mode of operation achieves an output signal identical to that from a linear CCD sensor except for an increased light sensitivity proportional to the number of horizontal rows in the TDI sensor. It can be appropriate in applications where both high speed and high sensitivity are required, providing a cheaper alternative to the combination of a line sensor and an image intensifier. An application of this technique to capturing images of web systems moving at up to 5 m/s is described in Ref. [10].

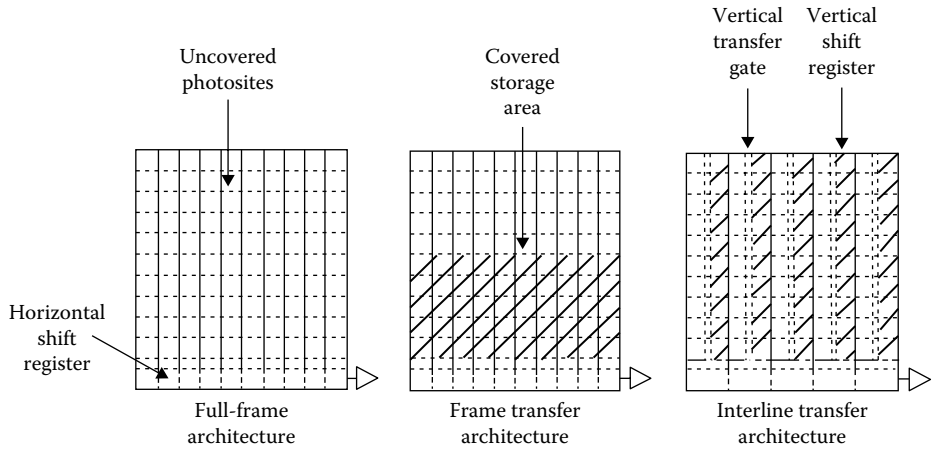


FIGURE 45.5 The basic architectures used in area CCDs. Full frame (FF) devices fully utilize the available area for image detection and therefore achieve the largest number of pixels, currently about 5000 × 5000, but need an external shutter. Frame transfer (FT) devices are normally used at standard video rates without an external shutter. The InterLine Transfer (ILT) architecture allows exposure times down to less than 0.1 ms. The lower fill factor of ILT devices compared with FF and FT devices results in higher values of modulation transfer function at the Nyquist limit. This increases the visibility of aliasing artifacts in images containing high spatial frequencies. A low-pass optical filter must be used if aliasing is a problem.

The requirement for an external shutter is greatly reduced in the frame-transfer CCD sensor by the provision of a light-shielded storage section into which the entire 2-D image charge is shifted at a much higher rate (limited primarily by CTE considerations) than is possible in a full-frame CCD. The charge can then be read from the storage region during the next integration period without any further image smearing. Some sensors, such as the EG&G Reticon RA1102 devices, have the storage area divided into halves placed on opposite sides of the imaging areas. This improves performance by halving the maximum number of transfers required to reach the nearest storage region. The same reduction occurs automatically in sensors designed for interlaced video applications where each integration period corresponds to one video field, and only half the number of rows in the frame is required at any one time. For example, to produce an interlaced video frame containing 576 image lines (CCIR standard), a frame-transfer sensor with only 288 rows of storage is required. By changing the clock signals, the odd field can be displaced vertically by half a line width relative to the even field. This ensures that the odd and even lines contain different information and reduces aliasing, because the cell width is twice the separation between the lines in the frame. Some of the companies that produce frame-transfer CCD sensors and cameras are Cohu, Dalsa, EG&G Reticon, EEV, Kodak, Philips, and Thomson-CSF.

Image smear is virtually eliminated by the interline-transfer (ILT) architecture in which each column of photosites has an adjacent light-shielded vertical CCD shift register into which the charge is transferred by a transfer pulse. The contents of the entire vertical shift registers are then shifted simultaneously one pixel at a time into a horizontal shift register where they are then rapidly shifted to an output amplifier. This approach makes it easy to implement exposure control and achieve short integration times, but it introduces *dead space* between the active pixels, reducing the sensitivity of the image sensing area and increasing aliasing effects compared with frame-transfer sensors. For the latter, the fill factor, which is the percentage of the imaging area that is light sensitive, can be close to 100%, whereas it is usually less than 50% for ILT devices. Localized bright objects tend to produce vertical streaks in an ILT device because strong light can leak under the narrow light shield covering the vertical shift registers, causing image smearing similar to that in a full-frame device. This reduces the usefulness of ILT sensors for scenes containing pronounced highlights. For interlaced operation, two adjacent pixels, for example, 1 and 2 and 3 and 4, are transferred to a single shift register cell on one field, and in the next field, pixels 2 and 3, 4 and 5, etc., are transferred together. This is rather similar to the interlaced operation of a frame-transfer CCD. The primary advantages of the ILT sensor are low-noise and good-exposure control, providing true *stop-motion* control on every field, because all photosites integrate light over the same period of time. Many companies manufacture ILT CCD sensors and cameras, including Hitachi, NEC, Panasonic, Pulnix, and Sony.

The frame-transfer and ILT approaches both have performance advantages and disadvantages, and the hybrid frame-interline-transfer (FIT) approach [11] combines some of the advantages of both. This architecture includes a light-shielded field storage area between an interline imaging area and the horizontal output shift register. With this arrangement, the charge associated with the whole image is first transferred horizontally into the interline storage area, which facilitates exposure control. The charge is then transferred at maximum speed (as in the FT sensor) into the field storage area, which minimizes the occurrence of vertical streaking. For example, the NEC microPD 3541 array clocks the vertical registers at 100 times the normal rate for an ILT sensor. This gives a 20 dB improvement in streaking threshold and an overall threshold of 80 dB, making streaking effects less than other optical effects such as lens flare. On the other hand, the FIT approach does not improve the fill factor, and aliasing artifacts associated with ILT sensors and the noise levels are somewhat higher, and the CTE reduced, because of the higher clock frequencies. Manufacturers of FIT sensors and cameras include JVC, Panasonic, and Sony.

45.2.5 Photodiode Arrays

Because photodiode arrays generally have less extensive electrode structures over each sensing element compared with CCD arrays, the spectral response is smoother and extends further at the blue end of the

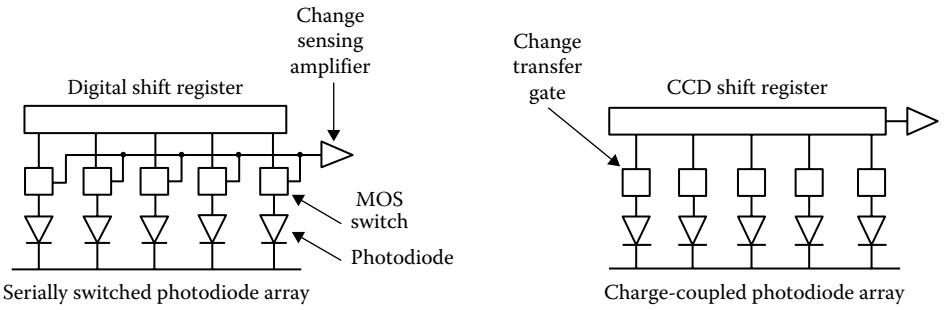


FIGURE 45.6 Readout architectures associated with photodiode array sensors. In the SSPD approach shown, each photodiode is connected in turn via an MOS switch to the sensing amplifier. This eliminates smearing, because the charge does not move between photosites. The device noise and speed characteristics are poorer than CCD sensors because of nonuniformity in the MOS switches and the large capacitance of the sensing bus. The CCPD approach combines the excellent blue-end response of the photodiode with the low-noise transfer characteristics of the CCD shift register to produce very high-quality linear devices.

spectrum. The peak quantum efficiency is also higher ranging from 60% to 80% compared with 10% to 60% for photogates, leading to almost twice the electric output power for a given light input power. Photodiode arrays would therefore appear to be attractive alternatives to CCD array sensors, but in practice, CCD sensors have lower noise levels because they do not have reverse-bias leakage current. A photodiode consists of a thin surface region of p-type silicon formed in an n-type silicon substrate. A negative voltage applied to surface electrode reverse biases the p-n junction and creates a depletion region in the n-silicon containing only immobile positive charge. When the electrode is isolated, the p-n junction is left charged and is effectively a charged capacitor. Light penetrating into the depletion region creates electron–hole pairs, which, with dark current, discharge the capacitor linearly with time. The penetration depth in silicon increases with wavelength, and the depletion region should be wide enough to absorb all wavelengths of interest. Dark current and most of the noise sources operating in the photodiode increase with reverse bias but can be reduced by employing a thicker p-type region, which allows a wide depletion region to be achieved with low bias voltage. However, this also degrades the blue-end response of the photodiode. To achieve good blue and UV response along with low bias voltage operation, a three-layer structure comprising thin p-type, intrinsic, and n-type substrates is employed. The intrinsic layer is so pure that the depletion region extends halfway across it at zero bias and extends right across it at a small reverse-bias voltage. This structure provides photodiodes with excellent linearity, noise performance, and speed of response at low operating voltages. At the end of an integration period, the states of charge of the photodiodes are measured, and image sensors based on two different types of readout approach are commercially available. These are the serially switched (sometimes called self-scanned) photodiode (SSPD) arrays and the charge-coupled photodiode (CCPD) arrays illustrated in Figure 45.6.

45.2.6 Serially Switched Photodiode Arrays

Associated with each photodiode in the sensor array is a metal-oxide semiconductor (MOS) switch, which, when turned on by a digital voltage level applied to its control line, connects the photodiode to a readout amplifier. Each control line is connected to one of the outputs of a digital shift register and, shifting a bit through the register, sequentially reads the charge on each photodiode in the array. This type of readout is very flexible, and random readout is achieved by replacing the digital shift register by an address decoder connected to an address bus as in the SR series linear arrays from EG&G Reticon. The 2-D arrays of photodiodes are connected in a configuration similar to a cross-point switching

matrix with a switch and diode at each cross point and separate vertical and horizontal shift registers. To scan the array, the vertical shift register turns on a complete row of switches, which causes the photodiodes in the corresponding row to dump their charge into vertical signal lines. These are in turn connected to the output amplifier by a set of horizontal switches controlled by the horizontal shift register. Switched arrays of photodiodes are made using processes similar to those employed in the manufacture of dynamic RAM, and this along with their simple structure and small size yields higher densities and lower manufacturing costs. However, the SSPD approach has one shortcoming, namely, the large capacitances of the readout buses. This reduces the speed of the device, increases the noise, and reduces the dynamic range, which are all generally significantly worse than those of CCD sensors. The dynamic range, for example, is typically 250–300 compared to more than 2500 for commercially available CCD sensors. SSPD arrays also suffer from fixed-pattern noise due to nonuniformities in the switches, which can be as high as 20% of the saturated output. The SSPD approach does have unique advantages:

- The readout method is extremely flexible and can be programmable.
- It is relatively free from blooming effects, because the sensors can be optimally isolated from each other.
- The blue response is good, which is particularly important for spectroscopic applications and for studio quality color television cameras.

Manufacturers supplying SSPD-based sensors include EG&G Reticon, Hamamatsu, and VLSI Vision.

45.2.7 Charge-Coupled Photodiode Arrays

Although individual photodiode elements are superior to individual CCD photogate elements in terms of dynamic range, the dynamic range of the SSPD is limited by fixed-pattern noise, switching noise, and the difficulty of transferring charge packets over the high capacitance output bus. Replacing the horizontal digital shift registers by a low-noise CCD shift register significantly improves the dynamic range of the device, and the resulting hybrid design is the CCPD array. During operation, the charges on an entire row of photodiodes are simultaneously transferred via MOS switches into the analog CCD shift register and then shifted out in sequence to the output amplifier. By providing a reset pulse (from a second vertical digital shift register) to eliminate all prestored charge in one row a fixed interval before the charge is transferred to the CCD shift register, it is relatively easy to implement an electronic equivalent of a focal-plane shutter [12] and control exposures in increments of one horizontal scan period. However, images of moving objects will be subject to distortions similar to those found using photographic cameras with focal-plane shutters. In the case of 2-D arrays, there are inefficiencies in the movement of small charge packets over the vertical transport lines, which get worse the smaller the packet. The solution adopted is to turn a small charge into a large one by precharging the lines with a priming charge equal to 10%–15% of the maximum charge capacity before transferring the photocharge to the CCD in a process known as charge-primed transfer. This requires several MOS switches forming a sequence of transfer gates and storage capacitors between each vertical transport line and CCD cell and depends critically on correct phasing of a large number of timing signals for proper operation. The CCPD approach appears to be most successful for linear arrays, which do not require a vertical transport bus. Manufacturers supplying linear CCPD-based sensors include Dalsa and EG&G Reticon.

45.3 Image Intensifiers

When the available solid-state image sensors do not have enough sensitivity for the scene illumination, it is necessary either to use a more sensitive tube camera or to amplify the available light. Companies supplying intensified cameras include Cohu, Kodak, Pulnix, and Philips. Intensified cameras are

standard video cameras fitted with image intensifiers that increase the number of available photons falling onto the image sensor. In addition to allowing operation over a wide range of lighting of the order of 0.0001 lux from overcast starlight to 100,000 lux of full sunlight, intensified cameras can be used to stop motion in nanosecond frame times, count photons, and observe images with very high intrascene dynamic ranges. However, this increased flexibility is achieved at a considerable increase in cost, usually some loss in resolution, reduced lifetime, and increases in geometric distortion, lag, and smear. Image intensifier tubes (IITs) are usually classified as generation I, II, and III devices.

45.3.1 Generation I Tubes

In a generation I intensifier, illustrated in Figure 45.7, the incoming photons strike a multialkali photocathode within a vacuum tube, and the resulting photoelectrons emitted into the vacuum are accelerated and focused onto a phosphor coating on the rear wall of the tube. The impact of the high-energy electrons on the phosphor creates many electron–hole pairs that in turn generate light at a higher level than that striking the photocathode. The electron gain depends strongly on the accelerating voltage and also on the photocathode and phosphor materials, and it is typically around 200 with a 12 kV accelerating voltage. These tubes tend to be bulkier than later-generation tubes and have relatively low gains, although tubes can be cascaded to achieve higher gains. They are also relatively intolerant of bright sources and rapid changes of scene illumination, with increased tendency toward highlight blooming. The light output from the intensifier tube is transferred to the solid-state sensor by either a relay lens or a fiber-optic coupler. The latter is shorter and much more efficient (typically 60% compared with 6%) but requires a sensor whose protective glass or quartz window has been replaced by the fiber-optic coupler resting against the sensor surface. The narrow gaps between couplers and between the coupler and sensor are filled by index-matching optical grease. The spectral response of the photocathode should be chosen to match the scene requirements, while the narrower phosphor response should be optimally matched to the image sensor for greatest sensitivity [13]. In general, for largest output, a phosphor with the maximum persistence that is acceptable in terms of image lag should be selected. But, if the scene includes object motion, a shorter-persistence phosphor may be necessary to avoid image smear. For night vision, P20 and P39 phosphors are frequently used. The former emits 550 nm light with very high luminous efficiency (65%) and relatively short persistence (0.2 ms), while the latter has a persistence of about 80 ms, which reduces high-frequency jitters. For high-speed imaging, the yellowish-green P46

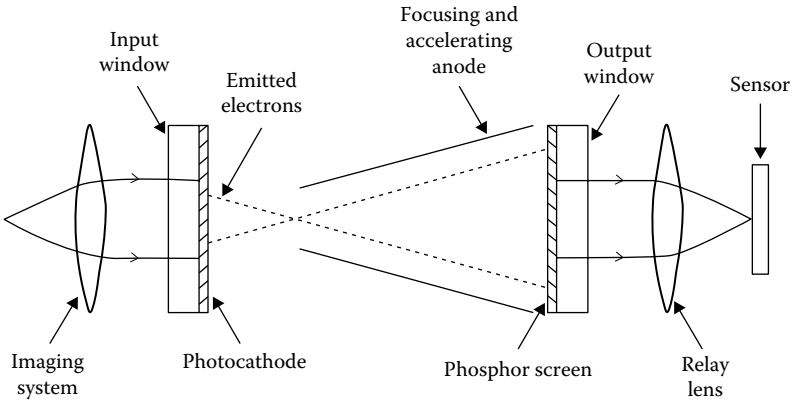


FIGURE 45.7 A schematic diagram showing the major components of a generation I IIT coupled to a sensor. Combined with a CCD sensor, the sensitivity of a SIT tube camera is achieved but with much higher reliability and greatly reduced lag, geometric distortion, and power consumption. An example is the Cohu 5512 generation I optical fiber coupled intensified CCD camera, sensitive down to 0.00015 lux and with 460 TV line horizontal resolution, currently costing about $12,000.

phosphor has a persistence of only 0.16 μs, while the purplish-blue P47 phosphor has a persistence of only 0.08 μs. However, both these phosphors have luminous efficiencies of only 3%. The resolution of an intensified camera is generally quoted in line pairs per millimeter (lp/mm) and is the harmonic mean of the individual resolutions of the intensifier (typically 20–30 lp/mm), the coupler (typically 80 lp/mm), and the image sensor (typically 20–50 lp/mm).

45.3.2 Generation II and III Tubes

Generation II intensifiers are similar to generation I devices except that gain is achieved using a microchannel plate (MCP) instead of an accelerating potential difference in vacuum (see Figure 45.8). Generation III intensifiers are similar to generation II except that the multialkali photocathode is replaced by a gallium arsenic solid-state structure. The resulting device has double the gain of a type II device and improved SNR and resolution. The heart of both types is an MCP disk about 25 mm diameter and 0.5 mm thick, consisting of an array of millions of glass tubes (called *microchannels*) with holes about 10 μm diameter coated with a secondary electron emitting substance. The two faces of the MCP are coated with conducting layers, allowing a voltage (typically 50–900 V) to be applied between the entrance and exit of each microchannel. An energetic electron from the photocathode enters a microchannel and produces secondary electrons on striking the wall. These electrons are accelerated by the axial electric field and, in turn, dislodge more secondary electrons. This process is repeated many times down the channel. The number of electrons continues to increase until either the end of the channel is reached or current saturation occurs because the electron cloud is so dense that it inhibits further secondary emission. The electrons exit the MCP and may be accelerated further before striking a phosphor screen. The light generated is then transported to the image sensor by fiber-optic couplers. A thick phosphor layer increases light gain but allows electrons and light scattering in the phosphor, which causes some loss of resolution. Improved performance is achieved by integrating the phosphor into the fiber-optic bundle in a so-called *intagliated construction*. Each glass core in the fiber-optic bundle is etched out to a predetermined depth, leaving a small cup of cladding glass at the entrance to each fiber, which is filled with phosphor. Electrons striking the phosphor produce light that is confined to one fiber,

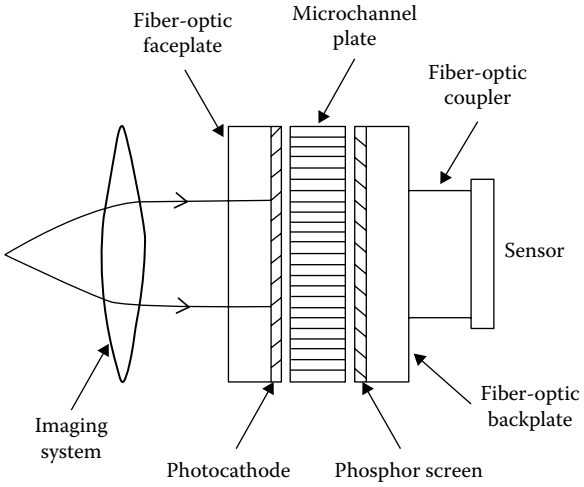


FIGURE 45.8 The basic components of a generation II or III intensified image sensor. These are typically two and four times more sensitive than generation I intensifiers. Narrow vacuum gaps between the photocathode, MCP, and phosphor screen achieve proximity focusing. An example is the Philips 1800 low-light video camera, which couples an XX1410 generation image intensifier with an FT800 1/2 in., 760 × 576 resolution frame-transfer CCD at a cost of around $8000.

minimizing cross talk and improving resolution. The overall light amplification achieved is of the order of 10^4, and higher gains can be achieved by stacking multiple MCPs and by increasing the accelerating voltages. The MCP has a nonlinear response, because the current density increases at an exponential rate until limited by saturation. The IIT can be gated in only a few nanoseconds by pulsing the cathode potential from the cutoff level (a few volts positive) to the normal operating voltage. This provides a high-speed shutter capability for stopping motion as well as a means of exposure control. Generation II devices have high gain and dynamic range and are easy to gate on and off, but they have relatively low output levels compared with generation I devices. The Kodak EktaPro 1000 high-speed video camera uses a generation II intensifier as a high-gain first stage coupled to a generation I intensifier to provide the required level of light output.

45.4 Fiber-Optic Scopes

It is sometimes necessary to acquire images within confined and inaccessible spaces. In some cases, this can be achieved with the aid of mirrors, but when the image must be conducted around bends, a fiberscope may be needed. This consists of a semirigid flexible sheathing around a bundle of optical fibers, each about 10 μm diameter aligned coherently so that fibers have the same relative location at entrance and exit of the bundle. An image formed on one end of the bundle therefore appears undistorted but dimmer at the other end. Between 20% and 35% of the bundle face area consists of fiber cladding and epoxy filler and does not transmit light. Light passing through fiber cores suffers Fresnel reflection losses of about 4% on entering and leaving and attenuation within the fibers of the order of 10%–15% for each meter of length. A 1 m fiberscope therefore transmits only about 40%–60% of the incident light. Fiberscopes range in diameter from about 2 to 18 mm and in length from about 0.6 to 3 m, with a minimum bend radius of about 75 mm for an 8 mm insertion diameter tube. To illuminate the object, a coaxial arrangement of fibers is often used, as in the flexible boroscope marketed by Edmund Scientific. The inner fibers are used for imaging, while fibers in the outer, noncoherent bundle are used to transport light to the object. Alternatively, a separate fiber-optic light guide may be used. Fiberscopes may be fitted with optics for viewing in the forward, oblique, side, and rearward directions but may not allow easy quantitative measurements of distances and areas because of distortions introduced by wide-angle optics. Fiberscopes are generally supplied with an eyepiece for direct observation, but C-mount couplers are often available to allow a CCD camera to be connected.

45.5 Components and Trends

A wide range of image sensors and cameras based on tube and solid-state technologies is now available, and a few examples are listed in Table 45.1 to give an indication of some of the currently available devices and their approximate prices, if purchased in the United Kingdom. It is advisable to contact local suppliers to determine actual prices and availability. Contact information for some companies is listed in Table 45.2. Manufacturers are continuing the incremental development of sensors with greater resolutions, readout rates, and signal-to-noise performances, and new application areas such as high-definition television (HDTV) and digital photography should cause prices to come down as the markets for these products expand. Some examples of this type of development are as follows:

- EG&G Reticon is currently developing a 1000 frames/s, 1024 × 1024 ILT CCD sensor for high-speed imaging.
- Thomson Components (France) has developed the THX-31163 CCD sensor specifically for use in HDTV. This chip is designed to output 1152 × 1260 pixel frames in 40 ms at a transfer frequency of 47 MHz.

TABLE 45.1 Cameras, Sensors, and Other Components

Manufacturer	Designation	Function	Specification
Cohu	4712	Monochrome TV camera	1/2 in. FT CCD, 754 × 484 pixels, 0.04 lux minimum, RS-170 output
Cohu	4110	Digital output camera	1/2 in. FT CCD, 755 × 484 pixels, 0.01 lux minimum, 8 bit per pixel RS-422 outputs
Cohu	4912	Monochrome TV camera	1/2 in. ILT CCD, 768 × 494 pixels, 0.02 lux minimum, RS-170 output
Cohu	2252	Color TV camera	1/2 in. microlens ILT CCD, CMYG mosaic filter, 0.3 lux minimum, NTSC/Y-C/RGB outputs
Edmund Scientific	A52999	Fiberscope	8 mm diameter, bend radius 75 mm, 60° field, 10–100 mm focus
EG&G Reticon	D Series	Line sensors	256, 512, 1024, or 2048 pixel CCPD arrays
EG&G Reticon	LC1911	Line scan camera	10 MHz data rate, D series sensors, RS-422 outputs
EG&G Reticon	RA1001	Area sensor	FF CCD, 1024 × 1024 pixels, dual outputs, 30 frames /s
EG&G Reticon	TD Series	TDI CCD array	1024 × 96 pixels, single output
Hamamatsu	S5464–1024Q	Spectroscopic line sensor	CMOS SSPD 1024 array, 2.5 mm by 25 μm pixels, quartz window
Kodak	DCS 410	Color digital camera	FF CCD, 1524 × 1012 pixels, mosaic filter, Nikon N90S camera
Kodak	DCS 460	Color digital camera	FF CCD, 3060 × 2036 pixels, mosaic filter, Nikon N90S camera
Panasonic	GP-US502	Remote head color camera	3 1/2 in. ILT CCDs, 768 × 494 pixels, 3 lux minimum, NTSC/Y-C/RGB outputs
Philips	XX1410	Intensifier	18 mm generation II
Philips	XQ3427	Tube	2/3 in. plumbicon, 400 TV lines at 55% modulation
Philips	XQ1270	Tube	2/3 in. vidicon, resolution 500 TV lines
VLSI Vision	Imputer 3	Intelligent camera	2/3 in. CMOS sensor, 512 × 512 × 8 bit, 32 bit RISC CPU
VLSI Vision	IDS	Development system	Software and hardware to create imputer 3 applications

- Manufacturers such as Kanimage, Kodak, and Leaf Systems supply digital cameras based on very high-resolution area array sensors. The Kodak professional range of color and monochrome digital cameras, for example, incorporate CCD full-frame-transfer sensors with resolutions ranging from 1024 × 1536 to 2036 × 3060 into standard 35 mm photographic cameras.

A separate trend is the development of smart cameras and smart sensors [14] for industrial and scientific applications. There are two basic motivations. One is to correct for imperfect behavior of sensors (such as pixel to pixel nonuniformity associated with MOS devices), and the other is to reduce the communications or processing bottlenecks associated with handling very large amounts of image data. A smart camera is a single-image acquisition subsystem incorporating image sensor, analog-to-digital conversion, and microprocessor that provides processed image data in a standard format such as SCSI or RS-423. Pulnix already supplies a series of smart line scan cameras that implement application functions such as object detection, size measurement, and go, no-go comparison. Another example is the imputer 3 [15] from VLSI Vision incorporating a complementary MOS (CMOS) area sensor with 512 × 512 digital resolution and an i960 32 bit reduced instruction set processor. Using VLSI technology, but not CCD technology, it is possible to integrate both image acquisition and low-level processing onto a single chip to achieve a kind of electronic retina. One way of constructing such chips

TABLE 45.2 Some Manufacturers of Image Sensors and Cameras

Burle Industries Inc.
Tube Products Division
1000 New Holland Avenue
Lancaster, PA 17601-5688
Tel: (800) 366-2875
Fax: (717) 295-6096

Dalsa Inc.
605 McMurray Road
Waterloo, Ontario, Canada N2V 2E9
Tel: (519) 886-6000
Fax: (519) 886-8023

Edmunds Scientific Company
International Dept.
Meg DiMinno
Barrington, NJ 08007
Tel: (609) 573-6263
Fax: (609) 573-6882

EG&G Reticon, Western Regional Sales
345 Potrero Avenue
Sunnyvale, CA 94086 4197
Tel: (408) 738-4266
Fax: (408) 738-6979

Hitachi Denshi (United Kingdom) Ltd.
14 Garrick Ind. Ctr.
Irving Way, Hendon
London NW9 6AQ
Tel: 0181 2024311
Fax: 0181 2022451

Panasonic Industrial Europe U.K.
Willoughby Road
Bracknell, Berks RG12 4FP
Tel: 01344 853087
Fax: 01344 853706

Pulnix America Inc.
1330 Orleans Drive
Sunnyvale, CA 94089
Tel: (408) 747-0300
Fax: (408) 747-0660

Thomson-CSF
Unit 4, Cartel Business Centre
Stroudley Road
Basingstoke, Hants RG4 0UG,
United Kingdom
Tel: 01256 843323
Fax: 01256 23172

Cohu Inc.
Electronics Division
5755 Kearny Villa Rd.
San Diego, CA 92123
Tel: (619) 277-6700
Fax: (619) 277-0221

Eastman Kodak Company
343 State Street
Rochester, NY 14650
Tel: (800) 235-6325
Internet: http://www.kodak.com

EEV Ltd.
Waterhouse Lane
Chelmsford, Essex CM12QU, United Kingdom
Tel: 01245 453652
Fax: 01245 492492

Hamamatsu Photonics (United Kingdom) Ltd.
2 Gladbeck Way
Windmill Hill, Enfield, Middx. EN2 7JA
Tel: 0181 367 3560
Fax: 0181 367 6384

NEC Electronics (United Kingdom) Ltd.
Sunrise Parkway
Linford Wd, Milton Keynes, Bucks. MK14 6NP
Tel: 01908 691133
Fax: 01908 670290

Philips Components Ltd.
Mullard House
Torrington Place
London WC1E 7HD
Tel: 071 5806633
Fax: 071 4362196

Sony Broadcast & Professional Europe
Image Sensing Products
Schipholweg 275
1171 PK Badhoevedorp
Amsterdam, the Netherlands
Tel: 020 658 1171
Tel: (United Kingdom) 01932 816300

VLSI Vision Ltd.
Aviation House
31 Pinkhill
Edinburgh EH 12 7BF, United Kingdom
Tel: 0131 5397111
Fax: 0131 5397141

is to integrate photodetector and processing element at the photosite, but this produces very poor fill factors. This can be alleviated to some extent by using microlenses to focus the light onto the sensitive region. An alternative approach is to implement the photodetection and processing as separate arrays on the chip. This allows a large fill factor but requires high-frequency data paths between the two arrays. The development of smart image sensors is still in its infancy, but nevertheless the creation of an artificial eye is now on the horizon.

References

1. T. N. Cornsweet, *Visual Perception*, New York: Academic Press, 1970.
2. CCIR, Characteristics of monochrome and colour television systems, Recommendations and Reports of the CCIR, Vol. XI, Part 1: Broadcasting Service (Television), Section IIA, 1982.
3. W. Booth, S. S. Ipson, and Y. Li, The application of machine vision and image simulation to the fabrication of knives used for cutting leather, in *Proceedings of the Second Asian Conference on Computer Vision* (*ACCV'95*), Singapore, pp. II574–II578, December 5–8, 1995.
4. R. K. Hopwood, Design considerations for a solid-state image sensing system, Reprint from *Proc. SPIE*, 230, 72–82, 1980; In *EG&G Reticon 1995/6 Image Sensing and Solid State Products*, Sunnyvale, CA.
5. S. F. Ray, *Applied Photographic Optics*, 2nd edn., Oxford, U.K.: Focal Press, 1994.
6. Depth of field characteristics using Reticon's image sensing arrays and cameras, Application note 127, in *EG&G Reticon 1995/6 Image Sensing and Solid State Products*, Sunnyvale, CA.
7. Practical radiometry, Application note in *Dalsa CCD Image Capture Technology, 1996–1997 Databook*, Dalsa Inc., Thousand Oaks, CA.
8. S. J. Lent, Pickup tubes and solid-state cameras, in K. G. Jackson and G. B. Townsend (eds.), *TV & Video Engineers Reference Book*, Part 3, Chapter 10 Oxford, U.K.: Butterworth-Heinemann, 1999.
9. L. Sheu and N. Kadekodi, Linear CCDs, advances in linear solid-state sensors, *Electronic Imaging*, Vol. 11, 72–78, August 1984.
10. S. C. Chamberlain and P. T. Jenkins, Capturing images at 1000 feet per minute with TDI, *Photonics Spectra*, 24(1), 155–160, 1990.
11. D. A. Rutherford, A new generation of cameras tackles tomorrow's challenges, *Photonics Spectra*, 23(9), 119–122, 1989.
12. A. Asano, MOS sensors continue to improve their image, *Advanced Imaging*, 42f–44f, 1989.
13. C. L. Rintz, Designing with image tubes, *Photonics Spectra*, 23(12), 141–143, 1989.
14. J. E. Brignell, Smart sensors, in W. Gopel, J. Hesse, and J. N. Zemel (eds.), *Sensors: A Comprehensive Review, Volume I: Fundamentals*, Weinheim, Germany: WCH Publications, pp. 1-16 1989.
15. O. Vellacott, VLSI Vision Review note, CMOS in camera, *IEE Review*, 40(3), 111–114, 1994.

Further Information

Burke M. W., *Image Acquisition: Handbook of Machine Vision Engineering, Vol. I*, London, U.K.: Chapman and Hall, 1996. Provides comprehensive coverage of lighting, imaging optics and image sensors.

Dalsa CCD Image Capture Technology, 1996–1997. Databook includes useful application notes and technical papers on image sensors and TDI.

46

Charge-Coupled Devices

John A. Nousek
The Pennsylvania State University

M.W. Bautz
The Pennsylvania State University

B.E. Burke
The Pennsylvania State University

J.A. Gregory
The Pennsylvania State University

R.E. Griffiths
The Pennsylvania State University

R.L. Kraft
The Pennsylvania State University

H.L. Kwok
The Pennsylvania State University

D.H. Lumb
The Pennsylvania State University

46.1 Introduction

Charge-coupled devices (*CCDs*) have become the detector of choice for sensitive, highly precise measurement of light over the electromagnetic spectrum from the near-infrared (IR) (<1.1 μm) to the x-ray band (up to 10 keV). Key advantages of CCDs over their predecessors (photographic emulsions and vacuum tube, electron beam readout devices such as vidicons and Silicon-intensifier target (SIT) tubes) are high quantum efficiency (QE), high linearity, large dynamic range, relatively uniform cosmetic response, low noise, and intrinsically digital image capture.

CCDs were initially designed as serial data storage media (an electronic analogy to the magnetic bubble memory units) in which charge packets were injected into linked capacitors to store data and read back by moving the packets back out of the device. When it was found that charge packets could be directly induced in the capacitors by exposing them to light, the CCD as light sensor was born.

Physically, *CCD operation* consists of four critical stages. First, an incident light photon must be photoabsorbed in the sensitive portion of the CCD chip (called the *depletion region*). At optical and IR wavelengths, the absorption results in a single electron being promoted into the conduction band (leaving a hole in the valence band); at shorter wavelengths, the photon has enough energy to make additional electrons via secondary ionizations by the photoelectron.

Second, the photon-induced electrons must be collected, via an electric field within the silicon, into localized regions near the front surface of the chip. The electric field is shaped by implanted dopants and by electric potentials applied to thin conducting strips (*gates*) that prevent the electrons from diffusing away. The resulting charge distribution corresponds to an electronic analog of the light intensity pattern shone on the CCD. The resolution of this pattern is governed by the size of the potential wells, which are designed to be periodic. Each well is called a *pixel* and corresponds to the minimum picture element detected by the CCD.

Third, after exposure is completed, the CCD charge pattern must be transferred out of the CCD. This is accomplished by modulating the potential applied to the CCD gates in such a way that no charge packets are mixed, but that each packet moves into the next pixel. The end pixel is transferred into a special pixel array called the *serial register*. Each movement of charge resulting from gate potential changes is called a clock cycle, and the serial register receives many clock cycles for each cycle of the full pixel array. The net result is a sequence of charge packets emerging from the serial register, each of which is directly proportional to the amount of light striking a particular location on the CCD.

Fourth, the emerging charges are converted into electric signals by a charge-sensitive preamplifier on the CCD chip. These signals are often digitized by electronics in the camera immediately outside the chip, but analog readouts that produce signals compatible with video standards are also used (the popular handheld video cameras are examples of this). Research-grade camera readouts are able to measure the charge pulses with accuracies as good as one or two electrons rms, if the CCD and electronics are cooled.

Currently available CCDs carry out these steps so well that they are nearly the ideal detector for precision low-light-level applications, especially in astronomy. Such an ideal detector would have perfect QE (i.e., convert every incident photon into detectable signal), no noise, unlimited dynamic range, linearity in response to incident intensity and position, and completely understandable characteristics.

CCDs have high QE because photons interact via photoabsorption in the depletion layer, which directly results in one or more electrons promoted into the conduction band of the silicon lattice, and are very efficiently collected by the CCD. The main obstacles to perfect QE are absorption of photons by the gate and insulator materials before they ever reach the depletion regions (or optically, by reflection off the front surfaces) or if the photon passes entirely through the depletion region without interacting.

There are many approaches to enhancing *CCD QE* for various applications. In soft x-ray and ultraviolet (UV) wavelengths, the gate and insulator layers on the front of the CCD absorb too much light. To solve this, CCDs are built with thin gates or thinned substrates and backside illumination. CCDs are also coated with phosphor coatings that down-convert UV light to longer wavelengths where the gate transmission is higher.

At hard x-ray and IR wavelengths, too much light can pass through the depletion region without interacting at all. The depletion region is the part of the CCD pixel that is swept clean of free charges during the readout process. The depletion region gets deeper if higher purity silicon is used and if higher voltage biases are applied during the readout.

Above 1.1 μm, photons do not have sufficient energy to promote electrons into the silicon conduction band, so other materials, such as germanium or a compound semiconductor such as InAs, InSb, or HgCdTe, must be used.

CCD noise results from four major factors: (1) thermal background noise, (2) charge transfer imperfections, (3) charge-to-voltage amplification noise, and (4) cosmetic imperfections in the CCDs due to, for example, microscopic shorts in the insulating layers of the CCD. Factor (1) results from a "dark current" of thermally excited electrons that accumulate in the pixels and can be eliminated by cooling the CCD (typically to −60°C to −120°C). Factor (2) results from traps that hold electrons long enough to

shift them into following charge packets identified with other pixels. Factor (3) is a fundamental limit related to the temperature and capacitance of the output mode (kT/C), but it can be suppressed by signal processing techniques such as "correlated double sampling" to the equivalent of a few electrons (1–5 in state-of-the-art cameras). Factors (2) and (4) can be greatly reduced by improved manufacturing technique, especially scrupulous contamination control during the process.

CCD dynamic range is set by the maximum charge packet that can be stored in a pixel. Termed "full-well capacity," this is set by the depth of the potential well. When the full-well capacity is exceeded, the image of a point source "blooms" as a result of charge leaking into surrounding pixels, and a trail of brighter pixels forms in the readout direction of the CCD due to charge incompletely transferring from pixel to pixel during a clock cycle. Modern CCDs have full-well capacities in excess of 10^5 electrons and can be designed even larger. (Note that larger full well also requires larger output capacitance, so a trade-off is generally required between blooming and low noise.)

CCD linearity in intensity response and position response is very good because the conduction band in the CCD has so many states that the very small injected photocharge does not affect subsequent photon interactions. The position linearity results from the photolithography of the manufacturing process, which must be accurate to less than 1 μm. The primary limitation on linearity results from imperfect *charge transfer efficiency* (CTE) in the process of clocking charge packets from pixel to pixel. At readout rates below 100,000 pixels/s, CTE imperfections have four causes:

1. *Design imperfections*: errors in CCD design can leave potential minima that are incompletely drained during clocking.
2. *Process-induced traps*: random cosmetic defects, presumably due to imperfections in manufacturing.
3. *Bulk traps*: lattice defects or impurities that introduce local potential minima, which temporarily capture electrons long enough to remove them from the original charge packet, but reemit them later.
4. *Radiation-induced traps*: similar to (3) but resulting from lattice defects caused by low-energy protons. This damage is most commonly seen by spacecraft CCD cameras.

46.2 CCD Operation and Data Reduction

In order to achieve ultimate *CCD performance* for a given goal, the CCD camera can be operated in special ways, and the postcamera data reduction can be optimized. Typical optical use of CCDs involves timed exposures, where the CCD pixels are exposed to light and the total charge integrated in pixels for a preselected time. At the conclusion of the integration, a shutter closes and the CCD is read out. As noise reduction limits the readout to about 100 kpixels/s, a large CCD (2048 × 2048 pixels) readout can take many seconds to complete.

To avoid the dead time associated with the closed shutter, some CCDs are made with framestore regions. The framestore is a pixel array equal to the integration region that is permanently blocked by a cover from any additional light. The pixels containing the charge pattern resulting from the integration are very rapidly clocked into the framestore region (typically requiring much less than 1 s) and then slowly clocked out into the readout region without moving the integration pixels.

Operationally, CCD reduction requires calibration exposures. These include bias frames, which are readouts with the same integration time but no light striking the CCD, and flat-field frames, which have a uniform illumination over the CCD. The bias frames are subtracted from the data frames to set the zero point corresponding to zero incident radiation. (Note that CCDs will accumulate charge due to thermal electrons and low-level shorts in the gates, even if no light hits the CCD.) The flat field allows correction for pixel-to-pixel sensitivity variations. Proper flat fielding can remove variations of arbitrary amplitude and spatial scale.

The CCD dark current bias can be reduced by cooling the CCD or by operating it in an inverted phase mode. In inverted phase operation, the gate electrode is given a suitable negative bias that attracts hole

carriers to the front surface of the CCD. These holes fill interface states at the Si–SiO_2 boundary between the conducting depletion region and the insulating layer under the gates. Suppression of these interface states dramatically lowers the dark current because the interface states are much more efficient at thermal electron promotion to the conduction band than the bulk material. Not all phases can be operated in inverted mode in a normal CCD because, without the restraining potentials provided by gates held at positive voltage, the pixel charge packets can intermingle. A special CCD called a multiphase pinned (MPP) device has extra implant doping that isolates the pixels even with all three phases inverted, yielding dark current so low that integration times up to minutes become possible in room-temperature MPP CCDs.

Other important uses of CCDs include cases where the CCD is continuously clocked, without any shutter. Suitable for high-light-level conditions, the effective integration time becomes the time to transfer a pixel charge across the source point spread function (PSF) on the CCD. This allows sensitive timing of source intensity changes.

A similar technique is called *drift scanning*, where the rate of clocking of pixels equals the rate of motion of the target across the CCD. Such a condition is common in astronomy, where a fixed detector on the Earth sees slow motion of stars in the field of view due to the Earth's rotation.

Drifts and instabilities in the camera electronics can be corrected using a technique called "overclocking." If the serial registers are clocked more times than there are physical pixels in the CCD, then the excess clocks will produce charge pulses corresponding to zero input light and zero dark current. The distribution of the overclock pulse is then a measure of the readnoise of the CCD chip-camera system, and the mean of the distribution sets the zero point of the energy to output voltage curve. Frequently, CCD cameras subtract the mean of the overclock pixels from all output values in a row (called "baseline restoration").

46.3 CCD Signal-to-Noise Ratios

To see how these characteristics of the CCD relate to measurement, it is instructive to study the signal-to-noise ratio (SNR) predicted for a given exposure time. In a single pixel illuminated by a source that contributes S_o counts (electrons) to the pixel, one also sees contributions from dark current (S_d) and background illumination (S_s, usually called the "sky" in astronomical usages), all in units of counts per pixel per second. The source contribution (S_o) can be expanded into the intensity of light from the source, I; the QE of the CCD, Q; and the integration time, t, to provide

$$S_o = I \times Q \times t \tag{46.1}$$

The camera readout contributes a randomly distributed but fixed Gaussian noise with variance N_r. The SNR of a particular pixel is then

$$\text{SNR} = I \times Q \times t = (I \times Q \times t + N_r^2 + S_d + S_s)^{1/2} \tag{46.2}$$

If the light from a source is distributed over a number of pixels, n (as might arise from a star viewed through a telescope with a PSF covering n pixels), then if the integral of S_o over the PSF is C_o and the integral of S_s over the pixels is C_s, then

$$\text{SNR} = \frac{C_o^{1/2}}{(1 + C_s C_o + n + r^2 C_o)^{1/2}} \tag{46.3}$$

Clearly, high Q and low r are desirable, and t should be chosen to make C_o greater than both C_s and r. It is worth noting that, in most optical applications (except for extremely faint sources), it is the uncertainty in the flat fielding (i.e., the corrections made for pixel-to-pixel sensitivity variations and background) that ultimately limits the achievable SNR.

46.4 CCD Structure and Charge Transport

46.4.1 CCD Structure and Its Potential Profile under Bias

A CCD is a semiconductor device operating under the principle that charges can be temporarily stored and transported along a string or array of metal oxide semiconductor (MOS) capacitors. The basic storage unit is called a pixel and is made up of several MOS capacitors. In almost all CCDs, charges are stored either directly at the oxide–semiconductor interface (surface-channel device) or deeper within an epitaxial layer (buried-channel device). Theoretically, a surface-channel CCD has a larger charge capacity, but it also is prone to noise arising from interface states at the boundary.

The CCD is operated by varying voltages applied to the surface electrodes. Typically, the CCD is kept for a long period in an integration state, where photon-induced electrons accumulate in the potential wells under the CCD pixels. After the integration finishes, the voltages are changed to transport the charge from one capacitor to the next. This sequence of moving charge packets by potential clocking is sometimes called "bucket-brigade" charge transfer.

The storage unit of the CCD is the MOS capacitor, and it is possible to deplete, invert, or form a surface accumulation layer in the MOS capacitor by simply changing the surface potential. CCDs operate in the so-called "deep depletion" mode when the surface layers are fully depleted. For a buried-channel CCD with an *n*-type buried layer within a *p*-type epilayer, this would require the application of a positive bias to the surface electrodes. The equations governing the 1D calculations of the potential distribution in the MOS capacitor are [1] as follows:

$$\begin{aligned}
&\frac{d^2\psi}{d_x^2} = 0 \quad -d < x < 0 \\
&\frac{d^2\psi}{d_x^2} = -q \cdot \frac{N_D}{\varepsilon_s} \quad 0 < x < t \\
&\frac{d^2\psi}{d_x^2} = q \cdot \frac{N_A}{\varepsilon_s} \quad t < x < x_p
\end{aligned} \tag{46.4}$$

where

ψ is the potential
N_A is the substrate acceptor density
N_D is the donor density in the epilayer
x_p is the depletion edge in the *p*-type substrate
ε_s is the semiconductor permittivity

Note also that d is the oxide thickness and the origin ($x = 0$) is located at the oxide–semiconductor boundary.

The boundary conditions will be

$$\begin{aligned}
&\psi(x = -d) = V_G \quad \left.\frac{d\psi}{d_x}\right|_{x=t^-} = \left.\frac{d\psi}{d_x}\right|_{x=t^+} \\
&\varepsilon_{ox} \cdot \left.\frac{d\psi}{d_x}\right|_{0^+} = \varepsilon_s \cdot \left.\frac{d\psi}{d_x}\right|_{0^+} \quad \psi(x = t^-) = \psi(x = t^+) \\
&\psi(x = 0^-) = \psi(x = 0^+) \quad \psi(x = t + x_p) = 0
\end{aligned} \tag{46.5}$$

and the solutions are

$$\psi = V_G - E_{ox} \cdot (x + d) \quad -d < x < 0$$
$$\psi = \psi_{max} \frac{-q \cdot N_D \cdot (x - x_n)^2}{(2\varepsilon_s)} \quad 0 < x < t \tag{46.6}$$
$$\psi = \frac{q \cdot N_A \cdot (x - t - x_p)^2}{(2\varepsilon_s)} \quad t < x < t + x_p$$

where

E_{ox} is the electric field in the oxide layer
V_G is the gate bias voltage
x_n is the position of the potential maximum

Figure 46.1 shows a typical potential profile across the MOS capacitor in a buried-channel CCD. Note the presence of a potential maximum, ψ_{max}, where the electrons will reside. It is given by

$$\psi_{max} = \psi_J \cdot \left(1 + \frac{N_A}{N_D}\right) \tag{46.7}$$

where $\psi_J = q \cdot N_A \cdot x_p^2/(2\varepsilon_S)$.

Since x_p increases linearly with N_D, the potential maximum ψ_{max} and hence the charge storage capability also increase for a heavily doped epilayer. The derivation of the potential profile in a surface-channel CCD will be similar, but with d set to zero.

46.4.2 Charge Transport

Charge transport in a CCD refers to the transfer of charges along the MOS capacitors (i.e., from one pixel to the next). There are several clocking schemes used, generally divided according to how many external voltage regions are applied per pixel, ranging from one (uniphase) to four phases.

Computation of charge transport in a CCD requires solving the 2D Poisson equation with appropriate boundary conditions to obtain the potential distribution within the pixel for each separate phase

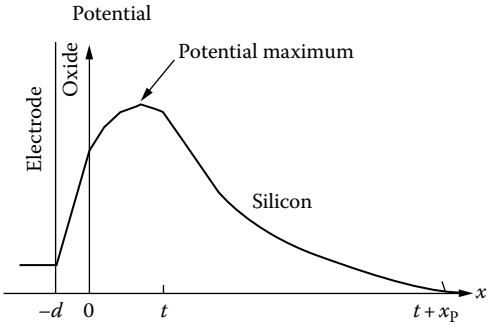


FIGURE 46.1 Potential profile across MOS capacitor.

within the clocking cycle. Numerical techniques using finite-difference or finite-element methods have been used. For charge transfer in an *n*-type epilayer, the electron flux, $F(x)$, is given by

$$F(x) = n_s(x,t) \cdot \upsilon(E(x)) - D(E(x)) \cdot (n_s(x + \Delta x, t) - n_s(x,t)) = \Delta x \tag{46.8}$$

where

$n_s(x, t)$ is the electron density
$E(x)$ is the electric field in the x-direction
$\upsilon(E(x))$ is the field-dependent velocity
$D(E(x))$ is the electron diffusivity

Figure 46.2 shows the time evolution of charges along a CCD. Some "smoothing" of the output charge profile is often observed, and, for high-speed devices, velocity saturation will be important. Computations of charge transfer are sometimes carried out using an equivalent circuit model for the CCD in a Simulation Program with Integrated Circuit Emphasis (SPICE)-type simulator. It has been shown to offer both faster computation time and the ability to include external support circuits into the model.

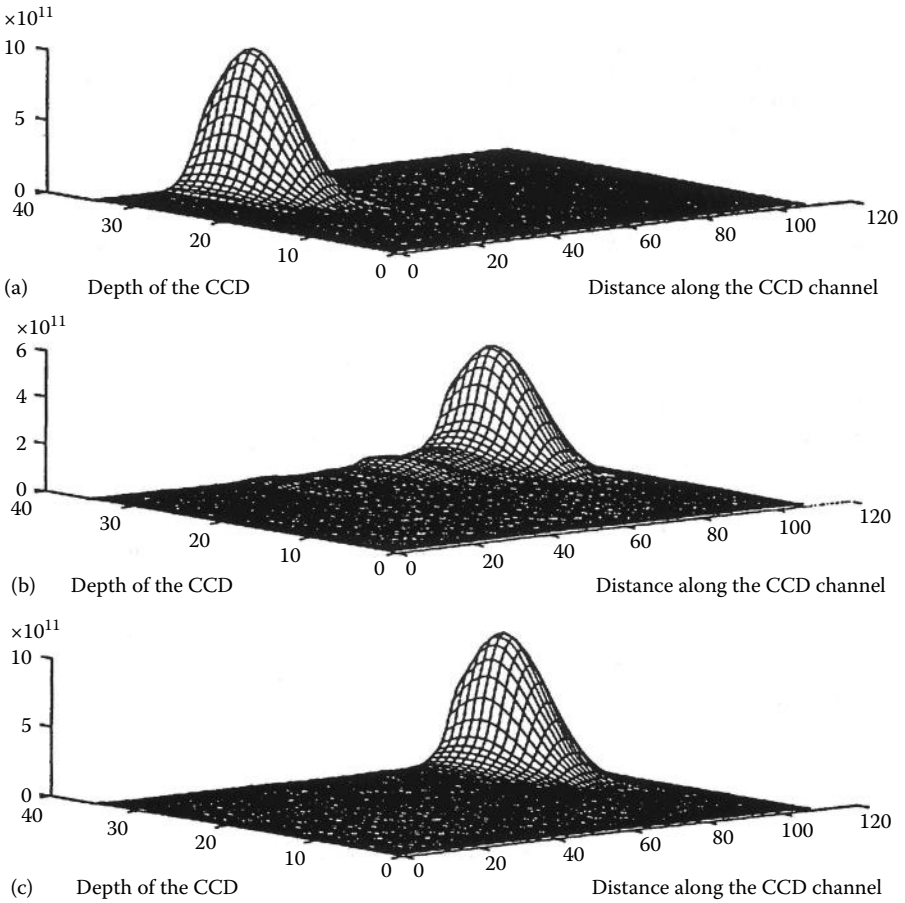


FIGURE 46.2 Time evolution of charges along a CCD: (a) t = 200 ps, (b) t = 400 ps, and (c) t = 600 ps.

CTE, η, is the key figure of merit in a CCD, at least comparable to QE. A common technique to determine η is to measure the amplitude response of the CCD using a network analyzer. η is then related to the amplitude response of the output, G_V, through the following relationship:

$$G_V = \exp\left\{-N_T \cdot (1-\eta) \cdot \left[1 - \cos\left(\frac{2\pi f}{f_c}\right)\right]\right\} \tag{46.9}$$

where
N_T is the number of transfers
f is the signal frequency
f_c is the clock frequency

For the relatively mature Si CCD technology, CTE of 0.99999 can be easily achieved. Note, however, that this analysis applies to sinusoid illuminated scenes and test signals, but does not apply to delta function injections of charge that are infrequently transferred, as applies to the x-ray photon-counting case.

46.4.3 Input and Output Structures

The input structure of the CCD usually consists of a *p–n* junction and two or more input gates as shown in Figure 46.3. Initially, when the input voltage V_{id} is pulsed low, charges will flow through the input (electrode) into the potential well created under the input gates G_1 and G_2. When V_{id} returns to its equilibrium value, charges will be withdrawn except those residing in the potential well formed by the potential difference between G_1 and G_2. Charge injection is now complete and a drop in the potential ϕ_1 at the transfer gate will allow charges to enter into the first CCD pixel. The input charge, Q_{in}, is approximately given by

$$Q_{in} = C_{ox} \cdot (V_{G_1} - V_{G_2}) \tag{46.10}$$

where
C_{ox} is the oxide capacitance
V_{G_1} and V_{G_2} refer to the voltage difference between G_1 and G_2

In CCD sensors and detectors, signal input normally relies on the incident photons liberating charge within the depletion region (after traversing either the front gates [front-illuminated] or the back surface [back-illuminated]). Charges so created will be collected in the potential wells formed by the clock voltages.

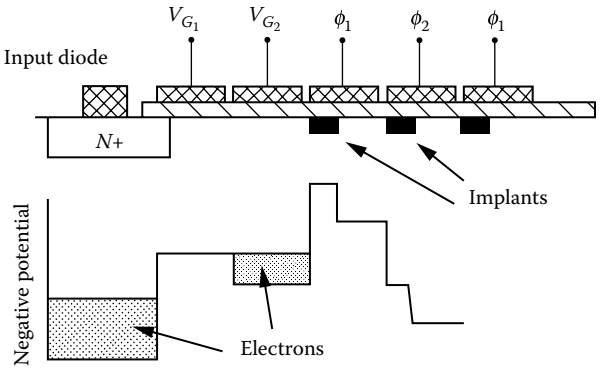


FIGURE 46.3 Input structure of a CCD.

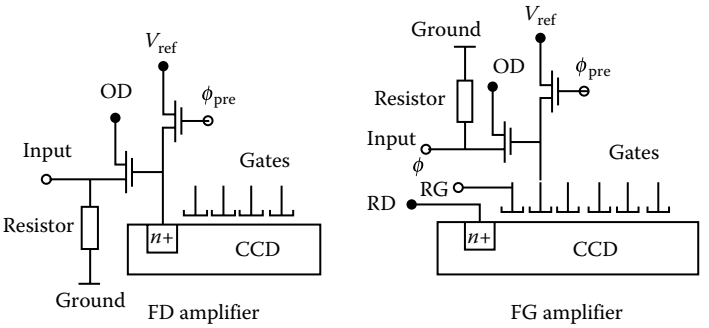


FIGURE 46.4 Output structures of a CCD.

Normally, the output structure of the CCD consists of a reset switch and a floating-diffusion amplifier (FDA) or floating-gate amplifier (FGA). These are shown in Figure 46.4. During sampling, the output charge flows into the floating node and a voltage is developed across the output capacitance, which is a combination of the depletion capacitance of the floating node and the input gate capacitance of the source follower. In general, the voltage output, V_{sig}, can be expressed as

$$V_{sig} = \frac{Q_{inj} \cdot A_v}{C_{fd}} \tag{46.11}$$

where

Q_{inj} is the injected charge
A_v is the small-signal voltage gain of the output amplifier
C_{fd} is the capacitance between the floating node and ground

Since both A_v and C_{fd} are sensitive to the output charge and temperature, this type of output structure is not generally linear. Improvements in the output structure can be achieved using either an FGA or Miller feedback at the output to reduce parasitic capacitances.

46.4.4 Noise in CCDs

CCD noise will degrade the SNR. The following are important noise sources in CCDs:

1. Thermal noise, or dark current, is due to thermally excited electrons that accumulate in the pixels during the integration. In applications where the clocking time is short compared to the integration time, the equivalent noise election, n_{th}, is given by

$$n_{th} = \sqrt{J_d \cdot A_E \cdot t_{int}} \tag{46.12}$$

where

J_d is the leakage current density
A_E is the area of the transport electrode
t_{int} is the integration time

A lowering of the operating temperature will normally reduce the leakage current density and hence the equivalent noise electron.

2. Bulk traps are the dominant noise sources for buried-channel CCDs, and the equivalent noise electron from this source is

$$n_{\text{trap}} = \sqrt{N_{\text{t}} \cdot V_{\text{sig}} \cdot Nt \cdot e^{-T_{\text{t}}/\tau_{\text{e}}}}\,[1 - e^{-T_{\text{t}}/\tau_{\text{e}}}] \tag{46.13}$$

where
V_{sig} is the volume of the charge packet under the transfer electrode
N_{t} is the density of the bulk trap states
T_{t} is the transfer time
τ_{e} is the electron emission time constant

It can be observed that the equivalent noise electron increases as the size of the charge packet increases, an indication that bulk traps do not generally affect the SNR in CCDs. In x-ray photon-counting applications, some effects of bulk trapping noise have been seen.

3. Both the input and output of a CCD involve charge storage across capacitors that will be affected by fluctuations in the circuit. The equivalent noise electron for this process is

$$n_{\text{inp}} = \sqrt{kT \cdot C_{\text{inp}}} \quad n_{\text{out}} = \sqrt{kT \cdot C_{\text{out}}} \tag{46.14}$$

where
kT is the thermal voltage
C_{inp} and C_{out} are the input and output capacitances

A special signal processing technique known as "correlated double sampling" can be used to reduce this kTC noise at the CCD output.

$1/f$ noise is also present in the output amplifier and is only significant at low frequencies. Methods to reduce $1/f$ noise include multiple sampling using distributed floating-gate amplifiers (DFGAs) [2] and noise cancellation techniques using alternate gain inversion (AGI) [3].

4. The maximum SNR of a CCD is limited by the maximum amount of charge that can be stored in a pixel, divided by the equivalent noise electron. Expressed in decibels, it is

$$\text{SNR} = 20\log(Q_{\text{max}} = Q_{\text{noise}}) \tag{46.15}$$

The maximum charge is

$$Q_{\text{max}} = C_{\text{e}} \cdot \psi_{\text{BH}} \tag{46.16}$$

where
C_{e} is the depletion capacitance associated with the electrode
ψ_{BH} is the depth of the potential well

46.4.5 CCD Power Dissipation

CCD power dissipation is primarily linked to the charging and discharging of the gate capacitances. This power is mainly consumed by the clock driver and is given by

$$P_{\text{clock}} = C_{\text{c}} \cdot V^2 \cdot f_{\text{c}} \tag{46.17}$$

where
C_{c} is the clock line capacitance associated with the electrodes
V is the clock voltage
f_{c} is the clock frequency

For f_{c} = 100 MHz and V = 5 V, typical values of P_{clock} are 50 μW per pixel.

In addition to P_{clock}, the on-chip power dissipation per pixel is approximately given by

$$P_{CCD} \approx \left(\frac{nq(f_c L)^2}{\mu_e + 2nqf_c} \right) \left(\frac{\psi_{BH} - nq}{2C_e} \right) \tag{46.18}$$

where

n is the number of electrons in a pixel
L is the pixel length
μ_e is the electron mobility

The overall power dissipation in a CCD is therefore the sum of P_{clock} and P_{CCD}.

46.5 CCD Applications to Light Sensing

46.5.1 Optical Imaging and Spectroscopy

The many fine properties of CCDs have made them the detector of choice for the recording and readout of images at both high and low light levels. The applications include (1) reconnaissance, both civilian and defensive, (2) scene and personnel monitoring, (3) robotics, and (4) astronomy. In astronomy, CCD cameras have found their way into almost all observatories in the world, both for direct readout of images and as the readout cameras for spectrographs.

Perhaps the most exacting application of CCDs has been to astronomy, a discipline in which the characteristics of the CCD need to be pushed to their limits in order to gain the optimum performance. As well as the attempt to optimize this performance in order to gain as much science as possible from the images, it has also been crucial to calibrate the devices as precisely as possible and to fully understand their behavior. In this sense, astronomy has been the "driver" for the scientific development of CCDs. Large-scale commercial applications, such as video recorders and monitoring cameras, have driven the need for improved manufacturing yield, blemish-free operation, uniformity in performance, and reliability.

In order to illustrate the importance of CCDs to exacting scientific requirements, this section will focus largely on the applications to astronomy.

The early demonstration of the performance of a commercial CCD at the focus of a telescope quickly led to the realization of their potential for both space and ground-based astronomical cameras and instruments. Fortunately, this development in the early 1970s was just in time for the first major optical observatory in Earth orbit, the Hubble space telescope (HST). Perhaps the most spectacular application of CCDs to visible light imaging has been within the cameras on board this telescope, that is, the wide field and planetary cameras built at the Jet Propulsion Laboratory. The CCDs in these cameras consisted of 800 × 800 pixel arrays, with each pixel of size 25 μm. The first version of this camera contained eight thinned, back-illuminated devices that were more or less fully depleted. An accumulation layer at the back surface was achieved by flooding the devices with UV light from a lamp internal to the instrument. This accumulation layer allowed the collection of charge resulting from the photoelectric absorption of blue light and accelerated the charge to the front-side potential wells. Furthermore, the devices were coated in a down-converting coronene phosphor that converted UV light into the yellow-green wavelength band, thus allowing a 20% efficiency in the UV. These devices thus had relatively high QE from the UV through to about 1 μm, vitally important for an instrument collecting photons from galaxies billions of light years away in the universe. As well as their high QE, these devices had high CTE and relatively low readout noise at the time of their development. All of these characteristics have since been superseded—at first by an improved version of the camera installed in the HST in late 1993 and then in the Advanced Camera for Surveys, to be

installed in 1999. These later instruments had larger format, higher QE, and lower readout noise. During the 1980s, devices were developed with these large formats and with readout noise levels approaching one electron.

Starting with modest arrays of size on the order of a few hundred pixels in the 1970s, the devices that became available in the 1990s were as large as 4096 pixels square. As well as in space, these devices have been employed at the prime foci of the world's largest telescopes to give images covering more than 10 arc minutes on a side while still sampling the atmospheric-limited resolution (arcsecond or subarcsecond) adequately. Such devices, when exposed to the sky for up to an hour through a broadband filter, can produce images with tens of thousands of objects for statistical studies in astronomy.

Typical pixel readout rates for astronomical cameras are of the order of 50 kHz, taking tens of seconds for a full readout. Integration times for astronomical CCD cameras can vary from fractions of a second to tens of minutes or even times in excess of an hour, usually limited by the background noise induced by cosmic-ray particles and their secondaries. These cosmic-ray "events" can be removed by taking at least two frames of data and cross-comparing them. This is important even on the ground, where the radioactive background and muon-induced events total about one per cm^2 per minute; but the cosmic-ray background in space is about one per cm^2 per second, so that camera exposures are rarely longer than 10–20 min. Sophisticated computer algorithms have been developed in order to remove the effects of this background radiation.

Of special importance to applications in space research is the packaging of the devices. Contamination of the cold CCD surface by as much as a monolayer of a heavy molecule will effectively render the CCD useless in the vacuum UV, for example. Careful attention must therefore be paid to the local instrument environment. The CCDs on the HST were in hermetically sealed packages, and the contamination problem was transferred to the lenses covering the CCD packages. In order to achieve optimum performance, the overall camera has to be designed to satisfy the exacting requirements of the CCD, from the point of view of thermal control and stability, absence of electronic interference, and rigorous attention to the elimination of water or heavy molecule contamination. For space application, it may also be advantageous to surround the CCD package with a cosmic-ray shield, such as the 1 cm (0.4 in.) of tantalum used in the CCD cameras on the HST and the Galileo mission to Jupiter. Such radiation shields have to be designed with caution, lest they introduce more secondary particles and induced radioactivity than the primary protons that they stop.

Finally, much effort is put into the processing of astronomical CCD data [4]. Special techniques have been developed to calibrate the pixel-to-pixel nonuniformity in QE (a function of wavelength), both on large and small scales. With some effort, this nonuniformity can be calibrated to levels below 1%. Although it may be easy to expose the CCD to a diffuse source such that there are at least 10,000 electrons per pixel, the corresponding flat-field calibration accuracy may be compromised at low light levels. For astronomical applications, especially in space, most of the pixels may receive only tens of electrons or fewer during an exposure. Low-level traps (including those caused by cosmic radiation) will then manifest themselves in the form of charge transfer inefficiencies over localized areas or columns. It may be important to calibrate the CCD at the same exposure levels as those typically encountered during the science observation. One way of improving this is to use hundreds or thousands of frames of data taken in different parts of the sky and to reject the astronomical objects in them. These "sky frames" can then be used to produce a "super sky flat," which is the average of the individual frames of data.

46.5.2 CCD X-Ray Imaging Spectroscopy

Scientific applications such as astronomy have been responsible for driving improvements in CCD technology. In the x-ray domain, for example, CCDs are starting to be employed for medical radiography, where the digital imaging capability and high sensitivity allow for lower patient doses and online image processing. As high-resolution readout detectors of dispersive x-ray spectrometers, CCDs are becoming

widely used in the new generation of high brilliance synchrotron beamlines. As spectrometers, they are also being considered in some applications as replacement for Si(Li) diodes, where the higher operating temperatures and improved resolution and efficiency down to x-ray energies below 1 keV lead to applications of interest to biological science and technology.

In contrast to optical imaging applications with many photons per charge packet, to measure x-ray spectral information directly with CCD, single photons per pixel per image frame are required. To use the CCD as a nondispersive x-ray spectrometer requires that the correspondence between the magnitude of the electron charge packet generated initially by the x-ray photon and the signal measured at the CCD output must be maintained. This places a very stringent requirement on the efficiency of charge transfer. However, even before the process of charge transfer is initiated, the physics of the charge collection may also degrade the energy measurement process.

The absorption process for a photon of energy E begins with the ejection of a photoelectron, of energy $E - E_B$, where E_B is the binding energy of the appropriate silicon atom electron shell. The range of the photoelectron may be only a fraction of 1 μm. In a few percent of cases, a silicon K shell fluorescence photon may be emitted, and this has a range of about 10 μm, so that there is a finite probability of the event energy splitting into more than one pixel.

Eventually, a proportion of the energy is converted to free electrons and the rest into phonons. The average energy to create an electron–hole pair is roughly constant at about 3.6 eV of incident photon energy per pair produced in silicon. If this free electron charge packet is created in the depletion layer of the CCD, it is promptly drifted under the influence of the electric field to the buried-channel collection site. During this drift time (t_d), the charge can laterally diffuse within a radius $\sim\sqrt{2Dt_d}$, where D is the diffusion constant. Except for the charge clouds originating deep in the depletion layer, this radius is small compared with the typical pixel size. If the charge is generated in a field-free layer *outside* the depleted volume, it will radially diffuse until a fraction recombines or reaches the depletion layer. The latter fraction then starts to drift with the same additional lateral drift as the depletion layer charge packets.

The pixel boundaries are loosely defined by the asymmetric fields created by electrode biasing schemes and surface-channel stop implants, and not hard physical boundaries. Hence, any lateral spreading processes may allow some splitting of the initial charge packet between pixels. Furthermore, charge *loss* may be experienced either by the partial recombination of events when liberated deep below the depletion layer or if the fraction of split charge is too low to be recognized against the device noise level.

In astronomical applications, in space-borne observatories, there is a continuous low-level background of charged particles. Rather than liberating point-like charge clouds, they liberate a population of signal electrons along their tracks throughout the silicon. They may be easily discriminated against x-rays if these tracks are highly skewed with respect to the silicon surface and cross many pixels. Also, if the track is long enough, the magnitude of total charge cloud generated may be large compared with the typical x-ray charge packet.

These features impose a requirement to perform event recognition and analysis. To perform this on ground would require transmission of all pixel data, including empty pixels, which for megapixel CCD frames generated on second time scales are quite infeasible. Performing this recognition on board first requires a comparison against some lower threshold. Selection of this level is critical—too high—and some split charge may be neglected with a consequent degradation of energy resolution. Too low a threshold and many spurious events will be counted. If there is a drift or change in the zero-energy signal upon which this threshold is applied, the energy scale may be misregistered, and/or the relative fraction of events selected at different x-ray energies may be unknowingly altered. Simulations show that for the potential energy resolution of CCDs, spectral analysis of cosmic plasmas will demand a calibration of relative detection efficiency versus energy to ~3%. If the threshold value is set at about 4σ times the readout noise, this calibration may be degraded by as little as a single digital bit of a commonly used 12 bit ADC in the CCD readout electronics. Thus, the real-time digital processing of events is required to be fast and complex.

A further complication of the event splitting is that to avoid pileup in CCD frames, the probability of having multiple events per pixel is made more stringent by requiring surrounding pixels to have no signal and allow this recognition process. Typically, an event rate of 1 photon per 200 pixels per CCD frame is therefore imposed, but this places a rather low limit on attainable count rate performance. Especially with high-resolution focusing optics of observatories such as AXAF [5], this can be much lower than for many previous experiments. If the core of the PSF of the mirror is not to be degraded by this effect, then special readout formats that reduce the imaged area, in order to accelerate frame readout times, must be employed.

46.6 Future Improvements to CCDs

46.6.1 Backside CCDs

The useful spectral range of conventional CCD detectors, although quite broad, is limited in part by the presence of the polysilicon gates on the front surface. As is illustrated in Figure 46.5, photons with wavelengths in the range ~3 nm < λ < ~400 nm have attenuation lengths in silicon less than the typical thickness of the gate structure (~0:4 μm). Photons in this spectral range are therefore absorbed before they can enter the photosensitive volume of the detector. For conventional CCDs, the detection efficiency is no more than a few percent in this band. If the gates can be made thinner by an order of magnitude or avoided altogether, then the detection efficiency can be improved dramatically in both the very soft x-ray and the UV. Efforts following the former approach are described later. The latter strategy, which requires illumination of the back surface of the CCD (i.e., the surface opposite the gates), is discussed here.

There are a number of techniques that can be used to produce back-illuminated devices [6–8], but they must all deal with two constraints: lateral diffusion of the photoelectrons before they are captured in the buried channel of a specific pixel and the tendency for photoelectrons to recombine at the back surface of the Si.

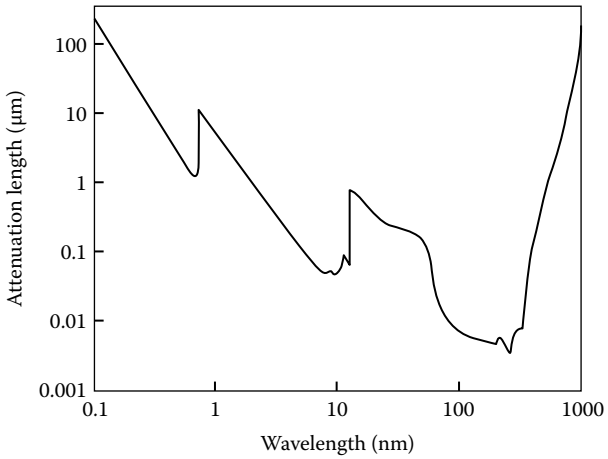


FIGURE 46.5 Attenuation length in silicon in the soft x-ray through near-IR spectral range. The polysilicon gates of front-illuminated devices, which are typically ~0.4 μm thick, are strongly absorbing at wavelengths for which the attenuation length is less than the gate thickness, viz. wavelengths (~3 mn < λ < ~400 nm). Thin-gate and back-illuminated CCDs, with much thinner dead layers, offer improved detection efficiency in the soft x-ray and UV portions of the spectrum. (From Edwards, D.F., in *Handbook of Optical Constants of Solids*, E.P. Palik, ed., Academic Press, New York, p. 547, 1985; Henke, B.L. et al., Low energy X-ray diagnostics, D.T. Attwood, B.L. Henke, eds., *AIP Conference Proceedings No. 75,* American Institute of Physics, New York, p. 340, 1981.)

The lateral diffusion of electrons can lead to a loss of resolution of the CCD since they may cross the boundary between one pixel and another, leading to an erroneous assignment of the origin of the photon that created the electron. This problem is minimized by thinning the *back-illuminated CCD* to thicknesses that are on the order of the depletion depth of the Si, so the least drift occurs before the electron is captured in the potential well of the buried channel. This thinning is accomplished by a combination of methods, including mechanical grinding and polishing and wet etching. Handling of back-illuminated CCDs after thinning, however, is a problem since the remaining Si is between 10 and 100 μm thick and is not strong enough to support rough handling, or even the intrinsic stresses arising from the initial fabrication of the CCD. To circumvent this problem, rim thinning or frame thinning can be used; in the first case, a rim several millimeters wide around the circumference of the wafer is protected during the thinning process, while in the latter, a region 1 or 2 mm wide around each device is protected.

Lateral diffusion can also be reduced by imposing an electric field through the thickness of the thinned membrane; but if this is done by imposing an electrode on the back surface, then the device will have a reduced QE due to absorption in this layer. The presence of a depletion region in the CCD will also give rise to an electric field, which reduces the lateral diffusion of electrons; but if the desired depletion region is on the order of tens of micrometers, in order to image high-energy x-rays or IR photons, then the induced field will be small, due to the low doping levels in the silicon and Gauss' law.

Although the lateral spread of photoelectrons is a problem, the loss of carriers to recombination at the back surface presents a greater technical difficulty. The recombination velocity can be minimized by introducing a strong electric field at the back surface and by passivating this surface. There are several methods employed to generate this electric field: one is by implanting carriers (p-type for the case where photoelectrons are to be captured) and annealing the sample to activate these carriers [7], a second is to deposit a thin layer of metal with the appropriate work function difference (typically Pt or another transition metal) [6], and a third method is to grow a heavily doped epitaxial layer on the back surface [9]. In the implant method, strong electric fields can be generated and the annealing process can help to passivate the back surface, but high-temperature annealing of the wafer can present a problem if metallization is present, unless excimer laser annealing is employed [10]. Deposition of a thin metallic layer can lead to long-term instability of QE and does not appear to yield the highest QEs. Deposition of an epitaxial layer relies on the use of sophisticated equipment and surface cleaning methods.

In addition to introducing an electric field on the back surface, it is necessary to electrically passivate that surface to lower the electron recombination velocity. Again, there are several approaches. The flash gate technique is probably the simplest, relying on inserting the sample in a steam atmosphere, but it is not stable over time [6]. Another method is to grow a high-quality oxide on the back surface [7], but this requires the use of high temperatures, which could melt metallizations and allow any dopant layer to diffuse too deeply into the silicon.

Although the aim of the backside treatment process is always to maximize the fraction of charge collected from the vicinity of the back surface, there are subtle differences in requirements that depend on the application. In the UV (100 nm < λ < 400 nm), the extremely small attenuation lengths (see Figure 46.5) place a premium on nonnegligible (more than ∼50%) charge collection efficiency very close to the back surface. Therefore, in UV sensors, the field in the immediate vicinity of the back surface is extremely important. Moreover, the real part of the index of refraction of silicon is quite high in the UV and optical bands, and Fresnel reflections limit the external detection efficiency that can be achieved. Thus, significant improvements in the UV response of back-illuminated CCDs have been obtained via the application of antireflection coatings as a part of the backside treatment process [8].

In the very soft x-ray range (2 nm < λ < 100 nm), the attenuation length can be larger by a factor of 10 than in the UV. Moreover, the range of the secondary photoelectrons is large enough to be comparable to the attenuation length [11], so the physics of the secondary ionization process becomes important. Thus, in the very soft x-ray regime, it is the average properties of the device over scales of a few hundred nanometers that determine performance. However, for spectroscopic applications,

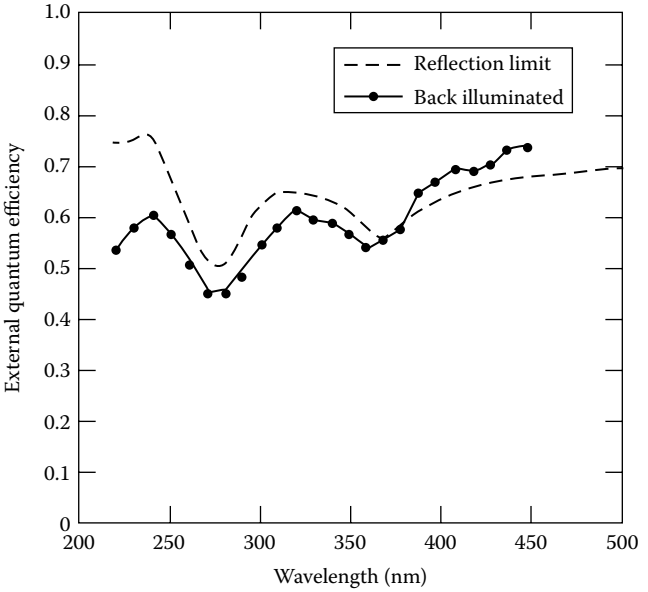


FIGURE 46.6 External QE of back-illuminated CCD using an implanted and annealed backside. For comparison purposes, the external QE of an ideal detector is shown, after modeling for the antireflection coating present on the back-illuminated detector.

the theoretical limit of device performance cannot be reached unless the charge collection efficiency substantially exceeds ~90%. Thus, in the soft x-ray, one requires very high collection efficiency over a relatively large volume in the vicinity of the back surface.

Any backside fabrication techniques will probably require compromises to be made, but it is possible to achieve very high external QE in the UV, as shown in the accompanying Figure 46.6 and in the soft x-ray region where 77% QE has been achieved at 277 eV [12]. Further improvements are necessary to achieve good results in the vacuum UV and theoretical limits for energy resolution in the soft x-ray. These improvements are possible, but will probably depend on the continued development of sophisticated processing tools.

46.6.2 Thinned Gate CCDs

Recently, a new type of CCD was developed that was optimized for spectroscopy below 1 keV. This detector, called a thin-gate CCD (TGCCD), contained two novel features: FGAs (or "skippers") and thin gates (electrodes) [2]. The goal was to develop a detector that had Fano-limited energy resolution with nonnegligible QE over the 200–1000 eV bandpass. This section briefly describes this detector and compares it with more conventional CCDs.

Most conventional CCDs use an FDA to read out the signal charge. The readnoise of the FDA has improved from several hundred electrons (rms) in the 1970s to less than three electrons in the last few years [13]. The energy resolution of CCDs over the entire soft x-ray bandpass is now limited by factors other than readnoise. It is desirable to push the readnoise down from a few electrons to a fraction of an electron because it allows photons of lower energy to be detected in single-photon-counting mode.

One method to further reduce the readnoise is to read out the same charge packet multiple times. The signal charge is destroyed when read out through an FDA. In the FGA, however, an insulating gate is placed between the charge transfer channel and the output node. Gates around the output node are used to read out and withdraw the signal charge multiple times. The readnoise is then reduced by the square root of the number of readouts. Effective readnoises of less than 1 electron have been demonstrated

using the FGA [2,14], and photons of energy 66 eV (Al L_α) have been detected using the TGCCD. If the readnoise could be pushed to a fraction of an electron, it would be possible to detect optical photons in the single-photon-counting mode.

46.6.3 Lower Readnoise

The other novel feature of the TGCCD is its electrode structure designed to maximize soft x-ray QE. Most conventional front-side-illuminated CCDs have several hundred nanometers of SiO_2/Si_3N_4 insulator and Si electrode on the surface. These layers strongly limit the QE of the device below 1 keV, but Fano-limited energy resolution is routinely achieved. The conventional solution to increase the low-energy QE has been to etch away the back surface of the device and illuminate it from behind (a backside-illuminated CCD). The difficulty with this technique is that the energy resolution is seriously degraded due to charge diffusion and poor charge collection.

The TGCCD is a front-side-illuminated device with a unique gate structure that gives reasonable soft x-ray QE (0.20 at 277 eV), while retaining the good energy resolution of front-side illumination. The detector is a three-phase CCD. The first two phases (comprising only 1/3 of the active area of the device) are covered with 400 nm of polysilicon (or just poly) in the usual way. A third layer of poly is deposited on top of the first two and acts as a bus for a fourth (40 nm thick) layer of poly that defines the third clock phase. Two-thirds of the active area of the detector is therefore covered with only the insulator (100 nm of SiO_2 and Si_3N_4) and this fourth layer of poly. The measured energy resolution of this device is 34 eV at C K_α (277 eV). This is approximately a factor of 3 better than the best energy resolution obtained with the ACIS backside-illuminated CCDs [5]. Simulations show that an astrophysical spectrum obtained with TGCCD contains as much or more information than one obtained with a backside-illuminated CCD [15].

46.6.4 Radiation Damage

CCDs are often used in environments in which they are subject to ionizing radiation, and this radiation can degrade their performance. In particular, radiation damage plays an important role in the performance of most space-borne detectors. The Van Allen belts contain protons and electrons with energies of tens of MeV. These particles can produce damage by several mechanisms. Protons can generate vacancy-interstitial pairs in the Si lattice, which diffuse freely, even at cryogenic temperatures, and it is possible for the vacancy to form a complex with P atoms in the *n*-Si buried channel that traps electrons. The trap reduces CTE by preventing signal electrons from leaving the pixel during the clocking cycle. Electrons can also produce vacancies (although with much lower cross section than the protons), but they can also create dangling bonds at the SiO_2–Si interface above the buried channel, leading to a shift in operating voltage in the CCD and contributing to noise.

The P–V complexes remove electrons from a charge packet in the conduction band of silicon. In the worst-case limit, when the illumination levels are so low that each charge packet transferred encounters each trap in its empty trap state, the degradation of CTE is approximately 5×10^{-7} rad^{-1} (radiation doses are expressed here in total ionizing rads in Si, although it should be noted that the fraction of proton energy that goes into production of vacancy-interstitial pairs, the nonionizing energy loss, is about 1/2000 of the total ionizing dose [16].) This degradation is significant: If the device is subjected to 1000 rad $year^{-1}$ and there are 1000 transfers necessary to clock out the charge, then 40% of the charge packet would be lost due to proton-induced defects. At higher illumination levels, the traps greatly complicate the device response, particularly in single-photon-counting applications. In this case, for example, the spectral response of the CCD becomes a function of the spatial distribution of the incident radiation.

This loss of CTE can be circumvented in several ways, including the use of increased shielding, a fat zero, annealing out the damage, implanting a narrow trough along the direction of transfer in the buried channel, or operating the device at a different (generally lower) temperature.

Relatively thick shielding, up to the equivalent 4–5 cm (1.6–2 in.) of aluminum, can provide significant attenuation (factors of several) of proton dose encountered in the radiation belts [17]. A fat zero is a uniform charge added to all the pixels of the CCD, usually induced by flooding the CCD with a low light level. Although a fat zero can reduce the CTE loss by filling the traps so they cannot be occupied by electrons from the charge packet of interest, they also add to the noise of the system. The P–V complex will dissociate at temperatures around 135 °C [18], but the annealing is incomplete and may interfere with other requirements or constraints of the satellite mission. Using photolithography, a narrow trough or notch can be included in the buried channel parallel to the direction of transfer during fabrication, so charge packets are kept constrained and are less likely to interact with traps generated randomly across the buried channel [19]. It is not difficult to introduce a channel that is approximately 2 μm wide and can handle a charge packet of tens of thousands of electrons; this can decrease the CTE loss to 7×10^{-8} rad^{-1} or less. Reducing the operating temperature slows the kinetics of trap emptying [20], so a single electron can be lost from the charge packet and fill the trap for the integration time of the frame. The effectiveness of this process depends on the clock speed and integration time, as well as the ability to operate the CCD at temperatures that approach –150 °C, but CTE loss can be as low as 10^{-8} rad^{-1} if troughs are included with low-temperature operation. Cosmic rays and electrons can also cause damage by displacing Si atoms from the lattice, but the flux of cosmic rays is much lower than protons in low Earth orbits, and the electrons are much less effective in creating a vacancy. The approaches to hardening against proton damage will also be effective with these latter two particles.

The major effect of electrons on CCDs is to shift the flat-band voltage of the CCDs and MOSFETs, but for doses around 1000 rads, this change is on the order of 10 mV. If further hardening of the gate oxide is desired, it can be accomplished using established methods [21]. The passage of electrons and other energetic charged particles through a device also creates free carriers that constitute an interfering signal. One electron–hole pair is created for each 3.65 eV absorbed in the Si, and charged particles can deposit charge packets of up to a full-well level and more, depending on the species, energy, and the active depth of the device. In astronomy, such events are an annoyance, even in terrestrial observatories, and are removed by comparing images of the same scene. In some applications, such as x-ray spectroscopy, these events are more serious because they can masquerade as desired signals unless careful analysis of the data is performed to exclude them. On the other hand, this sensitivity to charged particles can actually be used to advantage. The known conversion between particle energy deposition and liberated charge means that CCDs can be used as spectroscopic detectors, and such an application has been proposed for inertial-confinement fusion diagnostics [22].

Radiation hardening of CCDs has made rapid progress in the last few years, allowing devices to have much longer lives for scientific applications. Fabrication and operational changes have increased the hardness to displacement damage, and the process requirements of high-quality ICs have led to a gate dielectric that is quite robust to ionizing radiation.

References

1. C.J. Wordelman and E.K. Banghart, *Charge-Coupled Devices, Optoelectronic Devices*, J. Piprek (ed.), Springer, 2005, pp. 343–380, available at http://link.springer.com/chapter/10.1007%2F0-387-27256-9_12#, accessed on September 8, 2013.
2. R.P. Kraft, D.N. Burrows, G.P. Garmire, J.A. Nousek, J.R. Janesick, and P.N. Vu, Soft x-ray spectroscopy with sub-electron readnoise charge-coupled devices. *Nucl. Instr. Meth.*, A361, 372–383, 1995.
3. Y. Matsunaga and S. Ohsawa, Analysis of low signal level characteristics for high-sensitivity CCD charge detector. *Trans. Electron Devices*, 39, 1465–1468, 1992.
4. G.H. Jacoby (ed.), *CCDs in Astronomy*, Astronomical Society of the Pacific, San Francisco, CA: 1990.
5. R.P. Kraft, D.N. Burrows, G.P. Garmire, and J.A. Nousek, Thin-gate front side-illuminated versus back side-illuminated charge-coupled devices for X-ray astronomy. *Astrophys. J. Lett.*, 466, L51–L54, 1996.

6. J.R. Janesick, D. Campbell, T. Elliott, and T. Daud, Flash technology for charge-coupled-device imaging in the ultraviolet. *Opt. Eng*., 26: 852–863, 1987.
7. B.E. Burke, J.A. Gregory, R.W. Mountain, J.C.M. Huang, M.J. Cooper, and V.S. Dolat, High-performance visible/UV CCD imagers for space-based applications. *Proc. SPIE*, 1693, 86–100, 1992.
8. M.P. Lesser, Improving CCD quantum efficiency, *Inst. Astronomy VIII, Proc. SPIE*, 2198, 782–791, 1994.
9. M.E. Hoenk, P.J. Grunthaner, F.J. Grunthaner, R.W. Terhune, M. Fattahi, and H.-F. Tseng, Growth of a delta-doped silicon layer by molecular beam epitaxy on a charge-coupled device for reflection limited ultraviolet quantum efficiency. *Appl. Phys. Lett*., 61, 1084–1086, 1992.
10. C.M. Huang, B.E. Burke, B.B. Kosicki, R.W. Mountain, P.J. Daniels, D.C. Harrison, G.A. Lincoln, N. Usiak, M.A. Kaplan, and A.R. Forte, A new process for thinned, back-illuminated CCD imager devices. *Proceedings of the International Symposium on VLSI Technology, System, and Applications*, New York: IEEE, 1989, pp. 98–101.
11. F. Scholze and G. Ulm, Characterization of a windowless Si(Li) detector in the photon energy range 0.1–5 keV. *Nucl. Instr. Meth*., A339, 49–54, 1994.
12. G. Prigozhin, M. Bautz, S. Kissel, G. Ricker, S. Kraft, F. Scholze, R. Thornagel, and G. Ulm, Absolute measurement of oxygen edge structure in the quantum efficiency of X-ray CCDs, *IEEE Trans. Nucl. Sci*., 44, 970–975, 1997.
13. D.N. Burrows, G.D. Berthiaume, M.A. Catalano, G.P. Garmire, C. Larkin, F. Marks, J.A. Nousek, and G.M. Weaver, Penn State imaging x-ray spectrometer, in *EUV, X-ray, and Gamma-Ray Instrumentation for Astronomy and Atomic Physics*, C.J. Hailey and O.H.W. Siegmund (eds.), Bellingham, Washington, DC; *Proc. SPIE*, 1159, 92–104, 1989.
14. C.E. Chandler, R.A. Bredthauer, J.R. Janesick, J.A. Westphal, and J.E. Gunn, Sub-electron noise charge coupled devices, *SPIE Symposium on Electronic Imaging,* San Francisco, CA; *Proc. SPIE*, 1242, 238–251, 1990.
15. M.C. Weisskopf, S.L. O'Dell, R.F. Eisner, and L.P. Van Speybroeck, Advanced X-ray astrophysics facility (AXAF): An overview. *Proc. SPIE*, 2515, 312–329, 1995.
16. M.J. Cantella, B.E. Burke, J.A. Gregory, D.C. Harrison, E.D. Savoye, and B.-Y. Tsaur, Large silicon staring-array sensors, in *Infrared Focal Plane Arrays IIIA*, J.A. Jamieson (ed.), Washington, DC: Ballistic Missile Defense Organization, 1994, DOI: 10-1-10-76.
17. K.C. Gendreau, M.W. Bautz, and G.R. Ricker, Proton damage in X-ray CCDs for space applications: Ground evaluation techniques and effects on flight performance. *Nucl. Inst. Meth*., A335, 318–327, 1993.
18. M. Hirata, H. Saito, and J. Crawford, Effect of impurities on the annealing behavior of irradiated silicon. *J. Appl. Phys*., 38, 2433–2438, 1967.
19. T.S. Villani, W.F. Kosonocky, F.S. Shallcross, J.V. Groppe, G.M. Merais, J.T. O'Neill III, and B.J. Esposito, Construction and performance of a 320 × 244-element IR-CCD imager with PtSi Schottky-barrier detectors. *Proc. SPIE*, 1107, 9–21, 1989.
20. K.C. Gendreau, G.Y. Prigozhin, R.K. Huang, and M.W. Bautz, A technique to measure trap characteristics in CCDs using X-rays. *IEEE Trans. Electron Devices*, 42, 1912–1917, 1995.
21. P.V. Dressendorfer, Effects of radiation on microelectronics and techniques for hardening, in *Ionizing Radiation Effects in MOS Devices and Circuits*, T.P. Ma and P.V. Dressendorfer (eds.), New York: John Wiley & Sons, 1989, p. 333.
22. B.E. Burke, R.D. Petrasso, C.-K. Li, and T.C. Hotaling, Use of charge-coupled device imagers for charge-particle spectroscopy, *Rev. Sci. Inst*., 68, 599–602, 1997.
23. D.F. Edwards and R.H. White, *Semiconductors: Gallium Antimonide*, E.D. Palik (ed.), San Diego, CA: Academic Press, 1998, pp. 597–601.
24. B.L. Henke, Low energy x-ray spectroscopy with crystals and multilayers, in *AIP Conference Proceedings No. 75 on Low Energy X-Ray Diagnostics*, Monterey, CA, D.T. Attwood and B.L. Henke (eds.), New York: American Institute of Physics, 1981, p. 85.

47

Densitometry Measurement

Joseph H. Altman
Pittsford, New York

47.1 Introduction

As the name indicates, densitometry is the measurement of optical density (OD). OD, in turn, can be broadly defined as a measure of the attenuation of radiant flux by some sort of optical element that can be transmitting, reflecting, or both. Densitometry is most widely applied in photographic science, and measurements in this field are covered by a four-part ANSI/ISO standard [1]. The treatment here follows that document. The measurement of transmission and reflection from the standpoint of optics has been discussed by Palmer [2]. This chapter uses the single word density, without the adjective, and uses the words "light" and "flux" interchangeably (although density can be measured in the UV and infrared [IR] also, of course).

It turns out that, in practice, the measured density of a given sample can be affected significantly both by the design of the equipment and the nature of the sample. Therefore, the first rule of practical densitometry is that the sample must be measured in a way that is meaningful with respect to its intended use.

47.2 Monochrome Transmission Density

It is convenient to start by discussing transmitting samples, without the color aspect. Such samples can be either black-and-white photographic films or filter layers of various sorts. If light falls on such an element, a fraction of the flux is reflected from the first surface. The remainder penetrates the surface, and a part of this flux is absorbed. What is neither reflected nor absorbed exits the layer, as shown in Figure 47.1. Let these fractions be denoted, *r, a*, and *t*, where lowercase *t* indicates internal transmittance. If the element is in air, reflection can occur at the rear surface also, but this can be neglected. By the conservation of energy, one obtains

$$r + a + t = 1 \tag{47.1}$$

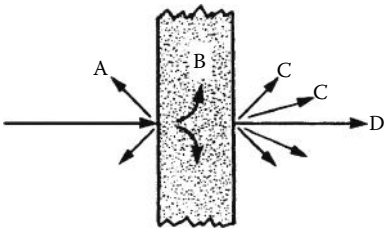


FIGURE 47.1 Cross section of emulsion layer showing (A) reflection, (B) absorption, (C) diffuse transmission, and (D) specular transmission.

The fractions r and t can be measured independently by spectrophotographic techniques, so that a is readily determined. This fraction is termed the absorptance of the layer, and its measurement is also discussed by Palmer [2].

Some applications require the measurement of absorptance, but in most practical work, what is important is the fraction of the incident light that exits the element, regardless of where attenuation takes place. Therefore, transmittance is defined by

$$T = \frac{\Phi_e}{\Phi_i} \tag{47.2}$$

where e and i refer to emergent and incident flux, respectively. Note that all losses are lumped together. In silver halide photographic layers, the first-surface reflection is usually small; in thin-film filters, a significant proportion of the total loss may occur by reflection. Following Hurter and Driffield [3], the English pioneers of densitometry, one can define the transmission density D

$$D = -\log_{10} T = \log_{10} \frac{1}{T} \tag{47.3}$$

Density is therefore a nonlinear dimensionless quantity ≥ 0. McCamy has published an extensive review of densitometry and sensitometry [4].

The operation of a practical densitometer can be discussed with reference to Figure 47.2. The system consists of a light source, an aperture to define the area being measured, a sensor, logging circuitry, and a suitable readout display. In many instruments, the output is fed to a computer.

To measure transmittance/density, a reading is first taken with no sample over the aperture. This reading is a measure of the incoming light Φ_i. A second reading with the sample in place then determines Φ_e. It should be clearly understood that what the system actually does is to sense the flux passing through the aperture in the two cases and then displays the negative logarithm of the ratio. This fact

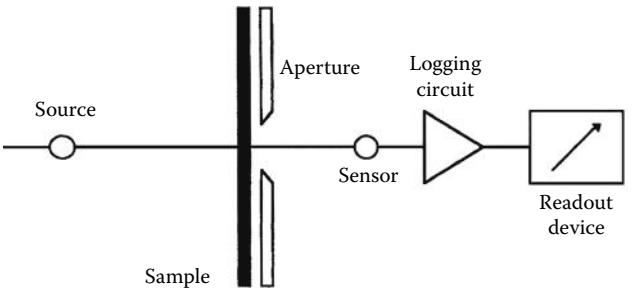


FIGURE 47.2 Schematic of transmission densitometer.

explains some of its operating characteristics. Starting from this basis, for example, McCamy [5] has treated measurement errors for the case when the sample is not uniform ("wedged") over the aperture.

The basic procedure also means that any source having enough power to operate the sensor can be used, and filters and apertures can readily be interchanged; it is only necessary to "zero" the instrument before making a measurement. The measuring aperture is usually of the order of 1 mm in diameter. If the area of this aperture is reduced to, say, 0.1 mm^2, the device becomes a microdensitometer. Such instruments present special problems not discussed here but treated in the literature [6].

Now consider the complications introduced by the nature of samples, especially photographic samples. Developed black-and-white photographic layers consist of discrete grains of silver metal dispersed in a thin layer of gelatin. Such grains act as scatterers, which means that some of the photons exiting the layer will not be traveling in their original directions. Such photons are identified by "C" in Figure 47.1. Clearly, the measured transmittance/density will depend on how much of this scattered light reaches the sensor. The effect of scattering can be quite large, and therefore in order to obtain values that are both reproducible and meaningful, it is necessary to control what is termed the "geometry" of the system, that is, the angular subtenses of the influx and efflux beams at the sample.

In Figure 47.3, Θ_i and Θ_e are these half angles, respectively, referred to the normal to the element surface. The current standard [1b] prescribes two configurations, the first of which is as follows: $\Theta_i = 5°$; $\Theta_e = 90°$. The reverse of this configuration, that is, $\Theta_i = 90°$ and $\Theta_e \leq 5°$, yields the same density readings and is also permitted by the standard. This case is termed "diffuse density" and is the most important practical case because most commercial densitometers conform to this mode. Essentially, all the light exiting the sample, whatever its direction, reaches the sensor and is evaluated. This configuration simulates the case of contact printing a photographic negative or of viewing a transparency on a diffuse illuminator. The practical construction of such an instrument is described later.

The second case occurs when the sample is illuminated and the emergent flux collected by lenses of finite aperture. In principle, in this case, the angles Θ_i and Θ_e can vary between, say, 5° and 90° and may differ from one another, but the standard specifies two sets of optics with matched apertures of *f*/1.6 and *f*/4.5. These relative apertures are representative of motion picture projectors and microfilm readers, respectively. The corresponding half angles are 18.2° and 6.4°. (For a lens in air, *f*/no. = $1/(2\sin\Theta)$, where Θ is the half angle shown in Figure 47.3.) This case is termed "projection" density.

It is useful to define a ratio

$$Q' = \frac{\text{Projection density}}{\text{Diffuse density}} \tag{47.4}$$

which has been termed the "effective Callier coefficient." As would be expected, projection density is greater than diffuse density for a given sample so that $Q' = 1$. The exact ratio between the two densities

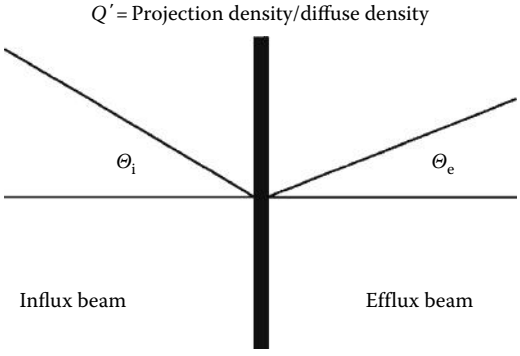


FIGURE 47.3 Angular subtenses of influx and efflux beams at sample.

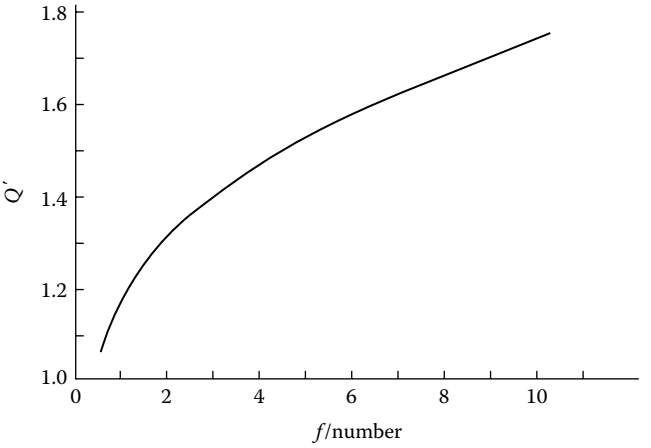


FIGURE 47.4 Variation of the effective Callier Q'-factor with the f-numbers of the influx and efflux beams, which were matched.

depends on the scattering characteristics of the sample and the f-numbers involved. Only limited information on Q' has appeared in the literature [7] but, as Figure 47.4 shows, this parameter behaves as might be expected.

The dependence of the effective density on aperture shown in Figure 47.4 may be important in practice. Amateur photographers are aware that a B&W negative usually produces a more contrasty print in a condenser enlarger than it does in a diffuser enlarger, and this behavior can easily be deduced from Equation 47.4. It may also be important in microdensitometry, since these instruments measure a projection density and a wide range of apertures may be encountered in various instruments. In such cases it is usually recommended that the data be reduced to diffuse density. This is easily done by measuring a sufficiently large sample in both projection and diffuse densitometers. It might be remarked that the scattering characteristics of individual films can vary significantly, and Q' should be measured for the samples at hand. Except for very low densities, Q' is relatively unchanged as the diffuse density varies.

For completeness, it is noted that earlier editions of density standards listed two other forms of density. The first of these was the case of angle Θ_i = angle $\Theta_e \leq 5°$. This case was termed "specular density" and was intended to simulate the use of the attenuating element in a collimated beam. The second case was the case of angle Θ_i = angle $\Theta_e = 90°$. This type of density was termed "doubly diffuse" and was intended to simulate the use of a negative in certain contact printers. Neither of these densities is important in modern practice; discussions of them can be found in the literature [8].

In a color film, the "grains" are actually tiny volumes of dyed gelatin. Since the index of refraction of such a "grain" is only slightly different from the surround, it does not act as a scatterer. Thus, the density differences found on color films as a function of densitometer geometry are usually negligible and often ignored in practice.

In recent years, the photographic industry has introduced a new type of grain, the "tabular" grain, whose thickness is much smaller than the dimensions in the plane parallel to the coating surface. It has been demonstrated that, in the undeveloped state, these grains scatter light significantly less than the older "3-D" types of grains. No studies of developed T-grains have appeared in the literature, but it seems possible that developed B&W layers consisting of T-grains may scatter less than the traditional types of layers, thus reducing the sensitivity of density measurement to instrument geometry. (The data of Figure 47.4 were obtained with "3-D" grains.)

One can now turn to the optical configuration of a practical diffuse densitometer. From an optical standpoint, a good way to collect all the flux exiting the sample would be to use an integrating sphere, and instruments have been built using such spheres. However, in an actual contact-printing

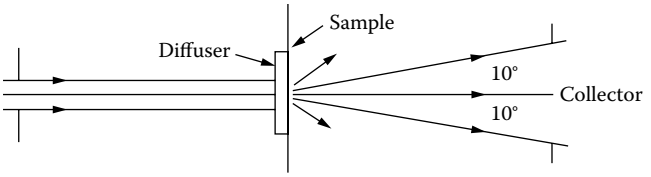


FIGURE 47.5 Optical system for measuring ISO diffuse density.

case, there will be some light reflected back to the negative from the paper surface. Any retroreflection from an integrating sphere will of course be quite different from that coming from a paper surface. To represent the contact-printing case, the present ANSI/ISO standard specifies that, in diffuse densitometers, the emulsion side of the sample shall be in contact with opal glass, which both acts as the diffuser and provides the desired back reflection. The optics of a diffuse densitometer is shown in Figure 47.5.

In the system shown, the sample is illuminated by diffuse light, which is the reverse of what is shown in Figure 47.3; however, as noted, diffuse density optics are reversible. Weaver studied the differences between opal-glass and integrating-sphere densities and found that because of the interreflections, the values were slightly lower in the opal-glass case [9]. The difference was ~0.04 density unit for very low densities, decreasing to 0.01–0.02 for samples whose densities were between 2.0 and 3.0. The opal-glass construction was also adopted in the standard because such instruments are easier to build and maintain. Care must be taken, however, to ensure that the diffuser meets the standard's specifications.

In actual operations, it is too time-consuming to zero the instrument before each reading. Likewise, the fraction Φ_e/Φ_i is rarely calculated specifically; the instrument simply provides a reading based on the amount of flux reaching the sensor. Because of the possibility of electronic drift however, it is advisable to zero the instrument periodically, unless it is known to be dependably stable. Likewise, the readings produced for some accurately known calibration sample of high density should be checked, a procedure known as "sloping" the instrument. Standard samples are available for this purpose; see the Appendix 47.A.

47.3 Monochrome Reflection Density

By definition, the reflectance factor

$$R = \frac{\Phi_r}{\Phi_o} \tag{47.5}$$

where

Φ_r is the flux returned to the sensor from the sample

Φ_o is the flux returned by a "perfectly reflecting, perfectly diffusing material located in place of the specimen" [1d]

Reflection density is then

$$D_R = -\log_{10} R = \log_{10} \frac{1}{R} \tag{47.6}$$

The optical configuration specified in the standard is shown in Figure 47.6. With reference to the normal to the specimen surface, one beam subtends an angle of 0° ± 5° and the other an angle of 45° ± 5°. Also, the 45° beam is annular. As in the case of transmission optics, the influx and efflux beams are interchangeable. When the influx beam is at 45°, the system is termed "45/0"; when the efflux beam is at 45°, the designation is "0/45." The angle between the two beams was selected to avoid the specular reflection from the surface. The 45° beam is made annular to minimize the effects of any texture pattern that might be embossed on the sample surface. The standard also specifies that when measured, samples

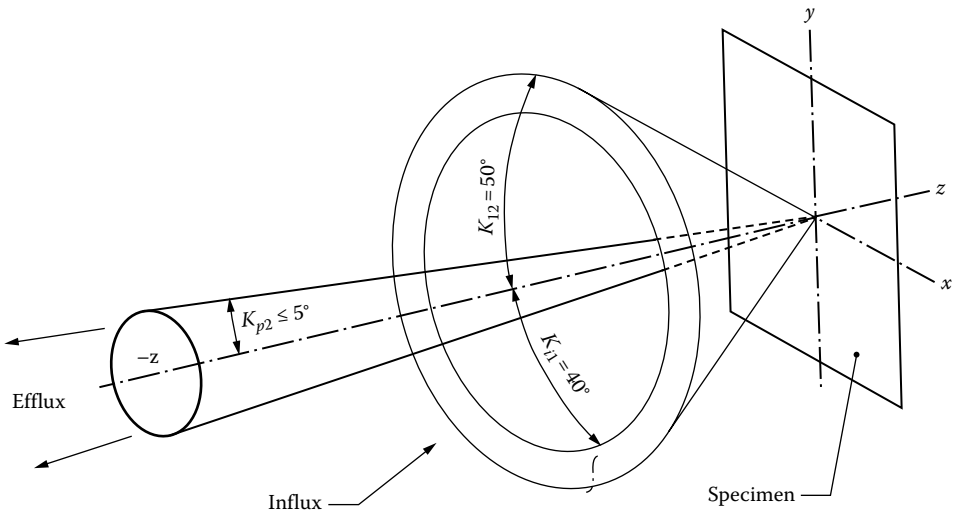


FIGURE 47.6 Annular geometry for measurement of reflection density.

shall be backed by a black diffusing material having a reflection density not less than 1.5. This procedure is specified to improve reproducibility, especially for thin samples.

It will be obvious that a reflection densitometer cannot be "zeroed" by taking a reading with no sample in place. Instead, as prescribed by the standard [1d], calibration samples that have been measured on other primary instruments are used both to zero and slope the densitometer. Formerly, coatings of MgO_2 or $BaSO_4$ were used as reference "whites," but these layers are difficult to use. An ASTM standard [10] discusses the preparation of reference whites from pressed powders. For routine work, "plaques" consisting of stove enamel of various gray levels on a metal substrate are widely used. These plaques are very durable, but their physical form is quite different from actual paper samples. Photographic paper strips that have been calibrated in primary instruments can also be used. Physically, these strips are exactly like the samples to be read, but they are very easily damaged and must be used with care. Any scratch in the surface can cause a specular reflection to reach the sensor, which will produce a false reading.

Actually, the reflection of light from a photographic paper is more complicated than might appear at first glance. A typical B&W paper is shown in cross section in Figure 47.7. As with film samples, the emulsion layer consists of Ag grains suspended in gelatin. The support is a diffuse reflector. It turns out that reflection to the sensor occurs in three ways, as sketched. (The specular reflection is omitted in Figure 47.7.) Note that the component r_2 traverses the emulsion layer twice and r_3 at least four times and that r_1 does not traverse

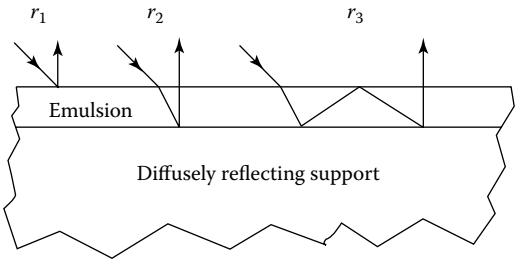


FIGURE 47.7 Cross section of photographic paper showing different reflection paths for influx beam: r_1, surface reflection; r_2, direct reflection from support to sensor; and r_3, one or more internal reflections before reaching sensor.

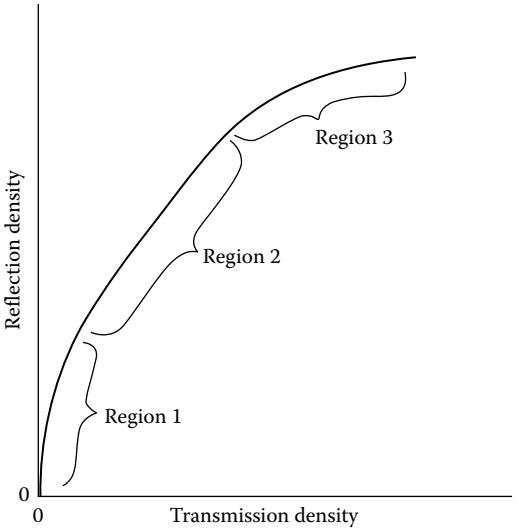


FIGURE 47.8 Densities measured on a film sample, both by transmission and after cementing the emulsion side to a diffusing reflector.

the layer. For low densities, all three components contribute to the reflection reading. As the population of grains in the emulsion layer grows, the multiply-reflected beam r_3 rapidly becomes negligible. With continued increases in the grain population, the component r_2 also reduces to insignificance, and only the r_1 component remains. This component, being a surface reflection, remains essentially constant at about $0.005I_0$.

The mechanism previously described can be confirmed by measuring the densities of a test strip by transmission and then cementing the strip to a diffuse reflector and measuring the same densities by reflection. The results of such an experiment are shown in Figure 47.8. This curve has three regions, as expected. In the first region, the slope is greater than 2, corresponding to the rapid loss of the r_3 component. In region 2, $\Delta D_R = 2\Delta D_T$; in this region, most of the photons reaching the sensor are those that have traversed the layer twice. Finally, in the third region of the curve, the r_2 component has become negligible, and only the surface component r_1 remains significant. Since, as noted, this component $\sim 0.005I_0$, the maximum reflection density attainable in a photographic paper should be about 2.0–2.3, and this is found to be the case for actual papers. (We have assumed a glossy paper. If the paper surface is textured, as it is in some products, specular reflections may be directed to the sensor from microareas, and the maximum density attainable on the sample will drop, in some cases significantly.) The curve of Figure 47.8 also becomes very important in color reflection work, as is discussed later.

47.4 Color Transmission Densitometry

So far, we have assumed either that the measuring beam was monochromatic or that the sample was neutral, that is, that transmittance or reflectance was constant across the spectrum of interest. Actually, the assumption of neutrality is reasonable for many B&W materials. However, of course, the densitometry of color images is extremely important, and this aspect is discussed later.

Before doing so, however, it will be useful to make a brief digression into the structure of photographic color materials [11]. Although an oversimplification, we can consider such materials to be three separate films coated one over the other on a single support. From the top down, the spectral sensitivities of these three layers are adjusted to record the blue, green, and red regions of the visible spectrum, respectively. In accordance with the principles of subtractive color reproduction, the images recorded in these layers are formed of yellow, magenta, and cyan dyes (again from the top down). The problem is

to densitometer these dyes. As before, we start with transmitting samples. Referring back to Figure 47.2, for any wavelength λ, the reading of the instrument with no sample in place will be

$$KS_\lambda s_\lambda F_\lambda = KJ_\lambda \tag{47.7}$$

where

K is the proportionality constant
S_λ is the spectral power of the influx beam at λ
s_λ is the relative spectral sensitivity of the receiver at λ
F_λ is the transmittance of any filters in the beam at λ

(The transmission of the instrument optics may be significant, but this can be included with S_λ, since it is a fixed characteristic of the device.) The product J is termed the "response" or the "spectral product" of the instrument.

If a sample whose transmission at wavelength λ is T_λ is now placed over the aperture, the reading will be

$$KS_\lambda s_\lambda F_\lambda T_\lambda = KJ_\lambda T_\lambda \tag{47.8}$$

and the measured transmittance reduces to T_λ, as it should. Thus, in the monochromatic case, the measured value is independent of the system response. When the influx beam contains two discrete wavelengths, the no sample reading becomes

$$K(J_1 + J_2) \tag{47.9}$$

and with the sample in place, it will be

$$K(J_1T_1 + J_2T_2) \tag{47.10}$$

The measured transmittance

$$T = \frac{J_1T_1 + J_2T_2}{J_1 + J_2} \tag{47.11}$$

Note that when the sample is neutral, Equation 47.11 again reduces to the case where the transmittance is independent of the instrument response. But when the sample is not neutral, Equation 47.11 can no longer be simplified. In other words, in the general case the transmittance/density values depend on the system response as well as on the sample.

For continuous bands of radiation, Equation 47.9 becomes

$$T = \frac{\int J_\lambda T_\lambda d\lambda}{\int J_\lambda d\lambda} \tag{47.12}$$

and

$$D = -\log_{10}\left[\frac{\int J_\lambda T_\lambda d\lambda}{\int J_\lambda d\lambda}\right] = \log_{10}\left[\frac{\int J_\lambda d\lambda}{\int J_\lambda T_\lambda d\lambda}\right] \tag{47.13}$$

The integration limits are set by the distributions.

From the previous analysis, it is clear that in order to obtain meaningful density values, the spectral characteristics of the densitometer must be made equal to those of the receiver that will "view" the sample in actual use. Furthermore, of course, the actual receiver will vary from application to application.

The densitometer characteristics are set by inserting appropriate filters in the beam. A number of spectral distributions are specified in Ref. [1c] as follows:

1. For reflection densitometry, the influx spectrum shall be ICI Illuminant A, which is essentially 2856 K.
2. For transmission densitometry, the influx spectrum shall be Illuminant A modified by an IR-absorbing filter, as tabulated in the standard.
3. For samples to be viewed by an observer, the spectral characteristic of the system shall match the relative luminosity function $V(\lambda)$. This case is termed "visual" density.
4. For color negative films, a set of three distributions in the red, green, and blue regions of the spectrum is specified. This set is identified Status *M* and approximates the spectral sensitivities of the three layers of color print materials.
5. For measuring red, blue, and green densities of color materials other than color negatives, a set of three responses identified Status *A* is provided.
6. To evaluate color images to be used in graphic arts processes, two sets of distributions are specified and are identified as Status *T* and Status *E*.
7. In the microfilm industry, prints are often made onto diazo or vesicular films. A narrow distribution centered at 400 nm is specified for this application, and the resulting densities are termed ISO printing densities type 1.
8. A spectral distribution covering the range $\lambda \approx 360$–540 nm is provided for measuring samples to be printed onto B&W photographic papers. Such densities are designated ISO printing densities type 2.
9. Status *I* response consists of three passbands centered at 420, 535, and 625 nm ± 5 nm. This set is used in evaluating graphic arts materials such as process inks on paper.
10. A narrowband response centered at 800 ± 20 nm is provided for measuring effective densities to *S*–1 photosurfaces, such as those used in optical sound systems. Densities measured with this response are identified "ISO type 3."

These responses are shown in Figure 47.9a through e, which is reproduced from the current standard.

The next problem in color densitometry stems from the fact that the three layers are superimposed and from the nature of the dyes. The spectrophotometer curves for a typical dye set are shown in Figure 47.10. In this figure, the lower three curves refer to the dyes measured individually, while the upper curve shows the result of superimposing them. The values for the individual dyes are termed "spectral analytical densities," and those for the tripack are termed "spectral integral densities." Since superimposed densities may be considered to add, at any wavelength the spectral integral value is the sum of the three analytical values. Note that all three of the individual dyes absorb light outside the spectral regions in which they are supposed to work.

This unwanted absorption can cause problems in color reproduction and, more to the point for this chapter, in densitometry, because straightforward measurements of a real film at any wavelengths yield integral values. In many cases, these integral values are required. In the manufacture and processing of films, however, it may become necessary to determine the densities in an individual layer. It is not sufficient to coat and measure a layer by itself, since layers coated in a tripack may respond differently from the same layers operated singly.

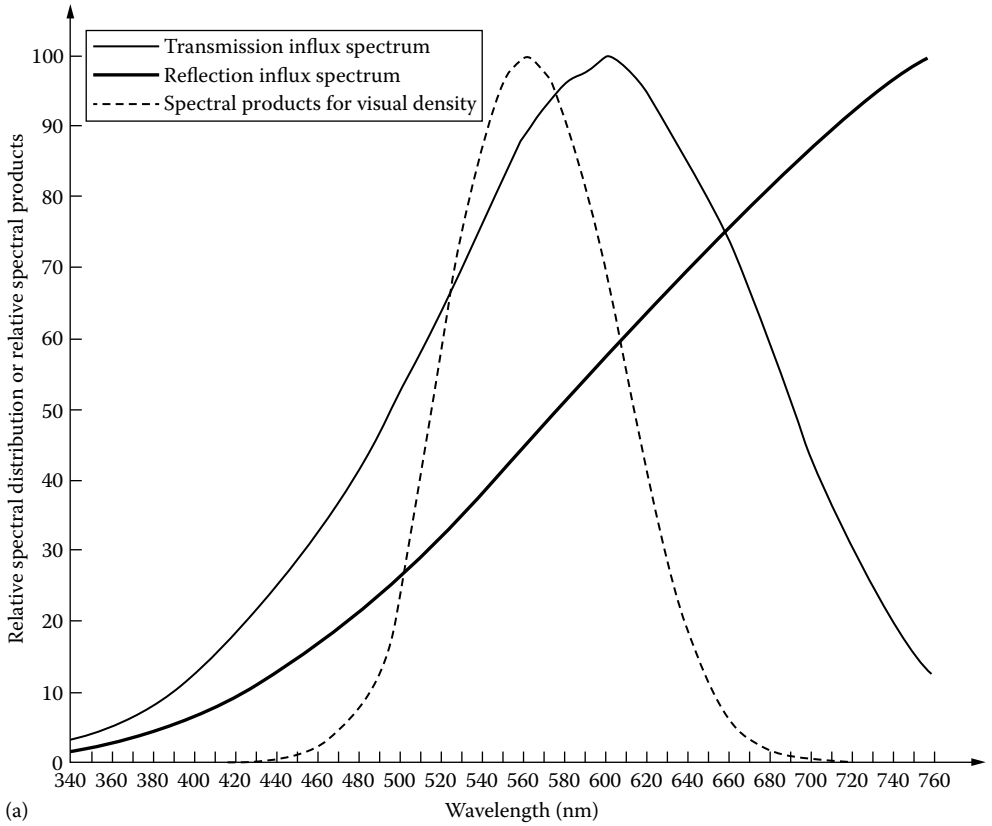
This problem is solved by taking advantage of two rules called the "additivity rule" and the "proportionality rule." The first of these is merely the rule that densities add. The proportionality rule is an extension of Beer's law. Consider an element, such as a glass cell or a layer of a color film, containing a dye whose concentration can be varied. Beer's law states that

$$T_\lambda = e^{-\beta_\lambda c} \tag{47.14}$$

where

c is the dye concentration

β_λ is the extinction coefficient at wavelength λ



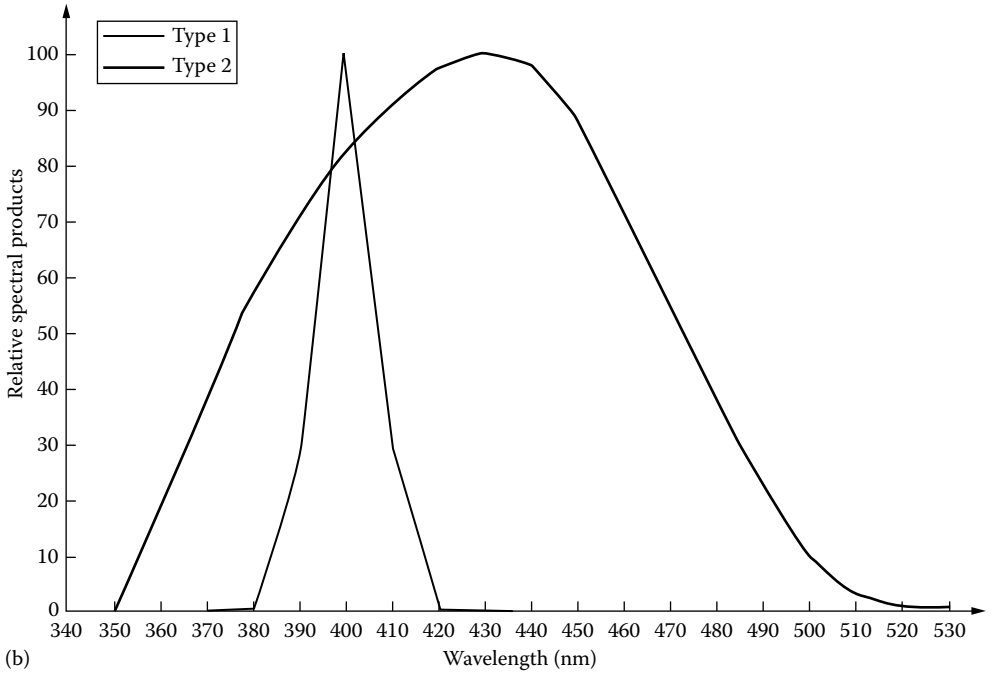
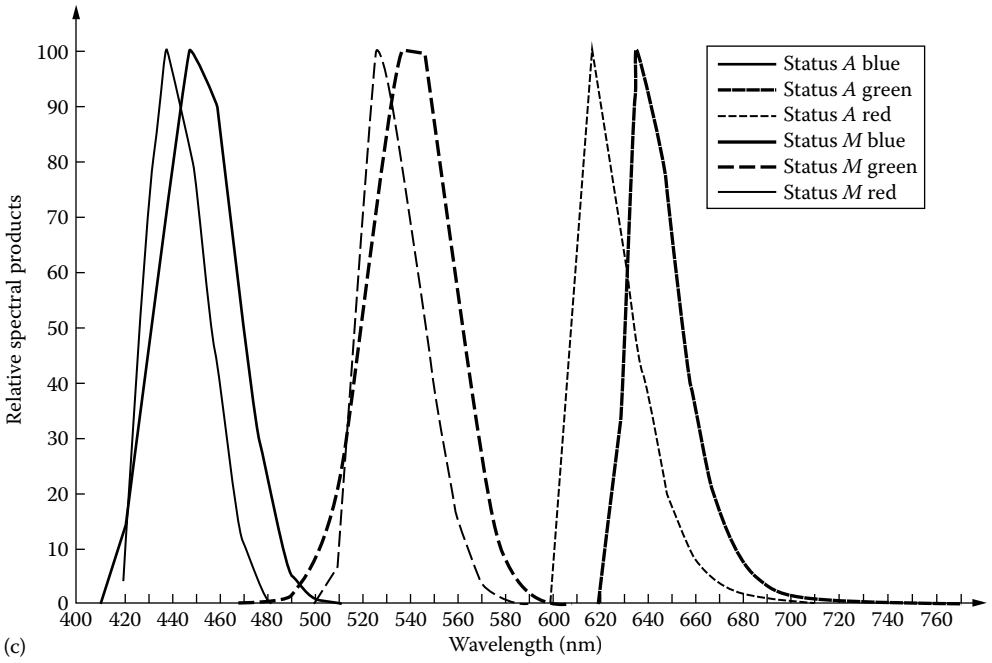


FIGURE 47.9 (a) Spectral products for ISO density and relative spectral power distributions for influxes. (b) Spectral products for type 1 and 2 densities.



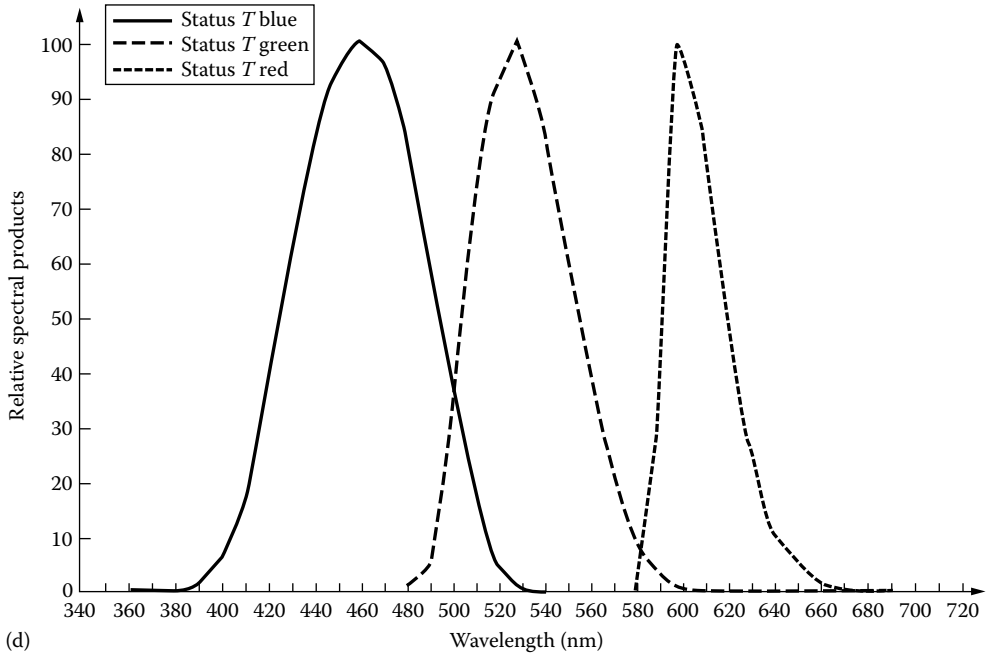


FIGURE 47.9 (continued) (c) Spectral products for Status *A* and *M* densities. (d) Spectral products for Status *T* densities.

(*continued*)

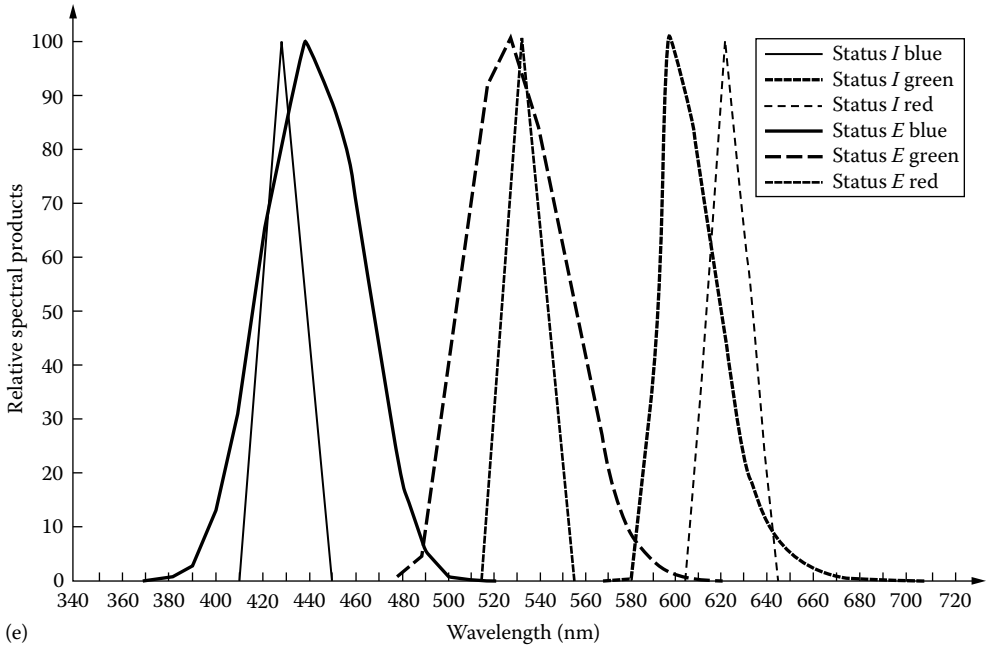


FIGURE 47.9 (continued) (e) Spectral products for Status *E* and *I* densities.

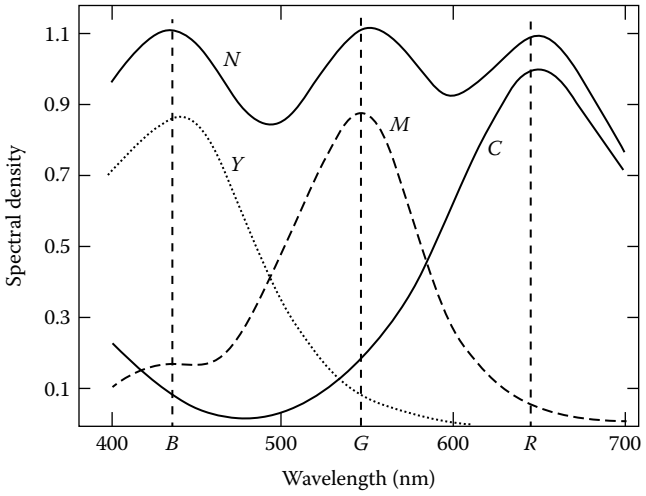


FIGURE 47.10 Spectral density distributions of the *C*, *M*, and *Y* components as functions of wavelength for a hypothetical color film and of the composite absorber (*N*).

The coefficient β varies with wavelength, of course, but will be constant at a given wavelength (unless changes in dye concentration produce chemical reactions.) Strictly speaking, Beer's law applies to the internal transmittance, but in the case of photographic layers, the surface reflectance is negligible. From Equation 47.14, it follows that

$$\ln T_\lambda = -\beta_\lambda c \tag{47.15}$$

and therefore,

$$D_\lambda = 0.434\beta_\lambda c \tag{47.16}$$

Consider now the density at two wavelengths for any given dye sample. Clearly,

$$\frac{D_{\lambda 1}}{D_{\lambda 2}} = \frac{\beta_{\lambda 1}}{\beta_{\lambda 2}} \tag{47.17}$$

Thus, for a given dye sample, the ratio of the densities at any two wavelengths is fixed regardless of dye concentration (i.e., assuming Beer's law holds), and this is a statement of the proportionality rule.

As noted, straightforward density measurements on a real sample yield integral densities. The proportionality rule can be used to derive analytical densities in the following manner [12]. Assume spectral densities for simplicity. With reference to Figure 47.10, we select three wavelengths at or near the peak absorptances of the three dyes. Consider first the blue wavelength. By the additivity rule, the blue integral density is

$$D_b = Y_b + M_b + C_b \tag{47.18}$$

where Y_b, M_b, and C_b are the blue analytical densities of the yellow, magenta, and cyan dyes, respectively. But by the proportionality rule, M_b/M_g = constant, where M_g is the analytical density of the magenta dye measured at the peak wavelength in the green. The value of this ratio is readily determined from a spectrophotometric curve of the dye. Denoting such ratios a_n, for measurements at the three peak wavelengths, we have

$$\begin{aligned} D_b &= Y_b + a_1 M_g + a_2 C_r \\ D_g &= a_3 Y_b + M_g + a_4 C_r \\ D_r &= a_5 Y_b + a_6 M_g + C_r \end{aligned} \tag{47.19}$$

This constitutes a set of three equations with three unknowns. The solution has the form

$$\begin{aligned} Y_b &= b_1 D_b + b_2 D_g + b_3 D_r \\ M_g &= b_4 D_b + b_5 D_g + b_6 D_r \\ C_r &= b_7 D_b + b_8 D_g + b_9 D_r \end{aligned} \tag{47.20}$$

where Y_b, M_g, and C_r are the desired values of the three analytical densities. The b-coefficients are algebraic combinations of the a-coefficients, which can be reduced to numerical values since the a-coefficients are known. In practical work, the determination of the a-coefficients is usually done using data for a number of different dye density levels. The off-peak density for each level is plotted against the peak density, giving a straight line whose slope is a good measure of the a-coefficient. Thus, for example, the coefficient a, in Equation 47.19, would be found by plotting M_b versus M_g for several levels. The values of the coefficients can also be found by doing narrowband densitometry on the samples instead of spectrophotometry.

So far, we have been assuming spectral densities. By the first law of densitometry, however, any sample should be measured with a system whose spectral response is equal to that used in actual practice—in other words, color negative samples should be measured with the Status *M* filters as described above. However, the analysis previously described works satisfactorily with the status filters, which are relatively narrowband, and much practical densitometry is performed in this manner.

Finally, because of its practical importance, we must introduce one more type of color density: the "equivalent neutral density" (END). END is defined as the neutral density that a given dye

deposit—yellow, magenta, or cyan—will produce when it is combined with the correct amounts of the other two dyes, whatever those amounts may be. It is important in practical film building, because one of the most important properties of a color film is its ability to produce a neutral as a neutral—in other words, to image a gray object as a gray that shows no residual tint of another color. Such a film is said to be "balanced," and good balance is considered extremely important by discriminating users. When the ENDs of the three layers of a color film are equal, the image will indeed be balanced, and color film sensitometry may be done in terms of ENDs. The END is a form of analytical density because it is a property of an individual layer. Originally, ENDs were measured in a special densitometer, but they are now calculated from the normal analytical densities by an equation of the form

$$\text{END} = m\,[\text{analytical spectral density}] + K \tag{47.21}$$

where the constants m and K are determined by comparing the visual densities of satisfactory neutral images with the corresponding analytical spectral densities. In some cases, a second-order equation has been found to yield better results [12b].

47.5 Color Reflection Densitometry

Color reflection densitometry follows the previously mentioned general principles set forth for the B&W reflection case. Since color prints are for the most part intended to be viewed by a human observer, in many cases a visual integral density provides the needed information. The film builder or process controller, however, may need information about the analytical densities of his layers, and here the approach used for transmitting samples does not work. The reason why it fails is that, for such samples, Beer's law fails, as is shown by the curve of Figure 47.8. If the law held, the reflection density would be proportional to the dye concentration (see Equation 47.11).

Pinney and Vogelsong [13] have described a method of obtaining reflection analytical densities in such cases. The method involves the empirical determination of a calibration curve similar to Figure 47.8 for the material at hand, relating the reflection densities of a dye deposit to the transmission densities. The reflection integral densities of the sample are then measured and converted to transmission densities. The analytical values for these transmission densities are calculated using the method previously described for such samples. Finally, the derived transmission values are converted back to reflection analyticals by going through the calibration curve in the reverse direction.

One last comment on the densitometry of color print papers should be added. In the commercial production of color prints, it may be extremely important to measure chemical stain in what should be the white areas of the picture. The passbands of the Status *A* filters are not well placed to monitor such stain, and it may be necessary to make additional measurements.

47.6 Densitometry of Halftone Patterns

Halftone patterns are used in the printing industry to reproduce continuous-tone images using two levels of ink: either ink or no ink. Various gray levels are produced by printing a pattern of repetitive "dots" too small to be resolved by the unaided eye. Essentially, the size of the dots is varied to produce a gray scale. Halftone patterns to be measured may be transmitting or reflecting, monochrome or color. A number of standards for densitometry in the graphic arts have been approved or are under development [14]. These standards provide much practical guidance and are consistent with ISO 5. Densitometry in this field is an example of the basic mechanism by which the instrument operates—comparing the flux reaching the sensor with and without the sample in place. When the

sample is a halftone, the flux reaching the sensor depends on the fractions of the area that are dense and clear. Adopting the notation of Ref. [13b] and assuming a transmitting sample, the measured density will be

$$D_t = -\log_{10}\overline{T} = -\log_{10}[fT_s + (1-f)T_b] \tag{47.22}$$

where

D_t is the measured density of pattern
$\overline{T}$ is the average pattern transmittance
f is the fraction of pattern area that is dense
T_s is the transmittance of "solid" areas (essentially equal to dot density)
T_b is the transmittance of clear areas

It is interesting to calculate the measured density of a halftone pattern, half of which is perfectly dense ($D = \infty$; $T = 0$) and half of which is perfectly clear ($D = 0$; $T = 1$). The calculated value for $D \approx 0.3$; and this value will indeed be found if such a pattern is densitometered.

If D_t, T_s, and T_b are known, Equation 47.22 provides an easy way to determine the fractional area covered by the pattern dots with an ordinary densitometer, and this is often done in graphic arts work. In practical work, T_s and T_b are measured in terms of density also, and this can be done satisfactorily on large areas. Since by definition $T = 10^{-D}$, the solution of Equation 47.22 for f can be written

$$f = \frac{1-10^{-(D_t - D_b)}}{1-10^{-(D_s - D_b)}} \tag{47.23}$$

In this form, the equation is known as the Murray–Davies equation.

47.7 Summary

A densitometer measures the flux reaching the sensor with a sample in place and calculates the ratio of this value to that obtained either with no sample in place (in the transmission case) or with a reference white in place (in the reflection case). The instrument then displays the negative log of this value.

The observed reading may depend significantly on the characteristics of the instrument, that is, on the

1. Angular substance of the influx beam
2. Angular substance of the efflux beam
3. Spectral power distribution of the influx beam
4. Spectral sensitivity of the receiver

For a complete specification of the density in a given case, these four parameters should be indicated. ISO 5.1 (Ref. [1a]) provides a standardized notation.

47.8 Appendix

A. Sources of Densitometers

Gretag–Macbeth
617 Little Britain Road
New Windsor, New York 12553-6148

X-Rite, Inc.
3100 44th Street SW
Grandville, MI 49418

Camag
Sonnenmattstr 11
CH-4132 Muttenz
Switzerland

Note: Makers of densitometers are listed in the *Photonics Buyers Guide*, published annually by Laurin Publishing Co., Inc., Berkshire Common, P.O. Box 4949, Pittsfield, MA 01202-4949.

B. Aids to Densitometry
   1. Status and Other Filters
      Gretag–Macbeth
      X-Rite, Inc.
      Eastman Kodak Co.
      Scientific Imaging Products
      343 State Street
      Rochester, New York 14650
   2. Standard Reference Materials, Calibration Samples
      Gretag–Macbeth
      X-Rite, Inc.
      Graphic Communications Association
      100 Dangerfield Road
      Alexandria, VA 22314

      PSI Associates
      3000 Mount Read Boulevard
      Rochester, New York 14616

      Lucent Technologies
      235 Middle Road
      Henrietta, New York 14467

      National Institute of Standards and Technology
      Gaithersburg, MD 20899
      Eastman Kodak Co. provides densities on film calibrated to ±5% or ±0.02 density unit, whichever is greater.

References

1. International Standard, *Photography Density Measurements*, International Organization for Standardization, Geneva, Switzerland.
   a. Part 1: Terms, Symbols and Notations IS05/1-1984.
   b. Part 2: *Geometric Conditions for Transmission Density* IS05-2, 3rd edn., 1991.
   c. Part 3: *Spectral Conditions*. IS05-3, 2nd edn., 1995.
   d. Part 4: *Geometric Conditions for Reflection Density*. IS05-4, 2nd edn., 1995.
2. J. L. Palmer, *The Handbook of Optics,* 2nd edn., New York: McGraw-Hill, 1995, Chapter 25. This chapter contains an extensive bibliography.
3. F. C. Hurter and V. C. Driffield, *J. Soc. Chem. Ind. London*, 9, 455, 1890.
4. C. S. McCamy, History of sensitometry and densitometry, in E. Ostroff (ed.), *Pioneers of Photography*, Springfield, VA: The Society for Imaging Science and Technology, 1987, Chapter 17, pp. 169–188.
5. C. S. McCamy, *J. Opt. Soc. Am*., 66, 350, 1976.
6. J. C. Dainty and R. Shaw, *Image Science*. New York: Academic Press, 1974, Chapter 9.
7. H. C. Schmitt and J. H. Altman, *Appl. Opt*., 9, 871, 1970.

8. J. H. Altman, *The Sensitometry of Black and White Materials*, in T. H. James (ed.), *The Theory of the Photographic Process,* 4th edn., New York: MacMillan, 1977, Chapter 17.
9. K. S. Weaver, *J. Opt. Soc. Am*., 40, 534, 1950.
10. *Standard Practice for Preparation of Pressed Powder White Reflection Transfer Standards*, ASTM E-259-93, West Conshohocken, PA: American Society for Testing Materials, 1997.
11. R. W. G. Hunt, *The Reproduction of Colour*, 6th edn., Chichester, West Sussex, UK: Wiley, 2004.
12. (a) R. M. Evans, W. T. Hanson, and W. L. Brewer, *Principles of Color Photography*, New York: John Wiley & Sons, 1953; (b) P. Kowaliski, *Applied photographic theory*, in T. H. James (ed.), *The Theory of the Photographic Process,* 4th edn., New York: MacMillan, 1977, Chapter 18.
13. J. E. Pinney and W. F. Vogelsong, *Phot. Sci. Eng*., 6, 367, 1962.
14. American National Standards, New York: American National Standards Institute. A number of additional standards were under development by the Committee for Graphic Arts Technology Standards (CGATS) as of late 1997.
    a. CGATS.4-1993: *Graphic Technology, Graphic Arts Reflection Densitometry Measurements, Terminology, Equations, Image Elements and Procedures*.
    b. CGATS. 9-1994: *Graphic Technology, Graphic Arts Transmission Densitometry Measurements, Terminology, Equations, Image Elements and Procedures*.

48

Colorimetry

Robert T. Marcus
Datacolor International

48.1 Introduction

Imagine how dull a world without color would be. Until the 1960s, many products were available in only a limited variety of colors. Consumers demand a variety of colors, and the materials available today allow manufacturers to meet those demands.

For many centuries, color was controlled by master color matchers adjusting the color of their products visually in natural daylight. It was about the 1950s when color-measuring instruments began to make a significant impact in the manufacturing process, and by the 1970s they were commonly used in most industries. Color-measuring instruments are used mostly for quality control but also for computerized color-matching systems, process control, and evaluation of raw materials and as an aid to solving color problems.

While it is common to speak of a red car or a red light, color is a perception—not an intrinsic property of the object or the light. Color perception is influenced by the light source, the reflectance or transmittance properties of the object, the eye, and the brain. Color is one aspect of appearance. Gloss and texture are other aspects.

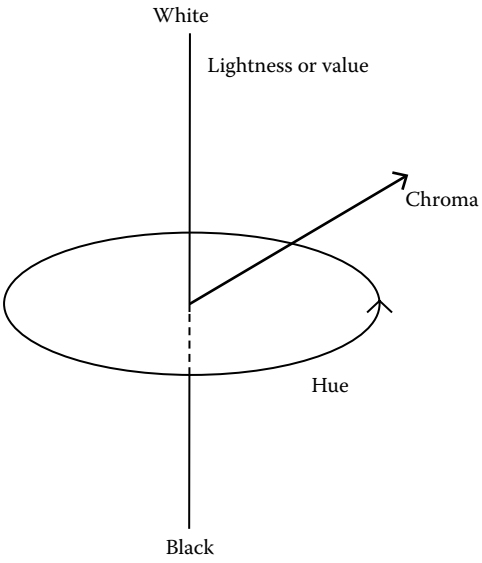


FIGURE 48.1 The three dimensions of color.

Color perception is 3D, that is, three terms are needed to describe a color. Hue, lightness (sometimes called value), and chroma are one set of terms often used to describe color. Hue distinguishes blue from green from yellow, etc. Lightness distinguishes light colors from dark colors—for example, a light-blue fabric from a dark-blue fabric. Chroma, the most difficult of the three terms to understand, describes how different a color is from gray—for example, distinguishing a pastel green from a bright green. If the two greens are of the same hue (one is neither yellower nor bluer than the other) and have the same lightness, then the pastel green would be described as having a low chroma and the bright green as having a high chroma. Figure 48.1 is a diagram illustrating the three dimensions of color.

Daylight, fluorescent lamps, and incandescent lamps are widely used light sources for color evaluation. The perceived color of an object changes as the light changes. The human visual systems attempt to compensate for the change in light source and hold the color constant. Light booths provide standardized and controlled light sources for the visual evaluation of color. Most light booths contain a simulated daylight lamp, a cool white fluorescent lamp, an incandescent lamp, and an ultraviolet lamp for detecting fluorescence. By use of a switch, a light booth can be used to examine how the perception of a colored material changes with different light sources.

Two objects may appear the same when viewed under one light source but different when viewed under another light source. This effect, called metamerism, is one of the major industrial problems for color matching. Metamerism usually occurs when attempts are made to produce objects that are the same color but made out of different materials, such as trying to match the dyed textile interior of a car with paint on the exterior and the plastic on the dashboard.

Colorimetry, the measurement of color, attempts to quantify the perception of color. The Commission Internationale de l'Éclairage (International Commission on Illumination [CIE]) is a voluntary organization of scientists and engineers from all over the world who are interested in light and color. The recommendations constituting modern colorimetry were first published by the CIE in 1931 and have been regularly updated since then [1].

Electromagnetic radiation (x-rays, gamma rays, light, and radio waves) irradiates the Earth constantly. Visible light is the name given to the electromagnetic radiation that the human eye perceives. The wavelength of visible light ranges from about 380 to about 780 nm. Sunlight is a mixture of all the wavelengths of light. Water vapor can spatially separate the light into its various wavelengths—the rainbow. Prisms and diffraction gratings can also spatially separate light into its component wavelengths.

Light sources can be described numerically by their spectral power distribution, the relative amount of power the source emits at each wavelength of interest. A light source may emit power at wavelengths below (ultraviolet "light") or above (infrared "light") those of visible light. Ultraviolet radiation is important to colorimetry because it can cause fluorescence. Infrared radiation is the basis of "heat lamps" but is not important in colorimetry. Spectroradiometers measure the spectral power distribution of light sources.

Illuminants and sources are sometimes confused. Sources are actual physical entities that produce visible radiation, whereas an illuminant may only be a numerical table of values of a spectral power distribution. Initially, the CIE recommended three light sources for colorimetry in 1931. Source A, which is still in use, is an incandescent, tungsten filament light. An illuminant is the spectral power distribution of a light source. Thus, Illuminant A is the spectral power distribution of Source A. An illuminant may be defined, even when a source for that illuminant does not exist. Examples of illuminants without sources are the D series of illuminants recommended by the CIE. The D illuminants represent various phases of daylight. Illuminant D65 represents average daylight and is the most common illuminant used in colorimetry. No sources were recommended for the D series of illuminants.

Fluorescent lights have great commercial importance. Cool white fluorescent lamps are the most common light sources in offices in the United States. There are a variety of other types of fluorescent lamps used in stores and offices. The CIE also recommended a series of illuminants to represent fluorescent lamps. F2 represents cool white fluorescent lamps. Figure 48.2 shows the spectral power distributions of Illuminants A, D65, and F2.

Color vision and perception is complex and has been extensively studied. Ninety-two percent of men and 99.5% of women have "normal" color vision. The eye's lens focuses images on the light-sensitive retina. Rod cells make up the majority of the retina and are sensitive to low levels of illumination (night vision). Cone cells provide color vision and are located in a small area of the retina called the *foveal pit*. There are three types of cone cells. One type of cone cells has peak sensitivity to blue light, one type to green light, and one type to red light. Signals from the cone cells are transmitted to the brain where they are processed into color perceptions.

Color perception is an extremely complex phenomenon. For example, the background on which a material is viewed can have a major effect on the perceived color of that material. The ambient light to which the eye becomes adapted also influences the color of materials. Basic colorimetry, the topic of this chapter, provides the rather simple concept of dealing with the measurement of single independent colors. Most of industrial color control is adequately described using basic colorimetry. Advanced colorimetry attempts to use physical measurements to describe the perceived color of a material when viewed in a complex scene.

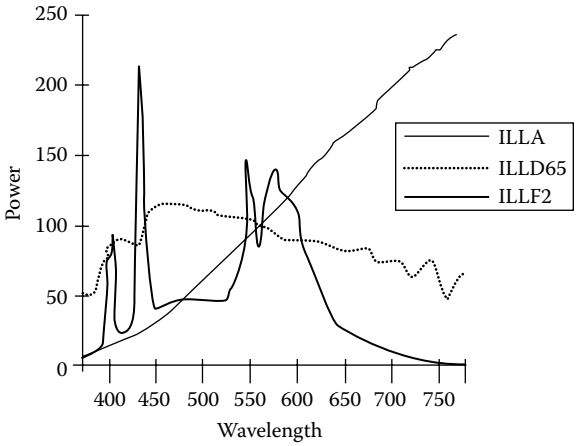


FIGURE 48.2 The spectral power distribution of CIE standard Illuminants A, D65, and F2.

All colors are perceived by stimulating combinations of the three cones. Computer monitors and color television tubes produce colors by lighting combinations of red, green, and blue phosphors. In 1931 when the CIE was developing the system for modern colorimetry, they transformed the data from real experiments that had human observers match each wavelength of visible light with combinations of red, green, and blue lights to three mathematical imaginary "lights" labeled *X*, *Y*, and *Z*. All colors can be matched by varying amounts of *X*, *Y*, and *Z*. The amounts of each *X*, *Y*, and *Z* imaginary light that must be mixed together to match a color are called the tristimulus values (see "CIE Standard Observers," explained in Section 48.3).

Most objects absorb, transmit, or reflect (scatter) light. Transparent objects absorb and transmit light. Opaque objects absorb and reflect light. Light sources emit light. Fluorescent objects absorb, reflect, and emit light. Translucent or hazy objects absorb, transmit, and scatter light. Measuring the color of fluorescent, translucent, and hazy objects is difficult and will be covered in later sections.

Objects are characterized by the amount of light they reflect or transmit at each wavelength of interest. Most spectrophotometers measure reflectance factors rather than reflectance. Reflectance is the amount of light reflected from an object compared to the amount of light illuminating that object. The reflectance factor is the amount of light reflected from an object compared to the amount of light reflected from a perfect diffuser under the same conditions. A perfect diffuser is a theoretical material that diffusely reflects 100% of the light incident upon it. The term reflectance is often used in a general sense, or as an abbreviation for, reflectance factor. Such usage may be assumed unless the term reflectance is specifically required by the context [48]. Spectrophotometers designed for color measurement usually measure reflectance or transmittance at 10 or 20 nm intervals throughout the visible spectrum.

When the reflectance or transmittance of an object is low, the object absorbs most of the incident light; when it is high, the object reflects or transmits most of the incident light. Figure 48.3 shows the reflectance curve of a printed yellow ink. Note that the yellow ink absorbs light in the blue portion of the spectrum and reflects light in the yellow and red portions of the spectrum. A green plastic (Figure 48.4) absorbs light in the blue and red portions of the spectrum and transmits light in the green portion.

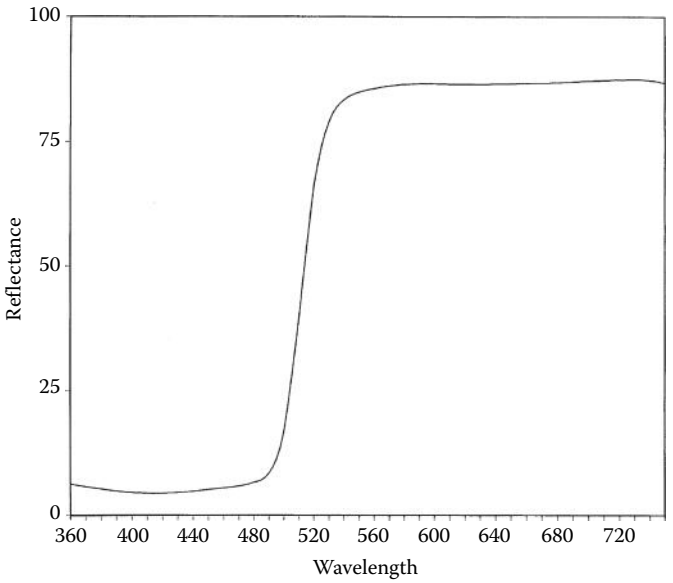


FIGURE 48.3 The spectral reflectance curve of a printed yellow ink.

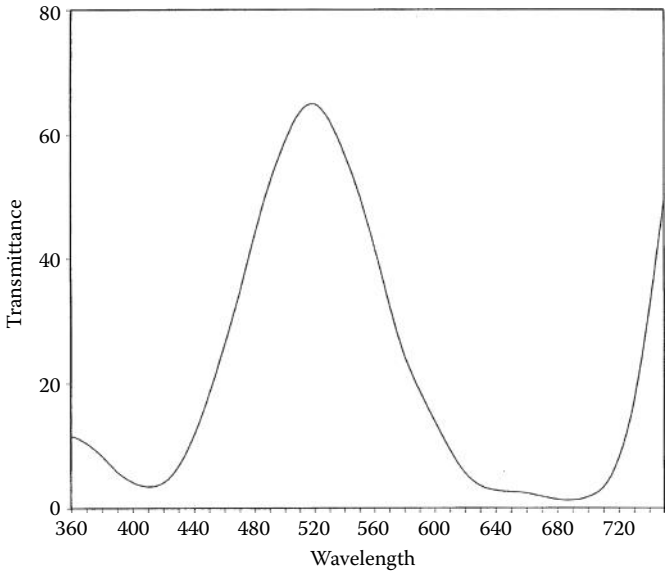


FIGURE 48.4 The spectral reflectance curve of a green transparent plastic.

The tristimulus values of an object can be calculated by combining the reflectance or transmittance of the object with the spectral power distribution of an illuminant and the color-matching functions of a standard observer. The yellow ink's tristimulus values for Illuminant D65 and the 10° standard observer are $X = 66.62$, $Y = 69.72$, and $Z = 7.03$. Those for the green plastic are $X = 21.03$, $Y = 39.40$, and $Z = 24.06$. An object's tristimulus values will change with the illuminant. For example, the tristimulus values for the yellow object for Illuminant A and the 10° standard observer are $X = 91.15$, $Y = 77.83$, and $Z = 2.80$.

Pairs of objects are said to match when their tristimulus values are the same. However, since the calculation of the tristimulus values included the source and observer as well as the object, when one of these changes, the objects may no longer match, that is, they may have different tristimulus values. Metameric colored objects have the same tristimulus values for the illuminant under which they match but different tristimulus values for illuminants where they do not match.

Color measurements are most often made to determine quantitatively whether or not the colors of two objects or batches of material are the same. But what happens when they do not match? Color difference equations were developed to quantify the difference. Starting with the objects' tristimulus values, a color difference equation will calculate the total color difference, ΔE or DE, and its component parts—the differences in lightness (ΔL or DL), chroma (ΔC or DC), and hue (ΔH or DH) or the differences in lightness, yellowness–blueness (Δb or Db), and redness–greenness (Δa or Da). Numerical color differences may be used for setting tolerances for quality control applications, to answer the question: is the match close enough?

48.2 Standardized Light Sources

48.2.1 Color Temperature

Light sources may be described by their color temperature. A block of carbon would look completely black when its temperature was at absolute zero, 0 K. When the carbon block is heated to about 2850 K, it looks yellow—about the same color as an incandescent light bulb. Heat the block to 5000 K, and it looks whitish. At 7500 K, the block would have the bluish color of north sky daylight. A full (blackbody)

radiator is a theoretically perfect emitter and absorber of radiation that changes color like the carbon block just described. The color temperature of a light source is the temperature of a full radiator that has the same color as the light source. When a light source does not exactly match the color of a full radiator, the correlated color temperature is used to describe the light source. The correlated color temperature is the temperature of a full radiator whose color is closest to the source.

Daylight varies during the day—redder in the morning and evening and bluer at noon. Average daylight (diffuse skylight without direct sunlight) has a color temperature of 6500 K. North sky daylight is preferred by many people for the visual evaluation of color and has a color temperature of 7500 K.

48.2.2 CIE Recommendations

Source A has a color temperature of 2856 K. The CIE recommended daylight illuminants are referred to by the prefix D, followed by the first two digits of their color temperature. Thus, CIE Illuminant D65 has a color temperature of 6500 K. Illuminant D50 has a color temperature of 5000 K and is preferred by the graphic arts community. Illuminant D75 would be used for north sky daylight having a color temperature of 7500 K. Illuminant F2 has a color temperature of about 4100 K.

48.3 CIE Standard Observers

The CIE adopted two standard observers based on color-matching experiments. The CIE 2° standard observer was recommended in 1931, and the CIE 10° standard observer was recommended in 1964.

In the CIE experiments, observers having normal color vision matched spectrum colors with combinations of red, green, and blue light. Figure 48.5 illustrates the experimental setup used in the standard observer experiments. One-half of a circular field was illuminated with the spectrum color, while the other half was illuminated with a mixture of red, green, and blue lights. The observer adjusted the amounts of red, green, and blue until the mixture matched the spectrum color. Unfortunately, not all of the spectrum colors could be matched with combinations of the red, green, and blue lights used in the experiment. In those cases, one of the lights had to be moved so that it illuminated the test field. By "diluting" the spectrum color with one of the lights, the resultant color could be matched with the remaining two lights. The amount of light used to dilute the spectrum color was considered

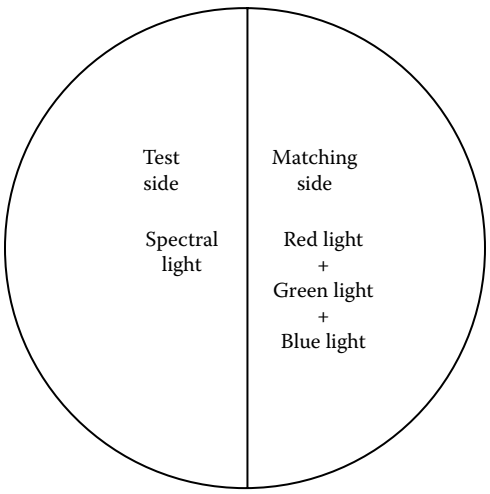


FIGURE 48.5 Experimental field of view for determining the CIE standard observers.

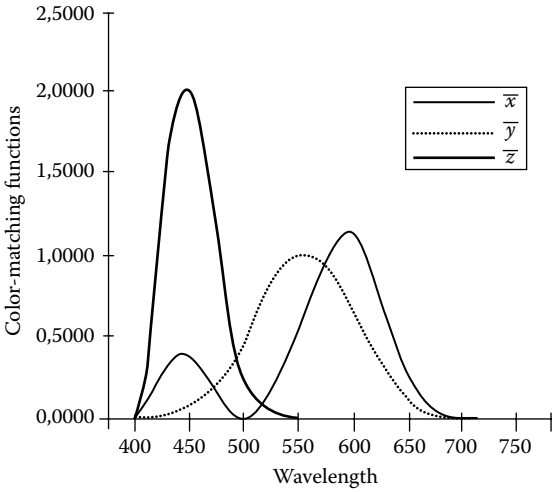


FIGURE 48.6 The color-matching functions for the CIE 10° standard observer.

to be a negative amount. To avoid having color-matching functions with negative amounts of light, three "imaginary" lights (X, Y, and Z) were created by performing a mathematical transformation. Color-matching functions are the amounts of X, Y, and Z required to match the colors of the spectrum and are used in the calculation of tristimulus values. Figure 48.6 shows the color-matching functions, ($\bar{x}$, $\bar{y}$, and $\bar{z}$) for the 10° standard observer.

In the first experiments, the circular field viewed was projected on the foveal pit and subtended a solid angle of 2°. This is about the size of a dime held at arm's length. The 2° standard observer developed from these experiments is very useful when viewing small fields, such as the signal lights of ships or small colored chips. Industrial color matchers view larger fields, such as two 5 × 12 in. panels. The 10° standard observer should be used when viewing larger fields. Observers for those experiments viewed a 10° visual field, which is about the size of a fist held at arm's length.

The standard observers represent combinations of the color-matching functions of a number of observers. Standard observer color-matching experiments are tedious and difficult to do. Few people can do them with reproducibility. Individual's color-matching functions are likely to vary from that of a CIE standard observer. Although these differences do not generally present a problem, they can affect the evaluation of metameric samples.

48.4 Calculating Tristimulus Values

Tristimulus values for reflecting samples are calculated from the following equations:

$$X = k \int_{380}^{780} R_\lambda S_\lambda \bar{x}_\lambda d\lambda \cong k \sum_{380}^{780} R_\lambda S_\lambda \bar{x}_\lambda \Delta\lambda \tag{48.1}$$

$$Y = k \int_{380}^{780} R_\lambda S_\lambda \bar{y}_\lambda d\lambda \cong k \sum_{380}^{780} R_\lambda S_\lambda \bar{y}_\lambda \Delta\lambda \tag{48.2}$$

$$Z = k \int_{380}^{780} R_\lambda S_\lambda \bar{z}_\lambda d\lambda \cong k \sum_{380}^{780} R_\lambda S_\lambda \bar{z}_\lambda \Delta\lambda \tag{48.3}$$

$$k = \frac{100}{\int_{380}^{780} S_\lambda \bar{y} d\lambda} \cong \frac{100}{\sum_{380}^{780} S_\lambda \bar{y} \Delta\lambda} \tag{48.4}$$

where

S_λ is the relative spectral power distribution of the illuminant at wavelength λ
R_λ is the reflectance factor of the sample at wavelength λ
$\bar{x}_y$, $\bar{y}_\lambda$, and $\bar{z}_\lambda$ are the color-matching functions of the observer at wavelength λ

For transmitting objects, substitute the transmission factor, T_λ, of the sample at wavelength λ for the reflectance factor. The factor k normalizes Y so that it will equal 100.00 for a perfect reflector or transmitter, that is, R_λ or T_λ is 100.00 at all wavelengths of interest. The summations in the previous equations are only valid if the reflectance or transmittance of the sample is measured at wavelength intervals of 1 or 5 nm from 380 to 780 nm.

Many commercial spectrophotometers measure wavelength intervals of 10 or 20 nm. To accurately calculate tristimulus values of samples measured with those instruments, the following equations should be used:

$$X = \sum W_{x\lambda} R_\lambda \tag{48.5}$$

$$Y = \sum W_{y\lambda} R_\lambda \tag{48.6}$$

$$Z = \sum W_{z\lambda} R_\lambda \tag{48.7}$$

in which $W_{x\lambda}$, $W_{y\lambda}$, and $W_{z\lambda}$ are weighting factors designed for 10 and 20 nm wavelength intervals. T_λ can be substituted for R_λ for transmitting samples. Although a number of weighting factors have been developed over the years [2–7], those recommended by ASTM in E 308 Standard Practice for Computing the Colors of Objects by Using the CIE System [7] are recommended.

Tristimulus values for a light source can be calculated easily from measurements taken at 1 or 5 nm intervals from the following equations:

$$X = \frac{k_m}{\sum \bar{y}_\lambda} \sum S_\lambda \bar{x}_\lambda \tag{48.8}$$

$$Y = \frac{k_m}{\sum \bar{y}_\lambda} \sum S_\lambda \bar{y}_\lambda \tag{48.9}$$

$$Z = \frac{k_m}{\sum \bar{y}_\lambda} \sum S_\lambda \bar{z}_\lambda \tag{48.10}$$

where

S_λ is the relative spectral power of the light source at wavelength λ
k_m is the maximum spectral luminous efficacy function (683 lm W^{-1})
$\bar{x}_y$, $\bar{y}_\lambda$, and $\bar{z}_\lambda$ are the color-matching functions for the standard observer

48.5 Reflectance Measurements

48.5.1 Specular and Diffuse Reflectance

Specular (sometimes called *regular*) reflection is the mirrorlike reflection from an object. If you shine a beam of light on a mirror, it will be entirely reflected at the same angle on the opposite side of a normal plane to the mirror's surface. However, if you shine a beam of light on a pellet of compressed barium sulfate ($BaSO_4$) powder, it will enter the surface, be scattered a number of times, and exit the pellet in many directions. This is called diffuse reflection. Glossy and semiglossy materials contain a combination of specular and diffuse reflection as shown in Figure 48.7. Observers usually discount the specular reflection when visually evaluating the color of a material.

48.5.2 Illuminating and Viewing Geometries for Reflectance

Instruments designed for measuring the color of reflecting objects consist of an illuminator, a sample holder, and a receiver. The CIE has recommended four illuminating and viewing geometries for making reflectance measurements.

Bidirectional geometries illuminate the sample with a narrow beam of light and view the sample with a receiver having a narrow entrance field. In the most commonly used bidirectional geometry, the sample is illuminated at a 45° angle (+2°) from the sample's normal and viewed along the sample's normal (±10°). The other recommended geometry illuminates the sample along its normal and views the sample at 45° from its normal. Bidirectional geometries measure only diffuse reflection and can be sensitive to the surface texture of the sample. The two bidirectional geometries produce equivalent results [8,9].

In the most common diffuse geometry, the illuminator includes an integrating sphere. An integrating sphere is a hollow metal sphere coated with a highly reflecting white coating with openings for the light source, the sample, and the receiver. The instrument's light source projects a beam of light onto the integrating sphere's wall. The light is reflected many times by the sphere's wall, and the sample is illuminated from all angles. When measuring in the diffuse/normal mode, *d*/0, the receiver views the sample along its normal. When the receiver is positioned in this manner, the specular reflection is directed back toward the light source and is not measured.

For some applications, it is useful to remove the specular reflection off of the sample's surface from the reflection of the light that is reflected back from the interior of the sample. This can be done by moving the receiver 6° or 8° from the sample's normal as shown in Figure 48.8. A specular inclusion port is placed at the equal angle on the opposite side of the sample's normal. By placing a white plug having the same reflectance as the sphere wall in the port, the specular reflection can be included in the measurement. By using a light trap at the specular inclusion port, some or all of the specular reflection can be excluded from the measurement. The narrower the specular reflection peak, that is, the glossier

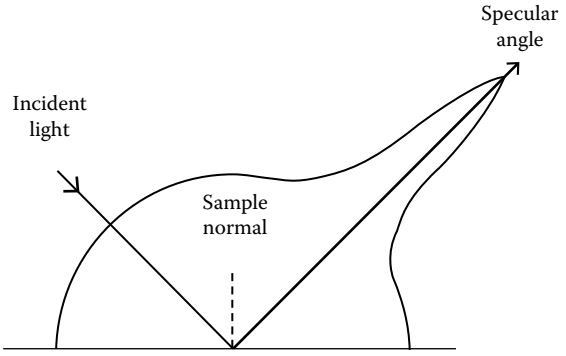


FIGURE 48.7 A semiglossy material showing a combination of specular and diffuse reflection.

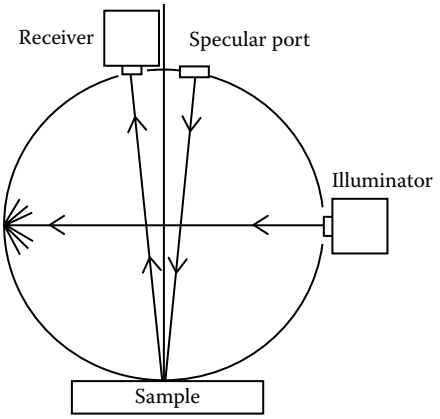


FIGURE 48.8 The diffuse/normal illuminating/viewing geometry with the ability to include or exclude the specular surface reflection.

the material, the more specular reflection is excluded. Specular-included measurements are normally abbreviated as SCI, while specular-excluded measurements are abbreviated as SCE. This geometry is referred to as diffuse/near normal but still abbreviated as *d*/0. When the specular reflection is included, this geometry is sometimes referred to as total/normal, *t*/0.

The last of the CIE recommended illuminating and viewing geometries is the normal/diffuse or near normal/diffuse, 0/*d* and 0/*t*. This is the reverse of the *d*/0 geometry. The illuminator illuminates the sample along its normal or slightly off of its normal. For this geometry, the integrating sphere is part of the receiver. Light reflected from the sample is captured by the integrating sphere, and the remaining optics of the receiver views the sphere wall. The two diffuse geometries produce equivalent results [8,9].

48.5.3 Monochromatic and Polychromatic Illumination

An instrument's illuminator may illuminate the sample with either a narrow band of wavelengths, 1–10 nm wide, called monochromatic illumination, or a wide band of wavelengths, usually simulating a daylight illuminant and called polychromatic illumination. For nonfluorescent samples, either illumination method can be used with equivalent results. However, for fluorescent samples, only polychromatic illumination can produce valid results.

48.5.4 Sample Texture and Bidirectional Geometries

Bidirectional illuminating and viewing geometries can be very sensitive to surface texture and any polarization of reflected light. Keeping the illuminator and receiver in the same plane maximizes this sensitivity. To reduce or eliminate this sensitivity, circumferential or annular illumination (or receiving) can be used. When an illuminator provides light (or the receiver possesses responsivity) at many points distributed uniformly around a 45° cone centered at the sample's normal, we refer to circumferential illumination (or viewing). When the illuminator provides light continuously and uniformly around the cone, we refer to annular illumination (or viewing).

48.5.5 Which Illuminating and Viewing Geometry Is Best?

"Which illuminating and viewing geometry is best for color measurement?" is a difficult question to answer [9–12]. For matte samples, all of the geometries produce equivalent results. For high-gloss samples, the diffuse geometries with the specular reflection excluded provide measurements very close

to the bidirectional geometries. For semigloss samples, the problem becomes more complex. The two bidirectional geometries are similar to the way a person evaluates color visually and are often thought to agree better with visual evaluation. On the other hand, the diffuse geometries measured with the specular reflection included minimize the effect of the sample's texture and gloss and are quite useful for computerized color matching. Rather than purchasing multiple instruments, most users select the geometry most suited to their major needs and compromise on other measurements.

48.5.6 Spectrophotometers

Spectrophotometers are used to measure an object's reflection characteristics throughout the visible spectrum. Double-beam spectrophotometers monitor a reference standard to compensate dynamically for fluctuations in source output, detector response, and atmospheric absorption to increase the instrument's stability [13]. Improvements in electronics and optics have allowed single-beam spectrophotometers to achieve the stability of double-beam instruments [14]. Single-beam instruments with an integrating sphere require a correction for the reduction of sphere efficiency caused by sample absorption [14].

Reflectance measurements are referred to as if they take place at a single wavelength. In actuality, a spectrophotometer has a finite spectral bandwidth or bandpass centered about that wavelength. Some instruments will have a bandpass as narrow as 1 nm, while others may exceed 20 nm. Bandpass is important because it influences the color measurement. The extent of this influence depends on the sample being measured [15,16]. The spectral bandpass should be equal to the wavelength increment used in the calculation of tristimulus values to obtain the best color measurement results [17].

Spectrophotometers must be standardized before making reflectance factor measurements. The high point of the measurement scale is set by measuring a white standard of known reflectance factor [18]. The zero point is set by measuring a light trap or a black calibration standard. Single-beam, integrating sphere spectrophotometers may also require a sphere wall calibration, which is often done with a gray standard.

Some spectrophotometers allow the user to vary the size of the area measured. For most applications, the largest area possible should be measured. A number of documentary standards exist for making and reporting reflectance measurements [9,17,19–25].

48.5.7 Colorimeters and Spectrocolorimeters

Colorimeters were developed in the 1920s–1930s as a less expensive alternative to spectrophotometers for quality control and color difference applications. They are simple to use and directly measure a sample's tristimulus values or related color coordinates. Three or four filters modify the light source and attempt to duplicate a standard illuminant/standard observer combination. Because of the difficulty in matching the CIE illuminant and standard observer functions, they are less accurate than spectrophotometers for determining a sample's tristimulus values. Colorimeters determine the color difference between two samples more accurately than they determine tristimulus values, and they are often called color difference meters. Since only one standard illuminant/standard observer combination is usually possible, colorimeters cannot be used to determine metamerism. Standards also exist for making colorimeter measurements [24–26].

A new class of instruments, spectrocolorimeters, began to appear in the 1980s. Spectrocolorimeters are spectrophotometers that only output tristimulus values or related color coordinates. They are less expensive and often have fewer options (such as variable area of view) than fully functional spectrophotometers. However, because they are spectrophotometers, they are capable of measuring metamerism.

48.5.8 Sample Preparation

Accurate color measurement is often dependent on sample preparation. Ideally, a sample for reflectance measurement is flat, has a uniform gloss and texture, is opaque, and is nondirectional. Always strive for sample preparation techniques that are reproducible. Consult standard test methods [24], standard practices [19,27–29], books, and articles [30–33] for help and advice on sample preparation.

48.6 Transmittance Measurement

48.6.1 Regular and Diffuse Transmittance

When a beam of light passes through a "transparent" material along its normal, the intensity of the beam will be decreased by absorption, but the direction of the beam will be unchanged. This is called regular transmittance. When a beam of light passes through a hazy or translucent material along its normal, the material scatters light and spreads the beam. This is called diffuse transmittance.

48.6.2 Illuminating and Viewing Geometries for Transmittance

The CIE has recommended three illuminating and viewing geometries for transmittance measurements. [1] Most transmittance measurements are made with spectrophotometers designed for analytical chemistry applications that use a normal/normal, 0/0, geometry. The illuminator directs the incident beam along the sample's normal, and the receiver views the sample along its normal. Only regular transmittance can be measured accurately using this geometry.

Regular and diffuse transmittance can be measured using the normal/diffuse, 0/*d*, geometries (or the equivalent diffuse/normal, *d*/0), which the CIE also recommended for reflectance measurements. Regular transmittance is measured by keeping the sample as far away as possible from the integrating sphere, whereas diffuse transmittance is measured by placing the sample in contact with the sphere. Instruments designed specifically for transmittance measurements would have only a sample or illuminator port and a receiver port. Instruments designed for reflectance measurements can be used for making transmittance measurements by placing a white material in the sample port and using the specular-included mode of measurement. The illuminator or receiver port would then serve as the sample port.

The third geometry recommended by the CIE, diffuse/diffuse, *d*/*d*, is not often used for industrial color measurement. One integrating sphere is used to illuminate the sample, and a second integrating sphere is used to view the sample.

The two bidirectional reflectance geometries (45/0 and 0/45) have also been used for making regular transmittance measurements.

48.6.3 Monochromatic and Polychromatic Illumination

Instruments for measuring transmittance may have either monochromatic or polychromatic illumination. The transmittance of nonfluorescent samples can be measured using either type of illumination, but fluorescent samples can only be measured using polychromatic illumination.

48.6.4 Standardizing Instruments for Transmittance Measurements

Three techniques exist for setting the high end of the measurement scale, a transmittance factor of 1.0. Each technique produces different results, so it is important to document the method used.

Setting the instrument to read a transmittance factor of 1.0 with no sample in the sample compartment is the easiest technique. The transmittance measurements will then be relative to air. This technique is often used when solid samples are being measured.

Transmittance measurements of solid materials may also be made relative to a clear blank of similar material. To make these measurements, the clear blank is placed in the sample compartment before the instrument is standardized.

When liquids are to be measured, a holder containing solvent or nonabsorbing liquid of the same refractive index as the liquid to be measured should be placed in the sample compartment before standardizing the instrument. This eliminates any effects of the holder's transmittance.

48.6.5 Sample Preparation

Ideal samples will be flat with parallel sides. Liquid sample holders should be made from optical glass. Because transmittance will change with sample thickness and the concentration of colorant, it is important that sample holders and blanks used for standardization have the same path length or thickness as the sample. When two different samples must be compared, they should be prepared using the same technique and have the same thickness.

It is extremely difficult to measure the transmittance of curved materials because the curvature of the object may act as a lens and deflect the incident light away from the receiver. If the curvature is not too great, it may be possible to make a diffuse transmittance measurement.

48.7 Color Difference Calculations

A number of equations have been developed over the years for calculating the color difference between two objects [1,10,24,30,34–36]. One of the most common equations is the CIELAB recommended by the CIE in 1976 [1]. In CIELAB, L^* represents the difference between light (where $L^* = 100$) and dark (where $L^* = 0$). A^* represents the difference between green ($-a^*$) and red ($+a^*$), and b^* represents the difference between yellow ($+b^*$) and blue ($-b^*$).

First, the CIELAB coordinates, L^* (lightness), a^*, b^*, C^*_{ab} (chroma), and h_{ab} (hue angle), are calculated with Equations 48.11 through 48.15:

$$L^* = 116 f\left(\frac{Y}{Y_n}\right) - 16 \tag{48.11}$$

$$a^* = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \tag{48.12}$$

$$b^* = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \tag{48.13}$$

$$C^*_{ab} = [a^{*2} + b^{*2}]^{1/2} \tag{48.14}$$

$$h_{ab} = \arctan\left[\frac{b^*}{a^*}\right] \tag{48.15}$$

in which X, Y, and Z are tristimulus values and the subscript n refers to the tristimulus values of the perfect diffuser for the given illuminant and standard observer; $f(X/X_n) = (X/X_n)^{1/3}$ for values of (X/X_n) greater than 0.008856 and $f(X/X_n) = 7.787(X/X_n) + 16/116$ for values of (X/X_n) equal to or less than 0.008856 and the same with Y and Z replacing X in turn. The hue angle is 0° along the $+a^*$ axis, 90° along the $+b^*$ axis, 180° along the $-a^*$ axis, and 270° along the $-b^*$ axis.

The total color difference, ΔE^*_{ab}, and its components—the lightness difference, ΔL^*; the chroma difference, ΔC^*_{ab}; and the hue difference, ΔH^*_{ab}—are calculated using Equations 48.16 through 48.21:

$$\Delta L^* = L^*_{trial} - L^*_{standard} \tag{48.16}$$

$$\Delta a^* = a^*_{trial} - a^*_{standard} \tag{48.17}$$

$$\Delta b^* = b^*_{trial} - b^*_{standard} \tag{48.18}$$

$$\Delta E^*_{ab} = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2} \tag{48.19}$$

$$\Delta C^*_{ab} = C^*_{ab\ trail} - C^*_{ab\ standard} \tag{48.20}$$

$$\Delta H^*_{ab} = [(\Delta E^*_{ab})^2 - (\Delta L^*)^2 - (\Delta C^*_{ab})^2]^{1/2} \tag{48.21}$$

A negative value of ΔL^* means the trial is darker than the standard, and a negative value of ΔC^*_{ab} means the trial has a lower chroma than the standard. When the hue angle h_{ab} of the trial is greater than that of the standard, the sign of ΔH^*_{ab} is positive, and vice versa. The total color difference and its components can then be used in setting color tolerances [19,37].

Researchers have been making modifications to the CIELAB color difference equation in an attempt to have the color difference more closely correlate with visually observed color differences. The CMC(*l*:*c*) color difference equation [38–40] has gained great acceptance in the textile industry and is being tested in other areas. Starting with the CIELAB color differences, the CMC(*l*:c) color differences are calculated with Equations 48.22 through 48.28:

$$\Delta E = \left[\left(\frac{\Delta L^*}{l S_L} \right)^2 + \left(\frac{\Delta C^*_{ab}}{c S_C} \right)^2 + \left(\frac{\Delta H^*_{ab}}{S_H} \right)^2 \right]^{1/2} \tag{48.22}$$

$$S_L = \frac{0.040975 L^*}{1 + 0.01765 L^*} \tag{48.23}$$

unless $L^* < 16$, in which case $S_L = 0.511$,

$$S_C = \frac{0.0638 C^*_{ab}}{1 + 0.0131 C^*_{ab}} + 0.638 \tag{48.24}$$

$$S_H = (FT + 1 - F) S_C \tag{48.25}$$

$$F = \left(\frac{(C^*_{ab})^4}{(C^*_{ab})^4 + 1900} \right)^{1/2} \tag{48.26}$$

$$T = 0.36 + \text{abs}[0.4 \cos(35 + h_{ab})] \tag{48.27}$$

unless h is between 164° and 345° when

$$T = 0.56 + \text{abs}[0.2\cos(168 + h_{ab})] \tag{48.28}$$

in which the notation "abs" indicates the absolute (i.e., positive) value of the term inside the square brackets. When $l = c = 1$, the formula quantifies the perceptibility of color differences. Optimum values of l and c for quantifying the acceptability of a color match must be determined for the material being measured. The textile industry has found the optimum values to be $l = 2$ and $c = 1$ [39].

In 1994, the CIE recommended a new color difference equation, CIE94 [36], which is similar to the CMC equation. CIE94 color differences are calculated using Equations 48.29 through 48.33. The perceived color difference, ΔV, is related to the measured color difference through an overall sensitivity factor, k_E:

$$\Delta V = \left(\frac{1}{k_E}\right)\Delta E^*_{94} \tag{48.29}$$

$$\Delta E^*_{94} = \left[\left(\frac{\Delta L^*}{k_L S_L}\right)^2 + \left(\frac{\Delta C^*_{ab}}{k_C S_C}\right)^2 + \left(\frac{\Delta H^*_{ab}}{k_H S_H}\right)^2\right]^{1/2} \tag{48.30}$$

$$S_L = 1 \tag{48.31}$$

$$S_C = 1 + 0.045 C^*_{ab} \tag{48.32}$$

$$S_H = 1 + 0.015 C^*_{ab} \tag{48.33}$$

The overall sensitivity factor, k_E, is used to account for variation in the illuminating and viewing conditions. A person in the textile industry who is using CMC(2:1) and would like to compare the results with CIE94 would set $k_L = 2$ and $k_C = k_H = 1$, that is, CIE94(2:1:1).

The improvement in correlating with visual assessments of color difference could result in either the CMC(*l*:*c*) or the CIE94 replacing the CIELAB color difference equation. Hunter and Harold [30] detail many of the older color difference equations, many of which are still in use. A history of the development of color metrics was written by Richter [41].

48.8 Special Cases

48.8.1 Fluorescent Samples

Fluorescent materials not only reflect light but also emit light. Light absorbed at some wavelengths is emitted at longer wavelengths. The amount of light emitted depends on the intensity and the spectral power distribution of the source. Because of the emission of light, spectrophotometers that illuminate a fluorescent material with monochromatic light cannot be used, because the light emitted at the longer wavelengths will be measured as if it had been reflected from the material. Reflectance measurements of the material illuminated by polychromatic light will include the emitted light at the proper wavelengths. If the instrument's light source is a good representation of the illuminant, the measurement will be indicative of the observed color. Techniques have been recommended by the CIE for assessing the quality of daylight simulators [42]. Measured reflectance factors at the wavelengths of emittance may be greater than 1.0. Special fluorescent calibration standards must be used to accurately measure these materials [43]. The 45/0 or 0/45 bidirectional geometries are recommended for measuring fluorescent materials [44,45].

The complete analysis of a fluorescent material requires that the emittance be separated from the reflectance. Spectrofluorimeters were designed to analyze fluorescent samples. These instruments illuminate the sample with monochromatic light. Reflected and emitted rays pass through a second monochromator to isolate the receiver wavelengths. By viewing the sample at the same wavelength as it is illuminated, the true reflectance of the sample can be determined. Either the excitance spectra or the emittance spectra can be studied by setting each monochromator at different wavelengths. Techniques have also been developed to correct spectrophotometric measurements of fluorescent samples [45,46].

48.8.2 Metallic and Pearlescent Samples

Materials that contain metallic and/or pearlescent pigments are goniochromatic, that is, they change color with the illuminating and viewing geometry. Goniospectrophotometers are needed to measure these materials. A goniospectrophotometer illuminates (or views) the sample at a fixed angle, usually 45°, and views (or illuminates) the sample at three or more angles. The position of the receivers (or illuminators) is described by the specular angle, the viewing angle measured from the direction of the specular reflectance, which is equal and opposite the angle of illumination. In a goniospectrophotometer that illuminates the sample at 45° and views at three angles, one of the viewing angles would be near the specular reflection at about 25°, the second at the CIE recommended bidirectional angle of 0°, and the third would be far away from the specular reflectance at about 70°.

Goniospectrophotometry is still in its infancy and the CIE recommendations and ASTM standards for making these measurements are still under development.

48.8.3 Retroreflecting Samples

Highway signs and high-visibility clothing are examples of retroreflectors. Light shining on a retroreflector is returned in directions very close to the illumination angle. Most retroreflectance measurements are made in connection with highway safety. There are only a few specially built instruments for measuring retroreflection. A projector is usually used to illuminate the sample, and a teleradiometer or a radiometric telecolorimeter is used to view the sample. Tristimulus values can be calculated from the retroreflectance factors. Standard practices and test methods exist for specifying the illuminating and viewing geometries and making retroreflectance measurements [47–51].

48.8.4 Lamps, Light Sources, and Displays

Spectral radiometers and radiometric colorimeters were designed to measure lamps, light sources, and displays. These instruments are similar to spectrophotometers and colorimeters but do not need an illuminator, because the sample being measured emits light.

Lamps and light sources can either be measured directly or by measuring the reflectance of a stable white reflecting surface being illuminated by the lamp or light source. To measure televisions, computer monitors, and other similar devices, the emitted light must usually be imaged directly on the optics of the spectral radiometer or radiometric colorimeter [52,53].

48.8.5 Hazy and Translucent Materials

Hazy and translucent materials are measured by placing them in contact with an integrating sphere and measuring their diffuse transmittance. A haze index can be calculated from four diffuse transmittance measurements on a reflecting spectrophotometer [54]. Two of the measurements are made with a white material in the reflectance sample port, and two measurements are made with a light trap in the port. Special instruments called *hazemeters* were designed to make haze measurements.

TABLE 48.1 Color-Measuring Instruments

Instrument Type
Colorimeters
Goniospectrophotometers
Hazemeters
Radiometric colorimeters
Retroreflectometers
Spectral analyzers
Spectrofluorimeters
Spectrophotometers
Spectroradiometers

TABLE 48.2 Instrument Manufacturers

BYK-Gardner, United States
Rivers Park II, 9104 Guilford Road
Columbia, MD 21046–2729
(301) 483–6500
Spectrophotometers, colorimeters, hazemeters

ColorTec
28 Center Street
Clinton, NJ 08809
(908) 735–2248
Spectral analyzers

Datacolor International
5 Princess Road
Lawrenceville, NJ 08648
(609) 924–2189
Spectrophotometers, goniospectrophotometers

GretagMacbeth
617 Little Britain Road
New Windsor, NY 12553–6148
(914) 565–7660
Spectrophotometers, goniospectrophotometers

Labsphere, Inc.
P.O. Box 70, Shaker Street
North Sutton, NH 03260–0070
(603) 927–4266
Spectroradiometers, radiometric colorimeters, spectrofluorimeters

Minolta Corporation
101 Williams Drive
Ramsey, NJ 07446
(201)825–4000
Spectrophotometers, spectroradiometers, colorimeters, radiometric colorimeters, spectrofluorimeters

X-Rite, Inc.
3100 44th St. SW
Grandville, MI 49418
(616) 534–7663
Spectrophotometers, spectrocolorimeters, goniospectrophotometers

Color Savvy, Ltd.
305 S. Main Street
Springboro, OH 45066
(513) 748–9160
Colorimeters, spectrophotometers

Datacolor International
5 Princess Road
Lawrenceville, NJ 08648
(609)924–2189
Spectrophotometers, goniospectrophotometers

Gamma Scientific Co.
8581 Aero Drive
San Diego, CA 92123
(619) 279–8034
Spectroradiometers, retroreflectometers

Hunter Associates Laboratory, Inc.
11491 Sunset Hills Road
Reston, VA 20190
(703) 471–6870
Spectrophotometers, colorimeters

Light Source, Inc.
4th floor, 4040 Civic Center Drive
San Rafael, CA 94903
(415) 446–4200
Spectrophotometers, spectroradiometers

Photo Research Inc.
9330 DeSoto Avenue
Chatsworth, CA 91311
(818)341–5151
Spectroradiometers

48.9 Instrument Manufacturers

Costs of color-measuring instruments vary significantly. The more accurate instruments with the best repeatability and reproducibility can be expected to cost more. Table 48.1 lists the price ranges for various types of color-measuring instruments. There is a classification in the table called spectral analyzers. Although these instruments may provide information that makes them appear to be colorimeters or spectrophotometers, they use measurement techniques not traditionally associated with those instruments. For example, several use LEDs instead of more traditional light sources. Since the instruments vary so widely in capability, the purchaser must ensure that a particular instrument has the capability to make the measurements desired with sufficient accuracy, repeatability, and reproducibility.

Table 48.2 provides contact information for a number of manufacturers of color-measuring instruments. The manufacturers listed in this table design instruments specifically for color measurement. Thus, manufacturers of analytical spectrophotometers are excluded even if their instruments have color-measuring capabilities. The table is also limited to manufacturers with a major presence in the United States.

Defining Terms

For a more extensive collection of terms relating to color and appearance, the reader should refer to ASTM E 284 Standard Terminology of Appearance [55].

Chroma: Attribute of color used to indicate the degree of departure of the color from a gray of the same lightness.

CIE: The abbreviation for the French title of the International Commission on Illumination, Commission Internationale de l'Éclairage.

CIE standard observers: The ideal colorimetric observer data adopted by the CIE to represent the response of the average human eye, when light-adapted, to an equal-energy spectrum. The standard observer adopted in 1931 was developed from data obtained with a 2° field of vision and is commonly called the "2° standard observer." The standard observer adopted in 1964 was developed from data obtained with a 10° annular field of vision and is commonly called the "10° standard observer."

CIE tristimulus values: Amounts of the three mathematical lights necessary in a three-color additive mixture required for matching a color in the CIE System. They are designated *X*, *Y*, and *Z*. The illuminant and standard observer color-matching functions must be designated.

Colorimetry: The science of color measurement.

Hue: The attribute of color perception by means of which a color is judged to be red, orange, yellow, green, blue, purple, or intermediate between adjacent pairs of these, considered in a close ring (red and purple being an adjacent pair.) White, black, and grays possess no hue.

Illuminant: A mathematical description of the relative power emitted by a real or imaginary light source at each wavelength in its emission spectrum.

Lightness: (1) The attribute of color perception by which a nonself-luminous body is judged to reflect more or less light. (2) The attribute by which a perceived color is judged to be equivalent to one of a series of grays ranging from black to white.

Metamerism: Property of two specimens that match under a specified illuminator and to a specified observer and whose spectral reflectances or transmittances differ in the visible wavelengths.

Perfect reflecting diffuser: Ideal reflecting surface that neither absorbs nor transmits light, but reflects diffusely, with the radiance of the reflecting surface being the same for all reflecting angles, regardless of the angular distribution of the incident light.

Reflectance: Ratio of the reflected radiant or luminous flux to the incident flux in the given conditions. The term reflectance is often used in a general sense or as an abbreviation for reflectance factor. Such usage may be assumed unless the previous definition is specifically required by the context.

Reflectance factor: Ratio of the flux reflected from the specimen to the flux reflected from the perfect reflecting diffuser under the same geometric and spectral conditions of measurement.

References

1. CIE, *CIE Publication 15.2, Colorimetry,* 2nd edn., Commission International de l'Éclairage (CIE), Central Bureau of the CIE, Vienna, Austria, 1986. Available from USNC/CIE Publications, TLA–Lighting Consultants, Inc., Salem, MA.
2. W. H. Foster, Jr., R. Gans, E. I. Steams, and R. E. Steams, Weights for calculation of tristimulus values from sixteen reflectance values. *Color Eng.*, **8**(3), 25–47, 1970.
3. E. I. Stearns, The determination of weights for use in calculating tristimulus values. *Color Res. Appl.*, **6**, 210–212, 1981.
4. E. I. Stearns, Calculation of tristimulus values and weights with the revised CIE recommendations. *Text. Chem. Color.*, **17**(8), 162/53–168/59, 1985.
5. H. S. Fairman, The calculation of weight factors for tristimulus integration. *Color Res. Appl.*, **10**, 1199–1203, 1985.
6. F. W. Billmeyer, Jr. and H. S. Fairman, CIE method for calculating tristimulus values. *Color Res. Appl.*, **12**, 27–36, 1987.
7. ASTM, ASTM E 308 Standard Practice for Computing the Colors of Objects by Using the CIE System. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
8. F. W. Billmeyer, Jr. and R. T. Marcus, Effect of illuminating and viewing geometry on the color coordinates of samples with various surface textures. *Appl. Opt.*, **8**, 1763–1768, 1969.
9. ASTM, ASTM E 179 Standard Guide for Selection of Geometric Conditions for Measurement of Reflection and Transmission Properties of Materials. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
10. R. W. G. Hunt, *Measuring Colour*, 2nd edn., Ellis Horwood Limited, Chichester, England: 1991.
11. D. C. Rich, The effect of measuring geometry on computer color matching. *Color Res. Appl.*, **13**, 113–118, 1988.
12. T. J. Mabon, Presented at the *Regional Technical Conference*, Chiang Mai, Thailand; Color measurement of plastics: Which geometry is best. *Color Tolerances: Measuring up to Today's Standards*, Society of Plastics Engineers, Inc., Cherry Hill, NJ: September 14–16, 1992.
13. J. C. Zwinkels, Errors in colorimetry caused by the measuring instrument. *Text. Chem. Color.*, **21**(2), 23–29, 1989.
14. R. H. Stanziola, H. Hemmendinger, and B. Momiroff, The spectro sensor: A new generation spectrophotometer. *Color Res. Appl.*, **4**, 157–163, 1979.
15. D. Strocka, Presented at the *2nd AIC Congress "Colour 73,"* York, England, July, 1973; Are intervals of 20nm sufficient for industrial colour measurement? In R. W. G. Hunt (ed.), *Colour 73*, John Wiley & Sons, New York: 1973.
16. H. Schmelzer, Influence of the design of instruments on the accuracy of color-difference and color-matching calculations. *J. Coat. Technol.*, **58**(739), 53–59, 1986.
17. ASTM, ASTM E 1164 Standard Practice for Obtaining Spectrophotometric Data for Object-Color Evaluation. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.

18. E. C. Carter, F. W. Billmeyer, Jr., and D. C. Rich, *Guide to Material Standards and Their Use in Color Measurement*, ISCC Technical Report 89-1. Available from The Inter-Society Color Council, Reston, VA.
19. SAE, SAE J1545 *Recommended Practice for Instrumental Color Difference Measurement for Exterior Finishes, Textiles, and Colored Trim*, Society of Automotive Engineers, Troy, MI.
20. ASTM, ASTM E 429 Standard Test Method for Measurement and Calculation of Reflecting Characteristics of Metallic Surfaces Using Integrating Sphere Instruments. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
21. ASTM, ASTM E 805 Standard Practice for Identification of Instrumental Methods of Color or Color-Difference Measurement of Materials. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
22. ASTM, ASTM E 1331 Standard Test Method for Reflectance Factor and Color by Spectrophotometry Using Hemispherical Geometry. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
23. ASTM, ASTM E 1349 Standard Test Method for Reflectance Factor and Color by Spectrophotometry Using Bidirectional Geometry. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
24. AATCC, AATCC Test Method 153 Color Measurement of Textiles: Instrumental. *AATCC Technical Manual/1993*, American Association of Textile Chemists and Colorists, Research Triangle Park, NC.
25. NPES, American National Standard CGATS.5 Graphic Technology—Spectral Measurement and Colorimetric Computation for Graphic Arts Images, NPES The Association for Suppliers of Printing and Publishing Technologies, Reston, VA.
26. ASTM, ASTM E 1347 Standard Test Method for Color and Color-Difference Measurement by Tristimulus (Filter) Colorimetry. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
27. ASTM, ASTM D 3925 Standard Practice for Sampling Liquid Paints and Related Pigmented Coatings. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
28. ASTM, ASTM D 3964 Standard Practice for Selection of Coating Specimens for Appearance Measurements. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
29. ASTM, ASTM D 823 Standard Practices for Producing Films of Uniform Thickness of Paint, Varnish, and Related Products on Test Panels. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
30. R. S. Hunter and R. W. Harold, *The Measurement of Appearance*, 2nd edn., John Wiley & Sons, New York: 1987.
31. R. L. Connelly, Sr., Preparation and mounting textile samples for color measurement. In G. Celikiz and R. G. Kuehni (eds.), *Color Technology in the Textile Industry*, American Association of Textile Chemists and Colorists, Research Triangle Park, NC: 1983.
32. C. Wilson and E. I. Stearns, The spectrophotometric reflectance measurement of small samples. In G. Celikiz and R. G. Kuehni (eds.), *Color Technology in the Textile Industry*, American Association of Textile Chemists and Colorists, Research Triangle Park, NC: 1983.
33. E. I. Stearns and W. B. Prescott, Measurement of translucent cloth samples. In G. Celikiz and R. G. Kuehni (eds.), *Color Technology in the Textile Industry*, American Association of Textile Chemists and Colorists, Research Triangle Park, NC: 1983.
34. G. Wyszecki and W. S. Stiles, *Color Science, Concepts and Methods, Quantitative Data and Formulae*, 2nd edn., John Wiley & Sons, New York: 1992.
35. ASTM, ASTM D 2244 Standard Test Method for Calculation of Color Differences from Instrumentally Measured Color Coordinates. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.

36. CIE, *CIE Technical Report 116, Industrial Colour-Difference Evaluation*, Commission International de l'Éclairage (CIE), Central Bureau of the CIE, Vienna, 1995. Available from USNC/CIE Publications, TLA–Lighting Consultants, Inc., Salem, MA.
37. ASTM, ASTM D 3134 Standard Practice for Establishing Color and Gloss Tolerances. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
38. F. J. J. Clarke, R. McDonald, and B. Rigg, Modification to the JPC79 colour-difference formula. *J. Soc. Dyers. Colour.*, **100**, 128–132 and 281–282, 1984.
39. British Standards Institution, BS6923:1988 British standard method for calculation of small colour differences. Available from British Standards Institution, London, U.K.
40. AATCC, AATCC Test Method 173 CMC: Calculation of small color differences for acceptability. *AATCC Technical Manual/1993*, American Association of Textile Chemists and Colorists, Research Triangle Park, NC.
41. M. Richter, The development of color metrics. *Color Res. Appl.*, **9**, 69–83, 1984.
42. CIE, *CIE Publication 51, A Method for Assessing the Quality of Daylight Simulators for Colorimetry*, Commission International de l'Éclairage (CIE), Central Bureau of the CIE, Vienna, 1982. Available from USNC/CIE Publications, TLA–Lighting Consultants, Inc., Salem, MA.
43. D. Gundlach and H. Terstiege, Problems in measurement of fluorescent materials. *Color Res. Appl.*, **19**, 427–436, 1994.
44. ASTM, ASTM E 991 Standard Practice for Color Measurement of Fluorescent Specimens. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
45. R. A. McKinnon, Methods of measuring the colour of opaque fluorescent materials. *Rev. Prog. Color.*, **17**, 56–60, 1987.
46. F. W. Billmeyer, Jr., Metrology, documentary standards, and color specifications for fluorescent materials. *Color Res. Appl.*, **19**, 413–425, 1994.
47. ASTM, ASTM E 808 Standard Practice for Describing Retroreflection. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
48. ASTM, ASTM E 809 Standard Practice for Measuring Photometric Characteristics of Retroreflectors. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
49. ASTM, ASTM E 810 Standard Test Method for Coefficient of Retroreflection of Retroreflective Sheeting. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
50. ASTM, ASTM D 4061 Standard Test Method for Retroreflection of Horizontal Coatings. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
51. ASTM, ASTM E 811 Standard Practice for Measuring Colorimetric Characteristics of Retroreflectors Under Nighttime Conditions. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
52. ASTM, ASTM E 1336 Standard Test Method for Obtaining Colorimetric Data from a Video Display Unit by Spectroradiometry. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
53. ASTM, ASTM E 1455 Standard Practice for Obtaining Colorimetric Data from a Video Display Unit Using Tristimulus Colorimeters. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
54. ASTM, ASTM D 1003 Standard Test Method for Haze and Luminous Transmittance of Transparent Plastics. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.
55. ASTM, ASTM E 284 Standard Terminology of Appearance. *Annual Book of ASTM Standards*, American Society for Testing and Materials, West Conshohocken, PA.

Further Information

American Society for Testing and Materials, *ASTM Standards on Color and Appearance Measurement*, Sponsored by ASTM Committee E-12 on Appearance of Materials, 5th edn., American Society for Testing and Materials, Philadelphia, PA: 1996.

Billmeyer, F. W., Jr. and M. Saltzman, *Principles of Color Technology*, 2nd edn., John Wiley & Sons, New York: 1981.

Color Research and Application, E. C. Carter (ed.), Periodical, Currently Volume 8, Issue 4, 2013. John Wiley & Sons, New York.

Special Topics, The Inter-Society Color Council, Reston, VA, http://www.iscc.org/, accessed on August 20, 2013.

49

Polarization Measurement

Soe-Mie F. Nee
U.S. Naval Air Warfare Center

49.1 Basic Concepts of Polarization

Polarization of light is a property of electromagnetic (EM) waves, which include heat, microwaves, radio waves, and x-rays. An EM wave has orthogonal electric and magnetic fields associated with it that vibrate in directions perpendicular to the direction of propagation. The *electric field* of a sinusoidal EM wave, in particular, can always be decomposed into two orthogonal components; each component has an *amplitude* and a *phase*. The amplitude is the maximum value of the field component, and the light *intensity* is proportional to the square of the amplitude. The phase, referred to a fixed position or time, tells what part of the cycle the electric field is vibrating in. G. G. Stokes pointed out in 1852 that these two orthogonal components do not interfere in amplitude but are additive according to vector algebra [1]. When the two orthogonal components are in phase, the EM wave is *linearly polarized*. When the two orthogonal components have the same amplitude and a relative phase of 90°, the EM wave is *circularly polarized*. In general, an EM wave has arbitrary amplitudes and phases for the two orthogonal fields and is elliptically polarized. The concept of polarization ellipse and the descriptions for the polarization of an EM wave in terms of Jones vectors and Stokes vectors are given in the subsection "Polarization of an EM Wave" and also in Refs. [1–9].

Light is composed of an ensemble of EM waves. A group of EM waves traveling in the same direction can have some linearly polarized waves, some circularly polarized waves, and some elliptically polarized waves. When they are combined, resulting light can be unpolarized, partially linearly polarized, or partially elliptically polarized. Unpolarized light occurs when there are no fixed directions of the electric field and also no fixed phase relations between the two orthogonal field components. In general, light is

partially polarized and can be decomposed into unpolarized light and elliptically polarized light. These concepts are described in terms of Stokes vector in the subsection "Polarization of Light" and also in Refs. [1–6].

Polarized light can be produced by passing light through a polarizer. An ideal polarizer transmits only light whose electric field is parallel to the transmission axis of the polarizer and rejects light with the orthogonal field. Polarization of light can be observed by stacking two polarizers together and turning one with respect to the other. The transmitted light intensity through these two polarizers will vary, and at some particular positions, it will vanish. In this case, light is linearly polarized after passing through the first polarizer. When the second polarizer is turned until its axis is perpendicular to the axis of the first polarizer, light cannot pass through the second polarizer. The transmitted intensity varies according to the square of the cosine of the angle between the two polarizers [5,9–12]. Real polarizers are not perfect and transmit light with minimum intensity $I_{\min}$ when the polarizer axis is perpendicular to the polarization of purely linearly polarized incident light. This is caused by the small depolarization of a polarizer [12]. Depolarization is a mechanism that turns polarized light into unpolarized light and is the opposite effect of polarization. The maximum transmitted intensity $I_{\max}$ occurs when the polarizer axis is parallel to the incident polarization direction. The extinction ratio of a polarizer is defined as $I_{\min}/I_{\max}$. Other relations for polarizers can be found in Refs. [9–12].

Besides the polarizer, another basic element in polarization measurements is the phase retarder or wave plate. A phase retarder changes the relative phase between the two orthogonal fields of an EM wave [4–11]. The change of relative phase between the two orthogonal components is called the *phase retardation* or *retardance*. The retardance of a quarter-wave retarder is 90°, and that of a half-wave retarder is 180°. Circularly polarized light can be generated by passing linearly polarized light through a quarter-wave plate whose axis is at 45° with respect to the incident linear polarization direction. A half-wave plate may change the polarization direction of linearly polarized light. In general, a phase retarder changes linearly polarized light into elliptically polarized light. Representations of the optical response of polarizers, retarders, and other materials in terms of the Müeller matrix and Jones matrix are given in the subsection "Polarization by the Response of a Medium" and also in Refs. [4–8,12–24].

Polarization is generated by the anisotropic response of materials and/or anisotropic geometry of systems. The mechanisms for producing polarization include preferential absorption in a dichroic material, reflection and transmission at oblique incidence, double refraction in a birefringent material, diffraction by grating or wires, and scattering by particles [2–11,13]. These properties can be utilized to make polarizers and phase retarders. For example, a wire-grid polarizer is made of parallel fine conducting wires. When light is incident on a wire-grid polarizer with the grid period smaller than the wavelength, the electric field parallel to the wires is shorted and absorbed so that only the electric field perpendicular to the wires passes through the polarizer. In a dichroic polarizer, anisotropic molecules are aligned in a preferential direction so that absorption is very different for the two orthogonal directions referred to the alignment direction. The nonpreferential field is absorbed by the molecules in the medium, while the preferential field passes through the medium [5,8]. In a prism polarizer, the two orthogonal fields are separated by double refraction in a birefringent crystal, and the unwanted polarization is deflected away by the special geometry of a prism. A material is birefringent when it has different refractive indices for different field directions. When a light beam passes through a birefringent slab, phase retardation is generated. Birefringent slabs can be used to make phase retarders or wave plates. Different kinds of polarizers and retarders are described in detail in Refs. [5,9–11].

When special arrangements of polarizers and phase retarders are combined with a light source and a detector, polarized light can be generated and analyzed. Such an optical system is called a *polarimeter* or an *ellipsometer*. The subsection "Principles of Polarimetry" discusses the generation and analysis of polarized light and the operational principles for polarizer–sample–analyzer (PSA) ellipsometry and polarizer–compensator–sample–analyzer (PCSA) ellipsometry using the intensity approach associated with Stokes vectors and Müeller matrices [4,17–23,25,26]. A phase retarder is also called a *compensator* because it was introduced into a polarimeter to compensate the phase change by a sample. The

intensity approach was chosen because intensity, but not electric field, is measured in most experiments and also because the electric field approach cannot treat depolarization. However, the electric field approach is convenient to use for highly polarized light when depolarization does not cause appreciable errors in the measurement. Discussion of ellipsometry using the electric field approach can be found in Refs. [4,15,16,27–33].

Polarization effects are widely applied in modern optical technologies. The electro-optic modulator and shutter are based on tunable birefringence by applying a high voltage across a birefringent crystal to modulate the phase of transmitted light and hence to achieve intensity modulation [4,7,9,34,35]. Liquid crystal displays use similar principles [36]. Birefringence can also be modulated by the photoelastic effect [37,38]. The magneto-optical readout for laser disks utilizes the magneto-optical Kerr effect that generates phase retardation upon reflection from magnetic materials [34,39]. Other applications of polarization are fiber optics, nonlinear optics, material characterization, medical optics, and many other fields. All of these applications utilize the anisotropic nature of materials or the anisotropic geometry of systems. This chapter is concerned with the application of polarization on material characterization. In this application, a polarimeter or ellipsometer is used to measure optical properties and surface properties of materials and thin films [40–48]. In the subsection "Polarization Instrumentation and Experiments," different components of polarimeters are discussed using an example of an automated reflection null ellipsometer (NE), and two sample experiments are described to measure birefringence of a birefringent slab and the optical constants of a material.

49.2 Polarization of an Electromagnetic Wave

The electric field $E(z, t)$ of a monochromatic EM wave propagating along the z-direction with a frequency ω and an angular wave number k can be decomposed into two orthogonal components E_x and E_y and represented by

$$\begin{aligned} E(z,t) &= \hat{x}E_x(z,t) + \hat{y}E_y(z,t) \\ E_x(z,t) &= a_x e^{i(\omega t - kz + \delta_x)} \\ E_y(z,t) &= a_y e^{i(\omega t - kz + \delta_y)} \end{aligned} \tag{49.1}$$

where a_x and δ_x are the amplitude and phase, respectively, for E_x, and a_y and δ_y are for E_y [1–4]. k is related to wavelength λ by $k = 2\pi/\lambda$. In vacuum, $k = \omega/c$. Let $\delta = \delta_y - \delta_x$ be the relative phase between E_y and E_x. Then Equation 49.1 can be simplified to

$$\mathbf{E}(z,t) = (\hat{x}a_x + \hat{y}a_y e^{i\delta}) e^{i(\omega t - kz + \delta_x)} \tag{49.2}$$

49.2.1 Polarization Ellipse

It is often convenient to express **E** in terms of a complex variable. The observed field is actually the real part of **E**. The projection of Re[**E**(z, t)] with $\delta_x = 0$ onto the xy-plane at $z = 0$ is given by

$$\mathbf{E}(0,t) = \hat{x}a_x \cos\omega t + \hat{y}a_y \cos(\omega t + \delta) \tag{49.3}$$

The loci of **E**$(0, t)$ with $a_x = 3$, $a_y = 2$, and different values of δ are shown in Figure 49.1. For $\delta = 0$, the locus of **E** is a line with a slope of a_y/a_x. The EM wave is linearly polarized when E_x and E_y are in phase with each other. In other cases, the loci are ellipses, which are called *polarization ellipses*, and the EM wave is elliptically polarized. The instantaneous electric field can be visualized by drawing an arrow from the origin to a point on an ellipse. The electric field direction rotates in the clockwise direction for $0 < \delta < 180°$, while in the counter clockwise direction for $-180 < \delta < 0°$. When $\delta = \pm 90°$, the axes of the ellipse correspond

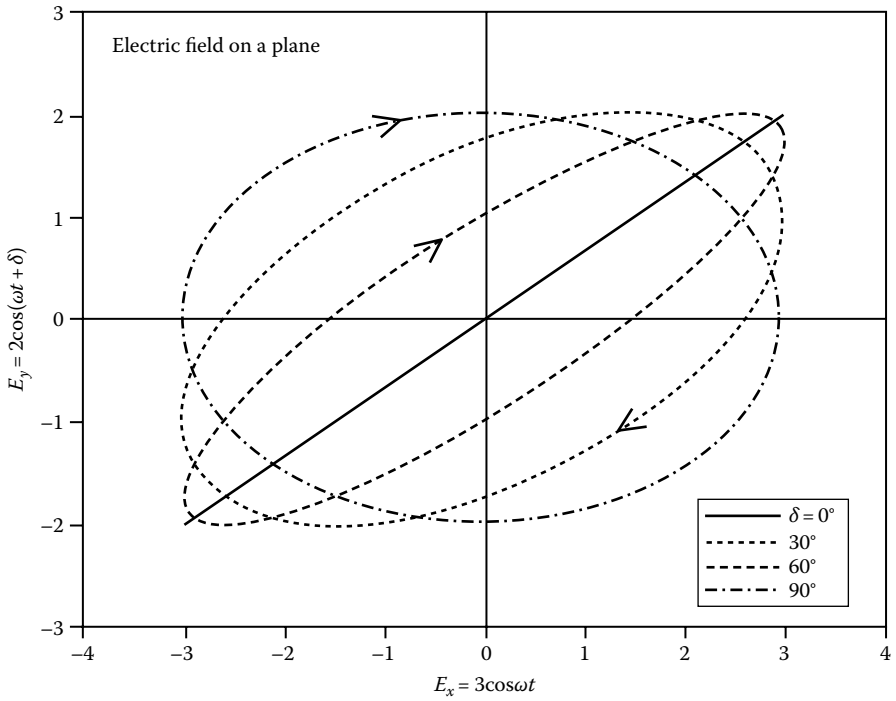


FIGURE 49.1 Projection of the electric field of an EM wave with amplitudes $a_x = 3$, $a_y = 2$ and different values of phase retardation δ onto the $z = 0$ plane. Most of these loci are ellipses and reduce to lines or circles in special cases. The electric field changes in the clockwise direction for δ between 0 and 180°.

to the x and y coordinate axes. If $a_x = a_y$, this ellipse then becomes a circle, and the EM wave is circularly polarized. The convention in ellipsometry defines the right-handed circularly polarized wave as the one whose field rotates in the clockwise direction with $\delta = 90°$ [4,5]. A left-handed circularly polarized wave thus corresponds to $\delta = -90°$ for a counterclockwise-rotating electric field.

In Figure 49.1, a polarization ellipse is specified by a set of three parameters: a_x, a_y, and δ. The ellipse can also be specified by the other set of three parameters: the major axis a, the minor axis b, and the orientation angle ϕ of the major axis measured from the x-axis. Figure 49.2 shows the geometry of an ellipse with these parameters. Parameters a and b can also be expressed in terms of the ellipticity e and ellipticity angle ε, defined by $e = b/a = \tan\varepsilon$. For linear polarization, $\delta = 0$, $b = 0 = e = \varepsilon$, and $\tan\phi = a_y/a_x$. For $\phi = 0$, the major and minor axes of the ellipse always correspond to the coordinate axes. For circular polarization, $a = b = a_x = a_y$ and $\delta = \pm 90°$. Right-handed circularly polarized light has a positive ellipticity with $e = 1$ and $\varepsilon = 45°$, and left-handed circularly polarized light has a negative ellipticity with $e = -1$ and $\varepsilon = -45°$. In general, ϕ represents the orientation of the ellipse and ε indicates the shape of the ellipse and the direction of field rotation. Refs. [2–7] give more details about this subject.

49.2.2 Jones Vector and Stokes Vector

The electric field expressed in vector form in Equation 49.1 can also be expressed as a column matrix. A polarization ellipse depends on a_x, a_y, δ_x, and δ_y, but not on k and ω. By neglecting the common factor of $e^{i(\omega t - kz)}$ in both E_x and E_y, the Jones vector is defined as [4,7,8]

$$\mathbf{E} = \begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} a_x e^{i\delta_x} \\ a_y e^{i\delta_y} \end{pmatrix} \tag{49.4}$$

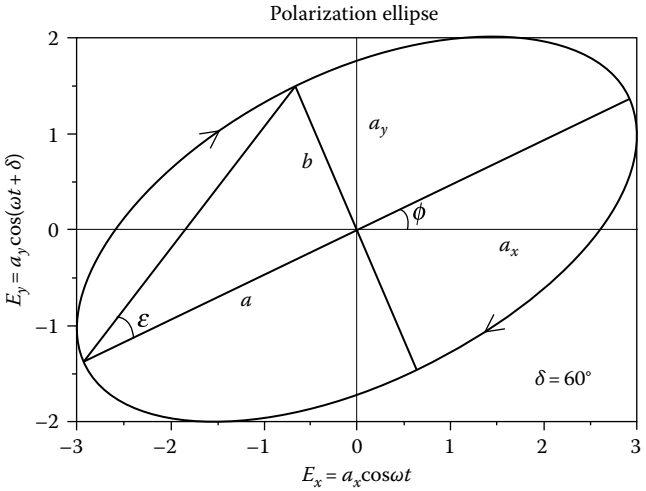


FIGURE 49.2 Characteristic parameters for a polarization ellipse. a_x and a_y are the field amplitudes in the x- and y-directions, and δ is the phase retardance; a and b are the major and minor axes of the ellipse, ϕ is the orientation of the major axis with respect to the x-axis, and ε is the ellipticity angle that is equal to $\tan^{-1}(b/a)$. A polarization ellipse can be characterized by (a_x, a_y, δ), (a, b, ϕ) or (I, ϕ, ε), where I is the intensity of the EM wave.

Both elements of a Jones vector are complex numbers. Jones algebra is convenient for describing perfectly polarized light. Since a light sensor measures only intensity but not electric field in most cases, the Stokes vector is more convenient to use in metrology. The Stokes parameters are four intensity-based parameters used to describe the polarization state of light, represented by S_0, S_1, S_2, S_3 or by I, Q, U, V [1–6,14]. The Stokes vector is the set of these Stokes parameters, defined as [4,14]

$$\mathbf{S} = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} \langle E_x E_x^* \rangle + \langle E_y E_y^* \rangle \\ \langle E_x E_x^* \rangle - \langle E_y E_y^* \rangle \\ \langle E_x E_y^* \rangle + \langle E_y E_x^* \rangle \\ i\langle E_x E_y^* \rangle + i\langle E_y E_x^* \rangle \end{pmatrix} \tag{49.5}$$

For an EM wave, the average bracket in Equation 49.5 represents the time average. $\langle E_x E_x^* \rangle = I_x$ is the intensity of the component of light linearly polarized in the x-direction. Similarly, $\langle E_y E_y^* \rangle = I_y$. All of the Stokes parameters are real numbers and are measurable. For an ensemble of EM waves, the average brackets represent both time and ensemble averages.

49.2.3 Perfectly Polarized Light

Stokes vectors for different polarization states of light can be evaluated using Equation 49.5. Table 49.1 lists the Jones vectors and Stokes vectors for different states of perfectly polarized light. The Jones vector **E** and Stokes vector **S** expressed in terms of ellipticity angle ε and orientation angle ϕ are

$$\mathbf{E} = a \begin{pmatrix} \cos\varepsilon\cos\phi - i\sin\varepsilon\sin\phi \\ \cos\varepsilon\sin\phi + i\sin\varepsilon\cos\phi \end{pmatrix} \tag{49.6}$$

TABLE 49.1 Jones Vectors and Stokes Vectors for Different Polarization States for Perfectly Polarized Light

Polarization	Linear	Linear	Linear	Circular	Elliptical
Direction	0°	90°	±45°	Right/left	
Phase	0°	0°	0°	±90°	δ
Jones vector	$\begin{pmatrix}1\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\end{pmatrix}$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\\pm1\end{pmatrix}$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\\pm\mathrm{i}\end{pmatrix}$	$\begin{pmatrix}a_x\\a_y e^{i\delta}\end{pmatrix}$
Stokes vector	$\begin{pmatrix}1\\1\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\\pm1\\0\end{pmatrix}$	$\begin{pmatrix}1\\0\\0\\\pm1\end{pmatrix}$	$\begin{pmatrix}a_x^2+a_y^2\\a_x^2-a_y^2\\2a_xa_y\cos\delta\\2a_xa_y\sin\delta\end{pmatrix}$

$$\mathbf{S} = I\begin{pmatrix}1\\\cos 2\varepsilon\cos 2\phi\\\cos 2\varepsilon\sin 2\phi\\\sin 2\varepsilon\end{pmatrix} \tag{49.7}$$

In Equations 49.6 and 49.7, the amplitude a and intensity I are totally separated from the angles ϕ and ε that determine the polarization state. The degrees of linear and circular polarization are $\cos 2\varepsilon$ and $\sin 2\varepsilon$, correspondingly. Stokes parameters for perfectly polarized light given in Equation 49.7 satisfy the identity

$$S_0^2 = S_1^2 + S_2^2 + S_3^2 \tag{49.8}$$

49.3 Polarization of Light

Light is composed of an ensemble of EM waves. A single EM wave has a certain electric field direction and phase. Unpolarized light can be visualized as an ensemble of EM waves with random field directions and phases. The field direction and phase for unpolarized light cannot be defined then. The description of light in terms of Jones vector is therefore inadequate to describe the polarization of unpolarized light. For an ensemble of many EM waves, the electric field components in the Stokes vector given by Equation 49.5 is the sum of the corresponding components for all waves. In particular, for an ensemble of incoherent EM waves, the Stokes vectors for individual waves are additive:

$$\mathbf{S} = \sum_{i=1}^{N} \mathbf{S_i} \tag{49.9}$$

For an ensemble of EM waves with identical $\phi_i = \phi$ and $\varepsilon_i = \varepsilon$, the resultant polarization is still the same as the individual wave, as indicated by Equations 49.6 and 49.7, regardless of whether these waves are coherent.

49.3.1 Unpolarized and Partially Polarized Light

If an ensemble consists of randomly oriented linearly polarized waves, all $\varepsilon_i = 0$ and ϕ_i are random, then $S_1 = S_2 = S_3 = 0$, according to Equations 49.7 and 49.9. Light is thus unpolarized, and $S = I(1, 0, 0, 0)$. If an ensemble consists of elliptically polarized waves with the same orientation $\phi_i = \phi$ or $\phi + \pi$ and perfectly

random ellipticity angle ε_i, then the Stokes vector is $I(1, 0, 0, 0)$, and light is also unpolarized. Thus, the Stokes vector **S** in Equation 49.9 already implies the sense of the ensemble average of polarization. The average brackets in Equation 49.5 can be considered as both the time average and ensemble average for incoherent waves.

For unpolarized light with $S_1 = S_2 = S_3 = 0$, Equation 49.8 does not hold. In general, light is partially polarized, that is, part of it is perfectly polarized and the rest is unpolarized [1–6]. Stokes parameters for arbitrary polarizations satisfy

$$S_0^2 \geq S_1^2 + S_2^2 + S_3^2 \tag{49.10}$$

The degree of polarization is given by

$$\mathcal{P} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \tag{49.11}$$

Perfectly polarized light has $\Pi = 1$, and unpolarized light has $\Pi = 0$. Partially polarized light has $0 < \Pi < 1$. The intensity of the polarized part is ΠI, and the intensity of the unpolarized part is $I(1 - \Pi)$. By the superposition concept, the Stokes vector for partially polarized light can be obtained from Equation 49.7 as

$$\mathbf{S} = I \begin{pmatrix} 1 \\ \mathcal{P}\cos 2\varepsilon \cos 2\phi \\ \mathcal{P}\cos 2\varepsilon \sin 2\phi \\ \mathcal{P}\sin 2\varepsilon \end{pmatrix} \tag{49.12}$$

The degree of linear polarization Π_L and circular polarization Π_C are

$$\begin{cases} \mathcal{P}_L = \mathcal{P}\cos 2\varepsilon = \dfrac{\sqrt{S_1^2 + S_2^2}}{S_0} \\ \mathcal{P}_C = \mathcal{P}\sin 2\varepsilon = \dfrac{S_3}{S_0} \end{cases} \tag{49.13}$$

49.3.2 Polarization by the Response of a Medium

49.3.2.1 Jones Matrix

To measure polarization, light must interact with a medium to give a response. The response of a polarizer is to pass one polarization and reject the orthogonal one. The response of a phase retarder is to change the relative phase between the two polarizations. A medium can be any optical component, a test sample, or any object under investigation. The response of a medium relates the state of output light to the state of incident light. The polarization state of light can be described by a complex vector EM field, a Jones vector, or a Stokes vector [4,5,7,8]. Let an incident EM wave be specified by a complex field or Jones vector (E_x, E_y) and the output field be (E_x', E_y'), the general relations between these two fields are

$$\mathbf{E}' = \begin{pmatrix} E_x' \\ E_y' \end{pmatrix} = \begin{pmatrix} r_{xx} & r_{xy} \\ r_{yx} & r_{yy} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} = \mathbf{J}\mathbf{E} \tag{49.14}$$

J in Equation 49.14 is a 2 × 2 matrix that relates the input Jones vector **E** to the output Jones vector **E**′ and is called the Jones matrix. The response of a medium is characterized by the elements of the Jones matrix, r_{xx}, r_{xy}, r_{yx}, and r_{yy}, which are all complex numbers.

49.3.3 Principal Coordinate System

Since the directions of an electric field are different in different rotated coordinate systems, $\{r_{ij}\}$ are not unique. For many symmetric media, there exists a coordinate system in which r_{xy} and r_{yx} are zero and r_{xx} and r_{yy} are called the eigenvalues for $\{r_{ij}\}$. This is the principal coordinate system or principal frame whose x- and y-axes are the two principal axes. Finding the principal frame is an eigenvalue problem. If the incident polarization is along one of the principal axis $\hat{x}$, then the output polarization is still along $\hat{x}$. In the principal frame, $\{r_{ij}\}$ are called the coefficients of response. For example, $\{r_{ij}\}$ may represent the reflection coefficients for a reflection response or the scattering coefficients for a scattering response, etc. In the principal coordinate system, Equation 49.14 is simplified to

$$\mathbf{E}' = \begin{pmatrix} E_x' \\ \\ E_y' \end{pmatrix} = \begin{pmatrix} r_{xx} & 0 \\ 0 & r_{yy} \end{pmatrix} \begin{pmatrix} E_x \\ \\ E_y \end{pmatrix} = \mathbf{J}(0)\mathbf{E} \tag{49.15}$$

J(0) is the diagonalized Jones matrix in the principal frame. For a polarizer, the principal axes are the transmission and extinction axes. The former is assigned to the x-axis. For a phase retarder, the principal axes are the fast and slow axes. The phase change for the EM wave with its field along the fast axis is larger than the slow axis. The fast axis is usually assigned to the x-axis. The principal Jones matrices **P** for a perfect polarizer and **c** for a perfect wave plate with a retardance τ are [4,5,7,8]

$$\mathbf{p} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} e^{i\tau/2} & 0 \\ 0 & e^{-i\tau/2} \end{pmatrix} \tag{49.16}$$

49.3.4 Ellipsometric Parameters

For reflection from a surface, the two principal axes are the s- and p-polarizations. The s-polarization field is along the y-axis, which is chosen to be perpendicular to the plane of incidence. The p-polarization field is along the x-axis, which is in the plane of incidence. The complex reflection coefficients for these two polarizations are designated as $r_p = r_{xx}$ and $r_s = r_{yy}$. The ellipsometric parameters ψ and Δ are defined by [4,15,16]

$$\frac{r_p}{r_s} = \frac{r_{xx}}{r_{yy}} = \tan\psi \exp(i\Delta) \tag{49.17}$$

Δ is the phase change between reflected and incident light. If the electric field direction of incident light is given by ϕ_0, then the field direction ϕ of reflected light can be obtained from $\tan\phi = \tan\phi_0 / \tan\psi$. At the Brewster angle, where $r_p = 0$ and $\psi = 0$, $\phi = 90°$, and reflected light is vertically polarized. Equation 49.17 can also be applied to transmissive systems. Since a vacuum does not change the polarization of light, the ellipsometric parameters are $\psi = 45°$ and $\Delta = 0°$. A perfect polarizer with the polarization along the x-axis has $\psi = 90°$ and $\Delta = 0°$, a perfect quarter-wave plate has $\psi = 45°$ and $\Delta = 90°$, and a perfect half-wave plate has $\psi = 45°$ and $\Delta = 180°$.

49.3.5 Müeller Matrix

The Jones calculus is convenient for perfectly polarized light and a nondepolarizing response [8]. If unpolarized light is incident on a sample, the Jones vector cannot describe the field direction and phase for unpolarized light. The Stokes vector and Müeller matrix are more convenient to use in treating polarization for general cases. A relation between the output Stokes vector $\mathbf{S}'$ and the input Stokes vector $\mathbf{S}$ is

$$\mathbf{S}' = \mathbf{MS} \tag{49.18}$$

The matrix $\mathbf{M}$ that relates the input and output Stokes vectors is called a Müeller matrix. $\mathbf{M}$ is a 4×4 matrix of real numbers.

For the general transformation of electric field given by Equation 49.14, the components M_{ij} of $\mathbf{M}$ can be derived from Equations 49.5, 49.14, and 49.18. The expressions for M_{ij} have been obtained by van de Hulst [13] and are also given as Equation 2.243) in Ref. [4]. In a measurement, the EM waves of output light may come from many different area or volume elements of a medium, so that statistical averages must be considered in the evaluation of M_{ij}. The ensemble average of $\mathbf{M}$ can still be expressed by the same expressions, with the ensemble average bracket applying to all M_{ij}. To make the Müeller matrix meaningful, the new subscripts of Equation 49.14 are reassigned as 1: *xx*, 2: *yy*, 3: *xy*, 4: *yx*. Subscripts 1 and 2 correspond to the copolarized response, and subscripts 3 and 4 correspond to the cross-polarized response. The ensemble average of any two of the response coefficients is called a *correlation function* for these coefficients. Let us define the self-correlation functions to be $2F_j$ and the cross-correlation functions to be $G_{jm} + i\,g_{jm}$ as follows:

$$\begin{cases} 2F_j \equiv \langle r_j r_j^* \rangle \\ G_{jm} + ig_{jm} \equiv \langle r_j r_m^* \rangle \end{cases} \quad j,m = 1, 2, 3, 4;\ j \neq m \tag{49.19}$$

The cross-correlation functions have the properties of $G_{jm} = G_{mj} = \text{Re}\ \langle r_j r_m^* \rangle$ and $g_{jm} = -g_{mj} = \text{Im}\ \langle r_j r_m^* \rangle$. All F_j, G_{jm}, and g_{jm} are real numbers. $\mathbf{M}$ is then

$$\mathbf{M} = \begin{pmatrix} F_1 + F_2 + F_3 + F_4 & F_1 - F_2 - F_3 + F_4 & G_{13} + G_{24} & g_{13} - g_{24} \\ F_1 - F_2 + F_3 - F_4 & F_1 + F_2 - F_3 - F_4 & G_{13} - G_{24} & g_{13} + g_{24} \\ G_{14} + G_{23} & G_{14} - G_{23} & G_{12} + G_{34} & g_{12} - g_{34} \\ -g_{14} + g_{23} & -g_{14} - g_{23} & -g_{12} - g_{34} & G_{12} - G_{34} \end{pmatrix} \tag{49.20}$$

The upper left quadrant of $\mathbf{M}$ corresponds to the self-correlation terms. The lower right quadrant corresponds to the cross-correlations between the two copolarized responses and between the two cross-polarized responses. The upper right and lower left quadrants correspond to the cross-correlations between the copolarized and cross-polarized responses.

49.3.6 Principal Müeller Matrix

For the Jones matrix in the principal frame given by Equation 49.15, the cross-polarized responses are zero, so that the M_{jm} in the upper right and lower left quadrants of Equation 49.20 are zero. $\mathbf{M}$ can be expressed in terms of ψ and Δ using Equations 49.17 and 49.20 as [4,17–19]

$$\mathbf{M} = R \begin{pmatrix} 1 & -\cos 2\psi & 0 & 0 \\ -\cos 2\psi & 1 & 0 & 0 \\ 0 & 0 & \sin 2\psi \cos\delta & \sin 2\psi \sin\Delta \\ 0 & 0 & -\sin 2\psi \sin\Delta & \sin 2\psi \cos\Delta \end{pmatrix} \tag{49.21}$$

where $R = (r_{xx}r_{xx}^{\star} + r_{yy}r_{yy}^{\star})/2$. For reflection, R is the average reflectance, and for transmission, R is the average transmittance. For a vacuum, $\psi = 45°$ and $\Delta = 0°$, **M** is a unit matrix. Using Equation 49.21 or Equations 49.16, 49.19, and 49.20 directly, matrix **P** for a perfect polarizer ($\psi = 90°$, $\Delta = 0°$) and matrix **C** for a perfect wave plate ($\psi = 45°$, $\Delta = \tau$) in the principal frame are obtained as [4,5,20]

$$\mathbf{P} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\tau & \sin\tau \\ 0 & 0 & -\sin\tau & \cos\tau \end{pmatrix} \tag{49.22}$$

49.3.7 Depolarization

A very interesting example is the perfectly random response. Analogous to the conditions used for incoherent scattering [17–19], the random response coefficients δr_j have the properties that

$$\begin{cases} \langle \delta r_j \rangle = 0 \\ \langle \delta r_j \delta r_m^{\star} \rangle = \langle |\, \delta r_j \,|^2 \rangle \delta_{j,m} \quad j,m = 1,2,3,4 \\ \langle |\, \delta r_j \,|^2 \rangle = \text{constant} \end{cases} \tag{49.23}$$

The first line states that all δr_j are each averaged to zero, so that they would not appear in the average of Equation 49.14. By substitution of r_j of Equation 49.19 by δr_j and using the conditions for δr_j given by Equation 49.23, the correlation functions F_j and $G_{jm} + i\,g_{jm}$ can be evaluated. The second line of Equation 49.23 states that all δr_j are uncorrelated with one another, so that all G_{ij} and g_{ij} are zero. The third line states that δr_j are isotropic, so that all F_j are the same. Eventually, all $M_{jm} = 0$ except $M_{00} = 4\,F_1$. The depolarization matrix **D** for a perfectly random response is then

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{49.24}$$

D is an ideal depolarizer as defined in Ref. [20]. The general Müeller matrix of Equation 49.20 satisfies the physical condition [14,24]:

$$\sum_{ij=0}^{3} M_{ij}^2 \le 4M_{00}^2 \tag{49.25}$$

The sum of all the squares of the elements of **D** is M_{00}^2. **D** satisfies the inequality of criterion in Equation 49.25. The matrix of Equation 49.21 is nondepolarizing such that output light is still perfectly polarized if incident light is perfectly polarized. The equality in Equation 49.25 holds for **M** of Equation 49.21. This section discusses optical components and samples that are nondepolarizing. Refs. [12,17–19] give more details about the Müeller matrices for samples that exhibit both polarization and depolarization properties.

49.3.8 Coordinate Transformation

In polarimetric measurements, polarizers and retarders are frequently rotated to desired positions. When a component is rotated, the incident field is not changed, but the representations of this field in the principal and laboratory coordinate systems are different. Transformations of the electric fields, Stokes vectors, Jones and Müeller matrices between these two coordinate systems or frames are basic exercises in polarimetry. Let the laboratory frame axes be x- and y-axes and the principal frame axes be x'- and y'-axes and the principal frame is rotated to an angle α with respect to the laboratory frame, as shown in Figure 49.3. The Jones vector (E'_x, E'_y) in the principal frame is related to (E_x, E_y) in the laboratory frame by

$$\mathbf{E}' = \begin{pmatrix} E'_x \\ E'_y \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} = r(\alpha)\mathbf{E} \tag{49.26}$$

The rotation matrix $\mathbf{r}(\alpha)$ is the 2×2 matrix in Equation 49.26 for transformation of Jones vectors. The inverse transform is given by $\mathbf{E} = \mathbf{r}^T(\alpha)\,\mathbf{E} = \mathbf{r}^T(\alpha)\mathbf{E}'$ where the superscript **T** denotes the transpose of a matrix. In Figure 49.3, $\mathbf{E}'$ appears to be turned by an angle of $-\alpha$, since the coordinate system is rotated by an angle α. The Faraday rotation matrix that rotates **E** by an angle of α is equivalent to $\mathbf{r}(\alpha)$.

One can substitute $r_1 = r_2 = \cos\alpha$ and $r_3 = -r_4 = \sin\alpha$ into Equations 49.19 and 49.20 to construct the rotation matrix $\mathbf{R}(\alpha)$ for transformation of a Stokes vector **S** to a coordinate system oriented at an angle α:

$$\mathbf{R}(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\alpha & \sin 2\alpha & 0 \\ 0 & -\sin 2\alpha & \cos 2\alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{49.27}$$

The transformations between Stokes vector $\mathbf{S}'$ in the principal frame and **S** in the laboratory frame are

$$\begin{cases} \mathbf{S}' = \mathbf{R}\ \mathbf{S} \\ \mathbf{S} = \mathbf{R}^T\ \mathbf{S}' \end{cases} \tag{49.28}$$

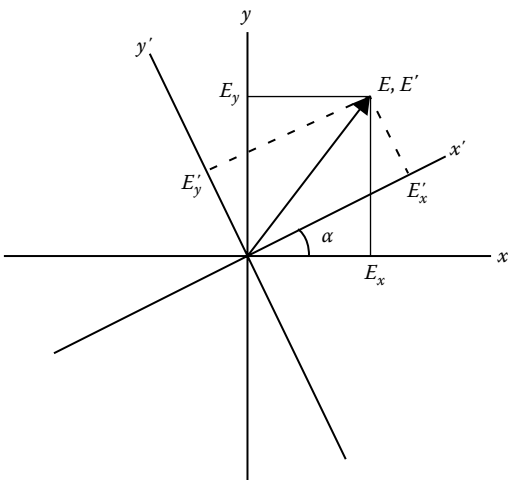


FIGURE 49.3 Coordinate transformation for the electric field components (E_x, E_y) in the laboratory system (x, y) and the components (E'_x, E'_y) in the principal coordinate system (x', y'). The principal frame is oriented at an angle a with respect to the laboratory frame.

The transformation of a Müeller matrix **M**(0) in the principal frame to **M**(α) in the laboratory frame can be obtained by a similarity transformation:

$$\mathbf{M}(\alpha) = \mathbf{R}^T(\alpha)\mathbf{M}(0)\mathbf{R}(\alpha) \tag{49.29}$$

Equation 49.29 can also be used for the transformation of Jones matrix **J**(0) in the principal coordinate frame to **J**(α) in the laboratory frame, provided that **M** is replaced by **J** and **R**(α) by **r**(α) in Equation 49.29.

The Jones matrix for a polarimetric component orientated at an angle α is

$$\mathbf{J}(\alpha) = \begin{pmatrix} r_{xx}\cos^2\alpha + r_{yy}\sin^2\alpha & \sin\alpha\cos\alpha(r_{xx} - r_{yy}) \\ \sin\alpha\cos\alpha(r_{xx} - r_{yy}) & r_{xx}\sin^2\alpha + r_{yy}\cos^2\alpha \end{pmatrix} \tag{49.30}$$

TABLE 49.2 Jones Matrices and Müeller Matrices for Perfect Polarizer and Wave Plates at Different Orientation Angles

Device	Angle	Jones Matrix	Müeller Matrix
Polarizer	0°	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
Polarizer	90°	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
Polarizer	±45°	$\frac{1}{2}\begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & 0 \\ \pm 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$\lambda/4$ plate	0°	$\begin{pmatrix} e^{i\pi/4} & 0 \\ 0 & e^{-i\pi/4} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$
$\lambda/4$ plate	±45°	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \pm i \\ \pm i & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \pm(-1) \\ 0 & 0 & 1 & 0 \\ 0 & \pm 1 & 0 & 0 \end{pmatrix}$
$\lambda/2$ plate	0°	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$
$\lambda/2$ plate	±45°	$\begin{pmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$

The Müeller matrix $\mathbf{P}(P)$ for a perfect polarizer oriented at an angle P and $C(C)$ for a perfect compensator with a retardance τ at an angle C are [4,6,21]

$$\mathbf{P}(P) = \frac{1}{2}\begin{pmatrix} 1 & \cos 2P & \sin 2P & 0 \\ \cos 2P & \cos^2 2P & \sin 2P \cos 2P & 0 \\ \sin 2P & \sin 2P \cos 2P & \sin^2 2P & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{49.31}$$

$$\mathbf{C}(C) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2C + \sin^2 2C \cos\tau & \sin 2C \cos 2C(1-\cos\tau) & -\sin 2C \sin\tau \\ 0 & \sin 2C \cos 2C(1-\cos\tau) & \sin^2 2C + \cos^2 2C \cos\tau & \cos 2C \sin\tau \\ 0 & \sin 2C \sin\tau & -\cos 2C \sin\tau & \cos\tau \end{pmatrix} \tag{49.32}$$

Equations 49.16, 49.22, and 49.30 through 49.32 can be used to calculate the Jones matrices and Müeller matrices for polarizers and wave plates at arbitrary orientations. Table 49.2 lists some of these matrices for the most frequently used devices.

For a light beam passing through successive components oriented at different angles, Müeller matrices or Jones matrices in the laboratory frame must be used for successive multiplications. According to Equation 49.18, the matrix $\mathbf{M}_1(\alpha_1)$ for the component that light first passes through should be placed at the extreme right, and the matrix $\mathbf{M}_n(\alpha_n)$ for the last component at the extreme left. The combined matrix $\mathbf{M}$ is

$$\mathbf{M} = \mathbf{M}_n(\alpha_n)\ldots\mathbf{M}_1(\alpha_1) \tag{49.33}$$

49.4 Principles of Polarimetry

Polarimetry is a method for measuring the polarization of light and the polarization response of materials. An optical system used for such purposes is called a polarimeter or an ellipsometer. To measure the polarization response of a sample, polarized light is generated and incident on the sample. By examining the polarization states of both incident and reflected or transmitted light, the characteristics of a sample can be determined. A schematic diagram of a polarimeter used to measure the polarization response of a sample is shown in Figure 49.4. The light source and polarizer are used to generate polarized light, and the analyzer and detector are used to analyze the polarization of light [4,20]. An analyzer is a polarizer used to analyze polarized light.

49.4.1 Analysis of Polarized Light

Measurement of polarization of light is essential in polarimetry, since polarized light to be examined is not limited to that generated in a laboratory. The instrument to measure the four Stokes parameters is called a photopolarimeter or a Stokesmeter. To measure linear polarization, pass the light beam through a linear analyzer oriented at angle A = 0°, 90°, and ±45° and measure the corresponding intensities I_x,

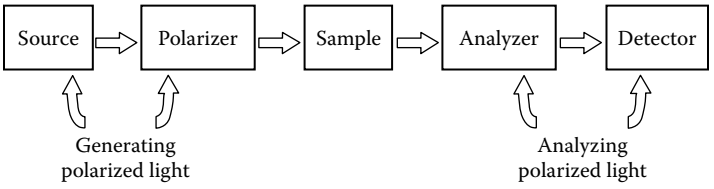


FIGURE 49.4 Schematic diagram of a polarimeter to measure polarization response. The light source and polarizer are used to generate polarized light, and the analyzer and detector are used to analyze the state of polarization of the light.

TABLE 49.3 Intensities I for a Light Beam with Stokes Parameters S_0, S_1, S_2, and S_3 Analyzed by Linear and Circular Analyzers

Analyzer	Linear	Linear	Circular
C (°)	NA	NA	0
A (°)	0, 90	45, −45	45, −45
Operation	0.5(1, ±1, 0, 0)	0.5(1, 0, ±1, 0)	0.5(1, 0, 0, ±1)
Intensity	I_x, I_y	I_+, I_-	I_R, I_L
I =	$(S_0 \pm S_1)/2$	$(S_0 \pm S_2)/2$	$(S_0 \pm S_3)/2$

A circular analyzer consists of a quarter-wave plate oriented at $C = 0°$, followed by an analyzer oriented at $A = \pm 45°$.

I_y, I_+, and I_-. To measure circular polarization, first pass the light beam through a quarter-wave retarder with $C = 0°$ then through an analyzer oriented at $A = \pm 45°$ and measure the intensities I_R and I_L. The pair of quarter-wave retarder and analyzer constitutes a circular analyzer. A detector measuring intensity corresponds to an operation given by a row vector $\mathbf{I} = (1, 0, 0, 0)$. The combined operation of a detector following an analyzer is $\mathbf{IA} = 0.5\,(1, \cos 2A, \sin 2A, 0)$. The operations for the linear and circular analyzers on a Stokes vector S and the intensities obtained are given in Table 49.3. The Stokes parameters can be obtained from the difference and sum of the intensities for these pair operations and are given by

$$\mathbf{S} = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I_x + I_y \\ I_x - I_y \\ I_+ - I_- \\ I_R - I_L \end{pmatrix} \tag{49.34}$$

Equation 49.34 is a general expression that is good for any polarization states and is also the operational principle for most Stokesmeters. The four-detector Stokesmeter designed by Azzam is an exception that contains no moving components and can measure the four Stokes parameters in real time [26].

49.4.2 Generation of Polarized Light

Characterization of polarization response of a sample requires incident polarized light whose polarization state is controllable. A convenient source is a laser, which may be constructed to emit polarized light directly without the help of extra devices. A half-wave plate may be used to rotate the laser polarization to a desired direction by placing the fast axis bisecting the new and old directions, as shown in Figure 49.5a. Light generated from a lamp and a monochromator is usually partially polarized [12]. To generate linearly polarized light at an angle P, a linear polarizer oriented at an angle P is placed behind the monochromator as shown in Figure 49.5b. To generate circularly polarized light, first generate linearly polarized light and then put a quarter-wave plate behind with the fast axis oriented at an angle of 45° or −45° with respect to the linear polarization as shown in Figure 49.5c. Such a combination of polarizer and quarter-wave plate is called a circular polarizer. When a phase retarder has an arbitrary retardance or is placed at an arbitrary angle relative to the polarizer, elliptically polarized light is then generated. Given an incident Stokes vector of (S_0, S_1, S_2, S_3), the Stokes vector $\mathbf{S}'$ for polarized light generated by the polarizers mentioned earlier can be obtained from Equations 49.18 and 49.31 through 49.33. The obtained $\mathbf{S}'$ are listed in Table 49.4. Note that $\mathbf{S}'$ is not directly proportional to S_0 unless incident light is unpolarized. Care must be taken in generating polarized light in an ellipsometer because incident light is rarely completely unpolarized.

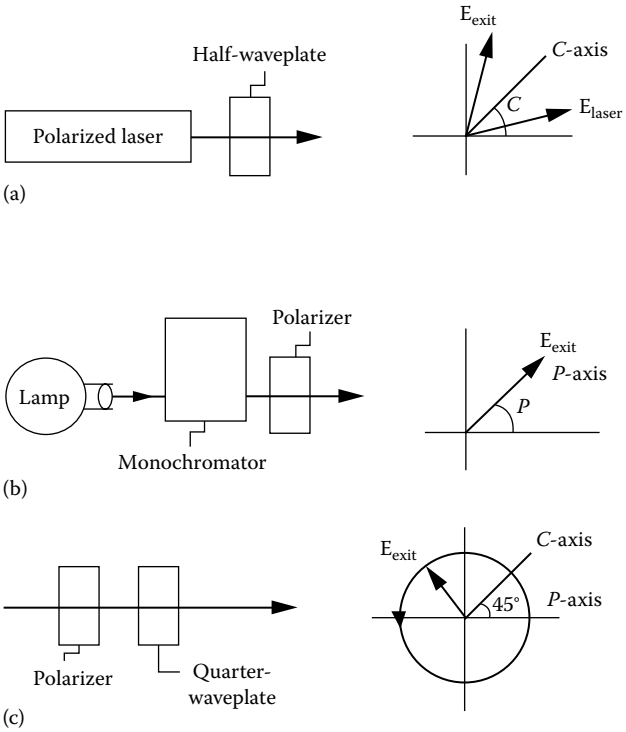


FIGURE 49.5 Generation of light linearly polarized at a desired direction using (a) a laser source and a half-wave plate and (b) a lamp, monochromator, and a polarizer, plus (c) generation of circularly polarized light using a polarizer and a quarter-wave plate.

TABLE 49.4 Stokes Vectors S' for Linearly and Circularly Polarized Light Generated by Specific Combinations of a Polarizer Oriented at an Angle *P* and a Quarter-Wave Retarder at an Angle *C*

Polarizer	Linear	Linear	Circular
P(°)	0, 90	45, −45	−45, 45
C(°)	NA	NA	0
Operation	**P**(P)**S**	**P**(P)**S**	**Q**(*C*) **P**(P) **S**
Stokes vector	$\mathbf{S}'_x, \mathbf{S}'_y$	$\mathbf{S}'_+, \mathbf{S}'_-$	$\mathbf{S}'_R, \mathbf{S}'_L$
S'	$\frac{S_0 \pm S_1}{2}\begin{pmatrix}1\\ \pm 1\\ 0\\ 0\end{pmatrix}$	$\frac{S_0 \pm S_2}{2}\begin{pmatrix}1\\ 0\\ \pm 1\\ 0\end{pmatrix}$	$\frac{S_0 \pm S_3}{2}\begin{pmatrix}1\\ 0\\ 0\\ \pm 1\end{pmatrix}$

The incident stokes vector is $S = (S_0, S_1, S_2, S_3)$.

49.4.3 Polarizer–Sample–Analyzer Ellipsometry

An ellipsometer is an instrument to measure the ellipsometric parameters ψ and Δ of a sample. It can be used for both reflection and transmission. An ellipsometer is usually referred to as the reflection system, and a polarimeter as the transmissive system [4]. Different ellipsometers are designed to measure different responses for different kinds of samples. It is important to know about the sample when

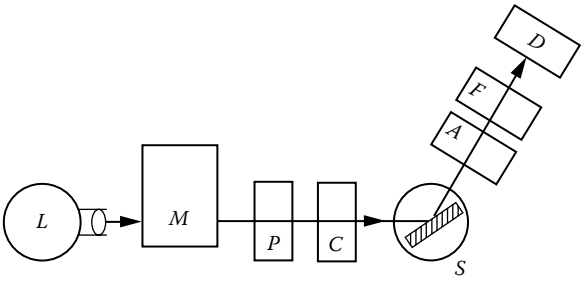


FIGURE 49.6 Schematic diagram of a PCSA ellipsometer.

designing an experiment. The simplest ellipsometer is a PSA ellipsometer. A more general one is a PCSA ellipsometer. Figure 49.6 shows a schematic diagram of a PCSA reflection ellipsometer. A PSA system can be visualized by removing the compensator *C* in Figure 49.6.

Let the oriented angles of the polarizer and analyzer be *P* and *A*, respectively, as measured from the plane of incidence. For unpolarized incident light and a sample matrix **M** given by Equation 49.21, the measured intensity *I* for a PSA ellipsometer is [22]

$$\begin{aligned} I(P,A) &= \mathbf{IA}(A)\,\mathbf{M}(R,\psi,\Delta)\mathbf{P}(P)\mathbf{S} \\ &= T_p T_a R I_0[1-\cos 2\psi(\cos 2P+\cos 2A)]+\cos 2P\cos 2A+\sin 2\psi\cos\Delta\sin 2P\sin 2A \end{aligned} \tag{49.35}$$

where

I_0 is the intensity of incident light

T_p and T_a are the transmittance of the polarizer and analyzer, respectively

If incident light is partially polarized with a Stokes vector (S_0, S_1, S_2, S_3), then the right-hand side of Equation 49.35 should be multiplied by the factor $(S_0 + S_1, \cos 2P + S_2 \sin 2P)/S_0$. In such a case, the dependence on *P* is more complicated. A good practice is to keep *P* fixed and vary only *A*. Many different ways can be devised to extract ψ and Δ from Equation 49.35, such as the Stokes polarimeter, null polarimeter, and rotating-analyzer ellipsometer (RAE).

For a Stokes polarimeter, *P* is set at 45° or −45°, and *A* at 0°, 90°, and ±45°. The ellipsometric parameters ψ and Δ can be solved from the four equations evaluated at these *P* and *A* positions via the relations

$$\begin{cases} \cos 2\psi = \dfrac{I_y - I_x}{I_y + I_x} \\ \sin 2\psi\cos\Delta = \dfrac{I_+ - I_-}{I_+ + I_-} \end{cases} \tag{49.36}$$

For a null polarimeter, set $P = \pm 45°$ and vary *A* to find the null positions. This method is excellent for reflection from transparent materials with Δ equal to 0 or π. The value of ψ is related to the null position $A_\pm$ as [22]

$$\tan 2A_\pm = \mp\tan 2\psi\cos\Delta \quad \text{for } P = \pm 45° \tag{49.37}$$

Average of $A_\pm$ at $P = \pm 45°$ can eliminate errors from the misalignment of analyzer and polarizer.

In a RAE, the analyzer is rotated at an angular frequency ω_r. Set $P = 45°$ or −45°, and $A = \omega_r t$. The measured intensity is

$$I = T_p R T_a I_0[1-\cos 2\psi\cos 2\omega_r t+\sin 2\psi\cos\Delta\sin 2\omega_r t] \tag{49.38}$$

The Fourier coefficients, being equal to $-\cos 2\psi$ and $\sin 2\psi \cos\Delta$, can be recovered from the demodulated signals or from a fast Fourier transform (FFT) technique [28,29]. However, the measured Δ cannot be distinguished from $-\Delta$. This system can be fully automated for real-time operation. With a white light source and monochromator, an RAE can serve as a spectroscopic ellipsometer [15,16,28,29].

49.4.4 Polarizer–Compensator–Sample–Analyzer Ellipsometry

In a PCSA or a PSCA ellipsometer, a compensator of retardance τ is inserted in front of or following the sample. Figure 49.6 shows a schematic diagram of a PCSA ellipsometer. For unpolarized incident light in a PCSA system, the measured intensity for general conditions of P, C, and A is given by [21]

$$\begin{aligned} I(P,C,A) &= \mathbf{I}\,\mathbf{A}(A)\,\mathbf{M}(R,\psi,\Delta)\,\mathbf{C}(C)\,\mathbf{P}(P)\,\mathbf{S} \\ &= T_p T_c R T_a I_0 [Y_0 + Y_1 \cos 2A + Y_2 \sin 2A] \end{aligned} \tag{49.39}$$

where

$$\begin{cases} Y_0 = \{1 - \cos 2\psi [\cos 2C \cos 2(P-C) - \cos\tau \sin 2C \sin 2(P-C)]\} \\ Y_1 = \{-\cos 2\psi + [\cos 2C \cos 2(P-C) - \cos\tau \sin 2(P-C)]\} \\ Y_2 = \sin 2\psi \{\cos\Delta [\sin 2C \cos 2(P-C) + \cos\tau \cos 2C \sin 2(P-C)]\} - \sin\tau \sin\Delta \sin 2(P-C) \end{cases} \tag{49.40}$$

If incident light is partially polarized, then I_0 in Equation 49.39 should be replaced by (S_0 + S_1 cos2P + S_2 sin2P). For a PSCA system, interchange P and A in Equations 49.39 and 49.40. These formulas can be used to design different kinds of PCSA ellipsometers by choosing different conditions and different types of modulation. For certain special conditions, Equations 49.40 can be greatly simplified.

49.4.5 Null Ellipsometry

Consider a PCSA NE in which the compensator is a perfect quarter-wave retarder ($\tau = 90°$) that is set at $C_{\pm} = \pm 45°$. The measured intensity is

$$I(P,C_{\pm},A) = T_p T_c R T_a I_0 [1 - \cos 2\psi \cos 2A \pm \sin 2\psi \sin 2A \cos\Delta \pm (2P - 90°)] \tag{49.41}$$

An NE is an instrument to find the null positions in order to determine ψ and Δ. The four zones that will null the intensity in Equation 49.41 and the null positions are listed in Table 49.5. The null positions of

TABLE 49.5 Null Positions of Polarizer Angle P and Analyzer Angle A for the Four Different Zones at $C = \pm 45°$ in a Null Ellipsometer

Zone	C(°)	P	A
1	−45	−45° + Δ/2	ψ
2	45	−45° − Δ/2	ψ
3	−45	45° + Δ/2	$-\psi$
4	45	45° − Δ/2	$-\psi$

Average $\psi = (A_1 + A_2 - A_3 - A_4)/4$, $\Delta = (P_1 - P_2 + P_3 - P_4)/2$.

A give ψ directly, and the null positions of P give Δ directly. Although ψ and Δ can be determined from measurements in only one zone, systematic errors caused by imperfect components, misalignment, and partially polarized incident light can be nonnegligible. By taking the average of four zones, many of these linear systematic errors can be cancelled [4,21,30,31]. To look for the null positions manually is a slow process; automation of the nulling process can speed up the measurements. Different methods can be used to automate the NE: (1) Both polarizer and analyzer are controlled by servomotors with the feedback from the detector to find the null intensity, (2) the intensity is digitized and fed into a computer that is used to find the null positions, and (3) Faraday rotators are used to rotate and modulate the polarization directions of light incident on and reflected from the sample to get the slopes of intensity versus angle until the slopes are zero at the null positions [4,27]. The advantage of NE is its simplicity in obtaining ψ and Δ. Also, its accuracy is unbeatable by other kinds of ellipsometry.

49.4.6 Phase-Modulated Ellipsometry

A phase-modulated ellipsometer (PME) uses a phase retarder whose retardance is modulative. For a PME, a good choice of P, C, and A in Equations 49.39 and 49.40 is $P = 0°$, $C = 45°$, and $A = 45°$. The intensity is then

$$I(\tau) = T_p T_c R T_a I_0 [1 - \cos 2\psi \cos\tau + \sin 2\psi \sin\Delta \sin\tau] \tag{49.42}$$

The retardance τ is modulated, and the modulated intensity is detected. If τ is modulated according to $\tau = \tau_o \cos\Omega_t$, then $I(\tau)$ can be expanded in a Fourier series with the Fourier coefficients, depending on $\cos 2\psi$, $\sin 2\psi \cos\Delta$, and the Bessel functions of τ_o. These coefficients can be recovered from the demodulated signals, and the values of ψ and Δ can then be solved. τ can be modulated by electro-optic modulation using the Pockels effect or piezoelectric modulation using the photoelastic effect [32,33]. These modulations are fast, and demodulation using a lock-in amplifier is convenient. The advantage of a PME is that it contains no moving components and offers real-time measurement.

49.5 Polarization Instrumentation and Experiments

Figure 49.7 shows a schematic diagram of a PCSA reflection null ellipsometer. The instrument is composed of five systems: the source system, polarimetric system, sample system, detection system, and computer system for automatic control, data acquisition, and processing.

The simplest source is a polarized laser. The monochromatic source system in Figure 49.7 is usually used in visible, ultraviolet, and infrared spectrometers. Light from a lamp source L is focused by a lens system or a spherical mirror system onto the entrance slit of a grating monochromator M. Light leaving the exit slit of M is collimated by another lens system and a set of iris apertures. The longpass filter F in the detector arm is used with a grating monochromator to remove undesired short-wavelength radiation. Choices of monochromators include a grating monochromator, prism monochromator, or Fourier transform spectrometer. A spectrometer is a necessary component for a spectroscopic ellipsometer. Synchrotron radiation is also a continuum source and is used to replace the lamp in the vacuum ultraviolet region [25]. The synchrotron radiation beam is intense and polarized. Grazing incidence reflection optics is usually used to avoid absorption in the components.

The polarimetric system shown in Figure 49.7 is a PCSA ellipsometer. For a PSA ellipsometer, remove the compensator C. For a PCSCA system, add another compensator in the detector arm [20]. It is better to mount the polarizers and phase retarders on automatic rotators so that their orientations can be easily aligned and read. For simple experiments, manual rotators can also do the job. Refs. [10,11] give detailed descriptions and references for different kinds of polarizers and phase retarders. Many of the well-known optical companies sell polarizers and wave plates (retarders) for use in the visible, ultraviolet, and near-infrared spectral regions. *The Buyers Guide of Laser Focus World* [49] and the *Photonics Buyers' Guide* [50] list companies that manufacture and sell polarizers, phase retarders, and polarimeters. Tables 49.6 and 49.7 list some companies that make these products. Commonly

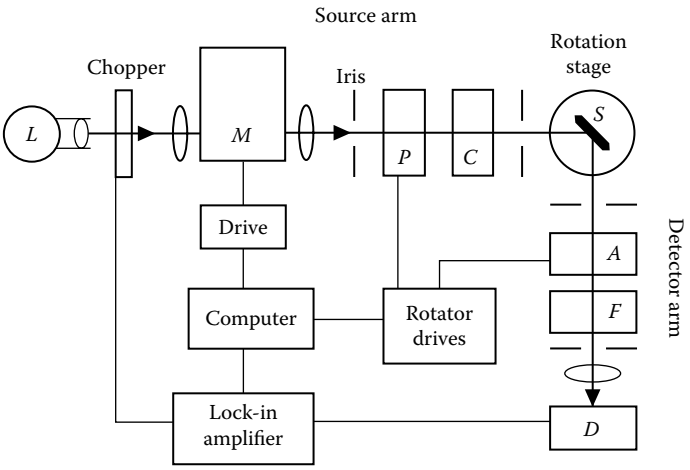


FIGURE 49.7 Schematic diagram of a PCSA reflection null ellipsometer, which is composed of five systems: the source system, polarimetric system, sample system, detection system, and computer system. The symbols are as follows: *L*, light source; *M*, monochromator; *F*, wavelength filter; *D*, detector. The symbols *P*, *C*, *S*, and *A* have their customary meanings.

TABLE 49.6 Companies That Make Polarizers and Phase Retarders

Company	Tel/Fax	Address
Cleveland Crystals	(216) 486-6100	19306 Redwood Ave., Cleveland, OH 44110
	(216) 486-6103	
Corning, Inc.	(607) 974-7966	POLARCOR Team, Advance Materials, HP-CB, Corning, New York 14831
	(607) 974-7210	
CVI Laser	(505) 296-9541	P.O. Box 11308, Albuquerque, NM 87192
	(505) 298-9908	
Hinds Instruments	(503) 690-2000	3175 NW Aloclek Drive, Hillsboro, OR 97124
	(503) 690-3000	
Karl Lambrecht	(312) 472-5442	4204 N. Lincoln Ave., Chicago, IL 60618
	(312) 472-2724	
Meadowlark	(303) 833-4333	P.O. Box 1000, 5964 Iris Parkway, Frederick, CO 80530
	(303) 833-4335	
Molectron, Inc.	(503) 620-9069	7470 SW Bridgeport Rd., Portland, OR 97224
	(503) 620-8964	
New Focus, Inc.	(408) 980-8088	2630 Walsh Ave., Santa Clara, CA 95051-0905
	(408) 980-8883	
Rocky Mountain Instrument	(303) 651-2211	1501 S. Sunset St., Longmont, Co 80501
	(303) 651-2648	
Special Optics	(201) 785-4015	P.O. Box 163, Little Falls, NJ 07424
	(201) 785-0166	
Tower Opt. Corp.	(201) 305-9626	130 Ryerson Ave., Wayne, NJ 07470
	(201) 305-1175	
II-VI, Inc.	(412) 352-1504	375 Saxonburg Blvd., Saxonburg, PA 16056
	(412) 352-4980	

TABLE 49.7 Companies That Make Ellipsometers

Company	Tel/Fax	Address
Gaertner Scientific	847-673-5006 847-673-5009	8228 McCormick Blvd., Skokie, IL 60076
Instrument SA, Inc.	908-494-8660 908-494-8796	6 Olsen Ave., Edison, NJ 08820
J. A. Woollam Co., Inc.	402-477-7501 402-477-8214	650 J. Street, Suite 39, Lincoln, NE 68508
Leonard Research	937-26-1222 937-426-3642	2792 Indian Ripple Rd., Beavercreek, OH 45440 P.O. Box 607, Beavercreek, OH 45434-0607
Rudolph Research	201-691-1300 201-691-5480	1 Rudolph Rd., P.O. Box 1000, Flanders, NJ 07836
SOPRA	(1) 47 81 09 49 (1) 42 42 29 34	26, rue Pierre Joigneaux, 92270 Bois-Colombes, France
Tencor Instruments	415-969-6767 415-969-6731	2400 Charleston Rd., Mountain View, CA 94043-9958

used polarizers in the visible are Glan prisms and dichroic sheets or plates. The best polarizer in the visible is the calcite Glan–Thompson prism, which has a very small extinction ratio and a large acceptance angle. It is more difficult to find a good broadband polarizer in the midinfrared spectral region (λ: 3 to 5 μm). Wire-grid polarizers are good in the long-wave infrared region ($\lambda > 8$ μm). For a broadband phase retarder, a Babinet–Soleil compensator is convenient, since it can be set to any retardance value using a micrometer adjustment. Inexpensive wave plates are good for laser sources. Phase modulation can be achieved by modulation of the birefringence of electro-optic or photoelastic retarders.

The sample system depends on the type of experiment to be performed. Components on the detector side of Figure 49.7 are normally mounted on a rail that is rotatable about the axis of the rotation stage. The sample holder *S* on the rotation stage should have enough degrees of freedom for easy alignment. In a reflection ellipsometer, the *x*-axes for *P*, *C*, and *A* should be well aligned to lie in the plane of incidence. A transmission polarimeter is much simpler to align.

The detection system consists of a detector and a noise suppression system. Diode detectors for use in the visible and near-infrared spectral regions are inexpensive. Photomultipliers for the visible and near-infrared regions have high sensitivity, and cooled semiconductor detectors give good performance in the midinfrared to far-infrared regions. In Figure 49.7, the noise suppression system includes a chopper and a lock-in amplifier. The chopper modulates incident light intensity, and the intensity detected by detector *D* is demodulated by a lock-in amplifier. This combination eliminates most broadband noise and greatly improves the signal-to-noise ratio. RAE and PME have their own modulations and do not need a chopper.

A computer system provides the automatic functions to control the polarizers, retarder, and monochromator and to acquire and process data. A computer system is essential for making accurate and rapid measurements. In Figure 49.7, the computer records the polarizer angle *P*, the analyzer angle *A*, the intensity *I* from the detector, and the wavelength λ of the monochromator. The data of $I(P)$ and $I(A)$ can be used to find the null positions of *P* and *A*. Then the computer controls the drivers to move *P* or *A* to the null positions. In a spectroscopic RAE, data of $I(t)$ are recorded as the analyzer is rotating. The computer uses the FFT program to find the Fourier coefficients, solves for ψ and Δ, records the results, and then drives the monochromator to a new wavelength and repeats the process. A spectroscopic PME uses similar computer process as RAE, besides the different modulation and demodulation.

49.5.1 Measurement of Birefringence

Retardance can be measured using a transmission PSA ellipsometer. A wave plate is a good sample for this experiment. The retardance δ for a birefringent slab of thickness d and principal refractive indices n_e and n_o is

$$\delta = \frac{2\pi d(n_e - n_o)}{\lambda} \tag{49.43}$$

in the absence of multiple reflections. For a wave plate, the value of ψ is close to 45°, and the Δ value is equal to δ. Use a lamp source with a monochromator to scan the wavelength. Put a polarizer at 45° and an analyzer at −45° with respect to the fast axis of the wave plate [40]. When λ is scanned, the transmitted intensity I through the PSA ellipsometer will vary, in proportion to $(1 + \cos\delta)$ according to Equation 49.35. I is a maximum when δ is 0° and is a minimum when $\delta = 180°$. From the wavelengths at which a maximum or a minimum intensity occurs, the birefringence $n_e - n_o$ can be determined.

49.5.2 Measurement of Optical Constants

A reflection ellipsometer can be used to measure optical constants n and k of materials. Light is incident obliquely on a sample at an angle of θ. The refractive index n and the extinction coefficient k can be calculated from the measured ψ and Δ values using the following formula [2,4,16,22]:

$$(n + ik)^2 = \sin^2\theta\left\{1 + \tan^2\theta\left[\frac{1 - \tan\psi\exp(i\Delta)}{1 + \tan\psi\exp(i\Delta)}\right]^2\right\} \tag{49.44}$$

If an automated system is not available, try an NE to obtain ψ and Δ manually. Automated systems such as an RAE or a PME can take data much faster. For transparent materials, Δ is either 0 or π; the simple null polarimeter with a PSA system offers satisfactory results [22]. The method is effective near the Brewster angle region. At the Brewster angle θ_B, $\psi = 0$, $n = \tan\theta_B$ according to Equation 49.44. For metals whose ψ is large when $\Delta = -90°$, the principal angle ellipsometry (PAE) can be used [2]. At the principal angle θ_p, $\cos\Delta = 0$, Equation 49.44 can be simplified to

$$(n + \mathrm{i}k)^2 = \sin^2\theta_P[1 + \tan^2\theta_P(\cos 4\psi + i\sin 4\psi)] \tag{49.45}$$

In PAE, the principal angle θ_P is searched and then ψ is measured. The Stokes polarimeter with a PSA system is suitable to search θ_P and measure ψ for PAE.

Determination of optical constants using reflection ellipsometry is subject to errors caused by surface roughness, natural oxides, and surface contamination [15,16,41–48]. These effects can be corrected by assuming that there is an effective layer on the surface and then using a least-square regression method to fit the ellipsometric data to the appropriate model [16,43–48]. Ellipsometry is also used to measure refractive index and thickness of a thin film on a bulk substrate whose optical constants are known [4,15,16,27]. Other applications of polarimetry and ellipsometry can be found in recent proceedings about polarization [51–53].

Acknowledgments

The author would like to thank Dr. J. M. Bennett for her review of the original manuscript and her suggestions especially for the introduction. This work was supported partially by the "Polarizer Standard and Metrology Program" funded by the US Naval Warfare Assessment Division.

References

1. G. G. Stokes, On the composition and resolution of streams of polarized light from different sources, *Trans. Cambridge Phil. Soc*., 9, 399–416, 1852, also as paper 10 of ref. [8].
2. M. Born and E. Wolf, *Principles of Optics*, 5th edn., Oxford, U.K.: Pergamon Press, 1975, Chapters 1, 2, 13, 14.
3. J. D. Jackson, *Classical Electrodynamics*, 2nd edn., New York: John Wiley & Sons, 1975, Sections 7.1–7.5.
4. R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light*, 2nd edn., Amsterdam, the Netherlands: North Holland, 1987.
5. W. A. Shurcliff and S. S. Ballard, *Polarized Light*, Princeton, NJ: van Nostrand, 1964.
6. K. Serkowski, Polarization of starlight, in Z. Kopal (ed.), *Advances in Astronomy and Astrophysics*, Vol. 1, New York: Academic Press, 1962, pp. 289–352.
7. G. R. Fowles, *Introduction to Modern Optics*, 2nd edn., New York: Holt, Reinhart and Winston, 1975, Chapter 6.
8. W. Swindell (ed.), *Polarized Light*, *Benchmark Papers in Optics*, Vol. 1, Stroudsburg, PA: Dowden, Hutchinson & Ross, 1975.
9. F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th edn., New York: McGraw-Hill, 1976, Chapters 24–28.
10. J. M. Bennett, Polarization, in M. Bass, E. S. Van Stryland, D. R. Williams, and W. L. Wolfe (eds.), *Handbook of Optics*, Vol. I, Chapter 5, pp. 5.1–5.30, New York: McGraw-Hill, 1995.
11. H. E. Bennett and J. M. Bennett, Polarization, in W G. Driscoll and W. Vaughan (eds.), *Handbook of Optics*, Chapter 10, pp 10.1-10.14, New York: McGraw-Hill, 1978.
12. S. F. Nee, C. Yoo, T. Cole, and D. Burge, Characterization of imperfect polarizers under imperfect conditions, *Appl. Opt*., 37, 54–64, 1998.
13. H. C. van de Hulst, *Light Scattering by Small Particles*, New York: Wiley, 1957, Chapter 5.
14. K. Kim, L. Mandel, and E. Wolf, Relationship between Jones and Müeller matrices for random media, *J. Opt. Soc. Am. A*., 4, 433–437, 1987.
15. D. E. Aspnes, Spectroscopic ellipsometry of solids, in B. O. Seraphin (ed.), *Optical Properties of Solids: New Developments*, Amsterdam, the Netherlands: North Holland, Chapter 15, pp.801–846, 1976.
16. D. E. Aspnes, The accurate determination of optical properties by ellipsometry, in E. D. Palik (ed.), *Handbook of Optical Constants of Solids I*, Orlando, FL: Academic Press, pp. 89–112 1985.
17. S. F. Nee, Polarization of specular reflection and near-specular scattering by a rough surface, *Appl. Opt*., 35, 3570–3582, 1996.
18. S. F. Nee, The effects of incoherent scattering on ellipsometry, in D. H. Goldstein and R. A. Chipman (eds.), *Polarization Analysis and Measurement*, Proceedings of SPIE, San Diego, CA; *Proc. SPIE*, 1746, 119–127, 1992.
19. S. F. Nee, Effects of near-specular scattering on polarimetry, in D. H. Goldstein and D. B. Chenault (eds.), *Polarization Analysis and Measurement II*, Proceedings of SPIE, San Diego, CA; *Proc. SPIE*, 2265, 304–313, 1994.
20. R. A. Chipman, Polarimetry, in M. Bass, E. S. Van Stryland, D. R. Williams, and W. L. Wolfe (eds.), *Handbook of Optics*, Vol. II, New York: McGraw-Hill, 1995, Chapter 22.
21. S. F. Nee, Error reductions for a serious compensator imperfection for null ellipsometry, *J. Opt. Soc. Am. A*, 8, 314–321, 1991.
22. S. F. Nee and H. E. Bennett, Accurate null polarimetry for measuring the refractive index of transparent materials, *J. Opt. Soc. Am. A*, 10, 2076–2083, 1993.
23. R. M. A. Azzam and N. M. Bashara, Ellipsometry with imperfect components including incoherent effects, *J. Opt. Soc. Am*., 61, 1380–1391, 1971.

24. E. S. Fry and G. W. Kattawar, Relationships between elements of the Stokes matrix, *Appl. Opt.*, 20, 2811–2814, 1981.
25. J. Barth, R. L. Johnson, and M. Cardona, Spectroscopic ellipsometry in the 6–35 eV region, in E. D. Palik (ed.), *Handbook of Optical Constants of Solids II*, Boston, MA: Academic Press, Chapter 10, pp. 213–245, 1991.
26. R. M. A. Azzam, Arrangement of four photodetectors for measuring the state of polarization of light, *Opt. Lett.*, 10, 309–311, 1985; U.S. Patent 4,681,450 1987.
27. R. M. A. Azzam, Ellipsometry, in M. Bass, E. S. Van Stryland, D. R. Williams, and W. L. Wolfe (eds.), *Handbook of Optics*, Vol. II, New York: McGraw-Hill, 1995, Chapter 27.
28. D. E. Aspnes, Fourier transform detection system for rotating-analyzer ellipsometers, *Opt. Commun.*, 8, 222–225, 1973.
29. D. E. Aspnes, High precision scanning ellipsometer, *Appl. Opt.*, 14, 220–228, 1975.
30. R. M. A. Azzam and N. M. Bashara, Unified analysis of ellipsometry errors due to imperfect components, cell-window birefringence, and incorrect azimuth angles, *J. Opt Soc. Am.*, 61, 600–607, 1971.
31. D. E. Aspnes, Measurement and correction of first-order errors in ellipsometry, *J. Opt Soc. Am.*, 61, 1077–1085, 1971.
32. S. N. Jasperson and S. E. Schnatterly, An improved method for high reflectivity ellipsometry based on a new polarization modulation technique, *Rev. Sci. Instrum.*, 40, 761–767, 1969.
33. J. C. Kemp, Piezo-optical birefringence modulators: New use for a long-known effect, *J. Opt. Soc. Am.*, 59, 950–954, 1969.
34. J. Wilson and J. F. B. Hawkes, *Optoelectronics: An Introduction*, Englewood Cliffs, NJ: Prentice-Hall, 1983, Chapters 1, 3.
35. T. A. Maldonado, Electro-optic modulators, in M. Bass, E. S. Van Stryland, D. R. Williams, and W. L. Wolfe (eds.), *Handbook of Optics*, Vol. II, New York: McGraw-Hill, 1995, Chapter 13, pp. 13.1–13.35.
36. S. T. Wu, Liquid crystals, in M. Bass, E. S. Van Stryland, D. R. Williams, and W. L. Wolfe (eds.), *Handbook of Optics*, Vol. II, New York: McGraw-Hill, 1995, Chapter 14.
37. G. L. Cloud, *Optical Methods of Engineering Analysis*, Cambridge, U.K.: Cambridge University Press, 1995, Parts I, II.
38. I. C. Chang, Acousto-optic devices and applications, in M. Bass, E. S. Van Stryland, D. R. Williams, and W. L. Wolfe (eds.), *Handbook of Optics*, Vol. II, New York: McGraw-Hill, Chapter 12, pp. 12.1–12.54, 1995.
39. M. Mansuripur, *The Physical Principles of Magneto-Optical Recording*, Cambridge, U.K.: Cambridge University Press, 1995, Chapter 6.
40. J. H. Shields and J. W. Ellis, Dispersion of birefringence of quartz in the near infrared, *J. Opt. Soc. Am.*, 46, 263–265, 1956.
41. D. K. Burge and H. E. Bennett, Effect of a thin surface film on the ellipsometric determination of optical constants, *J. Opt. Soc. Am.*, 54, 1428–1433, 1964.
42. C. F. Fenstermaker and F. L. McCrackin, Errors arising from surface roughness in ellipsometric measurement of the refractive index of a surface, *Surf. Sci.*, 16, 85–96, 1969.
43. D. E. Aspnes, Studies of surface, thin film and interface properties by automatic spectroscopic ellipsometry, *J. Vac. Sci. Technol.*, 18, 289–295, 1981.
44. M. E. Pedinoff and O. M. Stafsudd, Multiple angle ellipsometric analysis of surface layers and surface layer contaminants, *Appl. Opt.*, 21, 518–521, 1982.
45. J. P. Marton and E. C. Chang, Surface roughness interpretation of ellipsometer measurements using the Maxwell-Garnett theory, *J. Appl. Phys.*, 45, 5008–5014, 1974.
46. D. E. Aspnes, J. B. Theeten, and F. Hottier, Investigation of effective medium models of microscopic surface roughness by spectroscopic ellipsometry, *Phys. Rev. B*, 20, 3292–3302, 1979.
47. D. E. Aspnes, Optical properties of thin films, *Thin Solid Films*, 89, 249–262, 1982.

48. S. F. Nee, Ellipsometric analysis for surface roughness and texture, *Appl. Opt.*, 27, 2819–2831, 1988.
49. Anonymous, *Laser Focus World Buyers Guide*, *1996*, PennWell Publishing Co., Tulsa, OK.
50. Anonymous, *The Photonics Buyers' Guide*, *1996*, *Book II*, Laurin Publishing Co., Pittsfield, MA.
51. R. A. Chipman (ed.), *Polarization Considerations for Optical Systems II*, Vol. 1166, *Proceedings of SPIE*, San Diego, CA, 1989.
52. D. H. Goldstein and R. A. Chipman (eds.), *Polarization Analysis and Measurement*, Vol. 1746, Proceedings of SPIE, San Diego, CA, 1992.
53. D. H. Goldstein and D. B. Chenault (eds.), *Polarization Analysis and Measurement II*, Vol. 2265, Proceedings of SPIE, San Diego, CA, 1994.

50

Refractive Index Measurement

Gerry H. Meeten
Schlumberger
Gould Research

50.1 Introduction

Light is electromagnetic radiation of wavelength about 450–700 nm, in which region the human eye is sensitive and refractive index measurements are commonly made. The refractive index or indices of a substance describe an important part of its interaction with electromagnetic radiation. Refractive index is a dimensionless quantity, real for transparent materials and complex if there is absorption. It generally depends on the direction of light relative to axes of the material; such substances, for example, many crystals, are anisotropic and possess more than one refractive index. In general, the index is a tensor (3×3 matrix) with up to nine components. Many substances are isotropic, for example, liquids, glasses, and other noncrystalline materials, and one refractive index is sufficient. Some substances are optically inhomogeneous and possess refractive index fluctuations over distances comparable in size with a wavelength. They scatter light and appear milky or turbid. Many foodstuffs and drinks are optically heterogeneous. Refractometry is, fortunately, a robust technique in the face of material complexities. It has widespread application to many industries and materials. This chapter treats the refractive index as the principal measurand; the complications of real materials may require extra attention. Excluded ones are those strongly absorbing materials such as metals in which absorption are more important than refraction. The appropriate measurement technique is then ellipsometry, in which the amplitude and phase of the reflected light are measured, as the polarization and the angle of incidence is varied. Some refractometric methods are essentially ellipsometric, but for transparent samples transmission methods are also important. Refractive index measurements of high precision, for example, 10^{-4}–10^{-5}, are possible with relatively simple techniques. Although ocular instruments remain popular, semiconductor technology has been applied to refractive index measurement in recent years, improving the measurement speed and precision. One of the limits to improvement in precision is the sample's absorption or heterogeneity, a factor considered herein. Several online instruments exist, and the application of refractive index measurements to complex industrial fluids has been recently described [1]. Fiber-optic methods do not yet possess the precision of conventional methods but offer the possibility of remote sensing.

Refractive index is measured for many reasons. It is clearly important to know the refractive index of materials used for their clarity, such as glasses and solid plastics. In complex fluids such as drinks or foods, the refractive index is a measure of dissolved or submicronic material. The Brix scale relates refractive index to sugar concentration. Common industrial applications are to microemulsions to measure their oil/water ratio, to antifreeze to check the glycol/water ratio, and to inaccessible liquids such as the electrolyte of rechargeable cells [2]. The clinical applications of light have stimulated interest in biotissue refractometry [3], and refractometry is useful for the analysis of small samples of biofluids.

50.2 Physical Principles

The electromagnetic nature of light and the atomic origin of refraction are described herein only in enough detail to be useful to refractometry; detailed descriptions of them are readily found in optics textbooks [4].

Monochromatic light can be described as an electromagnetic radiation of frequency ν (Hz) and free-space wavelength λ_0. In free space (vacuum), the phase velocity c_0 is given by $c_0 = \nu\lambda_0$. In a medium or material substance, the phase velocity changes to $c = \nu\lambda$, the refractive index n of that substance then being defined as $n = c_0/c$. As the frequency is unchanged by entering the medium, $n = \lambda_0/\lambda$. In an electromagnetic wave, the vibration directions of the electric and the magnetic fields are 90° apart, and both fields vibrate at 90° to the wave's direction of propagation. In transparent and nonmagnetic materials, refractive index is determined primarily by the interaction of the electric field of the wave with the permittivity e of the medium. This quantity is written as $\varepsilon = \varepsilon_r\varepsilon_0$, where ε_r is the relative permittivity of the medium and ε_0 is the permittivity of free space being $10^7/(4\pi c_0^2)$F m^{-1}, or about 8.854 pF m^{-1}. In the wave theory of light, ε_r is determined by the strength of elastic interaction between the wave's electric field and the bound charges (electrons and protons) in the atoms and molecules of the medium. For nonmagnetic transparent materials, $n = \sqrt{\varepsilon_r}$ can be shown if both quantities are measured at the same frequency, and this relation is also closely true if the absorption is weak. Absorption can be ascribed to the inelasticity of the interaction of the wave's electric field with the bound charges, some of the wave's energy being turned to heat during propagation.

The time (t) and distance dependence of the electric field £ of a wave propagating in the z direction can be written as

$$E = E_0 \exp i(\omega t - k_0 z) \tag{50.1}$$

where

$\omega = 2\pi\nu$ is the pulsatance or angular frequency
$k_0 = 2\pi/\lambda_0$ is the free-space wave number
$i = \sqrt{-1}$

The field E in Equation 50.1 is complex, and the physical field is understood to be the real part of the complex one, that is, $E = E_0\cos(\omega t - k_0 z)$. The wave amplitude E_0 is the maximum excursion of E, both quantities being coparallel and normal to the direction of propagation z. The optical power per unit area (radiance) conveyed by the wave is proportional to $|E|^2$, or also $|E_0|^2$. The phase velocity c is given by ω/k, which the aforementioned relations show to be c_0/n. Waves in absorbing media are attenuated as z increases. This is described by a complex refractive index in Equation 50.1, that is, writing $n = n' - in''$ so that the field oscillates in space and time as before but decays exponentially with propagation distance z as $(-k_0 n'' z)$. The cause of the index n'' may be absorption, scattering, or both. It will be clear from context whether n refers to the real part of the complex quantity or to the complex quantity itself. For the most part, this chapter deals with transparent or only slightly absorbing materials, where n'' is very small.

Equation 50.1 shows that n determines the phase of a wave, and this is the basis of the interferometric methods of refractometry, most used for either highly precise measurements on liquids or solids or for gases where the refractive index is close to 1. Other common refractometric methods rely on the reflection or transmission of light at the interface between the sample and a material of known refractive index. Reflection and refraction of light at optically smooth interfaces are treated (Fresnel relations) in optics textbooks. A description follows that it should be sufficient for the experimental refractometrist.

Simple laws apply to the incident, reflected, and transmitted beams at a planar interface between two transparent media. For brevity, we consider only isotropic media. A parallel–collimated incident light beam produces a parallel–collimated reflected and transmitted beam. All beams are coplanar and define a plane of incidence, which is orthogonal to the planar interface. The interface normal also lies in the plane of incidence. Figure 50.1 shows the incident, reflected, and transmitted beams in the plane of incidence. In the first medium, the angle of incidence and reflection made to the interface normal are equal, both θ_1. The angle of transmission θ_2 into the second medium is given by Snell's law:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \tag{50.2}$$

where the subscripts refer to the first and second medium. If n_1 is known and the angles are measured, n_2 can be obtained. Equation 50.2 is the basis of deviation refractometry, $q \cdot v$.

If $n_1 > n_2$, Equation 50.2 shows that real angles of transmission do not exist for $\theta_1 > \arcsin(n_2/n_1)$; that is, θ_2 is imaginary. The incident and reflected beams are then of equal power, transmitted light is absent, and the reflection is said to be total. However, although in the second medium no wave propagates, there is an evanescent wave close to the interface, which is nonpropagating and therefore transmits no power. The critical angle is given by

$$\theta_c = \text{arc sin}\left(\frac{n_1}{n_2}\right) \tag{50.3}$$

(See Figure 50.2.) Reversing the light direction, Figure 50.3 shows that a beam traveling parallel to the interface in the least refractive index medium enters the second medium at a transmission angle also given by θ_c. A full analysis (Fresnel's equations) shows that in the region of the critical angle, the reflected or transmitted optical power changes very rapidly with angle of incidence. If there is no absorption, there is actually a discontinuity of reflected or transmitted radiance at θ_c. This makes θ_c easy to measure accurately and gives the critical angle or Abbe method of refractometry used by most commercial refractometers.

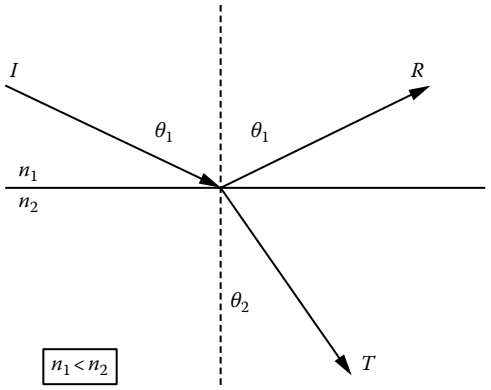


FIGURE 50.1 Light beams reflected (R) and transmitted or refracted (T) at the interface between two refractive indices, n_1 and n_2, the second being greater than the first.

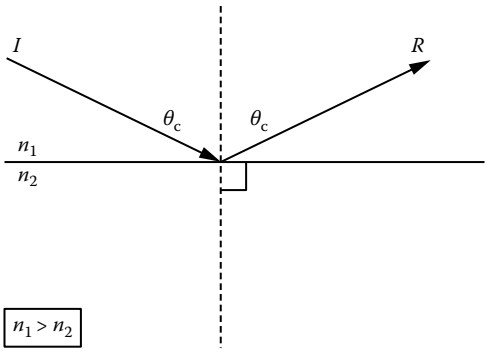


FIGURE 50.2 Total reflection at the critical angle of the incident beam (I) in the first medium of greater refractive index than the second. No light is transmitted into the second medium for angles exceeding the critical angle.

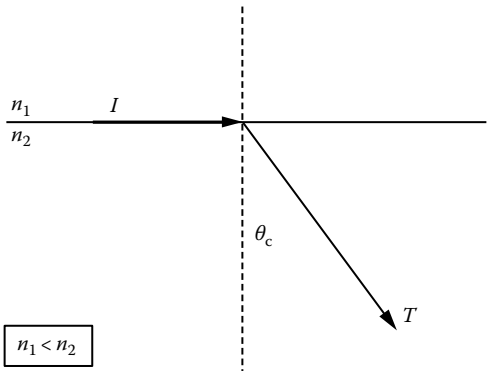


FIGURE 50.3 Total refraction of incident beam (I) transmitted as beam T into the medium of lesser refractive index at the critical angle. No light exists for angles in the second medium bigger than the critical angle.

The Fresnel laws of reflection at a plane interface between two media show that if the incident light is linearly polarized with its electric field in the plane of incidence, then there is no reflected light at the Brewster angle θ_B given by

$$\theta_B = \arctan\left(\frac{n_2}{n_1}\right) \tag{50.4}$$

(see Figure 50.4.) Measurements of θ_B usually in air (n_1 close to 1) will give n_2. The reflected optical power is zero at θ_B and it increases approximately quadratically for angles below and above θ_B. Consequently, the Brewster angle method is regarded to be of lower precision than the critical angle method and so is relatively less used. However, unlike the critical angle method, it does not require a reference medium of higher refractive index than the sample.

If the sample is a finely divided solid or has an irregular interface, none of the preceding methods is useful. Fresnel's equations show the reflectance of any shape of interface to be zero if the refractive index difference across the interface is also zero. In the method of refractive index match, the turbidity or scattering from a cloud of particles suspended in a liquid is measured, and the refractive index of the liquid is found, which gives minimum turbidity or scattering. The liquid should not swell or interact with the interior of the solid.

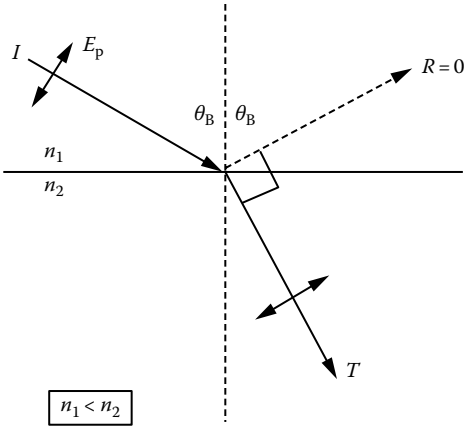


FIGURE 50.4 Light incident in the first medium at the Brewster angle, and polarized with its electric vector parallel to the plane of incidence, is transmitted (*T*) wholly into the second medium with zero reflection *R*.

50.3 Techniques

50.3.1 Interferometry

The sample, with parallel input and output faces, is put in the sample beam of a two-beam interferometer (e.g., Rayleigh, Michelson, or Jamin). From Equation 50.1, a sample of length z and refractive index n causes a phase lag of $2\pi nz/\lambda_0$, compared with $2\pi n_{air}z/\lambda_0$ for the reference beam in air. The phase difference between the two beams is thus

$$\delta = \frac{2\pi z(n - n_{air})}{\lambda_0} \tag{50.5}$$

The refractive index of air is 1 for approximate purposes or is known more precisely [5]. Sample insertion causes a step increment of δ in the phase difference, which is measured by adjusting a calibrated phase compensator in the reference beam to regain the zero-order fringe of white-light fringes. Equation 50.5 enables n to be calculated. For gaseous samples, use of the compensator may be avoided by allowing the gas to enter an initially evacuated tube with rigid and parallel window faces placed in the sample beam. As the gas enters slowly, the interference fringes passing a fiducial mark may be counted. Each whole fringe corresponds to a change in δ of π radian. The refractive index of a gas depends, through its density, on pressure and temperature. The number of fringes between a vacuum ($n = 1$) and the gas at a given pressure and temperature will then give the gas refractive index.

In interferometry, the effect of optical attenuation (via absorption or scattering) by a sample of complex refractive index decreases the fringe brightness but does not affect the phase. Thus, only the real part of the complex refractive index is measured.

About 0.1 of a fringe can be judged by eye; electro-optic methods can measure 10^{-6} of a fringe. The method's potential precision is very high. However, an interferometer is delicate and expensive, and good sample quality (homogeneity, face parallelism) is necessary. A difference in dispersion (wavelength dependence) of the sample and compensator refractive index can be a problem. For a sufficiently large number of fringes, the dispersion difference between the sample and compensator can make the white-light fringe pattern and zero-order fringe invisible [6]. One solution is to use a compensator more closely matched to the sample dispersion. A cheaper option is to reduce the sample length, with some loss of precision. A major use of the interference method is for gases, where $n - 1$ is about 1000 times less than $n - 1$ of solids and liquids. It is also used in solution differential refractometry, for example, to measure

the concentration in the eluted solvent relative to the pure solvent. A stable folded Jamin interferometer for refractive index measurements is described by Moosmüller and Arnott [7]. Liquid refractive index measurements using a Michelson interferometer are described by Richerzhagen [8].

50.3.2 Deviation Methods

Lateral and angular deviation methods make use of Snell's law (Equation 50.2). The incident beam is usually in air, of refractive index n_{air}.

Lateral deviation l occurs for a beam of light transmitted through a parallel-sided sample of thickness L (see Figure 50.5). If the incident angle is θ_{air}, Snell's law (Equation 50.2) gives

$$\frac{n}{n_{air}} = \left[1 + \left(\frac{\cos\theta_{air}}{\sin\theta_{air} - (l/L)}\right)^2\right]^{1/2} \sin\theta_{air} \tag{50.6}$$

If θ_{air} and L can be measured the most accurately, the precision of the method depends on the accuracy of measuring l. Modern electro-optic beam-displacement detection devices permit a precision of a few micrometers in l, giving a typical precision in n of about 0.001. The lateral deviation method is suitable for samples in sufficiently thick sheet form.

In the angular deviation method, parallel–collimated light is incident on one face of the sample of refractive index n in triangular prism form with vertex angle A, the vertex line of the prism being normal to the plane of incidence (see Figure 50.6). The exit beam deviates angularly from the direction of the incident beam. At the angle of minimum deviation D, the analysis is simplest, Equation 50.1, showing that

$$\frac{n}{n_{air}} = \frac{\sin((A+D)/2)}{\sin(A/2)} \tag{50.7}$$

For $A = 60°$, a 1 min (1′) uncertainty in D gives 0.0003 for the corresponding precision in n. Angular deviation requires the sample to be a less convenient shape than the parallel-sided sheet of the lateral deviation method. However, measurements of D and A can be made by simply using an ocular goniometer.

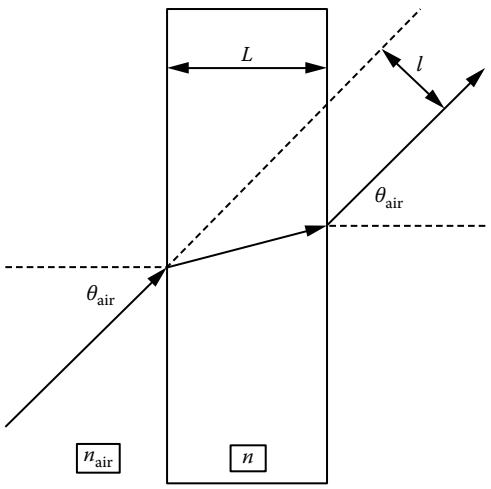


FIGURE 50.5 Lateral deviation l of a light beam on passage through a parallel-sided refractile sample slab of thickness L.

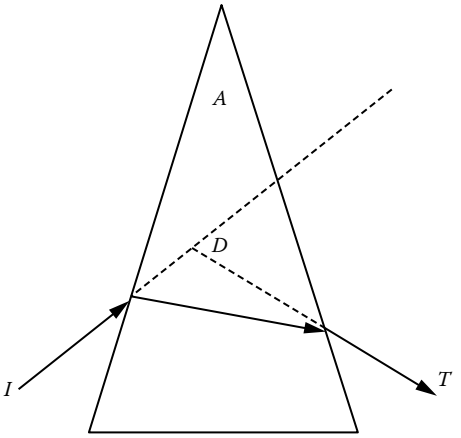


FIGURE 50.6 Angular deviation, D, of a light beam on passage through a prism of vertex angle A. If the interior beam forms an isosceles triangle with A at the vertex, then D is the angle of minimum deviation.

Lateral or angular deviation methods can be applied to liquid samples in a suitable cell with walls having parallel sides. Such cell walls will, however, give some lateral deviation but not cause angular deviation if used to make a hollow prism to contain a liquid sample. In the Hilger–Chance angular deviation refractometer [9], the liquid sample is contained in a glass 90° V-block of refractive index n_{glass}, with parallel outer faces. A parallel–collimated light beam incident normally on one outer face of the block suffers angular deviation D (not minimum), Equation 50.2, giving

$$n^2 = n_{glass}^2 + (n_{glass}^2 - \sin^2 D)^{1/2} \sin D \tag{50.8}$$

For a 1′ uncertainty in D, the precision in n is about 10^{-5}. Solid samples can be measured by coupling a prism of the sample into the V-block using a thin film of liquid of refractive index intermediate between that of the sample and the prism.

50.3.3 Critical Angle Method

In the region of the critical angle, Fresnel's equations show that the transmitted or reflected optical power varies discontinuously with angle. Thus, θ_c is accurately measurable and a refractive index precision of about 10^{-4}–10^{-5} is obtained. Liquid samples are placed directly onto one face of the refractometer prism of known refractive index. Solid samples need only one face to be flat and optically polished, which is coupled to face of the refractometer prism using a liquid film of intermediate refractive index. For a transparent sample, either the transmission or the reflection mode can be used, as θ_c is the same for both (see Figures 50.2 and 50.3). In the transmission mode, the sample–prism interface is illuminated diffusely through the sample, the critical angle then being the largest angle of transmission in the prism. In the reflection mode, useful for highly scattering or absorbing samples, the prism–sample interface is illuminated diffusely from within the prism. The critical angle is then where the reflectance changes discontinuously with angle of reflection within the prism, being total for $\theta \geq \theta_c$. In the transmission mode, the transmittance from the sample into the prism is zero for $\theta \geq \theta_c$ and finite for $\theta \leq \theta_c$, whereas in the reflection mode, typically a few percent of the light is reflected for $\theta \leq \theta_c$ and 100% is reflected for $\theta \geq \theta_c$. If the interface is illuminated with diffuse light, the transmission mode gives a dark–bright contrast at the critical angle, with a dim–bright contrast for the reflection mode. For ocular observation, this difference can make the transmission method preferable. Automatic instruments invariably use the reflection mode, this being necessary for samples that attenuate by absorption or scattering. For such samples, however, the discontinuity of reflectance is replaced by a continuous function with a maximum gradient

at an angle θ_{max} that differs increasingly from θ_c as the sample becomes more absorbing [10–12]. A separate measurement of the absorption index is required before θ_c can be obtained from θ_{max}. Attempts to improve the precision of critical angle refractometry of nontransparent samples are expected to become increasingly limited by this effect.

The critical angle method, however, is robust and simple to use. Despite the problems associated with optically attenuating samples, it is likely to persist as the preferred method in commercial refractometers.

50.3.4 Brewster Angle Method

Here, the angle of reflection θ_B is measured for which there is no linearly polarized light reflected from the sample surface. Equation 50.4 then gives the sample refractive index, where n_1 is the refractive index of the medium containing the incident and reflected light, usually air. The light is polarized such that its electric field vector vibrates parallel to the plane of incidence (see Figure 50.4). A linearly polarized laser is a suitable source of monochromatic light. Nonpolarized sources require polarizing with a Glan–Thomson prism or a quality dichroic polarizer. With a 1 mW HeNe laser and ocular judgment of the minimum reflectance, a precision in θ_B of about 0.1° can be obtained, giving a precision in the sample refractive index of about 0.005. Precision improvement by factors of 10–100 can be expected if photometric analysis replaces the eye. The Brewster angle method is then comparable in precision to the critical angle method. Unlike the critical angle method, however, it does not require a reference material of higher refractive index than the sample. If the sample is absorbing, the reflectance zero at θ_B is replaced by a reflectance minimum at an angle θ_{min}. It can be shown that a given absorption produces less error in the Brewster angle method than in the critical angle method [6,12]. Despite these good features of the Brewster angle method, a commercial Brewster angle refractometer is not known to this author.

50.3.5 Index Match Method

The refractive index match method does not require the sample to be in any special shape or form. Thus, it might be finely divided, that is, a powder. Assuming the particles to be isotropic and homogeneous, then the reflected radiance, the angular or lateral deviation, and the relative phase shift of transmitted light all become zero at refractive index match. If the sample is absorbing, the refractive index of the surrounding liquid that gives the minimum visibility still gives closely the real part of the sample's refractive index. Similarly, if the sample is heterogeneous or anisotropic, the method gives the average refractive index of the sample. The liquid should not dissolve, permeate, or react in any way with the solid sample.

For macroscopic samples, a precision of about 0.005 in n has been estimated [6] from the index match method, where a simple ocular estimation of match is employed. For a matching liquid of refractive index n_{liq}, the reflectance from an interface varies approximately as $(n - n_{liq})^2$. This quadratic dependence makes the sensitivity of the method decrease as the match is approached and makes it difficult in practice to judge whether n_{liq} should be increased or decreased to improve the match.

50.3.6 Microscopy and Scattering Methods

The refractive index of individual particles large enough to image in a microscope can also be measured by the index match method. Central or oblique illumination modes can be used corresponding to the Becke and Schroeder van der Kolk methods, respectively, which utilize null deviation at refractive index match. A precision of 0.0002 in n is quoted [13] for the Becke line method. Whether n_{liq} is lesser or greater than the sample's refractive index n is shown by the appearance of the image.

The phase difference between light transmitted through and around the particle is proportional to $n - n_{liq}$, so that the phase changes sign each side of match. This is used in measurement of refractive index of particles by phase-contrast or interference microscopy, in which the image is brighter or darker than the background according to whether $n > n_{liq}$ or $n < n_{liq}$. These methods can measure the refractive

index of particles below the conventional resolving limit, as the details of an object need not be fully resolved in order to discern match or otherwise. Users of microscopy in refractometry should refer to texts on phase-contrast and interference microscopy. Newer forms of optical microscopy (e.g., scanning confocal) have greater resolving power and can be used for refractive index matching of finer particles.

For a suspension of particles smaller than about 1 μm, the optical turbidity and light scattering are zero at refractive index match, leading to turbidimetric or nephelometric methods of measuring the particle refractive index. For particles of colloidal size, it is important to avoid aggregation or flocculation and suspension destabilization. Thus, chemical restrictions can preclude the attainment of match, particularly if the particle's refractive index is large compared with n_{liq}. In such cases, the turbidity E is measured for changes in n_{liq} that do not destabilize the suspension, and the value of n_{liq} is found by extrapolation to $E = 0$. Extrapolation to obtain this value of n_{liq} is difficult unless the experimental parameters can be arranged to have a linear relation. Theory and experiments described by Griffin and Griffin [14] showed that linearity is obtained if $n_{liq}\sqrt{E}$ is plotted versus n_{liq}.

50.3.7 Anisotropic and Heterogeneous Materials

If a plane face is made from a general optically anisotropic sample, and this face is then offered to the prism of a critical angle refractometer, two critical angles can be simultaneously observed. One can be enhanced over the other through the use of a linear polarizer rotated in the incident or refracted light path. The interpretation of these critical angles, in the general case, is not simple, and the reader is referred to analyses [15–17]. Similar complexity is visited upon the Brewster angle method if the sample is anisotropic, and again the reader is referred to Ref. [6] for works on ellipsometry.

Highly turbid samples, such as polymer latexes and other concentrated colloidal fluids, have been investigated using the critical angle method [1,18–20]. It appears that if the particle size is less than about half a wavelength, the critical angle method remains successful for volume fractions up to at least 0.5. For high concentrations of larger particles, there is a rapid degradation of the reflectance or transmittance discontinuity at the critical angle. This angle is replaced by a region where the reflectance or transmittance may vary in an unexpected way as the particle size and concentration increase.

The effect of heterogeneity on the reflection of polarized light, and on the Brewster angle, is also affected by sample heterogeneity [21,22]. Whereas for a homogeneous sample the parallel-polarized reflectance at the Brewster angle is zero, the effect of sample heterogeneity on the wavelength scale is to cause finite reflectance. Comparison of the Brewster angle and the critical angle methods for heterogeneous materials awaits investigation.

Gate and coworkers have measured the refractive index of emulsion paint films [23] and coated paper [24] by a non-Brewster angle reflection method. Both materials were strongly heterogeneous on the wavelength scale and so (in air) had a strong nonspecular or diffuse reflectance. Some specular reflection remained, however, and the method derives the refractive index of the surface of the film from the ratio of the specular reflectances at a non-Brewster angle for polarizations parallel and normal to the plane of incidence.

The application of refractometry to heterogeneous and absorbing samples is clearly a developing art, and although it is clear that refractometry offers valuable information, the refractometrist should use existing instruments critically.

50.4 Review of Refractometers

The author's initial review of refractometer manufacturers revealed about 30 companies, and the true number may be more than 50. A very large number of general- and special-purpose instruments are offered by each company. It is clearly not possible to review, compare, and advise on individual models. The aspiring refractometrist should do this according to needs and budget. The generalities that follow should assist.

Most instruments use the critical angle (or Abbe) method. They thus have a prism of known refractive index with a face to which the sample is offered. The larger manufacturers offer three types of critical angle instrument:

1. The portable handheld models, which require no power or battery and are customized and calibrated for Brix (sugar), urine, antifreeze, cutting oil, etc. Their refractive index precision is typically between 10^{-3} and 10^{-4}, but their range is small owing to their particular application. The user looks into an eyepiece and reads a scale on which is superposed the dark–light boundary caused by transmission or reflection in the region of angle each side of the critical angle, often termed the critical edge. The price varies greatly between models and manufacturers, typically U.S. $400. Some manufacturers offer digital readout portable models, of similar weight (less than 1 kg) and precision, and battery powered. These cost about U.S. $2000.
2. The laboratory (Abbe) refractometer, which has a refractive index precision of typically 10^{-4}–10^{-5} and covers a wide range of refractive index, typically 1.3–1.75 in a low refractive index instrument and 1.4–1.85 in a high refractive index model. Such instruments are calibrated so that the temperature can be varied using an external liquid bath, and the wavelength can be varied using external light sources. Very roughly, their weight, cost, and precision are about 10 times that of the portable instruments in (1) earlier. They require the user to estimate the position of the critical edge on a graduated scale viewed through an eyepiece.
3. The digital refractometer. Compared with those in (2) earlier, these have similar precision and general facilities, a cost typically a few times greater, but with the advantage of a digital or electronic readout of the data. The upper refractive index offered by several manufacturers appears to be restricted to about 1.6.
4. The process refractometer is a version of those described in (3) earlier, but where the prism is mounted to contact a fluid in a reactor or a pipe and viewed via a fiber-optic cable. Thus, high pressures and temperatures can be accommodated. The prism is usually of sapphire to avoid wear of the optical face, and the precision generally tends to be somewhat lower than those described in (2) and (3) earlier.

References

1. M. Mohammadi, Colloidal refractometry: Meaning and measurement of refractive index for dispersions; the science that time forgot, *Adv. Colloid Interface Sci.*, 62, 17–29, 1995.
2. J. E. Geake and C. Smalley, A simple linear hand-held refractometer, *J. Phys. E: Sci. Instrum.*, 16, 608–610, 1983.
3. H. Li and S. Xie, Measurement method of the refractive index of biotissue by total internal reflection, *Appl. Opt.*, 35, 1793–1795, 1996.
4. M. Born and E. Wolf, *Principles of Optics*, 6th edn., Oxford, U.K.: Pergamon, 1980.
5. P. E. Ciddor, Refractive index of air: New equations for the visible and near infrared, *Appl. Opt.*, 35, 1566–1573, 1996.
6. G. H. Meeten, Refraction and extinction of polymers, in *Optical Properties of Polymers*, Ed. G. H. Meeten, London, U.K.: Elsevier Applied Science, 1986.
7. H. Moosmüller and W. P. Arnott, Folded Jamin interferometer: A stable instrument for refractive index measurements, *Opt. Lett.*, 21, 438–440, 1996.
8. B. Richerzhagen, Interferometer for measuring the absolute refractive index of liquid water as a function of temperature at 1.064 μm, *Appl. Opt.*, 35, 1650–1653, 1996.
9. R. S. Longhurst, *General and Physical Optics*, 3rd edn., London, U.K.: Longmans, 1973.
10. G. H. Meeten, A. N. North, and F. M. Willmouth, Errors in critical-angle measurement of refractive index of optically absorbing materials, *J. Phys. E: Sci. Instrum.*, 17, 642–643, 1984.
11. P. R. Jarvis and G. H. Meeten, Critical angle measurement of refractive index of absorbing materials, *J. Phys. E: Sci. Instrum.*, 19, 296–298, 1986.

12. G. H. Meeten, Refractive index errors in the critical-angle and Brewster-angle methods applied to absorbing and heterogeneous materials, *Meas. Sci. Technol.*, 8, 728–733, 1997.
13. C. P. Saylor, *Advances in Optical and Electron Microscopy*, Vol. 1, Eds. R. Barer and V. E. Cosslett, London, U.K.: Academic Press, 1966.
14. M. C. A. Griffin and W. G. Griffin, A simple turbidimetric method for the determination of the refractive index of large colloidal particles applied to casein micelles, *J. Colloid Interface Sci.*, 104, 409–415, 1985.
15. F. Yang, G. W. Bradberry, and J. R. Sambles, A method for the optical characterization of thin uniaxial samples, *J. Modern Opt.*, 42, 763–774, 1995.
16. F. Yang, G. W. Bradberry, and J. R. Sambles, Critical edge characterization of the optical tensor of a uniaxial crystal, *J. Modern Opt.*, 42, 1241–1252, 1995.
17. F. Yang, J. R. Sambles, and G. W. Bradberry, A simple procedure for characterizing uniaxial media, *J. Modern Opt.*, 42, 1477–1458, 1995.
18. G. H. Meeten and A. N. North, Refractive index measurement of turbid colloidal fluids by transmission near the critical angle, *J. Meas. Sci. Technol.*, 2, 441–447, 1991.
19. J. E. Geake, C. S. Mill, and M. S. Mohammadi, A linear differentiating refractometer, *J. Meas. Sci. Technol.*, 5, 531–539, 1994.
20. G. H. Meeten and A. N. North, Refractive index measurement of absorbing and turbid fluids by reflection near the critical angle, *J. Meas. Sci. Technol.*, 6, 214–221, 1995.
21. E. K. Mann, E. A. van der Zeeuw, G. J. M. Koper, P. Schaaf, and D. Bedeaux, Optical properties of surfaces covered with latex particles: Comparison with theory, *J. Phys. Chem.*, 99, 790–797, 1995.
22. E. A. van der Zeeuw, L. M. C. Sagis, and G. J. M. Koper, Direct observation of swelling of non-cross-linked latex particles by scanning angle reflectometry, *Macromolecules*, 29, 801–803, 1996.
23. L. F. Gate and J. S. Preston, The specular reflection and surface structure of emulsion paint films, *Surf. Coat. Int.*, 8, 312–330, 1995.
24. L. F. Gate and D. J. Parsons, The specular reflection of polarised light from coated paper, *Transactions of the Tenth Fundamental Research Symposium on Products of Papermaking*. Oxford, U.K., Vol. 1, pp. 263–283, September 1993, Leatherhead, U.K.: Pira International 1994.

51

Turbidity Measurement

Daniel Harrison
John Carroll University

Michael Fisch
John Carroll University

51.1 Introduction

For nearly 50 years, turbidity measurements have been used to perform a wide variety of physical measurements. These include the determination of particle concentration per unit volume when the scattering cross section per particle is known, the determination of particle size (scattering cross section) when the concentration of particles is known [1–4], and the determination of some of the critical exponents associated with second-order and nearly second-order phase transitions [5–9].

The basic ideas necessary to understand turbidity measurements are fairly simple. In the absence of reflection losses, when a weak beam of light passes through a dielectric sample, the two processes most responsible for reducing the intensity of the transmitted beam are absorption and scattering. The reduction in transmitted light intensity due to scattering is called the sample's turbidity. Extinction includes the effects of both absorption and scattering. The Beer–Lambert or Lambert law describes the effects of both absorption and turbidity on the transmitted intensity. This law is written as

$$I_T = I_0 \exp -(\alpha + \tau)l \tag{51.1}$$

where
I_T is the intensity of the light transmitted through the sample
I_0 is the intensity of the light incident on the sample
α is the absorption coefficient per unit length
τ is the turbidity per unit length
l is the length of the light path in the sample

As discussed later, it is more correct to use powers rather than intensities; thus, this equation should be written as

$$P_T = P_0 \exp -(\alpha + \tau)l \tag{51.2}$$

The most general situation is to have an absorbing medium with absorbing and scattering particles embedded within it; however, here the focus is on the simpler case of nonabsorbing medium and consider the two most common cases: nonabsorbing particles in medium and scattering caused by fluctuations in the medium itself.

Instruments generally fall into two categories: commercial and laboratory constructed. The commercial units are of two general types: (1) attachments to spectrophotometers and (2) white-light turbidity meters, which operate under ambient conditions. Anyone interested in the former should consult the catalog of accessories for the instrument in question. Those interested in the latter should check under "nephelometers" or "turbidimeters" in scientific supply catalogs or on the web [10]. Units of this type perform their designed function admirably and are sufficient for routine work. However, they may need modification for the more specialized measurements performed in many research laboratories. For this reason, it is common to construct special laboratory instruments. These come in at least two types, most commonly either single-beam or dual-beam instruments. The single-beam instrument usually has an intensity-stabilized light source, whereas the dual-beam instrument corrects for drift in the light source and reflection losses by either electronically or mathematically taking the ratio of the transmitted light power and a reference beam power.

The balance of this chapter is divided into three parts. The first part briefly discusses the physical basis of turbidity measurements and demonstrates how such measurements may be used to infer the scattering cross section of the particles in the solution or the concentration of the scatters in the solution. The second section discusses turbidity of pure fluids and shows how certain critical exponents may be determined from such measurements. The last section discusses laboratory instruments and the relative trade-offs involved in such instruments.

51.2 Extinction and Turbidity: Particles in a Nonabsorbing Medium

Suppose that electromagnetic radiation is incident upon a slab of medium consisting of randomly positioned particles and that the transmitted radiation is detected as shown in Figure 51.1. For the present discussion, assume that the source and detector are in the medium. The radiation that is received by the detector will be less than that incident on the slab because of the presence of the particles in the medium—the particles have caused extinction (or attenuation) of the beam. The extinction of the radiation depends on two physical processes, scattering and absorption, whereas the turbidity depends only on scattering. In scattering, there is no change in the total energy of the radiation; rather, some of the

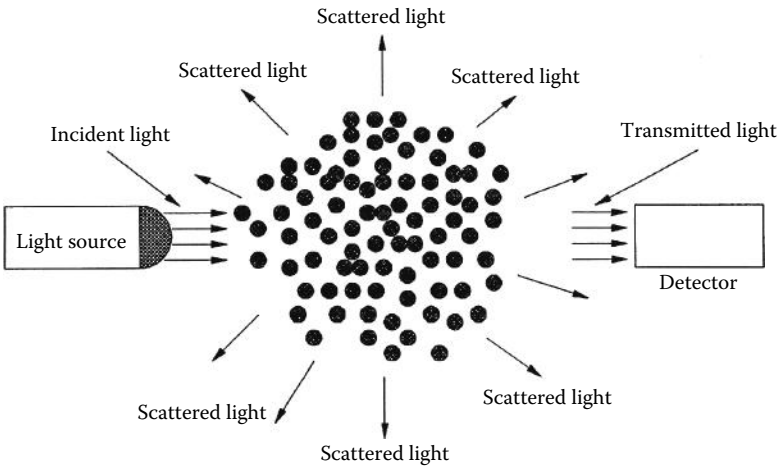


FIGURE 51.1 An idealized experiment indicating the physical basis of turbidity.

incident radiation is redistributed away from the incident direction. In absorption, some of the energy of the incident beam is transformed into other energy forms. This extinction of the beam is rather complicated; it will, in general, depend on the chemical composition, size, shape, number, and orientation of the particles; the chemical composition of the medium; and the frequency and polarization of the incident radiation [2,3].

In discussions of scattering, one often focuses on the *cross sections*, C_i, of the particles. On the basis of conservation of energy, one must write

$$C_{ext} = C_{scatt} + C_{abs} \tag{51.3}$$

where C_{ext}, C_{scatt}, and C_{abs} are the cross sections for extinction, scattering, and absorption, respectively. These cross sections all have the dimension of area, and in all cases a larger cross section (area) indicates a larger effect, and a smaller cross section a smaller effect. In general the cross section will depend on all the intensive quantities that describe the extinction. Finally, note that for nonabsorbing media, $C_{abs} = 0$.

To see how this relates to the measured turbidity, assume now that the medium is nonabsorbing, that all the particles are nonabsorbing and have the same scattering cross section (C), and that there are n of these particles per unit volume. Then, one obtains the following expression for the transmitted light power:

$$P_T = P_0 \exp -(nCl) \tag{51.4}$$

where

P_T is the transmitted power

P_0 is the incident power

We have switched from intensity to power because power is the quantity detected, and power is independent of the details of the spatial distribution of electromagnetic radiation. Upon noting this change and comparing this to Equation 51.2, when $\alpha = 0$, it is clear that the turbidity can be expressed as

$$\tau = nC \tag{51.5}$$

Thus, a measured turbidity can be related to C_{scatt} if n is known, or n may be determined if C_{scatt} is known. The difficulty is then in determining C_{scatt}. There are a number of approaches that can be used. For well-characterized particles, theoretical expressions may be determined [2]. In the absence of such theory, measurements may be made as a function of concentration and then, since $\tau = nC$, C may be determined. The generalizations to account for nonmonochromatic beams and extinction by a collection of noninteracting particles of the same type but different sizes are straightforward. The simplest generalization to include different sizes assumes that there is one parameter, ζ, that describes the distribution of particle size and the corresponding C_{scatt} [1]. That is, let $C_{scatt}(\zeta)$ be the scattering cross section for a particle characterized by the parameter (for instance, radius) ζ, and let the number per unit volume with parameter between ζ and $\zeta + d\zeta$ be $N(\zeta)d\zeta$, such that

$$n = \int_0^\infty N(\zeta)d\zeta$$

then

$$\tau = \int_0^\infty C_{scatt}(\zeta)N(\zeta)d\zeta \tag{51.6}$$

A similar integral exists when the incident radiation is not monochromatic, but, in this case, the average is over the wavelength variation of the incident radiation.

51.3 Turbidity Due to Density Fluctuations in Pure Fluids

It is well known that even in very pure and well-filtered liquids, in which there are essentially no particles, there is still some reduction in the detected power due to scattering [3]. The explanation is that there are always thermal fluctuations in the dielectric constant of the media, ε, and it is these fluctuations from the mean dielectric constant that lead to the scattering of light. Through application of thermodynamic and statistical mechanical principles [3], the following expression for the turbidity results:

$$\tau = \frac{8\pi^3}{2\lambda^4}\left[\rho\left(\frac{\partial\varepsilon}{\partial\rho}\right)_T\right]^2 kT\beta_T \tag{51.7}$$

where

ρ is the density
ε is the dielectric constant
k is Boltzmann's constant
T is the absolute temperature
λ is the wavelength of the incident light
β_T is the isothermal compressibility

Near a second-order phase transition, the fluctuations become correlated over larger distances, and this expression must be generalized to include the effects of this increase in the *correlation length*, ξ. In this case, the turbidity is given by [6]

$$\tau = A\pi\beta_T\left[\frac{2\alpha^2+2\alpha+1}{\alpha^3}\ln(1+2\alpha) - \frac{2(1+\alpha)}{\alpha^2}\right] \tag{51.8}$$

where

$$A = \frac{\pi^2}{\lambda^4}\left[\rho\left(\frac{\partial\varepsilon}{\partial\rho}\right)_T\right]^2 kT \quad \text{and} \quad \alpha \equiv 2\left(\frac{2\pi n}{\lambda}\xi\right)^2$$

where n is the (mean) index of refraction of the fluid. Both the isothermal compressibility and the correlation length exhibit approximate power-law behavior of the form

$$\beta_T, \xi \sim \left(\frac{|T_c - T|}{T_c}\right)^{-x_i} \tag{51.9}$$

near these phase transitions. Thus, a measurement of $\tau(T)$ allows the determination of the critical exponents, x_i, and the values of β_T and ξ far from the phase transition. This technique has been generalized to binary liquid mixtures, liquid crystals, and other systems. In these cases, the isothermal compressibility is replaced by the appropriate generalized susceptibility. There are also, in some cases, small corrections to the aforementioned expressions. These can also be found in Refs. [11–13].

51.4 Design of Laboratory Instruments

When more exact measurements are required than can be obtained using commercial instruments, and other enhancements such as mK temperature and/or pressure controlled sample chamber are required, it is common to construct a turbidity instrument. There are basically two types of instruments (single beam and dual beam). With both types of instruments, both ac and dc detection

schemes are possible. Generally, the long-term time stability of synchronous ac detection techniques is superior to dc techniques. This section discusses all of these techniques.

51.4.1 Single-Beam Instruments: Optics

The basic optical design of a single-beam instrument is shown in Figure 51.2. The light source is a low-power (1–5 mW) vertically polarized He–Ne laser. It is important to use a polarized laser. The polarization direction of an unpolarized laser wanders, and this can lead to time-dependent reflection coefficients from the various surfaces in this instrument and hence reduced performance. The beam passes through a laser intensity stabilizer. This device reduces long-term drift in the laser power to ≤0.1%, thus allowing single-beam operation. Table 51.1 shows several commercial intensity stabilizers.

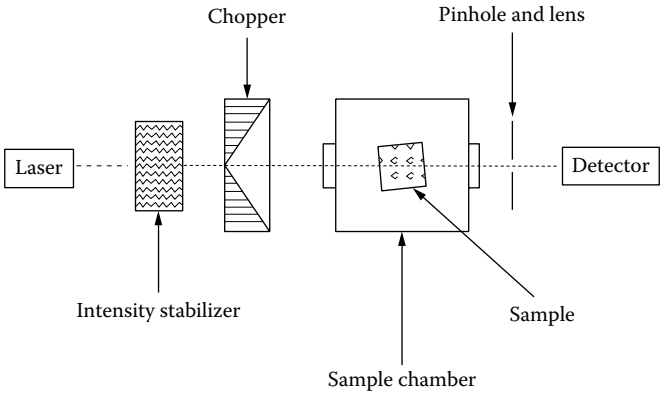


FIGURE 51.2 Block diagram of the optics for a single-beam turbidity apparatus.

TABLE 51.1 Commercial Laser Intensity Stabilizers and Stabilized Lasers

Manufacturer/Supplier	Model	Comment
Newport Corporation 1791 Deere Avenue Irvine, CA 92606 (949)-863–3144 Fax: (949) 253–1680	R-32734	1.5 mW intensity- and frequency-stabilized laser
Edmund Optics 101 East Gloucester Pike Barrington, NJ 08007 Phone: (856)-573–6250 (800)-363–1992 Fax: (856)-573–6295	NT59–939	
Thorlabs Newton, New Jersey 56 Sparta Avenue Newton, NJ 07860 (973)-579–7227 Fax: (973)-300–3600	HRS015	
Brockton Electro-Optics 50 Central Square Suite 3 Bridgewater, MA 02324 (508)559–9958 Fax (508) 861–0214	LPC-VIS	Outstanding for general use

To reduce any possible heating of the sample due to absorption, the beam is then passed through a neutral density filter, which attenuates the beam further to a level of 50–100 μW. The next item in the beam line depends on the detection scheme. For ac operation, a light chopper is inserted into the beam line, while for dc detection, nothing else is required before the sample. An output from the chopper is used for synchronous detection. The sample chamber then follows. This chamber requires an input and an output port and a design that precludes interference of any beam with other reflected beams. Other design details are determined by the materials to be studied and any required external factors that are to be controlled. After the beam has been transmitted through the sample, it passes through either a pinhole or a lens–pinhole combination [14,15]. This results in a small angle of acceptance for the detector that, in turn, reduces the scattered light at small angles from the transmitted beam that reaches the detector. A detector then completes the optical train. For most applications, a *pin* photodiode is an excellent, inexpensive detector; however, in some applications, other types of detectors may be superior.

51.4.2 Single-Beam Systems: Electronics

For both synchronous and dc detection, the output of the photodiode, which is a current device, must be transformed into a voltage using a current-to-voltage converter (CVC). It is important not to use just a resistor for this function. The simplest workable CVC uses an FET-input operation amplifier such as an LM-11 and one resistor. The circuit, also known as a transimpedance amplifier, is common and can be found in many sources [16].

When dc detection is used, the output voltage from the CVC is read on a voltmeter. The relationship between this voltage and the turbidity will be discussed later. The stability of dc operation is greatly improved by maintaining the photodiode and the electronics at a stable temperature. The drifts in detector efficiency and in component values associated with temperature changes can limit the accuracy of the instrument. Furthermore, the detector must be shielded from stray light, which means the instrument must be operated in a darkened room or under a light-tight cover. Such covers may cause excessive heating of the laser and the system. A "spike" filter that passes the wavelength of the laser may also assist in solving this problem.

When synchronous ac detection is used, the output from the CVC is the input to a lock-in amplifier that is synchronized to the chopper. Commercial units are available and can be used. One can also construct a simple lock-in using an Analog Devices AD-630 balanced modulator/demodulator, a phase shifter to ensure that the reference signal and the transmitted signal are in phase, and an *RC* filter. The output voltage from the RC filter is then read with a dc voltmeter. The necessary wiring diagrams are available from Analog Devices [17]. Synchronous detection has the advantage of having lower $1/f$ noise than dc circuits. Also, synchronous detection can discriminate against signals not synchronized with the reference and hence offers much greater immunity to room lights and other stray light sources. The advantages of such detection are discussed extensively in the literature [16,18].

Single-beam techniques are most useful for measuring changes in turbidity. To see this, note that the voltage output from the CVC or the lock-in, V_t, is given by $V_t = AP_T$, where A is an instrument constant that includes factors such as current-to-voltage gain, light power level to current conversion in the photodiode, reflection losses from the various interfaces, incident light power, and gain of the lock-in. Now, P_T is given by Equation 51.2; hence V_{dc}, the measured dc voltage, may be written as

$$V_{dc} = B\exp -(\alpha + \tau) \tag{51.10}$$

where $B = AP_0$ is another constant, and we have included absorption. B may be determined by putting a solution of known extinction in the instrument and solving Equation 51.10 for B. Then, knowing l, α, B, and V_{dc}, the turbidity may be found by solving Equation 51.10 for τ, that is

$$\tau = \frac{1}{l}\ln\left(\frac{V_{dc}}{B}\right) - \alpha \tag{51.11}$$

This technique is rather limited if a large number of different samples must be compared. For a given sample, the change in turbidity is easily measured as a function of externally controlled parameters. However, putting samples into and taking them out of the instrument usually produces small random errors that slightly change the instrumental constant A and degrade performance. In some cases, the difference in turbidity between two samples may be small and result in a situation in which one must deal with a small difference of large numbers. This situation may be rectified using dual-beam instruments.

For simple measurements, a turbidimeter device is available for GE. It is described as follows: "GE Sensing/NovaSensor Turbidity Sensors measure the amount of suspended particles, or turbidity, in the wash water of washing machines and dishwashers." There are devices with both turbidity and temperature sensors and only turbidity. It has the potential to be of use in simple one-beam instruments that have some special characteristic such as low battery drain or flow through water conditions. These devices require 5 V dc, although less appears to be acceptable, and draw 30 mA max. To operate, one needs a voltage source, two resistors, and a voltmeter [19].

51.4.3 Dual-Beam Instruments: Optics

The optical design of a dual-beam, two-detector instrument is shown in Figure 51.3. The first several parts of the optical train are the same as for the single-beam instrument, and the same comments apply in this case. The first real difference is that there need not be an intensity stabilizer in the dual-beam system. However, inclusion of the intensity stabilizer may improve system performance. The second difference is the inclusion of a beam splitter that splits the laser beam into two parts. We have found that a microscope slide works well as a beam splitter [20]. One of these beams (usually the reflected beam) becomes a reference beam that passes through a reference sample. The other passes through the sample as in the single-beam instrument. The beams from these two paths go through lens–pinhole combinations and fall on detectors. All of the comments that apply to the single-beam instrument also apply here. The associated electronics are also identical. Once more, both dc and ac detection may be used. Experience in our laboratory shows that the best results are obtained when the sample signal is simultaneously divided by the reference signal using an integrated circuit divider. A modification that uses two beams and one detector is shown in Figure 51.4. This modified design eliminates the need for absolute stability of two separate detectors but requires somewhat more complex optics.

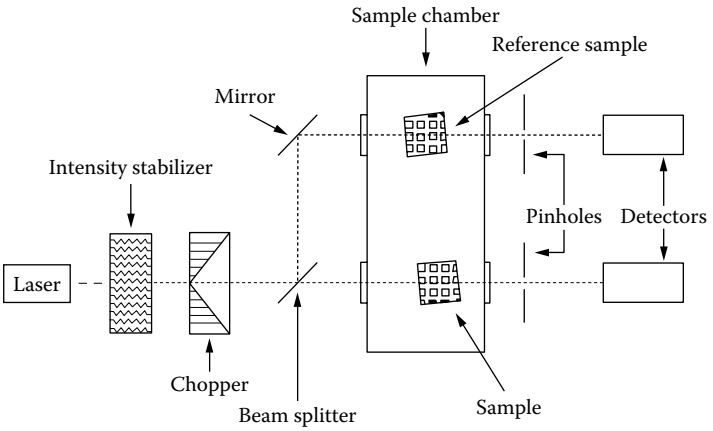


FIGURE 51.3 Block diagram of the optics for a dual-beam, two-detector turbidity apparatus.

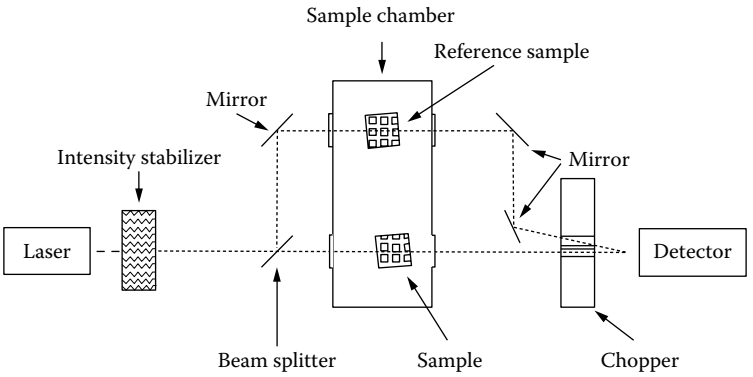


FIGURE 51.4 Block diagram of the optics for a dual-beam, one-detector turbidity apparatus.

The following analysis assumes that the measured output is the ratio of the sample signal voltage to the reference signal voltage. Let the voltage output from the detector's CVC be V_{sig} and the voltage output of the reference signal detector's CVC be V_{ref}. Then

$$V_{sig} = AP_T \tag{51.12}$$

and

$$V_{ref} = CP_R \tag{51.13}$$

where

A and C are the instrumental constants totally analogous to the instrumental constant A earlier
P_R is the power of the reference beam at the detector

It is important to realize that A and C are constant, but they need not be identical. The output voltage ratio, V_{out}, is given by

$$V_{out} = \frac{V_{sig}}{V_{ref}} = \frac{AP_T}{CP_R} = D\frac{P_T}{P_R} \tag{51.14}$$

where $D = A/C$ is a constant. Note that in this arrangement, common multiplicative noise on the two voltages will tend to cancel out. Now P_T can be replaced by Equation 51.1, and P_R is given as

$$P_R = f^{\star} P_0 \exp -(\alpha_R + \tau_R) l_R \tag{51.15}$$

where

α_R is the absorption
τ_R is the turbidity of the sample in the reference beam
l_R is the thickness of the sample in the reference beam
f is a constant determined by the beam splitter

Then, assuming that l_R is the same as l, the following expression results:

$$V_{out} = D\frac{P_0 \exp -(\alpha + \tau) l}{f^{\star} P_0 \exp -(\alpha_R + \tau_R) l} \tag{51.16}$$

and thus,

$$\tau = -\frac{1}{l}\ln\left(\frac{V_{\text{out}}}{D'}\right) + \tau_{\text{R}} - \alpha + \alpha_{\text{R}} \tag{51.17}$$

where $D' = D/f$. Often, this can be considerably simplified when the reference sample is a pure liquid because then τ_{R} is of order 10^{-5} and α_{R} is totally negligible. Normally, the reference sample is chosen so that its optical properties approach those of the sample in the case where the concentration of particles is low or the system is far from a phase transition. The constant D' may be determined by using the same sample in both the reference and the signal arms of the instrument—in this case, $V_{\text{out}} = D'$.

51.4.4 Dual-Beam Instruments: Electronics

The electronics for dc detection consist of two CVC converters, one for each detector. The voltage outputs from these may be read directly. However, experiments indicate that greater stability is achieved by simultaneously dividing these two signal outputs as discussed earlier. In our laboratory, this is accomplished using an Analog Devices AD-532 internally trimmed integrated circuit divider. These can be purchased with different accuracies. In the present application, they typically perform slightly better than the factory specifications indicate. The circuit details are available from Analog Devices [21].

The electronics for ac detection generally consist of two synchronous detectors followed by division of the sample signal by the reference signal. Once more, this may be accomplished using two Analog Devices AD-630 balanced modulator/demodulators or two commercial lock-in amplifiers. Division can be accomplished in the same way as in the dc technique section.

51.5 Limitations

Highly turbid samples, which transmit very little light, may be difficult to measure using a photodiode. In such cases, the photodiode in the sample arm may be replaced with a photomultiplier (PM). In this case, a chopper can be used to reduce the effects of the dark signal from the PM. The outputs, when the signal is present and when it is blocked, are both measured and averaged (preferably digitally), and the difference is computed. By counting for a fixed number of reference counts (requiring another PM), drift in the source may be eliminated. A Stanford Research Instruments photon counter [22] may be configured in this mode. Furthermore, unless extreme precautions are taken in this situation, erroneous results may occur because the scattered power is as large or larger the transmitted beam power.

References

1. H. C. van de Hulst, *Light Scattering by Small Particles*, New York: Dover, 1981.
2. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, New York: Wiley, 1983.
3. M. Kerker, *The Scattering of Light and Other Electromagnetic Radiation*, New York: Academic Press, 1969.
4. M. Cano-Sarabia, A. Angelova, N. Vensota, S. Lesieur, and J. Veciana, *J. Colloid Interf. Sci.*, 350: 10–15, 2010.
5. P. Calmettes, I. Laguës, and C. Laj, *Phys. Rev. Lett.*, 28: 478–180, 1972.
6. V. G. Puglielli and N. C. Ford, *Phys. Rev. Lett.*, 25: 143–147, 1970.
7. D. Beysens, A. Bourgou, and P. Clamettes, *Phys. Rev. A*, 26: 3589–3609, 1982.
8. J. Rouch, A. Safouane, P. Tartaglia, and S. H. Chen, *Phys. Rev. A*, 37: 4995–4997, 1988.
9. W. Schröer, *Contrib. Plasma Phys.,* 52: 78–88, 2012.

10. M. Sadar, *Turbidity and Other Sediment Surrogates Workshop*, April 30–May 2, 2002, Reno, NV, A review from 2002 is available.
11. R. F. Chang, H. Burstyn, and J. V. Sengers, *Phys. Rev. A*, 19: 866–882, 1979.
12. R. Pecora (ed.), *Dynamic Light Scattering*, New York: Plenum, 1985.
13. D. L. Sidebottom and C. M. Sorensen, *J. Chem. Phys*., 89: 1608–1615, 1988.
14. P. Walstra, *Br. J. Appl. Phys*., 16: 1187–1192, 1965.
15. R. O. Gumprecht and C. M. Sliepcevich, *J. Phys. Chem*., 57: 90–95, 1953.
16. P. Horowitz and W. Hill, *The Art of Electronics*, 2nd edn., New York: Cambridge University Press, 1989.
17. AD630 balanced modulator/demodulator, Product Details, Analog Devices, http://www.analog.com/en/rfif-components/modulatorsdemodulators/ad630/products/product.html, accessed on August 21, 2013.
18. R. A. Dunlap, *Experimental Physics Modern Methods*, New York: Oxford University Press, 1988.
19. Mouser electronics parts numbers 527-165D602P003 and 527-TSW-10. These cost roughly $5.00 in single quantities and have a tolerance of 20%. Their data sheet indicates that calibrated single units may be significantly better.
20. R. J. Nash and M. R. Fisch, *Rev. Sci. Instr*., 60: 3051–3054, 1989.
21. AD532 internally trimmed integrated circuit multiplier, Product Details, Analog Devices, http://www.analog.com/en/special-linear-functions/analog-multipliersdividers/ad532/products/product.html, accessed on August 21, 2013.
22. SR400 Dual Channel Photon Counter, Data Sheets, *Stanford Research Systems*, (408) 744-9040, http://www.thinksrs.com/products/SR400.htm, accessed August 21, 2013.

52

Laser Output Measurement

Haiyin Sun
Coherent Auburn Group

52.1 Introduction

A laser is a device that emits an optical beam. A laser beam carries a certain amount of optical power. Lasers emit beams in two different ways: continuous wave (CW) or pulsed. A CW laser emits a steady power as shown in Figure 52.1. CW laser power is measured in terms of watts. The power of commonly used CW lasers ranges from a fraction of milliwatts for a small helium–neon (He–Ne) laser to tens of kilowatts for a carbon dioxide (CO_2) laser. A pulsed laser emits a pulsed power as shown in Figure 52.2. The pulse can be repeated. The pulse duration and repetition rate vary for lasers of different types and are adjustable for some types of lasers. Pulsed laser power is more easily measured in terms of energy per pulse such as joules per pulse. The term "per pulse" is usually omitted. The energy of commonly used pulsed lasers ranges from a few picojoules for a semiconductor laser to tens of megajoules for a semiconductor laser array. The pulse repetition rate ranges from a single pulse for an excimer laser or an x-ray laser to hundreds of megahertz for a neodymium/YLF laser. The pulse duration ranges from tens of femtoseconds for a Ti/sapphire laser up to CW for a CO_2 or He–Ne laser.

A laser beam may contain more than one wavelength (monochromatic) component. One wavelength component may have an optical intensity different from that of another wavelength component. Optical spectral intensity is defined as power per unit wavelength. The optical intensity–wavelength profile of a laser beam is known as the "spectrum." Spectrum measurement usually means measuring the spectral profile. The absolute value of the optical spectral intensity is in fact not important. Figure 52.3 shows a typical spectrum of a multimode semiconductor laser. The wavelength components are inside several bands. Each band may be a "longitudinal mode," and the width of a band is known as the "mode linewidth." Figure 52.4 shows a typical spectrum of a single-mode He–Ne laser. There is only one longitudinal mode with a very narrow linewidth. The power of a laser beam can be calculated by integrating over the laser spectrum.

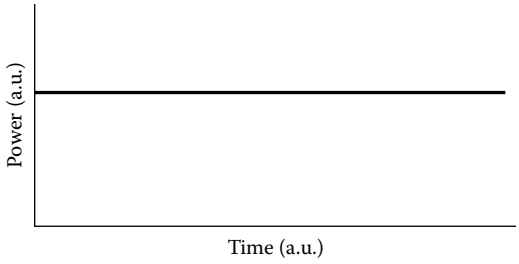


FIGURE 52.1 A CW laser emits a steady power.

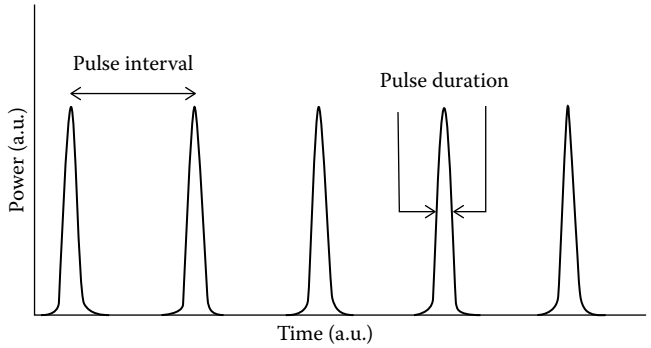


FIGURE 52.2 A pulsed laser emits a pulsed power.

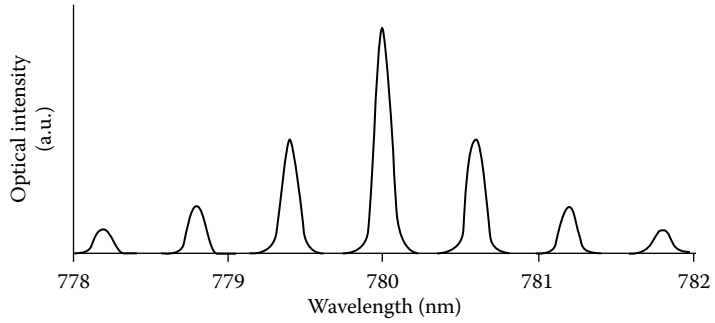


FIGURE 52.3 Spectrum of a multimode semiconductor laser.

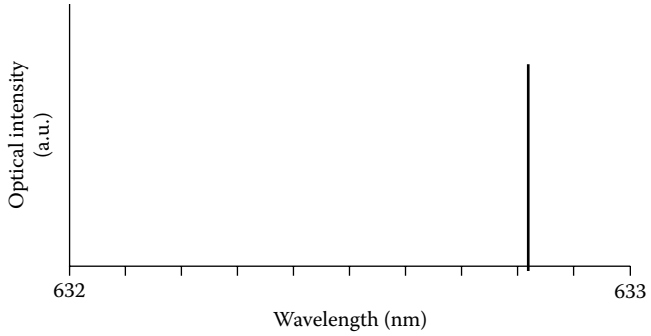


FIGURE 52.4 Spectrum of a single-mode helium–neon laser.

The wavelength of a laser is referred to as the *central wavelength* of the spectrum. But the term "central" is usually omitted. For example, the wavelength is 780 nm for the semiconductor laser shown in Figure 52.3 and is about 632.8 nm for the He–Ne laser shown in Figure 52.4. The wavelength of commonly used lasers ranges from 270 nm in ultraviolet for a neodymium/YAG laser, or even lower for an x-ray laser, to 10.6 μm in infrared for a CO_2 laser. Many lasers have a beam with single longitudinal mode and very narrow linewidth like that shown in Figure 52.4. The spectrum measurement for these lasers then virtually reduces to wavelength measurement.

Power, spectrum, and wavelength are the three most important parameters describing a laser. Laser manufacturers should provide information about these three parameters of their lasers. Some lasers have adjustable power, spectrum, or wavelength. Users often need to measure these parameters to ensure appropriate use of these lasers. This chapter discusses the principles and techniques involved in the measurement of laser power, spectrum, and wavelength.

Readers interested in knowing more about laser working principles and characteristics can read laser textbooks such as Ref. [1], available in many libraries.

52.2 Measurement of Laser Power

A laser power meter can measure laser power or energy. Figure 52.5 shows the scheme of a laser power meter. The three basic components of a laser power meter are a photodetector, an electronic conditioner, and a display device. The photodetector detects the laser beam under measurement and outputs an electrical response proportional to the laser power. The electronic conditioner processes the response and provides a digital or analog signal for displaying. The display device displays the measurement result in terms of watts or joules. To a large extent, the characteristics of the photodetector used determines the performance of a laser power meter. When selecting or using a laser power meter, the following issues are of primary concern:

1. *Spectral response range*: A photodetector has a certain spectral response range that limits the spectral range of a laser power meter. Some photodetectors have a wavelength-dependent spectral response. Calibration is necessary when using a laser power meter with such a photodetector to measure the power at different wavelengths.
2. *Power range*: The detection threshold and damage threshold of a photodetector usually determine the power range of a laser power meter. Incident laser power lower than the detection threshold can cause measurement error, while incident power higher than the damage threshold can cause permanent damage to the photodetector. Specifically, power range includes CW power range, peak power range, spatial power density range, single-pulse energy range, and spatial energy density range. An optical attenuator put in front of the photodetector can raise the damage threshold to as high as megawatt. However, an attenuator will also raise the detection threshold.

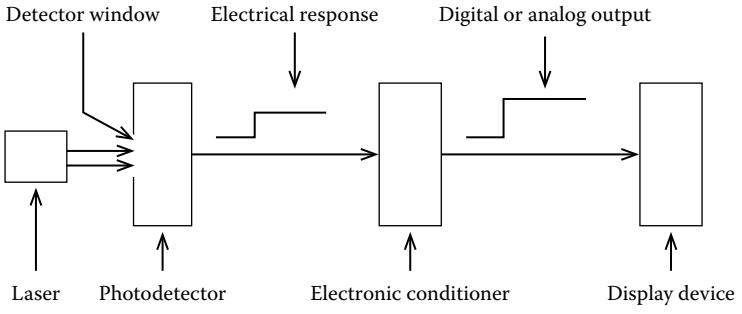


FIGURE 52.5 Scheme of a laser power meter.

3. *Response time*: A photodetector needs a certain time to respond to an incident laser beam. In order to determine the shape of a laser pulse, the detector response time must be shorter than the width of the pulse.
4. *Detector window size*: The input window size of commonly used photodetectors are from a few to tens of millimeters. When the size of a laser beam under measurement is larger than the detector window size, focusing optics must be used to reduce the beam size. Power loss caused by the focusing optics must be excluded to avoid measurement error.

52.2.1 Photodetectors

Three photodetectors are commonly used in laser power meters. They are thermopiles, photodiodes, and pyroelectric probes.

52.2.1.1 Thermopiles

A thermopile is usually a light absorber disk onto which a ring of thermocouples has been deposited. The absorber converts the laser power incident on it into heat and generates between the absorber and a heat sink a temperature difference proportional to the incident laser power. The thermocouples generate and output an electrical response proportional to the temperature difference. Thermopiles usually have a long response time of a few seconds, a broad and fiat spectral response range from 200 nm to 20.0 μm, and a power range of 1 mW–5 kW for CW lasers or 0.01–300 J for pulsed lasers. Because of the flat spectral response, calibration is independent of wavelength. Thermopiles are primarily used to measure moderate to high power output of CW lasers, moderate to high energy output of single-shot pulsed lasers, and the energy output of pulsed lasers with a repetition rate higher than 10 Hz.

52.2.1.2 Photodiodes

Silicon and germanium photodiodes are widely used as photodetectors. A photodiode absorbs the photons (laser beam) incident on it and utilizes the photon energy to create free-carrier pairs (electrons and holes). These free carriers form a response current in an external circuit. The responsivity of a photodiode in amps/watt is given by

$$R\left(\frac{A}{W}\right) = \eta_D \frac{e}{hv} \tag{52.1}$$

where

R is the current produced by the photodiode for per watt of incident power
η_D is the detection efficiency
e is the electron charge (1.6×10^{-19} C)
hv is the photon energy

Since a typical photon energy is 2 eV ≈ 3×10^{-19} J, e/hv is of the order of 0.5 A/W. Normal detection efficiencies η_D exceed 0.5 at optical wavelengths (500–800 nm), leading to typical responsivities of 0.25 A/W. Silicon photodiodes have a narrow spectral response range (400 nm–1.1 μm). The response peak is at about 800 nm. Germanium photodiodes also have a narrow spectral response range (800 nm–1.8 μm). The response peak is at about 1.5 μm. Photodiodes usually have a power range of 1 nW–50 mW for CW lasers or 1 pJ–1 μJ for pulsed lasers and a short response time of about 100 ms. Photodiodes are best for measuring low power output of CW lasers or low energy output of pulsed lasers. Because their spectral response is wavelength dependent, calibration is always required when measuring the power of lasers with different wavelengths. Manufacturers should attach a calibration data sheet to their laser power meters that use a photodiode as the detector. Several other semiconductor photodetectors with different spectral response ranges are available commercially. For example, the spectral response ranges are from

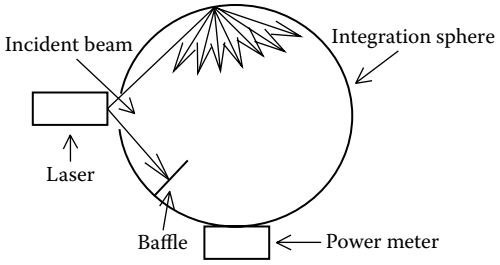


FIGURE 52.6 Scheme of an integration sphere.

1.0 to 3.6 μm, 1.0 to 5.5 μm, and 2 to 22 μm for indium arsenide photodetectors, indium antimonide photodetectors, and mercury cadmium telluride photodetectors, respectively.

52.2.1.3 Pyroelectric Probes

A pyroelectric probe uses a ferroelectric material that is electrically polarized at a certain temperature. The material is placed between two electrodes. Any change in temperature of the material caused by the absorption of laser power produces a response electric current in the external circuit. Pyroelectric probes are primarily used to measure the energy of pulsed lasers because they only respond to the rate of temperature change. Pyroelectric probes usually have a spectral response range from 100 nm to 100 μm, a response time as short as a few picoseconds, and a pulsed energy range from 10 nJ to 20 J.

52.2.2 Integration Spheres

Integration spheres are designed to collect the power of highly divergent laser beams, such as semiconductor laser beams, since these beams can overfill the input window of a photodetector and cause considerable measurement error. Figure 52.6 shows the scheme of an integration sphere. The hollow spherical cavity has a diffusive internal wall and at least two windows. The reflectivity of the internal wall is high and slightly wavelength dependent. The highly divergent laser beam under measurement is incident into the sphere from one window. The photodetector of a laser power meter is mounted on another window. A baffle is used to prevent the photodetector being directly hit by the incident beam. The sphere can collect all the incident laser power and convert the power into a diffusive radiation proportional to the power. The laser power meter measures the radiation and displays the laser power under measurement based on the calibration data of the integration sphere.

Readers interested in knowing more about laser power meters can read Refs. [2–4] or contact manufacturers. Chapters 43 and 44 provide more information about photodetectors.

52.3 Measurement of Laser Spectrum

An instrument that can measure a laser spectrum is known as an optical spectrum analyzer. An optical spectrum analyzer consists of two basic components: an optical device and a laser power meter. The optical device can select and output a certain wavelength band of a polychromatic laser beam incident on the device. The central wavelength of the output wavelength band can be scanned over a certain range by the scan of the optical device. The laser power meter measures the power of the optical device output. The width of the output wavelength band usually does not change. Therefore, the power measured by the power meter is proportional to the optical spectral intensity. Since the absolute value of the optical spectral intensity is not important in the measurement of spectrum, the laser power meter sometimes can be as simple as a photodiode combined with a voltmeter. As the optical device is scanned, the power meter outputs the power (optical spectral intensity) as a function of the wavelength and thereby measures the spectrum. Diffraction gratings and scanning Fabry–Perot interferometers (SFPIs) are two widely used optical devices for spectrum analysis.

52.3.1 Light Interference

Most optical spectrum analyzers make measurements utilizing light interference. The wave theory of light can explain interference. A single-mode, narrow-linewidth laser beam can be described by its electric field $E(z)$:

$$E(z) = A\exp[-i\phi(z) + i\alpha(t)] \tag{52.2}$$

where

A is the amplitude

$\phi(z) = 2\pi z/\lambda$ is the phase, λ is the laser wavelength, z is the coordinate in the direction of beam propagation

$\alpha(t)$ is a phase factor that varies fast with time t

The intensity of the laser beam is given by

$$I = |E(z)|^2 = A^2 \tag{52.3}$$

52.3.1.1 Two-Beam Interference

A beam splitter such as a partially transparent plate can split a laser beam into two described by

$$E_1(z) = A_1\exp[-i\phi(z) + i\alpha(t)] \tag{52.4}$$

$$E_2(z) = A_2\exp[-i\phi(z) + i\alpha(t)] \tag{52.5}$$

where A_1 and A_2 are the amplitudes. Let the two beams propagate through two different distances, z_1 and z_2, respectively, and then be recombined into one. The intensity of the recombined beam is

$$I = |E_1(z_1) + E_2(z_2)|^2 = A_1^2 + A_2^2 + 2A_1A_2\cos\frac{2\pi(z_2 - z_1)}{\lambda} \tag{52.6}$$

Equation 52.6 shows that I varies sinusoidally as $|z_2 - z_1|$ is varied. Such a phenomenon is known as *two-beam interference*. When $|z_2 - z_1| = m\lambda$ (m is any integer), I takes the maximum value of $(A_1 + A_2)^2$ and the situation is called "constructive interference." When $|z_2 - z_1| = (m + 1/2)\lambda$, I takes the minimum value of $(A_1 - A_2)^2$, and the situation is called *destructive interference*. Combining two beams with the same wavelengths but from two different lasers results in

$$I = A_1^2 + A_2^2 + 2A_1A_2\cos\left[\frac{2\pi(z_2 - z_1)}{\lambda} + \alpha_2(t) - \alpha_1(t)\right] \tag{52.7}$$

where $\alpha_1(t)$ and $\alpha_2(t)$ are the two different phase factors of the two lasers, respectively. For broadband lasers, $\alpha_1(t)$ and $\alpha_2(t)$ are uncorrelated, and they vary rapidly and randomly. The last term at the right-hand side of Equation 52.7 can contribute to I only in a time-averaged way, and the time average of this term is zero. For narrowband lasers, a beat is obtained at the differential frequency of the two lasers, and I varies sinusoidally in time for any fixed $z_2 - z_1$. The time average of the last term in Equation 52.7 is also zero. Thus, I always equals a constant of $A_1^2 + A_2^2$, no matter how $|z_2 - z_1|$ is varied.

52.3.1.2 Multibeam Interference

Interference can occur among any number of beams obtained by splitting one laser beam. Let a laser beam with unit amplitude be split into $N(N > 1)$ beams, the amplitudes of these N beams fall off progressively, and the phase difference between two successive beams be a constant $\Delta\phi$. These beams can be described by $E_k(z) = \rho^k\exp[-i\phi(z) - ik\Delta\phi + i\alpha(t)]$, where $k = 0, 1, \ldots, N - 1$ and $\rho < 1$. Then, these N beams

are recombined. Multibeam interference occurs among these beams. The intensity of the recombination of these beams is given by

$$I = \left| \exp[-i\phi(z) + i\alpha(t)] \sum_{k=0}^{N-1} \rho^k \exp(-ik\Delta\phi) \right|^2 = \frac{1 + \rho^{2N} - 2\rho^N \cos(N\Delta\phi)}{1 + \rho^2 - 2\rho\cos(\Delta\phi)} \tag{52.8}$$

Equation 52.8 shows that I is a function of ρ and N and is a periodic function of $\Delta\phi$. More information about the wave theory of light and light interference can be found in many advanced optics textbooks such as Ref. [5].

52.3.2 Diffraction Gratings

A diffraction grating can spatially disperse a polychromatic laser beam into its monochromatic components and is the most widely used optical device for analyzing optical spectrum covering a relatively wide range. Several different types of diffraction gratings are available. A planar reflective diffraction grating is a collection of many small, identical, and slit-shaped grooves ruled on a planar high reflective surface; Figure 52.7 shows two grooves of such a grating, an incident laser beam with three wavelength components $\lambda_1 < \lambda_2 < \lambda_3$ and three diffraction beams marked by $m = 0$, 1, and 2, respectively. The beams are in the plane defined by the grating normal and the groove normal. The grooves illuminated by the incident beam generate diffraction beams with the same amplitude since the grooves are identical. Multibeam interference occurs among the diffraction beams. The phase difference $\Delta\phi$ between two adjacent diffraction beams of a grating is given by

$$\Delta\phi = \frac{2\pi}{\lambda} d[\sin(\theta_1) + \sin(\theta_2)] \tag{52.9}$$

where

d is the groove period

θ_1 is the angle between the incident beam and the grating normal

θ_2 is the angle between the diffraction beams and the grating normal (θ_2 is not marked in Figure 52.7)

$d[\sin(\theta_1) + \sin(\theta_2)]$ is the path difference between two adjacent diffraction beams. Constructive interference among the diffraction beams occurs at

$$\Delta\phi = 2m\pi \tag{52.10}$$

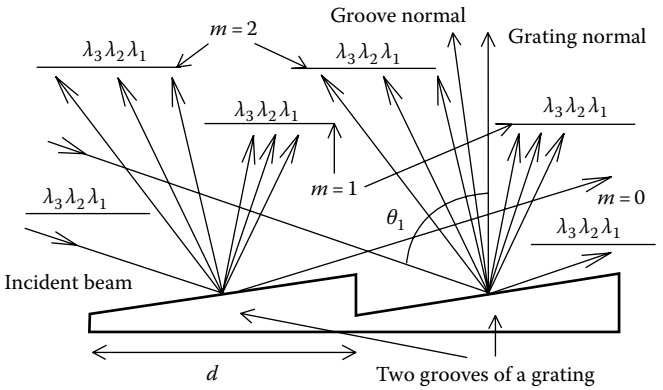


FIGURE 52.7 Two grooves of a planar reflective diffraction grating. The grating disperses an incident beam containing three wavelengths, $\lambda_1 < \lambda_2 < \lambda_3$. There are three diffraction beams with diffraction order $m = 0$, 1, 2.

where m is an integer known as the *diffraction order*. Combining Equations 52.9 and 52.10 results in

$$d[\sin(\theta_1) + \sin(\theta_2)] = m\lambda \tag{52.11}$$

Equation 52.11 is known as the *grating equation* and shows that for a given θ_1 and m, θ_2 is a function of λ. θ_2 gives the diffraction beams a propagation direction in which I has the maximum value. The angular dispersion resolution of a grating is defined by $d\theta_2/d\lambda$ and can be obtained by differentiating Equation 52.11. The result is

$$\frac{d\theta_2}{d\lambda} = \frac{m}{d\cos(\theta_2)} \tag{52.12}$$

where θ_1 is assumed to be a constant. Large $d\theta_2/d\lambda$ is often desired and can be obtained by the use of a small d, a large θ_2 close to 90°, and a large m. Usually, θ_2 must be smaller than 80° to maintain the proper functioning of the grating and $d \approx \lambda$; Equation 52.11 shows that the largest possible value of m is 2. Equations 52.11 and 52.12 also show that $m = 0$ leads to $\theta_2 = -\theta_1$ and $d\theta_2/d\lambda = 0$. That means that in the zero diffraction order, the grating does not disperse the incident beam, and all the diffraction beams propagate in the same direction, $\theta_2 = -\theta_1$. Commonly used diffraction gratings have a groove density from 300/mm ($d \approx 3.33$ μm) to 2400/mm ($d \approx 420$ nm). The corresponding dispersion resolution for $m = 1$ is from $d\theta_2/d\lambda = 4 \times 10^{-4}$ rad/nm to $d\theta_2/d\lambda = 3.4 \times 10^{-3}$ rad/nm. The intensity of the diffraction beams of a grating can also be described by Equation 52.8 with $\rho = 1$. That is,

$$I = \frac{\sin^2(N\Delta\phi/2)}{\sin^2(\Delta\phi/2)} \tag{52.13}$$

Equation 52.13 is plotted in Figure 52.8 for $N = 5$ and 20, respectively. Equation 52.13 shows that I is a periodic function of $\Delta\phi$. I reaches maximum of N^2 at $\Delta\phi = 2\,m\pi$ and zero at $\Delta\phi = 2k\pi/N$ (k is any integer, but $k \neq mN$). The fringe width $\Delta\phi_w$ of I is given by

$$\Delta\phi_w = 2\left[2m\pi - \frac{-2(mN-1)\pi}{N}\right] = \frac{4\pi}{N} \tag{52.14}$$

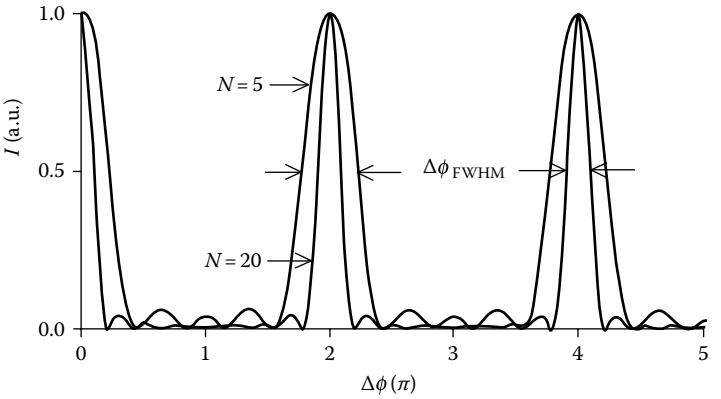


FIGURE 52.8 Multibeam interference intensity I is a periodic function of the phase difference $\Delta\phi$ between two adjacent beams. I is plotted for beam number $N = 5$ and 20, respectively. The full fringe width at half maximum power $\Delta\phi_{FWHM}$ of I decreases as N is increased.

Equation 52.14 shows that the resolution of a grating is proportional to N, since the resolution can be defined as $1/\Delta\phi_w$.

Most advanced optics textbooks (e.g., Ref. [5]) study gratings, and manufacturers' catalogs (e.g., Ref. [6]) give a product-oriented description of gratings.

52.3.3 Monochromators

Monochromators are instruments widely used for optical spectrum analysis. Various types of monochromators have been developed. Figure 52.9 shows the scheme of a simple monochromator. The laser beam under measurement is incident through the entrance slit, two concave spherical mirrors collimate the beam, the grating disperses the beam, then one concave spherical mirror focuses the beam on the exit slit, and a laser power meter measures the power of the beam passing through the exit slit. The grating is mounted on a rotator. The position of the two slits is fixed. For any given grating orientation, θ_1 and θ_2 are known. The wavelength λ of the diffraction beams passing through the exit slit can be calculated using Equation 52.11. By rotating the grating and recording the power measured as a function of the corresponding λ, we can measure the spectrum. The widths of the two slits are Δs_1 and Δs_2, respectively, and are adjustable. The three mirrors image the entrance slit on the exit plane. The image width $\Delta s_1'$ of the entrance slit is proportional to Δs_1; Equation 52.15 gives the measurement resolution of a monochromator:

$$\Delta s \frac{d\lambda}{ds} = \frac{\Delta s\, d\lambda}{f\, d\theta_2} \tag{52.15}$$

where

$d\lambda/ds$ is the linear dispersion resolution
Δs equals the larger of $\Delta s_1'$ and Δs_2
$d\lambda/d\theta_2$ is the inverse of the angular resolution of the grating used
f is the focal length of the focusing mirror

Equation 52.15 shows that reducing the widths of the two slits can increase the measurement resolution. However, the slits must be wide enough to allow enough laser power passing through for measurement. The grating rotation angular resolution also affects the measurement resolution of a monochromator. The resolution of commonly used monochromators is from 0.1 to 1 nm. Monochromators usually have a moderate price and measurement resolution. A product-oriented description of monochromators can be found in manufacturers' catalogs such as Ref. [7].

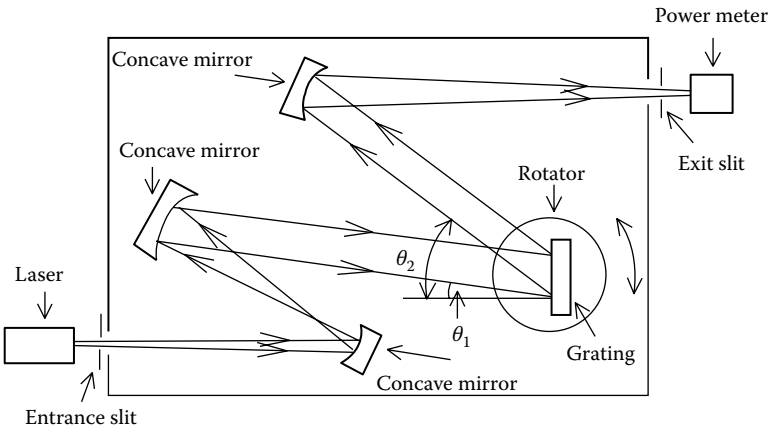


FIGURE 52.9 Scheme of a simple monochromator.

52.3.4 Scanning Fabry–Perot Interferometer

Another widely used optical spectrum analyzer is an SFPI. An SFPI consists of two slightly wedged transparent plates with flat surfaces as shown in Figure 52.10. The two inner surfaces of the plates are set parallel to each other and are high-reflecting coated. The distance D between the two inner surfaces can be adjusted by a piezoelectric device. The two outer surfaces have a small angle between them, so that reflections of the two outer surfaces cannot interfere with the reflections of the two inner surfaces. The medium between the two plates is usually air with a unit refractive index. The laser beam under measurement is collimated by a lens and incident on the SFPI. The beam transmitted through the SFPI is focused by another lens onto the photodetector of a laser power meter. When a beam with unit amplitude is incident at angle θ on the inner surface of an SFPI, multiple reflections take place at the two inner surfaces and produce a series of transmitted beams whose amplitudes fall off progressively. In Figure 52.10, only three incident rays with a few reflections are plotted. The phase difference between two successive transmitted beams is

$$\Delta\phi = \frac{4\pi D}{\lambda \cos(\theta)} \tag{52.16}$$

Interference occurs among the amplitude of the transmitted beams. The intensity of the combination of the transmitted beam is given by

$$I = \left| T \sum_{k=0}^{\infty} R^k \exp(-ik\Delta\phi) \right|^2 \tag{52.17}$$

where R and $T = 1 - R$ are the power reflectivity and transmission coefficient of the inner surfaces of the SFPI, respectively, and the reflectivity of the outer surfaces of the SFPI is neglected. The laser power meter measures the power of the transmitted beams. Note that Equation 52.8 can be reduced to Equation 52.17 by letting the amplitude $\rightarrow T$, $\rho \rightarrow R < 1$, and $N \rightarrow \infty$. Therefore, the I obtained in Equation 52.17 can also be described by the curve shown in Figure 52.8. The full width at half the maximum (FWHM) power of I can be found by solving Equation 52.17. The result is

$$\Delta\phi_{\text{FWHM}} = 4\sin^{-1}\left(\frac{1-R}{2R^{1/2}}\right) \tag{52.18}$$

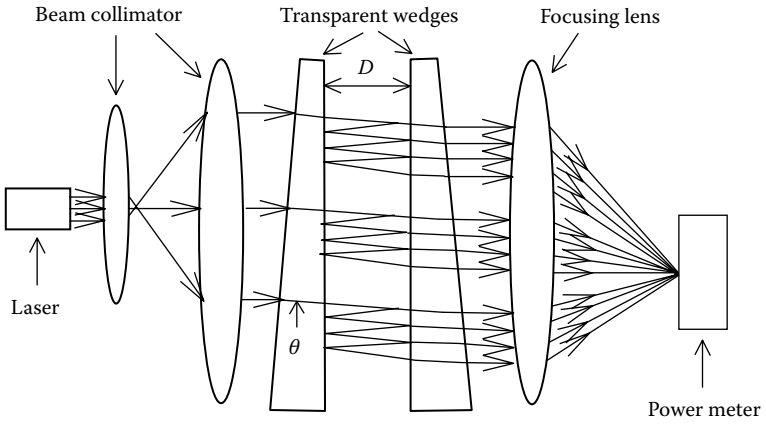


FIGURE 52.10 Scheme of an SFPI. Only three incident rays and a few reflections are plotted.

When R is close to 1, the right-hand side of Equation 52.18 is close to zero. We have

$$\Delta\phi_{\text{FWHM}} = \frac{2(1-R)}{2R^{1/2}} \tag{52.19}$$

Equation 52.19 shows that the fringe width of I reduces as R is increased, because larger R results in more reflections between the two inner surfaces of the SFPI. In an SFPI, the beam is usually arranged to incident at normal on the inner surface. θ becomes zero. Combining Equations 52.10 and 52.16 results in

$$\lambda = \frac{2D}{m} \tag{52.20}$$

where λ is the wavelength at which the peak of I appears. Figure 52.8 shows that an SFPI functions like a multibandpass filter. The central wavelengths of the bands (fringes) are given by Equation 52.20 and can be adjusted by adjusting D. The FWHM of the bands is given by Equation 52.19. The spacing between two adjacent bands in terms of wavelength can be obtained by differentiating Equation 52.20 with respect to m, eliminating m, and letting $\Delta m = 1$. The result is

$$\Delta\lambda = \frac{\lambda^2}{2D} \tag{52.21}$$

$\Delta\lambda$ is known as the free spectral range (FSR) of an SFPI. The laser beam under measurement must have a wavelength bandwidth smaller than the FSR to ensure that the beam can transmit only through one band. When D is scanned, the transmitted wavelength is scanned, and the laser power meter measures and outputs the beam spectrum. The measurement resolution of an SFPI is limited by $\Delta\phi_{\text{FWHM}}$. Most SFPIs have an $R > 0.95$ to reduce the $\Delta\phi_{\text{FWHM}}$. The characteristics of an SFPI can be described by the fringe finesse F, defined as

$$F = \frac{2\pi}{\Delta\phi_{\text{FWHM}}} = \frac{\pi R^{1/2}}{1-R} \tag{52.22}$$

where 2π is the period of the fringes given by Equation 52.10 in terms of phase. An SFPI usually has a resolution of FSR/100. For λ = 500 nm and D = 5 mm, the FSR is 0.025 nm, and the resolution is 0.00025 nm. Compared with monochromators, SFPIs are more expensive, and they have much higher resolution and a much smaller wavelength range.

More information about SFPIs can be found in many advanced optics textbooks such as Ref. [5]. A product-orientated description of Fabry–Perot interferometers can be found in manufacturers' catalogs such as Ref. [8].

52.4 Measurement of Laser Wavelength

For single-mode and narrow-linewidth lasers, the spectrum measurement reduces to wavelength measurement. Laser wavelength can be measured to a higher accuracy by the use of techniques that are not much different from those used to measure the spectrum. A laser wavemeter is an instrument designed to measure the wavelength without knowing the details of the spectrum.

52.4.1 Michelson CW Laser Wavemeter

Figure 52.11 shows the scheme of a widely used Michelson CW laser wavemeter that consists of a Michelson interferometer (MI), a laser power meter, and a computer system for data processing. The MI has two optical arms formed by a beam splitter and two mirrors, respectively. Mirror M_1 can be moved by a stepper motor driving system, and thereby the arm length z_1 can be changed. The position of mirror M_2 is fixed, and

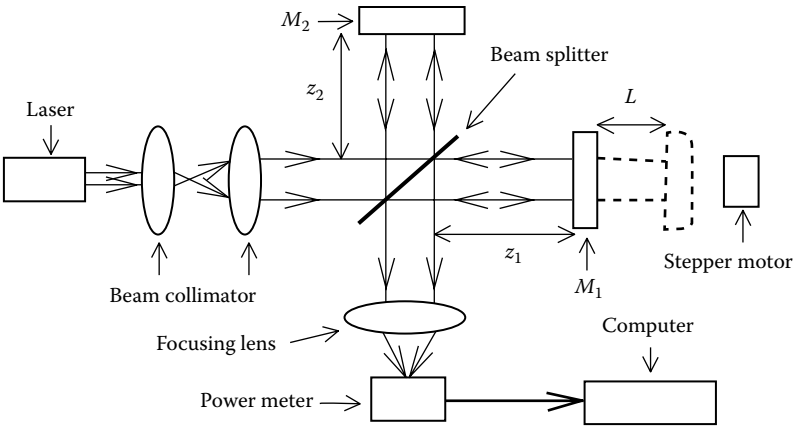


FIGURE 52.11 Scheme of a Michelson CW laser wavemeter.

the arm length z_2 cannot be changed. Two lenses collimate the laser beam under measurement. The beam splitter splits the collimated beam into two. The two beams propagate in the two arms and are reflected by the two mirrors back to the beam splitter, respectively. Then the beam splitter recombines the two beams. Interference occurs between these two beams. Another lens focuses the combined beams onto the photodetector of a laser power meter. The computer system processes the output data of the laser power meter. The intensity I of the two combined beams can be described by Equation 52.6. As M_1 is moved, z_1 is changed, the path difference $|z_2 - z_1|$ between the two optical arms varies, and I varies periodically. The computer counts the number of the varying period of I, known as *counting the fringes*. Equation 52.23 relates the wavelength λ under measurement, M_1 moving distance L, and the counted fringe number $m + \Delta m$ by

$$(m + \Delta m)\lambda = 2L \tag{52.23}$$

where
- m is an integer
- Δm is a fraction
- the factor of 2 is introduced because the beam round-trip distance is considered

A He–Ne laser with accurately known wavelength λ_H is used as a calibration source. For the He–Ne laser, there is the relation

$$(m_H + \Delta m_H) = 2L \tag{52.24}$$

where
- m_H is another integer
- Δm_H is another fraction

Combining Equations 52.23 and 52.24 results in

$$\lambda = \frac{m_H + \Delta m_H}{m + \Delta m}\lambda_H \tag{52.25}$$

For $L = 500$ mm and $\lambda \approx 500$ nm, Equations 52.23 and 52.24 give $m \approx 2 \times 10^6$ and $m_H \approx 2 \times 10^6$. If Δm and Δm_H can be counted to an accuracy of 0.1, Equation 52.25 can provide six significant digits. Thus, λ can be calculated to a relative accuracy of about 10^{-6}. The commonly used commercial Michelson CW laser wavemeters have a relative measurement accuracy from 10^{-4} to 10^{-7}. The measurement range is usually from 400 nm to 1.1 μm, limited by the spectral range of the silicon photodetector used.

Steadily moving M_1 over a distance of 500 mm or so can take several seconds, and the computer is counting fringes during the entire moving period of M_1. Several seconds is much longer than the pulse duration of most pulsed lasers. Thus, the computer will miss fringes in the period of time between two successive pulses when measuring the wavelength of a pulsed laser, and the measurement result will be erroneous. To measure the wavelength of pulsed lasers, the laser wavemeter must not have any moving parts and must be capable of taking data instantaneously.

More information about MIs can be found in many advanced optics textbooks such as Ref. [5]. Reference [9] presents a comprehensive study about the design and development of a MI CW laser wavemeter that achieves an accuracy of a few parts in 10^9.

52.4.2 Pulsed Laser Wavemeter

Figure 52.12 shows the scheme of a Fizeau pulsed laser wavemeter, which consists of a Fizeau wedge, a charge-coupled device (CCD) linear sensing array with 1024 pixels, a computer system for data processing, and an optical collimation system. The Fizeau wedge is made of a transparent material such as glass. The wedge has two flat surfaces. The angle α between the two surfaces is a few milliradians. The optical thickness $l(x)$ of the wedge is about 2 mm and varies linearly along the wedge. The optical system collimates the laser beam under measurement. The collimated beam is incident on the wedge. Both surfaces of the wedge reflect the incident beam. The two reflected beams form tens of spatial interference fringes on the CCD array. The CCD array detects the interference fringes and sends the data to the computer for processing. A Fizeau wavemeter does not have any moving parts and can measure the wavelength of pulsed lasers. It can be shown that the fringe period p is proportional to the wavelength λ under measurement by

$$p = F_1(\alpha)\lambda \tag{52.26}$$

where $F_1(\alpha)$ is a function of α. A He–Ne laser with accurately known wavelength λ_H is used separately as a calibration source. The He–Ne laser also forms spatial interference fringes on the CCD array with a fringe period p_H proportional to λ_H by

$$p_H = F_1(\alpha)\lambda_H \tag{52.27}$$

The data processing algorithm used in Fizeau wavemeters has two steps. λ is first calculated by combining Equations 52.26 and 52.27:

$$\lambda_1 = p\frac{\lambda_H}{p_H} \tag{52.28}$$

In Equation 52.28, the symbol λ_1 denotes the wavelength calculated in the first step because λ_1 is still not accurate enough. The data processing algorithm can calculate the fringe period p and p_H to an accuracy of

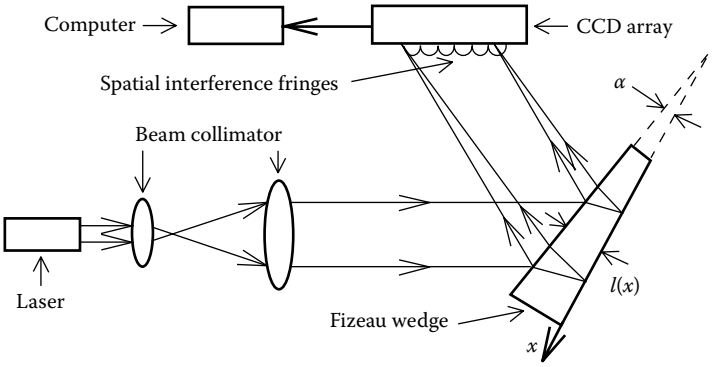


FIGURE 52.12 Scheme of a Fizeau pulsed laser wavemeter.

higher than $10^{-4}p$ and $10^{-4}p_H$ utilizing the 1024 sensing data provided by the CCD array. Thus, Equation 52.28 can provide λ_1 to an accuracy higher than $10^{-4}\lambda_1$, and λ must fall somewhere inside the range

$$\lambda_1 - 10^{-4}\lambda_1 < \lambda < \lambda_1 + 10^{-4}\lambda_1 \tag{52.29}$$

Any point on the CCD array corresponds to a point on the wedge with a certain optical thickness $l(x)$. It can be shown that at a given point on the CCD array, there is such a relation for λ and λ_H that

$$(m + \Delta m)\lambda = F_2[l(x)] \tag{52.30}$$

$$(m_H + \Delta m_H)\lambda_H = F_2[l(x)] \tag{52.31}$$

where

m and m_H are two integers and are unknown
Δm and Δm_H are two fraction orders at this point on the CCD array and can be measured
$F_2[l(x)]$ is a function of $l(x)$
$l(x)$ is the thickness of the wedge at the point corresponding to the point on the CCD array

In the second step of the data processing, λ is calculated by combining Equations 52.30 and 52.31. The result is

$$\lambda_2 = \frac{m_H + \Delta m_H}{m + \Delta m}\lambda_H \tag{52.32}$$

In Equation 52.32, we use the symbol λ_2 to denote the wavelength calculated, because λ_2 is still not necessarily equal to λ. A number of test m and m_H values are inserted into Equation 52.32 and result in a number of different λ_2 values. It can be shown that only one combination of m and m_H can result in a λ_2 that falls inside the range of Equation 52.29. This λ_2 is accepted as λ. m and m_H are of the order of 10^5, and Δm and Δm_H can be measured to an accuracy of 0.01. Therefore, Equation 52.32 can provide Δ to an accuracy of 10^{-6}. The commercial Fizeau pulsed wavemeters have a measurement accuracy of $10^{-5}\,\lambda$. The measurement spectral range is from 400 nm to 1.1 μm, limited by the CCD array, which is made of silicon material. More information about Fizeau wedges can be found in many advanced optics textbooks such as Ref. [5]. Readers interested in learning more about the design and development of Fizeau pulsed wavemeters will find Ref. [10] to be a good starting point.

52.5 Instrumentation and Components

Table 52.1 lists a few companies manufacturing photodetectors, laser power meters, diffraction gratings, monochromators, SFPIs, Michelson CW laser wavemeters, and Fizeau pulsed laser wavemeters. The price ranges of these instruments and components are also listed in Table 52.1. Table 52.2 lists the

TABLE 52.1 Manufacturers, Components and Instruments, and Price Ranges

Product Description	Manufacturer
Thermopile for laser power meters	Coherent, Newport
Photodiode head for laser power meters	Coherent, Newport
Pyroelectric probe head for laser power meters	Coherent, Newport
Laser power meter (including one detector head)	Coherent, Newport
Integration sphere	Newport, Oriel
Diffraction grating	Milton Roy
Monochromator	Milton Roy, Oriel
Scanning Fabry–Perot interferometer	Burleigh Instruments
Michelson CW laser wavemeter	Burleigh Instruments
Fizeau pulsed laser wavemeter	New Focus

TABLE 52.2 Addresses of Companies Listed in Table 52.1

Burleigh Instruments, Inc. Burleigh Fishers, NY 14453 Tel: (716) 924-9355	**Coherent, Inc.** Photonics Division 2303 Lindbergh St. Auburn, CA 95602 Tel: (530) 889-5365	**Milton Roy Instruments** 820 Linden Ave. Rochester, NY 14625 Tel: (716) 248-4000
New Focus, Inc. 2630 Walsh Ave. Santa Clara, CA 95051 Tel: (408) 980-8088	**Newport Corp.** 1791 Deere Ave. Irvine, CA 92714 Tel: (800) 222-6440 Tel: (203) 377-8282	**Oriel Corp.** P.O. Box 872, 250 Long Beach Blvd. Stratford, CT 06497

address of these companies. These two tables are far from exhaustive. Interested readers could consult two excellent books, *Laser Focus World Buyers Guide* [11] and *Photonics Buyers' Guide* [12]. These two books are published annually and contain up-to-date information about most optical and laser manufacturers in the United States and their products.

Defining Terms

Continuous-wave (CW) laser power: Laser power that does not vary with time.

Laser spectrum: Laser power–wavelength profile.

Laser wavelength: The spatial period of a laser electric field.

Pulsed laser power: Laser power that lasts only a short period of time.

References

1. J. Hecht, *The Laser Guidebook*, 2nd edn., New York: McGraw-Hill, Inc., 1992.
2. R. DeMeis, Choose the right detector and meter to test beam strength, *Laser Focus World*, 31(6): 105–113, 1995.
3. W. DeCosta, Power and energy meters meet measurement needs, *Laser Focus World*, 31(5): 199–203, 1995.
4. G. Shelmire, How to make accurate laser output measurements, *Laser Focus World*, 29(4): 241–248, 1993.
5. M. Born and E. Wolf, *Principles of Optics*, 6th edn., Oxford, U.K.: Pergamon Press, 1980.
6. C. Palmer (ed.), *Diffraction Grating Handbook*, 2nd edn., Rochester, NY: Milton Roy Instruments, 1993.
7. Laser Diode Modules, Resources, Product catalog, Oriel Corp., Stratford, CT. http://search.newport.com/?q=Fabryperot, accessed on August 21, 2013.
8. Monochromators and Spectrometers, Overview of all Products, McPherson, http://www.mcphersoninc.com/productlines.htm, accessed on August 21, 2013.
9. J. Monchalin, M. Kelly, J. Thomas, N. Kurnit, A. Szoeke, F. Zernike, P. Lee, and A. Javan, Accurate laser wavelength measurement with a precision two-beam scanning Michelson interferometer, *Appl. Opt.*, 20: 736–757, 1981.
10. B. Faust and L. Klynning, Low-cost wavemeter with a solid Fizeau interferometer and fiber-optic input, *Appl. Opt.*, 30: 5254–5259, 1991.
11. Anonymous, *Laser Focus World Buyers Guide*, Nashua, NH: PennWell Publishing Co., 1996.
12. Anonymous, *Photonics Buyers' Guide*, Pittsfield, MA: Laurin Publishing Co., 1996.

53

Optical Loss: Principles and Applications

M.M. Rad
Ciena

Halit Eren
Curtin University

Martin Maier
Institut National de la Recherche Scientifique

53.1 Introduction

Fiber optics is a fundamental part of most computer and communications networks and devices. This chapter focuses on the principle of fiber-optics for a variety of applications mainly in telecommunications and fiber-optic sensors. Starting from the very basic concepts of optics and photonic systems, the principles of the loss mechanism in fibers are explained in detail including extrinsic (bending, connector, and launching) losses and intrinsic (absorption and scattering) losses. The loss mechanisms of fiber-optics are explained in detail for sensor applications and traditional optical time domain reflectometry (OTDR) methods. Standard attenuation test techniques are also explained. Finally, this chapter provides useful references and list of the manufacturers (Table 53.1) for fiber-optics equipment and devices.

53.2 Basic Concepts

53.2.1 Light and Optics

Light or visible light is the portion of electromagnetic spectrum that is visible to the human eye. Visible light corresponds to the wavelength range from 390 to 750 nm. However, in physics, the term light often comprises the two adjacent regions of infrared (higher wavelengths) and ultraviolet (lower wavelengths), not visible to the human eye. Primary properties of light are intensity, propagation direction, wavelength (frequency), polarization, and phase. The light speed in vacuum is however one of the fundamental constants of nature and is about 300,000 km/s.

TABLE 53.1 List of Manufacturers

AxSys Communications
P.O. Box 571
Danielson, CT 06239-0571
Tel: (203) 774-4102
Fax: (203) 774-4783

Cuda Products Corp.
6000-T Power Ave.
Jacksonville, FL 32217-2279
Tel: (904)737-7611
Fax: (904) 733-4832

Fiber Options
80-T Orville Dr.
Bohemia, NY 11716-2533
Tel: (800) 739-9105, (516) 567-8320
Fax: (516) 567-8322

PHILTEC Inc.
P.O. Box 359
Arnold, MD 21012
Tel: (410) 757-4404
Fax: (410) 757-8138

EXFO Inc.
400 Godin Ave.
Québec City, Québec G1M 2K2, Canada
Tel: (418) 683-0211
Fax: (418) 681-3936

JDSU
430 N. McCarthy Blvd.
Milpitas, CA 95035
Tel:1 408 546 5000
Fax: 1 408 546 4300

OZ Optics
219 Westbrook Road,
Ottawa, Ontario K0A 1L0, Canada
Tel: 1-613-831-0981
Fax: 1-800-361-5415

Corning Inc.
One Riverfront Plaza
Corning, NY 14831
Tel: (607) 796-0624
Fax: (607) 796-0147

Sycamore Networks
100 Century Parkway
Suite 120, Mt. Laurel, NJ 08054
Tel: 856-359-9301
Fax: 856-359-9302

Chiu Technical Corp.
252 Indian Head Rd.
Kings Park, NY 11754
Tel: (516) 544-0606
Fax: (516) 544-0809

Fiberoptics Technology Inc.
13 Fiber Rd.
Pomfret, CT 06258
Tel: (800) 433-5248, (203) 928-0443
Fax: (800) 543-2558, (203) 928-7664

Incom Inc.
P.O. Drawer G.
Southbridge, MA 01550-0528
Tel: (508) 765-9151
Fax: (508) 765-0041

Vicon Fiber Optics Corp.
90 Secor Lane
Pelham Manor, NY 10803
Tel: (800) 828-2071, (914) 738-5006
Fax: (914) 738–6920

TeraXion Inc.
2716 Einstein St,
Québec City, Québec, Canada, G1P 4S8
Tel: (418) 658-9500
Fax: (418) 658-9595

Lee Laser Inc.
7605 Presidents Drive
Orlando, FL 32809
Tel: (407) 812-4611
Fax: (407) 850-2422

Avago Technologies
1350 West Trimble Road,
San Jose, CA 95131
Tel: (408) 435-7400
Fax: (408) 435-4172

Finisar Inc.
1389 Moffett Park Dr Sunnyvale
CA 94089
Tel: (408) 548-1000
Fax: (408) 541-6138

Optical Cable Corp.
5290 Concourse Drive,
Roanoke, VA 24019
Tel: (540) 265-0690
Fax: (540) 265-0724

TABLE 53.1 (continued) List of Manufacturers

Oclaro Inc. 2560 Junction Ave. San Jose, CA 95134 Tel: (408) 383-1400 Fax: (408) 919-1501	**International Fibercom Inc.** 3230 East Broadway Suite 200 Phoenix, AZ 85034 Tel: (602) 387-4000 Fax: (602) 276-0567
IPG Photonics Corp. 50 Old Webster Road Oxford, MA 01540 Tel: (877) 980-1550 Fax: (508) 373-1103	**Alliance Fiber Optic Products Inc.** Sales, 275 Gibraltar Drive Sunnyvale, CA 94089 Tel: (408) 736-6900 Fax: (408) 736-4882
MPB Communication Inc. 147 Hymus Blvd, Pointe-Claire Montréal, Québec H9R 1E9, Canada Tel: (514) 694-8751 Fax: (514) 694-6869	

The branch of physics, which involves the behavior and properties of light, is known as optics. The science of optics studies the interaction of light with matter and explores the construction of instruments that use or detect the related phenomena. The aforementioned properties of light are manipulated and detected using the measurement and instruments for a particular purpose. Applications vary from photographic imaging to high-speed data transmission via fibers. Depending on the application, a different optical signal processing scheme may be utilized to expand, condense, collimate, reflect, polarize, filter, diffuse, absorb, refract, or scatter the generated light. The desired information modulated on the light is detected by the appropriate instrument. Practical applications of optics are found in a variety of technologies including mirrors, lenses, telescopes, microscopes, lasers, fiber optics (telecommunication and sensors), and medicine.

53.2.2 Fiber Optics

During the past three decades, fiber optics has attracted considerable attention as the light transmission media. They are now widely used in telecommunication applications and enable us to transmit at longer distances and with higher bandwidth (data rates) compared to other forms of telecommunication. Fibers are also used for a variety of other applications, including sensors and fiber lasers.

In principle, an optical fiber acts as a waveguide or "light pipe" for the light propagating between the two ends of the fiber. In fact, a fiber-optic system is similar to the copper wire system that fiber optics is replacing. The difference is that fiber optics uses light pulses to transmit information down fiber lines instead of using electronic pulses inside the copper lines. It is a circular dielectric glass* waveguide, which consists of a core (central glass) surrounded by a concentric cladding material with a slightly lower (≈1%) refractive index. Because of *total internal reflection,* the light is confined to the core if the angular condition of total internal reflection is met. For this to happen, not leaking out, the light must strike the boundary with an angle greater than the critical angle $\theta_{c,}$ which occurs only when light enters the fiber within a certain range of angles $\phi_{max.}$ Hence, the acceptance and transmission of light depends greatly on the angle at which the light enters the fiber. The angle must be less than the *critical acceptance angle* of the particular fiber being used, as shown in Figure 53.1. This angle depends on the difference of the refractive index of the core and cladding. Light then travels along the fiber bouncing back and forth of the boundary.

* Some other polymers are also used.

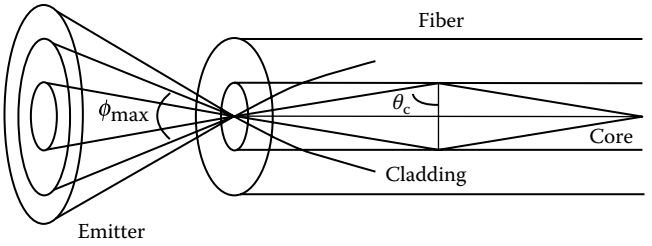


FIGURE 53.1 An optical fiber is a circular dielectric glass waveguide that can efficiently transport optical energy. It consists of a central glass core surrounded by a concentric cladding material. An important requirement for the connection of a light source is that sufficient amount of useful light must be coupled into the fiber. The core has a higher index of refraction than the cladding. Therefore, light is confined to the core only if the angular condition of total internal reflection is met. The acceptance and transmission of light depends greatly upon the angle at which the light rays enter the fiber, and it must be less than the *critical acceptance angle* of a particular fiber.

Theoretically, light propagation in fibers can be treated in a number of ways. For example, in treating propagation by modes, the fibers are viewed as optical waveguides (such as single-mode fibers), whereas treatment by rays is an approximate description of fibers with diameters much greater than the wavelength (such as multimode fibers).

There are three basic types of fibers: single-mode step index, multimode graded index, and multimode step index, as shown in Figure 53.2a through c, respectively. The characteristics of optical losses in these three types of fibers vary slightly due to differences in construction and the nature of propagation of light. For example, in the case of multimode step-index fibers, light striking the core-cladding junction at an angle greater than the angle of internal reflection passes through and becomes absorbed by the opaque jacket. This represents a significant source of attenuation, limiting the injection efficiency at the transmitting end.

Single-mode fibers, shown in Figure 53.2a, are used in transmitting broadband signals over large distances. By choosing a very small core diameter, these fibers allow only one mode of propagation through the fiber. Their attenuation is generally very small and their transmission band is large. Owing to material properties, low attenuations (~0.25 dB/km) can be expected for wavelengths around 1.3–1.6 μm. Additional attenuation arises from splices and fiber bending. The most common type of single-mode fiber has a core diameter of 8–10 μm and is designed for use in the near infrared (1.5 μm). The mode structure depends on the wavelength of the light used, so that this type of fiber actually supports a small number of additional modes at visible wavelengths. Single-mode fiber provides a higher transmission rate and up to 50 times more distance than multimode, but it also costs more. Single-mode fiber has a much smaller core than multimode fiber and virtually eliminates any distortion that could result from overlapping light pulses, thus providing the least signal attenuation and the highest transmission speeds of any fiber cable type. Single-mode fiber is used in many applications where data are sent at multifrequency (wavelength division multiplexing [WDM]), so only one cable is needed (single-mode on one single fiber).

Multimode graded-index fibers have medium size cores (50–100 μm) and refractive indices that decrease radially outward. The two optical materials with different refractive indices are mixed together in such a way that the index of refraction decreases smoothly with distance from the fiber axis. The graded index causes the light to gradually bend back and forth across the axis in sinusoidal manner when very small injection angles are used. Hence, no abrupt reflection from the core-cladding boundary occurs. This greatly reduces the light losses from the fiber and results in better bandwidth and power efficiency. Indeed, the index profile is chosen to minimize the difference in axial propagation speeds of the various rays in the fiber. Therefore, the graded-index fibers reduce the total dispersion experienced by light wave signals compared to step-index versions. They also allow the use of simpler splicing techniques. Light waves are dispersed into numerous paths, or modes, as they travel through the cable's core typically at 850 or 1300 nm. Typical multimode fiber core diameters are 50, 62.5, and 100 μm. In the United States, the most common size is 62.5 μm.

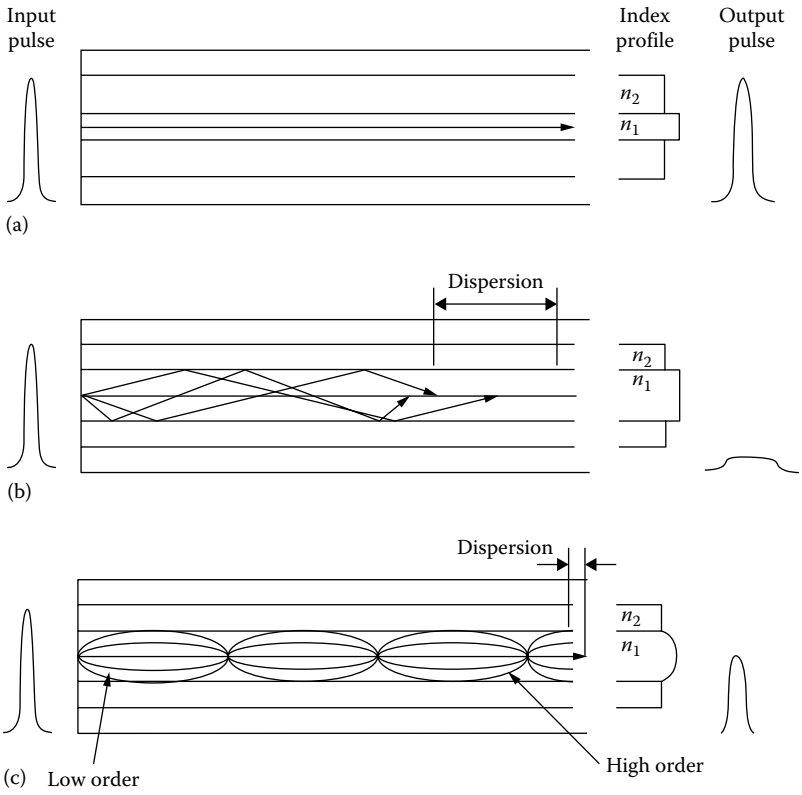


FIGURE 53.2 Three basic types of fibers: (a) single-mode step index that is used in transmitting broadband signals over large distances, (b) multimode step index in which refractive index changes between fiber and cladding and rays striking the core-cladding junction reflect back in the glass, and (c) multimode graded index, obtained by mixing two optical materials with different refractive indices together in such a way that the index of refraction decreases smoothly with distance from the fiber axis.

Special purpose fibers are another type of fibers usually made for a specific application. They are usually fabricated with noncylindrical core and/or cladding layer, usually with an elliptical or rectangular cross section. The most important examples of these fibers are polarization maintaining fibers, photonic crystal fibers, and plastic optical fibers (POFs).

There are two basic cable designs for fiber optics called loose-tube cables and tight-buffered cables. While the former is used in the majority of outside-plant installation such as in aerial, duct, and direct-buried applications, the latter is utilized for inside buildings. The modular buffer-tube design permits easy drop-off of groups of fibers at intermediate points, without interfering with other protected buffer tubes being routed to other locations. The loose-tube design also helps in the identification and administration of fibers in the system. Both single- and multifiber tight-buffered cables are available. Single-fiber tight-buffered cables are used as pigtails, patch cords, and jumpers to terminate loose-tube cables directly into optoelectronic transmitters, receivers, and other active and passive components. Multifiber tight-buffered cables are also available and are used primarily for alternative routing and handling flexibility and ease within buildings. There are a variety of fiber cable types such as distribution cable, indoor/outdoor tight buffer, indoor/outdoor breakout, aerial cables, hybrid and composite cables, armored cables, low-smoke zero halogen (LSZH), corning cable systems, freedom LTS cables, and Kron indoor/outdoor dry loose-tube cables. Depending on the application and the role of the fiber, different types of fibers are employed in an optical system.

Glass optical fibers are usually made from silica. Silica exhibits good transmission over a wide wavelength range and has extremely low loss near the infrared regime (~1.5 μm). In addition, for silica fibers

cleaving and fusion splicing is relatively effective. They also have high mechanical strength against pulling and bending. Notwithstanding that, other materials such as fluorozirconate, fluoroaluminate, and chalcogenide glasses and crystalline materials like sapphire are sometimes used for longer-wavelength infrared or other specialized applications. While bandwidth is the primary consideration in the use of fiber optics in communication applications, light attenuation characteristics are equally important. The overall quality of a fiber-optic light guide is determined by its light transmissivity, defined as the ratio of the output light to that put into the fiber.

53.2.3 Light Sources for Fiber Optics

The most commonly used optical transmitters are semiconductor devices such as light-emitting diodes (LEDs) and laser diodes. The difference between LEDs and laser diodes is that LEDs produce incoherent light, while laser diodes produce coherent light. The emitted light is incoherent with a relatively wide spectral width of 30–60 nm. LED light transmission is also inefficient, with only about 1% of input power, or about 100 μW, eventually converted into launched power, which has been coupled into the optical fiber. However, due to their relatively simple design, LEDs are very useful for low-cost applications. LEDs are suitable primarily for local-area-network applications with bit rates of 10–100 Mbit/s and transmission distances of a few kilometers. They are currently in use for local-area WDM networks.

A semiconductor laser emits light through stimulated emission rather than spontaneous emission, which results in high output power (~100 mW) as well as other benefits related to the nature of coherent light. The output of a laser is relatively directional, allowing high coupling efficiency (~50%) into single-mode fiber. The narrow spectral width also allows for high bit rates since it reduces the effect of chromatic dispersion. Furthermore, semiconductor lasers can be modulated directly at high frequencies. Laser diodes are often directly modulated, that is, the light output is controlled by a current applied directly to the device. For very high data rates or very long-distance links, a laser source may be operated continuous wave in an external modulation scheme. The choice of wavelength is also determined by the transmission characteristics of optical fibers.

53.3 Losses in Fiber Optics

Attenuation in fiber optics, also known as transmission loss, is the reduction in intensity of the light with respect to distance as a result of propagation through the fiber optic. By definition, signal attenuation is defined as the ratio of optical input power to the optical output power. Attenuation coefficients in fiber optics expressed in units of dB/km are an important factor, which limits the data transmission across large distances. It was shown that attenuation in optical fiber is caused primarily by both scattering and absorption. Scattering, also known as diffuse reflection, is a consequence of the reflection of light ray in random directions due to the total internal reflection at molecular level in the glass structure of the fiber core. These sources of fiber loss, also known as intrinsic fiber loss, are discussed in more detail in the next section.

The interpretation and treatment of "optical loss" is different for each individual instrumentation and measurement activity. It entirely depends on the purpose and method of measurements. For example, scattering of the light (radiation) from clouds is a useful property for determining atmospheric characteristics, whereas scattering of light in optical fibers may not be desirable due to resulting losses in power and decrease in efficiency. While many instrumentation systems (e.g., imaging) make use of scattering, absorption, refraction, and reflection as useful properties, in others mainly telecommunications, these are considered to be mere losses that cause undesirable attenuation.

This chapter concentrates on losses in fiber optics to clearly demonstrate the fundamental principles. Nevertheless, it is important to recognize that the same losses in fiber optics may not be regarded as losses in other applications. It will be shown here that losses in fiber optics have useful properties in determining the optical fiber characteristics.

53.3.1 Loss Mechanisms in Optical Fibers

Although due to the total internal reflection principle an optical signal is confined into the fiber core (i.e., no loss exists for the core-cladding interface), the light experiences attenuation as it travels through the fiber. This attenuation mechanism in fiber can be categorized into extrinsic and intrinsic losses. Intrinsic loss is related to the attenuation mechanisms occurring within the fiber core material only, whereas extrinsic attenuation is due to nonideal modifications of the fiber or external damages, which consequently change the behavior of the fiber-optic media. These mechanisms are explained in detail in the following subsections.

53.3.2 Extrinsic Fiber Losses

Extrinsic fiber losses are generated from sources outside the optical fiber core, which finally affect the transmission of light down the fiber. These losses are not functions of the fiber core materials; they are specific to geometry and handling problems. There are three basic types of extrinsic losses, namely, bending losses, launching losses, and connector losses.

53.3.2.1 Bending Losses

Bending losses are the result of distortion of the fiber cable from the ideal configuration and are classified according to the radius of curvature of the bending: microbend loss versus macrobend loss. The bending loss depends on the radius of the curvature; the smaller the radius of the curvature is, the greater is the bending loss. For instance, kinks or microbends are an important source of attenuation and generally impose higher bending loss compared to macrobends.

Microbends are generated by microscopic discontinuities and imperfections as a result of deformation or damage in the fiber structure. As exemplified in Figure 53.3, irregularly distributed undulations in the fiber with radius of curvature of a few millimeters and deviations from the mean line of a few micrometers can cause microbends. Imperfect coating and improper cabling procedures increase the loss of microbends. Light wave signals experience higher losses because low-order modes become coupled with high-order modes that are naturally more lossy. Microbending is usually distributed along the fiber.

Microbending is a critical feature of a fiber and can cause significant transmission loss. External forces (such as mechanical forces, stress, or crushing forces) or environmental factors (like temperature and pressure) can damage the smooth cylindrical surface of fiber or the cable jacket surrounding the fiber, hence affecting the quality of the fiber as a transmission medium. These damages can be caused by

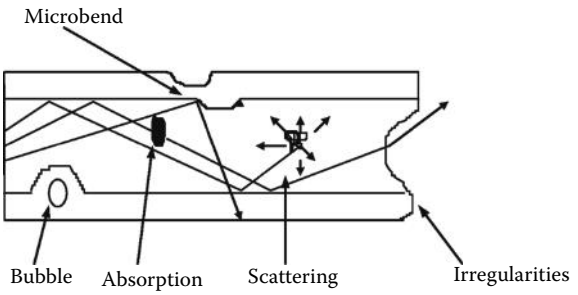


FIGURE 53.3 An important cause of attenuation is due to microbending of the fiber. This is due to irregularly distributed undulations in the fiber and from mechanical tensile forces when fiber is pressed against a rough surface. Absorption losses are caused by the presence of impurities such as traces of metal ions (e.g., Cu^{2+}, Fe^{3+}) and hydroxyl (OH^-) ions. Despite the careful manufacturing techniques, fibers may have inhomogeneous, disordered, amorphous structures. Power losses due to scattering are caused by such imperfections in the core material and irregularities between the junction and cladding.

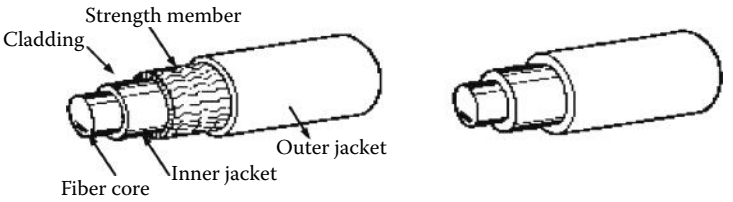


FIGURE 53.4 Some of the fiber losses may be avoided by careful cable constructions, eliminating excessive mechanical forces, and controlling the temperature variations of the cable assembly. This is achieved by a loose encasing of the fiber in a plastic sheath or by covering of the fiber with soft flexible materials. Most optical fibers constructed for communication purposes have inner and outer jackets for protection and strength.

poor or incorrect fiber processing during cabling, low temperature, and high pressure. It is interesting to mention that while the fiber may appear normal under standard conditions, improper handling and incorrect installation can severely cause microbending losses throughout the fiber. These losses may be avoided by careful cable constructions, protecting from excessive mechanical forces, and controlling the temperature variations of the fiber. Often this is achieved by a loose encasing of the fiber in a plastic sheath or by covering the fiber with soft flexible material, as shown in Figure 53.4.

Macrobend losses are observed when a fiber bend's radius of curvature is large compared to the fiber diameter. These bends become a great source of loss when the radius of curvature is less than several centimeters. In this situation, the optical path traveled by light is shorter at the inner side of the bend than the outer side. Thus, light can be lost when the angle of incidence (equivalently the fiber bend radius) exceeds the critical angle. In other words, part of the light is transformed to higher-order modes that are more lossy and radiates out of the fiber. This is more problematic for fibers with low numerical apertures, that is, larger critical angles. Usually improper cable handling, such as poor reeling and poor rolling, creates macrobending losses during network installation. Rather than microbending, macrobending is localized and the bend might not be simply identified on troubleshooting.

The fiber sensitivity to macrobending losses can be reduced by increasing the refractive index of the fiber core and/or increasing the overall fiber diameter. However, increases in the fiber core diameter increase the number of propagating modes in the fiber, which are more lossy leading to more sensitive fibers. Generally, extrinsic losses are less important for fibers with high numerical apertures. Bending losses also depend on the wavelength of light and generally are higher for longer wavelength; therefore, longer wavelengths in U band (1625–1675 nm) have been recommended by ITU-T for the fiber link monitoring application.

53.3.2.2 Launching Losses

Since the maximum transmission power is limited due to practical constraints (laser peak power and nonlinearities), it is generally desirable to maximize the power utilization generated by the light source connected to a fiber. In practice, more than half of the generated power by light sources may be lost at the interface between a laser diode and a single-mode optical fiber.

The term *launching loss* refers to an optical fiber not being able to propagate all the incoming light rays from an optical source. These occur during the process of coupling light into the fiber (e.g., losses at the interface stages). Rays launched outside the angle of acceptance excite only dissipative radiation modes in the fiber. In general, elaborate techniques have been developed to realize efficient coupling between the light source and the fiber, mainly achieved by means of condensers and focusing lenses. The focused input light beam needs to be carefully matched to the fiber for an efficient coupling.

Similarly, once the light is transmitted through the fiber, output fiber characteristics must also match the output target characteristics to be able to couple the largest proportion of the transmitted light. This can be done by suitably focusing lens arrangements at the output. There are also initial face (Fresnel)

losses due to reflections at the entrance aperture. The Fresnel losses are greater if the fiber/source is air coupled. Hence, most optical couplings to a fiber utilize index-matching materials to substantially reduce coupling loss.

53.3.2.3 Connector Losses

A fiber-optic cable may be terminated with connectors creating a temporary connection between two fibers and/or a fiber with a network component. However, splicing creates a permanent joint between the two fibers. Fiber terminations are of great importance because of their insertion loss; hence, they should be correctly installed and be protected against dirt and damage for a correctly selected connector. Ever since the introduction of fiber optics, more than 80 types of connector have been manufactured and tens of installation schemes have been proposed. A fiber-optic connector is basically a rigid cylindrical barrel surrounded by a sleeve that holds the barrel in its mating socket.

Significant insertion loss may arise in fiber connectors and splices of the cores of the joined fibers having unequal diameters or misaligned centers, or if their axes are tilted. As shown in Figure 53.5, mismatching of fiber diameters causes losses that can be approximated by $-10 \log(d/D)$. There are other connection losses such as offsets or tilts or air gaps between fibers and poor surface finishes. Some of these are illustrated in Figure 53.5. However, to take full advantage of fiber characteristics in transmission systems, the contribution of losses from other sources must be minimized. The attenuation $a_s(d)$ due to coupling efficiency may be expressed as

$$a_s(d) = -10 \text{ dB} \log \eta(d) \tag{53.1}$$

where $\eta(d)$ is the coupling coefficient.

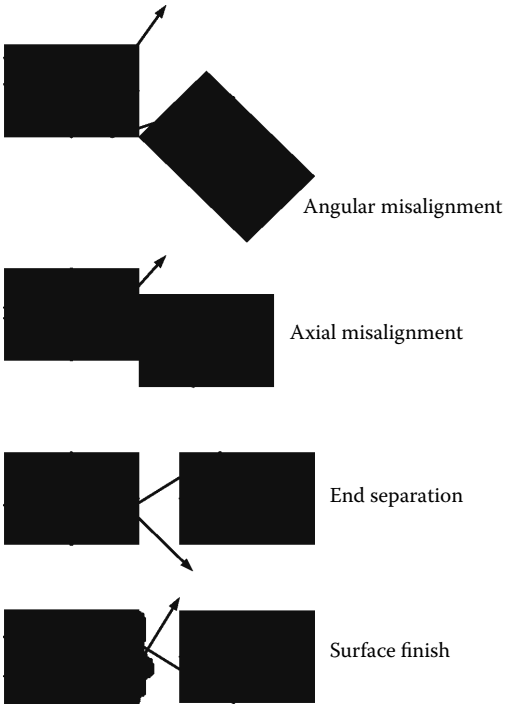


FIGURE 53.5 Some of the fiber losses may be avoided by careful cable constructions, eliminating excessive mechanical forces, and controlling the temperature variations of the cable assembly. This is achieved by a loose encasing of the fiber in a plastic sheath or by covering the fiber with soft flexible materials. Most optical fibers constructed for communication purposes have inner and outer jackets for protection and strength.

Different mating schemes such as "push and click," "turn and latch" ("bayonet"), or screw-in (threaded) can be used for connector implementation. Also depending on the type of fiber and application, various polishing profiles are used. For instance, for single-mode fiber, fiber ends are polished with a slight curvature, which is known as "physical contact" (PC) connection. However, polishing at an angle is called "angled physical contact" (APC) and has higher insertion loss compared to PC connections but suffers less from back reflection problems. Consequently, various connector different have been developed in different standard forms such as ferrule connector (or fiber channel [FC]), subscriber connector (square connector or standard connector [SC], straight tip [ST], lucent connector [LC]), or mechanical transfer registered jack (or media termination recommended jack [MTRJ]).

The loss leaking out of the angled surface is also known as gap loss. End gaps cause both insertion loss and return loss. The light emerging out of the connector spills over the receiving fiber core and is partially lost. In addition, the air gap between the fibers causes a reflection when the light faces the refractive index difference from the glass fiber to the air in the gap. This reflection (also known as the Fresnel reflection or back reflection) accounts for about 5% in typical flat polished connectors. In other words, no connector with an air gap can have less than 0.3 dB loss. Back reflection or simply optical return loss might be problematic in laser-based systems. To mitigate this problem, connectors use a number of polishing techniques to insure PC of the fiber ends and minimize back reflection. On mechanical splices, it is possible to reduce back reflection by using nonperpendicular cleaves, which cause back reflections to be absorbed in the cladding of the fiber.

Gap losses can also be viewed as a type of Fresnel loss as the fiber–air interface introduces Fresnel reflection losses. In this case, there are two major losses to be considered. The first takes place in the inner surface of the transmitting fiber, and the second occurs due to reflections from the surface of the second fiber. One way of eliminating these losses is by introducing a coupler that matches the optical impedances of the two materials. This arrangement results in matched reflection coefficients, which is analogous to matching of impedances.

In general, the positions and shapes of the fiber cores are controlled to tight manufacturing tolerances. Fibers having an attenuation of greater than 1 dB/km are rarely used in communication networks. Nevertheless, the attenuation of badly matched fibers may exceed 1 dB/km per connector or splice if they are badly handled during installation stages. Good coupling efficiency requires precise positioning of the fibers to center the cores. The simplest way to avoid connector losses is by splicing the two ends of the fibers permanently, either by gluing or by fusing them at high temperatures.

Splicing joins two fibers together and forms a continuous optical waveguide. Generally, splicing is realized by employing the arc fusion splicing method. In this technique, an electric arc is used to melt the fiber ends together. Mechanical fiber splices are additionally designed to be quicker and easier to install. Good splicing requires appropriate stripping, careful cleaning, and precision cleaving. A splice loss under 0.1 dB is typical. The complexity of this process makes fiber splicing much more difficult than splicing copper wire. All splicing techniques involve the use of an enclosure for protection afterward.

Similar to connector loss, splice loss is minimized when the two fiber cores are identical and perfectly aligned. The fiber ends must also be properly polished; a rough surface will scatter light and dirt can scatter and absorb light. Due to the small size of optical fibers, typical airborne dirt can be a major source of loss. Therefore, whenever connectors are not terminated, they should be covered to protect the ends from dirt. More importantly, one should never even touch the fiber end, since the oils on one's skin cause the fiber to attract dirt. It is also highly recommended to clean connectors with lint-free wipes moistened with isopropyl alcohol before connection and/or testing operations.

Usually fiber mismatches occur for two reasons: First, we occasionally need to interconnect dissimilar fibers, which produce difference in fiber index of the same nominal dimensions. Second, it is possible to have dissimilar fiber connections in optical systems designed for single- or multimode fibers. However, some system manufacturers provide guidelines on using various fibers, and some don't. For example, by connecting a smaller fiber (in diameter) to a larger one, the coupling loss will be minimal, limited only by the Fresnel loss. However, connecting larger fibers to smaller ones results in substantial losses.

53.3.3 Intrinsic Fiber Losses

Intrinsic fiber losses are those associated with the fiber-optic material itself and are considered the primary source of attenuation. There is no way to eliminate all impurities in fiber optics even with a precise manufacturing process. The total loss is proportional to the fiber length L and the attenuation coefficient α is expressed in decibels per kilometer of fiber. Intrinsic fiber losses are mainly caused by absorption and scattering because scattering and absorption occur when light wave signals interact with impurities in the fiber.

53.3.3.1 Material Absorption Losses

Absorption is the main source of intrinsic loss in optical fibers, which causes attenuation by converting optical power into vibration energy or other energy forms such as heat. Unlike scattering, absorption can be limited by controlling the amount of impurities during the manufacturing process. Note that as fiber is almost pure, it does not heat up because of absorption. In principle, absorption occurs when a light particle (i.e., photon) interacts with an electron and excites it to a higher energy level. In other words, it is caused by the material properties of the optical fiber and is usually explained by three important factors listed as follows:

- Imperfections in the atomic structure of the fiber material
- Intrinsic or basic fiber material properties
- Extrinsic fiber material properties (presence of impurities)

Intrinsic and extrinsic material properties are the main cause of absorption. Intrinsic absorption is caused by basic material properties of the fiber itself (i.e., imperfection, impurities, and undulation) and sets a minimal attenuation level due to total absorption loss. For this reason, silica (pure glass) is predominantly used in fiber optics because of its very low intrinsic absorption in the operational wavelength window of 700–1600 nm. Similar to most transmissive systems, the absorption loss depends on the fiber length with an exponential function of length. In silica glass, absorption is caused by the vibration of silicon–oxygen (Si–O) bonds. The interaction between the vibrating bond and the electromagnetic field of the optical signal causes intrinsic absorption. Hence, energy is transferred from the electromagnetic field to the bond. Absorption is also induced by the diffusion of hydrogen molecules into the glass fiber.

Extrinsic absorption is caused by impurities in the fiber such as iron, nickel, and chromium during fabrication. In this case, the ion transitions between energy levels cause absorption. Extrinsic absorption also occurs when hydroxyl ions (OH^-) are introduced into the fiber. Optical power is absorbed in the excitation of molecular vibrations of such impurities in the glass, as illustrated in Figure 53.6. One characteristic of absorption is that it occurs only in the vicinity of wavelengths corresponding to the natural oscillation frequencies or their harmonics of the particular material. In modern fibers, absorption losses are almost entirely due to OH^- ions. The fundamental vibration mode of these ions corresponds to $\lambda = 2.73$ μm and the harmonics at 1.37 and 0.95 μm. Dehydration techniques can be employed during the manufacturing process to reduce the presence of OH^- ions. As illustrated in Figure 53.6, absorption losses due to each type of impurity act like a band suppression filter, showing peak absorption at well-defined wavelengths. Absorption accounts for 3%–5% of fiber attenuation. In addition, the amount of water (OH^-) impurities in a fiber should be less than a few parts per billion as extrinsic absorption is only affected by the level of impurities present in the fiber.

53.3.3.2 Scattering Losses

The light traveling down the fiber core interacts with the silica molecules. Rayleigh scattering is the result of these elastic collisions between the light wave and the silica molecules in the fiber. Rayleigh scattering accounts for about 96% of attenuation in optical fiber. However, no attenuation occurs if the scattered light maintains an angle that only supports traveling in the forward direction within the core.

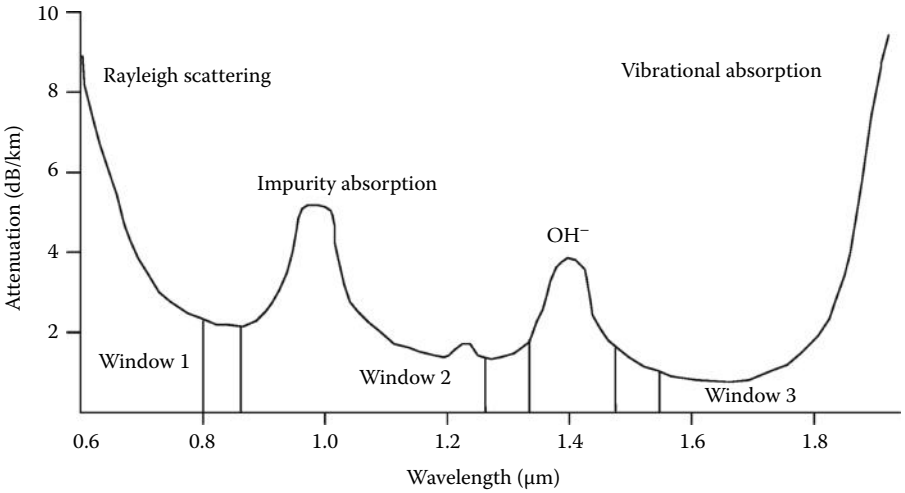


FIGURE 53.6 Attenuation characteristics of a typical optical fiber. The attenuation coefficient α varies against wavelength for all low-loss fused silica fiber due to Rayleigh scattering and impurity absorptions. The optical losses for wavelengths below 1 μm are mainly due to scattering. At longer wavelengths absorption losses are important, notably at 1.4 μm through absorption by OH^- ions. Above 1.6 μm, absorption due to impurities becomes dominant.

Otherwise, the light is diverted out of the core and attenuation occurs. In other words, depending on the incident angle, portion of the light propagates forward and the remaining portion deviates out and escapes from the fiber core. Some scattered light is reflected back toward the light source, which can be used in an optical time domain reflectometer for testing live (data carrying) fibers. The same principle applies to analyzing loss associated with localized events in the fiber such as splices, connectors, and bending.

Despite careful manufacturing techniques, most fibers are inhomogeneous and have disordered, amorphous structures, which cause power losses due to scattering caused by imperfections in the core material and irregularities between the junction and cladding. Generally, shorter wavelengths are scattered more than longer wavelengths. More precisely, wavelengths below 800 nm are useless for optical communication due to very high Rayleigh scattering attenuation. At the same time, propagation above 1700 nm is not possible due to high losses resulting from infrared absorption.

Inhomogeneities can be either structural or compositional in nature. In structural inhomogeneities, the basic molecular structure has random components, whereas in compositional inhomogeneity the chemical composition of the material varies. The net effect from either inhomogeneity is a fluctuation in the refractive index. As a rule of thumb, if the scale of these fluctuations is on the order of $\lambda/10$ or less, each irregularity acts as a scattering center. This is a form of Rayleigh scattering and is characterized by an effective absorption coefficient that is proportional to λ^{-4}. Rayleigh scattering can be caused by the existence of tiny dielectric inconsistencies in the glass. Because these perturbations are small with respect to the waves being propagated, light striking a Rayleigh imperfection scatters in all directions. Scattering losses are less at longer wavelengths, where the majority of the transmission losses are due to absorption from impurities such as ions. Rayleigh scattering losses are not localized, and they follow a distribution law throughout the fiber. However, they can be minimized by having low thermodynamic density fluctuations.

A small part of the scattered light propagates in the opposite direction. This backscattering has important characteristics and may be used for measuring fiber properties. Usually, the inhomogeneities in the glass are smaller than the wavelength λ of the light. The scattering losses in glass fibers approximately follow the Rayleigh scattering law; that is, they are very high for small wavelengths and decrease with increasing wavelength.

In general, optical losses in the glass cause the optical power in a fiber to fall off exponentially with the length L of the fiber:

$$P(L) = P(0)\ 10^{(-\alpha L/10\ \mathrm{dB})} \tag{53.2}$$

where

$P(0)$ denotes the optical power that couples to the fiber
$P(L)$ denotes the power remaining after length L
α is the attenuation coefficient indicating the rate of loss of optical power in dB/km

The product αL is called the *attenuation of the fiber*. An attenuation of 10 dB means that the optical power $P(L)$ at the end of the fiber is only 10% of the initial power $P(0)$. For instance, a 3 dB attenuation is equivalent to 50% and 1 dB to about 80% of $P(0)$. A typical attenuation coefficient α against wavelength λ is shown in Figure 53.6 for common low-loss fused silica fiber. The optical losses for wavelengths below 1 μm are mainly due to Rayleigh scattering. At longer wavelengths, absorption losses are important, notably at 1.4 μm through absorption by OH^- ions. Above 1.6 μm, absorption due to impurities becomes dominant. Because of attenuation, only limited wavelength ranges are appropriate for optical data transmission.

Although intrinsic fiber losses can be associated with the core index n_f, the core index plays an important role in determining the propagation time delay of optical signals. The propagation time delay t_p may be expressed by

$$t_p = n_f \frac{L}{c} \tag{53.3}$$

where

c denotes the velocity of light in the fiber
L denotes the fiber length

Another type of loss in optical fibers occurs due to the propagation of light at different angles.* For a given length of fiber, the high-order modes reflect more often and cover longer distances than the low-order modes. Therefore, high-order modes suffer from more losses, thus causing modal dispersion. Modal dispersion is one of the primary causes of rise time degradation for increasing fiber wavelengths. In addition, the propagation time varies with the index of refraction so that different wavelength components of the source spectrum have different travel times, thus causing chromatic dispersion.

53.4 Application of Fiber Loss Properties in Fiber Sensors

As mentioned earlier in this chapter, while fiber loss characteristics are undesirable for data transmission in telecommunication applications, the loss properties may be employed to develop high-accuracy sensing devices for a variety of applications. Fiber-optic-based sensors have significant advantages compared to conventional electric counterparts. They are highly sensitive (more than other technologies), configuration versatile (i.e., provide both localized and distributed sensing), chemically and physically passive, free of electromagnetic interference, and remotely powered. They also have a wide dynamic range and bandwidth, are available in extremely small size, are immune to electromagnetic interference and radio frequencies, and provide environmental ruggedness.

* Light propagating at shallow angles is called low-order mode, while light propagating at larger angles is called high-order mode.

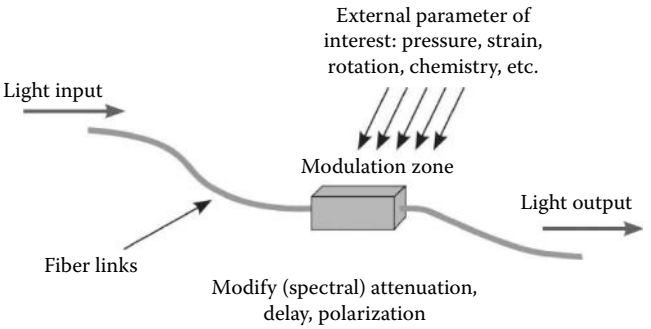


FIGURE 53.7 Operation principle of fiber sensors. In extrinsic fiber-optic sensors, the fiber mainly acts as a transmission medium for the modulated light beam by an environmental effect. In intrinsic fiber-optic sensors, the propagating light beam is modulated either directly or indirectly by the environmental effects.

The operation principle of fiber-optic sensors is depicted in Figure 53.7. As illustrated, in principle, any physical or chemical phenomenon, which can directly/indirectly change the loss properties of fiber transmission media, can be measured and controlled by monitoring the corresponding variations in the transmitted light wave power and/or rate. These include temperature, pressure, acoustics, linear and angular position, strain, voltage, current, liquid level, viscosity, torque, twisting, rotation, acceleration, and particle velocity. In these applications, the desired variable quantity to be monitored modulates the intensity, phase, polarization, wavelength, or transit time of light in the fiber.

Depending on the role of the fiber, fiber-optic sensors can generally be classified into two groups: intrinsic (or hybrid) fiber sensors and extrinsic fiber sensors (or all-fiber sensors). In intrinsic fiber sensors, fiber is utilized as the sensing element, where the propagating light is modulated by the environmental effects, whereas in extrinsic fiber sensors, the signals are relayed via fiber from a remote sensor (can also be fiber optics) to the optoelectronics, which process the signals. In its simplest form, a fiber-optic sensor is an intrinsic type, whose operation is based on intensity modulation. Spectral, interferometric, and distributed (multiplexing) mechanisms may also be employed for sensing application. Fluorescent-based fiber sensors are used in medical application and chemical sensing; fiber-optic gyros are used in navigation systems (most types of vehicles such as Boeing 777), pointing and tracking of satellite antennas.

Distributed sensing over very large distances is provided by intrinsic fiber-optic sensors. Figure 53.8 illustrates the general scheme of distributed sensing techniques based on backscattering of the launched optical signals to the fiber. Reflected and backscattered signals coming from different points in the fiber provide a continuous profile of temperature (external parameter of interest) versus distance. The spatial and temperature resolution as well as the reach length* of the system varies, depending on the application and technology. Some examples are provided in Figure 53.8.

In general, both OTDR and optical frequency domain reflectometry (OFDR) techniques may be employed for this purpose. For instance, distributed temperature sensing (DTS) is available by using a fiber that has an evanescent loss that varies with temperature or by analyzing the Raman scattering of the optical fiber. A continuous profile of temperature versus distance with special resolution of 1 m over very large distances such as 30 km within ±1 °C at a resolution of 0.01 °C with measurement times in order of minutes can be obtained. These so-called distributed sensor techniques are unique to fiber-optic technology. OTDR is addressed in greater detail in the next section.

Brillouin scattering (i.e., OFDR in fibers) is due to the interaction of light wave signals in a medium (such as crystal) with time-dependent optical density variations. From a quantum point of view, an incident light beam (photon) can be converted into a scattered light (photon) of slightly lower frequency (energy), usually propagating in the backward direction. Indeed, similar to sound waves, the backward propagating

* The longest distance for which the results are valid with acceptable accuracy.

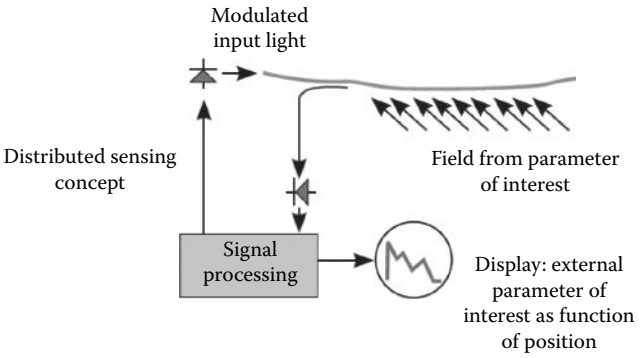


FIGURE 53.8 Distributed sensing using fiber-optic technology. Backscattering from different points of the fiber reflected back to the launch point at one fiber end can be measured. A continuous profile of temperature versus distance is obtained with special resolution of 1 m over very large distances such as 30 km within ±1 °C at a resolution of 0.01 °C with measurement times in order of minutes.

light experiences a Doppler shift, that is, frequency change. This frequency shift in the reflected light is known as *Brillouin frequency shift* (BFS) and can be measured by Brillouin spectrometers. The BFS for the reflected light is slightly lower than that of the launched light. BFS in fibers can be calculated from the effective refractive index n, the acoustic velocity v_a, and the vacuum wavelength λ as follows:

$$\nu_B = \frac{2nv_a}{\lambda} \tag{53.4}$$

In general, BFS depends on the material composition inside the fiber core, temperature, and pressure of the medium, which are exploited in fiber sensors.

53.5 Optical Time Domain Reflectometry

Time domain reflectometry, initially developed for conventional copper wires, is a well-known solution for infield monitoring of installed fiber links. OTDR is used to test fibers during fabrication, cabling, and warehousing; it is also used for troubleshooting deployed fiber infrastructures. The very first OTDR equipment has been implemented for optical carriers in long-distance transmission systems.

53.5.1 OTDR Principle

The principle of OTDR is based on the fact that as the optical pulse propagates through the fiber, both reflection and scattering result in a fraction of the optical signal being reflected back in the opposite direction. Rayleigh scattering and Fresnel reflections are physical causes of this behavior. Generally, the magnitude of the backscattered light depends on the Rayleigh scattering, fiber attenuation and imperfections, splices, and the launched light wave signal specifications (power and pulse width).

The OTDR equipment launches a series of short light pulses into the fiber and measures the backscattered light. A general schematic of an OTDR device is presented in Figure 53.9. The reflected signal arrives at the transmitter with a delay proportional to the corresponding relative distance with respect to the transmitter. The backreflected signals are decoupled from the fiber by means of a beam splitter and circulators installed inside the OTDR. The OTDR trace then gives the impulse response of the link under the test, that is, a plot of the power versus distance. OTDR is an optoelectronic device, which provides a full characterization of the fiber link under test.

Figure 53.10 presents a typical OTDR trace. Using this trace, an impulse response of the fiber link is provided by the OTDR device. For instance, the jumps in Figure 53.10 correspond to the insertion loss

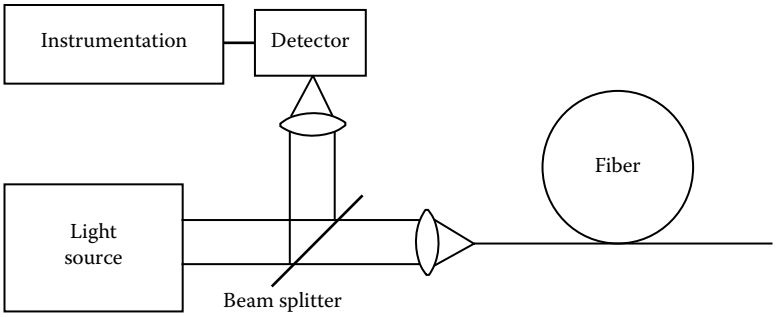


FIGURE 53.9 OTDR is used in link tests. The optical power is launched into the fiber, and the reflected power associated with Rayleigh scattering is measured from the same sending end. OTDR is usually offered with customized analysis software and optical modules to be integrated into a computer. OTDR complements attenuation measurements by measuring backscattering. This permits not only the attenuation of the complete fiber to be measured but also different attenuation coefficients of fiber segments as well as optical losses in connectors and splices.

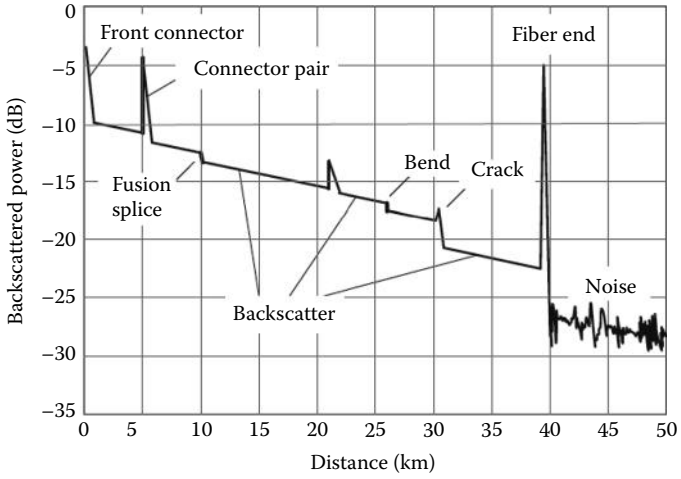


FIGURE 53.10 A typical OTDR trace, which provides full fiber link characterization. The jumps correspond to the insertion loss of network components as illustrated, whereas the power reflection peak at 40 km signifies the Fresnel reflections at the fiber–air interface (e.g., fiber cut). No backscattering is detected after the fiber end and the trace drops to the receiver noise. The power level decreases linearly from left to right (in logarithmic scale). From the location and height of the events, the positions and magnitudes of the local losses can easily be identified.

of network components as illustrated, whereas the power reflection peak at 40 km signifies the Fresnel reflections at the fiber–air interface. After the fiber end no backscattering is detected and the trace drops to the receiver noise, which is usually dominated by the thermal noise of the electronic circuit.

Generally, the OTDR trace can be used to extract information about link faults, including fiber misalignment, fiber mismatch, angular faults, dirt on connectors, macrobends, and/or breaks. These faults are usually referred to as events on the OTDR trace. An OTDR device is a distributed fiber sensing instrument and can be effectively employed for fiber link troubleshooting and fault root cause analysis. OTDR is also useful to reduce the operational cost of fiber repair while reducing the network downtime.

The backscattering factor depends on scattering and numerical aperture. This factor is usually very small in the range of 50–70 dB and is reduced proportionally by decreasing the transmitted pulse width. Despite of its low level, the reflected power is still detectable for several kilometers depending on the application. In many cases, light pulses with high-energy content are used along with sensitive receivers based on avalanche photodiodes (APDs), which provide both detection and amplification.

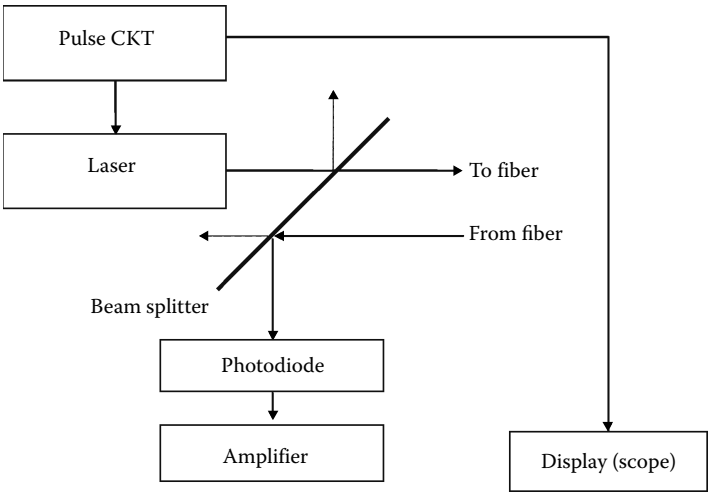


FIGURE 53.11 OTDR products usually come with a laser source and modules that allow for wavelength selection. The short, intense laser pulses with high peak power are coupled into the fiber. Backscattering echoes are detected from different regions of the fiber. The returned signal is decoupled from the fiber by means of a beam splitter to be further processed for analysis.

Assuming a constant attenuation coefficient throughout the fiber, in the OTDR trace the power level decreases exponentially (or linearly in logarithmic scale, i.e., decibel) from left to right. In practice, local events due to imperfect connectors, splices, etc. impose some discontinuities in the OTDR trace. These localized jumps (discontinuities) provide information about both the physical location and insertion loss of the corresponding events. A practical advantage of this technique is that it involves only one of the fiber ends. The reference trace acquired under normal conditions is very useful to determine where changes have occurred. Some companies keep the reference OTDR test results in case of future fiber failure for warranty claims.

Manufacturers usually provide OTDR products with customized software analysis to be integrated into a computer and optical modules, as illustrated in Figure 53.11. Therefore, OTDR products usually have significant computing ability with graphical display. Notwithstanding that, the proper operation and interpretation of the OTDR trace require technical training and experience.

53.5.2 OTDR Design and Considerations

Generally, OTDR is available for a variety of fiber types and wavelengths for different applications. Longer wavelengths, such as 1625 nm (U band), are highly recommended for troubleshooting as splice or connector losses impose higher losses and hence are more visible on OTDR traces. U band probe signals also provide a transparent scheme, where monitoring systems do not interfere with other data communication bands such as C and L bands. The dynamic range of an OTDR is limited by a combination of optical output power, transmitted pulse width, processing time at the receiver (or equivalently the electrical bandwidth), and receiver sensitivity. Higher launched power and better sensitivity are usually combined to improve the measurement range and are often fixed features of a particular instrument. The pulse width adjustment, however, implies certain trade-offs, depending on the desired application. A longer laser pulse width provides a better dynamic range and attenuation measurement resolution at the expense of higher distance resolution (i.e., spatial resolution). For example, using a long pulse width, it may be possible to measure attenuation over a distance of more than 100 km; however, for this application, only events physically located farther than 1 km (for 1 μs pulse width) are distinguishable in the OTDR trace. This scenario is useful for overall characterization of a link, but is useless for fault localization.

Similarly, for a short pulse width, we have a better distance resolution of events, while suffering from a smaller measuring range and attenuation measurement resolution. The "apparent measurement length" of an event is known as the "dead zone" on the OTDR trace. The dead zone is the minimum distinguishable distance between any two events in the OTDR trace, which in turn is a function of the transmitted pulse width, light wave speed in the fiber medium, and size of the event. Therefore, events physically located closer than the "dead zone" of an OTDR are not separable. As mentioned earlier, for the attenuation zone, where no reflecting event exists, the power level drops linearly (in dB) with the slope of fiber attenuation.

Most OTDR solutions employ averaging techniques to improve both the sensitivity and dynamic range. The sensitivity of an OTDR increases with the square root of the total integration time (i.e., averaging time). For instance, a 16-time increase in averaging time results in a sensitivity increased by a factor of four. However, the integration time is practically limited to few minutes. Advanced digital signal processing techniques may also be employed for further improvement in the OTDR trace quality and analysis.

While OTDR is able to identify events such as unacceptable losses, it is merely used for accurate end-to-end loss or return loss measurements. A separate end-to-end loss test with a light source and power meter is required for this purpose. In some applications, the OTDR capability to localize a fault may be improved via augmented (sometimes localized) modules such as red laser fault locator, clip-on identifier, or "cold clamp" optical cable marker.

Multiple joined fibers of different types and lengths can also cause an incorrect interpretation of the events on an OTDR trace. A well-known example is where a fiber joint shows an abnormal gain (jump) in the OTDR trace instead of loss. This is because connected (adjacent) fibers may have different properties such as backscatter coefficients and core diameters. Consequently, more light is reflected right after the fiber junction while the same light is propagating down the fiber. Placing the OTDR equipment on the other side of the same fiber will illustrate an abnormally high loss at that joint. In these cases, the combination of the two directional tests (taken from both ends of the link) should be utilized for proper interpretation of the OTDR data. For this reason, it is common OTDR practice to measure and combine the traces from both ends for a more accurate characterization of the fiber link.

53.5.3 OTDR for Branched Optical Networks

While providing an automatic monitoring system and full characterization of the fiber link, OTDR is ineffective for a point-to-multipoint (PMP) network (i.e., branched network) such as a passive optical network (PON). This is because the backscattering signal of each branch in a PMP network is partially masked by the others. In this case, the total power measured via OTDR is a linear sum of all powers coming from different branches. Useful information can nevertheless be extracted from the global backscattering trace. Actually, the OTDR trace does not take into account the fact that there are different paths because it only measures the total accumulated backscattered light coming from all the branches. The desired information about the network, however, is in the OTDR trace itself, which needs to be decoded and extracted. Extracting the desired information from the OTDR trace might require considerable off-line signal processing and detailed knowledge of the fiber network infrastructure.

The key point for analyzing an OTDR trace in the case of a branched network is to compare the current backscattering trace with a reference acquired from the PMP network under standard conditions. A network simulator is then required to obtain more significant data from the measured trace. The software developed for this purpose needs a self-learning module to develop and continuously update a database containing the reference patterns and an event table. During the in-service monitoring of the network, if any deviation in the measured trace (with respect to the standard trace) is observed, an alarm is sent to the service provider and the software starts to analyze the measured data. The simulator in the software then generates a variety of virtual network configurations to achieve the measured trace by using the information of the standard trace.

The accuracy of such software is strongly dependent on the accuracy of the simulator developed for the network and the uncertainties of both the training and measured data. Deviation in the backscattered coefficient of different fiber branches is another source of error. The situation is problematic when two or more branches are located at the same distance. In such cases, OTDR cannot distinguish different events and branches. The test system then should take separate end-to-end measures for each branch to differentiate them among individual branches in the overall OTDR trace. Recall that as the number of branches (i.e., network size) increases, the complexity of the analysis increases dramatically leading to a less reliable network analysis by the OTDR device.

Neglecting the complexity in the analysis of the OTDR trace in a PMP network, the huge loss incurred usually by passive splitters, located at the remote node (RN), leads to a significant drop in the measured power. For example, if there is a 1:32 splitter at the RN, the total backscattered light from each branch suffers 15 dB loss, as shown in Figure 53.12. The RN is then recognized as the fiber end and no useful information can be extracted for distances farther than the RN. In traditional OTDR, losses higher than 3–7 dB are usually identified as end-of-fiber. By simply modifying the OTDR analysis, testing can be performed through splitters with losses up to 20 dB. This kind of OTDR is usually referred to as PON-tuned OTDR.

In order to further reduce the complexity in the analysis of OTDR trace in a PMP networks, a variety of solutions have been proposed to distinguish individual fiber branches, among which the most well-known techniques are reference reflectors, fiber selectors, wavelength routers, chained branches, and dark fibers. In each of these techniques, a unique signature is assigned to each fiber branch to make it distinguishable from the others.

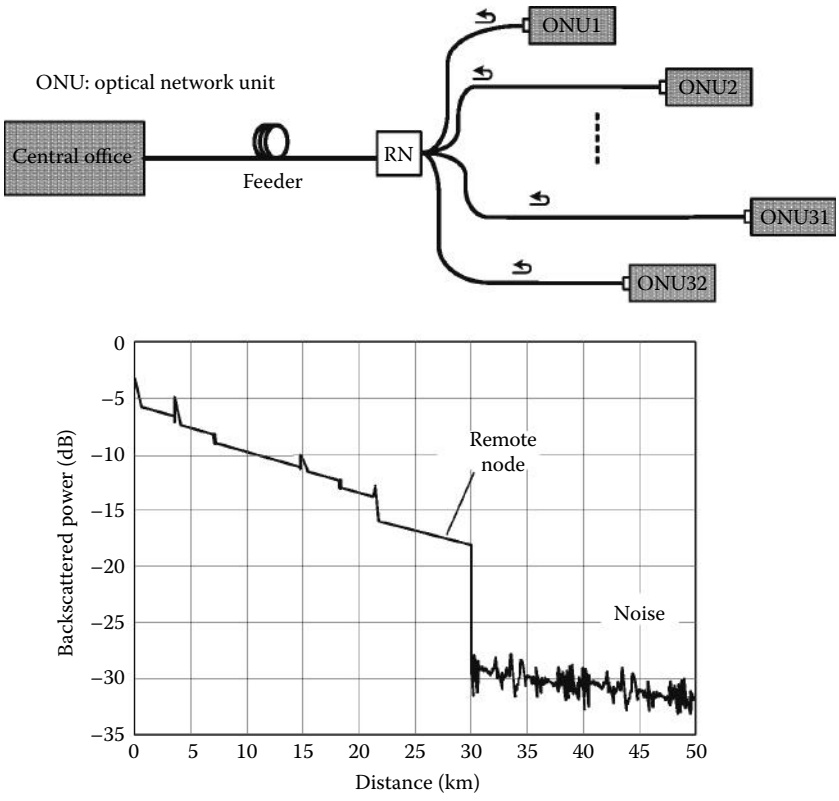


FIGURE 53.12 OTDR for a typical PMP network with a 1 × 32 splitter at the RN. The RN imposes 15 dB additional loss on the data measured via OTDR. The figure shows the corresponding OTDR trace of the network. It is observed that the power falls to the receiver noise level, and therefore, no useful information is extracted after the passive splitter.

53.5.4 OTDR Challenges for Branched Networks

The most critical components for OTDR-based monitoring technologies are optical selectors, filters, reflectors, and WDM devices. These devices should be cost- and dimension-effective (i.e., low cost and high density) to be able to manage a large amount of fibers in future access networks. Recall that while the ITU-2 recommendations propose the U band, 1625–1675 nm, for monitoring applications, the behavior of passive components is not very well investigated for this wavelength regime. The OTDR itself is the basic equipment for the automatic test system and needs to have suitable technical characteristics. These include spatial resolution, dynamic range, dead zone, wavelength stability and compatibility, launched power, and sensitivity.

These requirements should be met for OTDR to be an effective monitoring solution for future access networks. As the field of OTDR techniques and the corresponding key components are continuously advancing, these techniques are expected to be more reliable in the future. Recall that the cost is an important issue, but as OTDR is shared among many branches (customers) in the network, it is not the most critical one. Also note that the OTDR trace analysis in a PON is still challenging, especially for a large number of individual branches.

53.6 Standard Field Fiber-Optic Attenuation Test

In practical applications of optical fibers, it is necessary to have quantitative knowledge about the whole range of properties. The most important properties are length, attenuation, and bandwidth of the fibers, along with external diameter, core diameter, numerical aperture, and refractive index profile. The actual performance of installed links may be different from the desired performance. Therefore, it is important that both individual components and the entire assembled system undergo testing to verify compliance with the required operations. Additional testing may be conducted over the lifetime of the system to ensure continued functional operation over long periods of time.

There are two basic standard fiber-optic attenuation tests: the component acceptance test and the link test. The component acceptance test is performed prior to installation to verify the power and performance acceptance levels of each fiber. The acceptance testing of a functional module begins with power testing of the transmitter. The testing is done by attaching the module set into a reference link and verifying the data rates and bit error rates taken at the other end of the fiber, as shown in Figure 53.13.

Fiber acceptance testing also requires power testing, which can be done by launching a known power from a reference source, as explained in the OTDR section. The functional testing verifies power budget of coupling attenuation effects as well as bandwidth for the fiber. The attenuation and power transmission depend on the mode distribution within the fiber. In short fibers, higher modes dominate, whereas in long fibers, power is more concentrated in lower-order modes. Mode stripping is often used by introducing small bends into the fiber to correlate the performance of tested fiber to actual in-service performance.

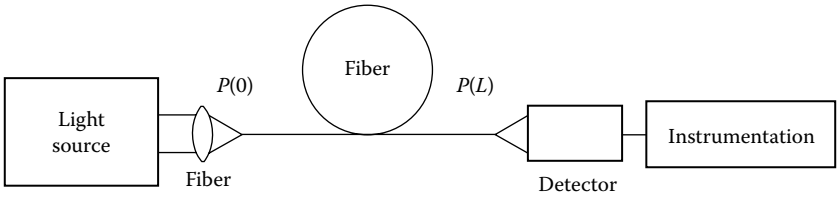


FIGURE 53.13 The measurement of transmission properties of laid optical fibers is obtained by end-to-end tests. The testing is done by attaching the module to a reference link and verifying the data rates and bit error rates taken at the other end of the fiber.

In addition to acceptance tests, fibers undergo other tests such as pull strength, breaking, humidity resistance, prolonged tension, and bend tests. Connectors, splitters, combiners, and fiber amplifiers need to be tested frequently to minimize power transmission losses.

In almost all fiber optics, low noise and very sensitive photodetectors are used in measurements. Semiconductor photodiodes made from silicon are suitable for measurements with wavelengths below 1 μm. For larger wavelengths, other detectors are used, such as those made of germanium. Semiconductor photodiodes work via the internal photoelectric effect; that is, they absorb photons of energy $h\nu$ containing the incoming light beam power P and emit a number of electrons proportional to the number of photons, creating a current given by

$$i_P = \frac{P\eta e}{h\nu} \tag{53.5}$$

where

h denotes Planck's constant
ν denotes the frequency of the absorbed light ($\nu = c/\lambda$)
η is a constant of proportionality or quantum efficiency ($\eta < 1$)
e denotes the charge of an electron

Photodetectors are used to measure the attenuation in fibers by measuring the optical power at the input $P(0)$ and power $P(L)$ at the end of the fiber. From Equation 53.2, the attenuation of the fiber can be calculated as

$$a_0 = \alpha L = \frac{-10 \log P(L)}{P(0)} \tag{53.6}$$

When photodetectors are used, Equation 53.6, can be written as

$$a_0 = \frac{-10 \log i_P(L)}{i_P(0)} \tag{53.7}$$

It is worth noting that optical power is proportional to current i_p, not to i_p^2.

LEDs or incandescent halogen lamps are suitable for use as light sources for attenuation measurements. If LEDs are used, several interchangeable ones are needed, suitable to different wavelengths to make measurements. On the other hand, halogen lamps yield a wide spectrum of radiation, from which light of various wavelengths can be filtered with a monochromator that uses optical filters, prisms, or diffraction gratings. With an arrangement of this kind, the attenuation of the fiber can be measured as a function of the wavelength of the light.

To measure small optical powers precisely for the purpose of determining attenuation, the photodetector is usually connected to a selective amplifier, and the light source is modulated at a low frequency (e.g., 400 Hz) using a rotating disk called a *chopper*. The selective amplifier amplifies only the similarly modulated components of the photocurrent so that detector noise and the influence of background light are suppressed. For this purpose, frequency-selective level meters or lock-in amplifiers are often used. In both cases, the frequency and phase of the modulation are fed back in the form of a reference voltage.

53.7 Out-of-Plane Scattering and Polarization Methods

The art of scatter measurements has long been evolved to established forms within the optics industry. Scatter methods provide extremely sensitive measurements in many diverse applications. There are two basic instruments developed for this purpose: the scatterometer and the polarimeter.

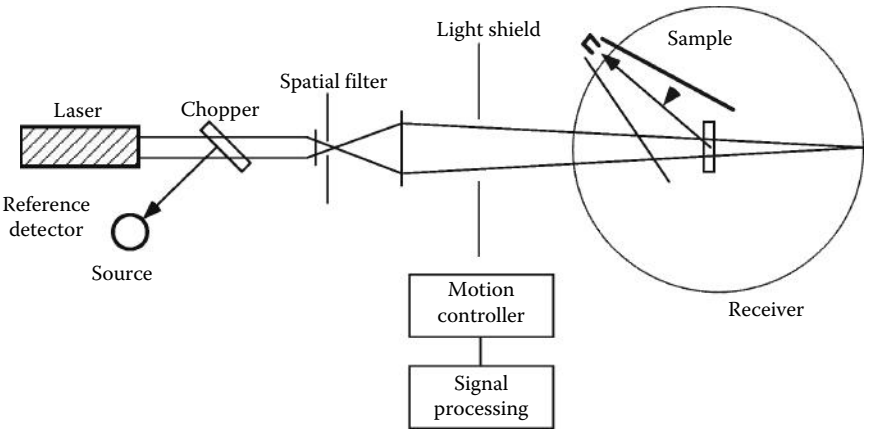


FIGURE 53.14 A typical scatterometer. In this particular type, the source is fixed in position and the sample is rotated to the required incident angle. The receiver is rotated about the sample in the plane of incidence. In other types, source and receiver may be fixed and the sample is rotated so that the scatter pattern moves past the receiver. Scatterometers are offered with full supporting software for analysis and appropriate software.

The basic setup for a scatterometer is shown in Figure 53.14. Measured scatter is a good indicator of surface quality as well as discrete defects. However, the scattered signals are generally small compared to the specular beams, and they can vary by several orders of magnitude for just a few degrees. Therefore, scatter measurements need sophisticated instrumentation and the signal processing is more complex than many other optical techniques. For successful applications, the system specifications and measurements must be given in terms of accepted, well-defined quantities. Scatter methods are used routinely as a quality check of optical components in fiber-optic applications. Conversion of surface scatter data to other required formats, such as surface roughness, is common practice. Out-of-plane measurements and polarization-sensitive measurements are areas currently experiencing rapid advancements.

Another instrument in common use is the polarimeter, which senses the polarization of scattered light. The scattering characteristics of a sample are generally described by its bidirectional reflectance distribution function. The reflectance distribution function is the ratio of scattered flux in a particular direction to the flux of an incident beam. The scattered light is often a sensitive indicator of surface conditions. A small amount of surface roughness may reduce the specular power by less than 1% while increasing the scattered power by orders of magnitude. Similarly, the retardance, attenuation, and depolarization of scattered light provide sensitive indicators of conditions, such as uniformity of refractive index, orientation of surface defects, texture, strain, subsurface damage, coating microstructure, and the degree of multiscattering. Among many other methods, the use of prisms helps realize polarization and depolarization of the scattered light.

Bibliography

Agrawal G., *Fiber-Optic Communication Systems*, 3rd edn., New York: John Wiley & Sons, 2001.

Bass M., *Handbook of Optics-Fundamentals, Techniques and Design*, 2nd edn., Vols. I and II, New York: McGraw-Hill, 1995.

Burke M. W., *Image Acquisition-Handbook of Machine Vision Engineering*, Vol. 1, Oxford, U.K.: Chapman & Hall, 1996.

DeCusatis B., and I. Kaminow, *The Optical Communications Reference*, Burlington, MA: Elsevier, 2009.

Derickson D., *Fiber Optics Test and Measurements*, Upper Saddle River, NJ: Prentice Hall, 2002.

Dutton H., *Understanding Optical Communications*, Upper Saddle River, NJ: IBM Redbooks, 1998.

Fischer R., B. Tadic-Galeb, and P. Yoder, *Optical System Design*, 2nd edn., New York: McGraw-Hill, 2008.
Forestieri E., *Optical Communication Theory and Techniques*, 1st edn., New York: Springer, 2004.
Gagliardi R. and S. Karp, *Optical Communications*, 2nd edn., New York: John Wiley & Sons, 1995.
Green L. D., *Fiber Optic Communication*, Boca Raton, FL: CRC Press, 1993.
Girard A., *FTTx PON Technology and Measurement*, EXFO Electro-Optic Inc., Quebic City, Canada, 2009.
Haus J., *Optical Sensors: Basics and Applications*, Weinheim, Germany, Wiley-VCH Verlag GmbH & Co, 2010.
Huber J., *Industrial Fiber Optic Networks*, Research Triangle Park, NC: Instrument Society of America, Instrumentation Systems, 1995.
Keiser G., *Optical Fiber Communications*, 4th edn., New York: McGraw-Hill, 2010.
Palias J., *Fiber Optic Communications*, 4th edn., Englewood Cliffs, NJ: Prentice Hall, 1998.
Pearson E., *Complete Guide to Fiber Optic Cable Systems Installation*, 1st edn., Clifton Park, NY: Delmar Cengage Learning, 1996.
Rastogi P., *Optical Measurement Techniques and Applications*, Boston, MA: Artech House, 1997.
Senior J. M., *Optical Fiber Communications: Principles and Practice*, 2nd edn., Englewood Cliffs, NJ: Prentice Hall, 1992.
Yin S., P. B. Ruffin, and F. T. S. Yu, *Fiber Optic Sensors*, 2nd edn., Boca Raton, FL: CRC Press, 2008.

VI

Chemical Variables

54

Interference and Selectivity in Portable Chemical Sensors

Denise Wilson
University of Washington

54.1 Introduction

Chemical sensors can be loosely classified into those that use light to extract information about chemical concentration (optical sensors) and those that do not (nonoptical sensors). Optical chemical sensors tend to use multiple stages of transduction and require large amounts of system and signal processing overhead to convert target analyte information into a useful output. These sensors often become complex systems or entire pieces of equipment in their final, commercialized form. For this reason, optical chemical sensors are not frequently found in portable, handheld, or field-monitoring units. Nonoptical chemical sensors, on the other hand, especially those that are integrated into an electronic device, are much better suited to portable and related applications. In the interests of evaluating chemical "sensors," where the bulk of the complete system is the sensor itself, this chapter focuses on nonoptical chemical sensors. While compact, low-cost, small in size, and suitable for field or portable monitoring applications compared to other types of chemical sensors, nonoptical sensors are often hampered by (a) insufficient selectivity or chemical interference where a sensor responds to nuisance analytes in addition to target analytes and (b) nonchemical interference (from both physical/environmental and electronic influences). Advances in materials, sensor design, and signal measurement have made a significant impact on interference and selectivity barriers; however, barriers to commercialization remain. This chapter explains the basic issues associated with selectivity and interference and highlights some of the most promising advances in enhancing selectivity and reducing interference that support continued commercialization of chemical sensor products.

54.2 Fundamentals of Selectivity

Most nonoptical chemical sensors rely on a direct interaction between an analyte and the active sensor area. For purposes of a discussion that focuses on improving selectivity (and thereby reducing chemical interference), these types of interactions can be loosely classified into three different categories:

- Chemical: The analyte binds chemically to a portion of the active sensor area. These reactions can be (1) lock and key, where the active sensor area reacts with only one analyte and is insensitive to all others; (2) semiselective, where the active sensor area reacts with a particular class of chemicals (e.g., aldehydes or alcohols); and (3) broadly selective, where the active sensor area reacts to a group of unrelated chemicals with a common property (e.g., reducing gases).
- Catalytic: The analyte binds chemically to a portion of the sensor area, but the reaction is made faster by the use of a catalyst or is otherwise assisted by the use of an enzyme. In the process of speeding up the reaction rate with one analyte (the target) as opposed to others (the interferents), the net response to the target analyte increases, thereby increasing selectivity.
- Electrochemical: A semiselective sensor area is made more selective by the use of electrochemistry where an applied potential in amperometric modes of operation limits the number of analytes that will bind with the active sensor area.

54.3 Chemical Interactions

Chemical sensors can respond with selectivity ranging from one-to-one (lock-and-key) to broadly selective (reducing gases) interactions with analytes of interest. Although most chemical sensors are developed to detect single, known analytes of interest, some chemical sensors are developed to differentiate one mixture from another; for example, it may be more relevant to distinguish the quality of one manufactured olive oil from the next rather than to discern specific chemical concentrations in the oil itself. The following discussion focuses on single analytes rather than on mixtures and, in so doing, evaluates different levels of interaction between sensor material and target analyte(s) that determine quality and selectivity of the final sensor system.

54.3.1 Lock-and-Key Chemical Interactions

In many commercial sensor applications, it is highly desirable to sense one analyte with one sensor. The ideal sensor has 100% specificity, responding to no other analytes aside from the molecule of interest. Lock-and-key chemical interactions provide this one-to-one specificity and are surprisingly common among many analytes of interest, paving the way for a wide range of highly specific chemical sensors. Affinity-based recognition elements involved in lock-and-key sensing specifically bind to individual analytes of interest or small groups of structurally related analytes.

Antibodies are by far the most common affinity-based recognition elements. They are both highly sensitive and highly selective. Polyclonal antibodies are cheap but require animals to be immunized to obtain them and are vulnerable to nonspecific binding (interference from undesirable analytes). Monoclonal antibodies do not have these specificity problems but are costly and time consuming (Van Dorst et al., 2010). Used as chemical sensors in nonclinical applications, sensors based on the use of both polyclonal and monoclonal antibodies are highly vulnerable to losing their highly specific binding affinities under changing environmental conditions (Van Dorst et al., 2010). Thus, while antibodies have the potential for a variety of lock-and-key chemical interactions, their use as chemical sensors outside clinical applications is impractical.

Phages provide similar levels of specificity to antibodies but with fewer limitations. Bacteriophages infect bacteria, using the host bacterial cell to replicate themselves. They indicate recognition of specific

bacteria by displaying specific peptides or proteins on their surface. Phages can be selected and genetically modified to obtain high specificity to specific bacterial agents. They are also inexpensive, do not require animal hosts to produce, and are stable in a wide range of environmental conditions. They have been used in amperometric (electrochemical sensor) designs to detect such important biological agents as *Escherichia coli* (Neufeld et al., 2005) and *Listeria monocytogenes* (Benhar et al., 2001) with detection limits of 1 cfu/mL and 500 cells/mL, respectively. In order to work properly as sensors, however, the target bacteria in these types of sensors must be immobilized, which make their use in portable sensor systems limited.

Nucleic acids provide similar specificity to phages, exploiting the complementary nature of DNA strands to obtain high affinity for specific types of bacteria. These sensors have some sensitivity limitations but have been integrated into electrochemical systems using voltammetry to detect chemicals such as oxytetracycline (Kim et al., 2009) and trinitrotoluene TNT (Ehrentreich-Forster et al., 2008) as well as biological toxins such as botulinum neurotoxin (Fang and Ho, 2009) and *Trichoderma harzianum* (Shafiquzzaman et al., 2012).

Antibodies, phages, and nucleic acids are, in principle, ideally suited to lock-and-key sensors but have limitations that often make them impractical for portable or field-monitoring applications. For these types of applications, the molecular imprintable polymers (MIPs) are often a good choice. MIPs essentially mimic the biological activity of antibodies but do so in a way that can be engineered to make them highly specific to a wide range of chemical analytes. A MIP is made by binding a monomer to a template molecule and then removing the template molecule, thereby leaving a "space" for recognition of target analytes of the same morphology (shape and size) as the template molecule (Figure 54.1). MIPs can be made using covalent or noncovalent bonds to create a binding site for a specific analyte. The noncovalent bonds (where a target molecule physically links to the MIP sensor) are most common because it provides the most flexibility in creating target sites (and thereby provides the widest range of sensors). Unlike antibodies, MIPs are highly stable, both thermally and chemically, producing shelf lives of up to 8 years (Van Dorst et al., 2010). To date, MIPs have been used to bind to such diverse analytes as uranium (Metilda et al., 2007), ephedrine (Mazzota et al., 2008), TNT (Riskin et al., 2008),

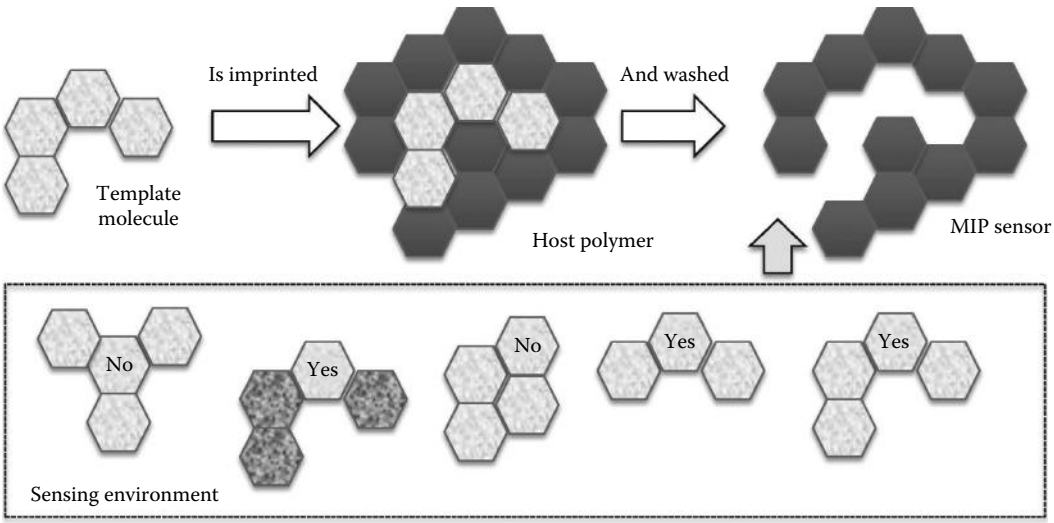


FIGURE 54.1 Molecular imprinted polymers as sensors. A monomer is bound to a template molecule, creating the sensor space (characterized by its shape). The template molecule is then removed via washing or similar process, leaving a space by which molecules of appropriate size and shape in the sensing environment can bond strongly (covalently) or weakly (physisorption) to the MIP sensor, creating a highly selective interaction between the sensor and the shape/size of target molecules.

phenols (Sergeyeva et al., 2010), atrazine (Pardieu et al., 2009), caffeine (Alizadeh et al., 2010), glucose (Fang et al., 2009), and melamine (Liang et al., 2009) using either voltammetry or other potentiometric electrochemical transduction mechanisms.

While the research on highly specific, lock-and-key chemical interactions for use in chemical sensing remains extremely active (see Van Dorst et al., 2010 for an excellent review of activity in phage, nucleic acid, and MIP-based interactions), a significant barrier to their use in commercialized sensors (especially portable or field based) is the difficulty in transducing the interaction into an electrical signal. Optical means to transduction predominate and often require complex systems overhead that lead to high-cost systems that can only be used in a laboratory. Even sensors based on electronic devices using electrochemical transduction often involve complex voltage sweeps (voltammetry) that keep system overhead high and portable or field deployment out of reach. Thus, sensors with less specificity remain popular for portable sensors. These sensors rely on semiselective or broadly selective chemical interactions that are described next.

54.3.2 Broadly Selective Chemical Interactions

On the other end of the spectrum to lock-and-key chemical interactions lie those chemical sensors that respond broadly to a range of chemicals. By far the most common types of nonoptical chemical sensor that uses broadly selective chemical interactions are those based on metal–oxide–semiconductors. Metal–oxides can be used as chemical sensors in pure form, doped, or modified by the addition of a catalyst as described in the next section. A metal–oxide sensor whose resistance changes as a function of analyte concentration (a chemiresistor) works as an n-type semiconductor with an excess of electrons in the conduction band to conduct current. The surface of a bulk metal–oxide chemical sensor contains oxygen molecules that have not bonded with the metal in the semiconductor, creating dangling bonds that are then free to interact with analytes or molecules in the surrounding air (Figure 54.2). As these

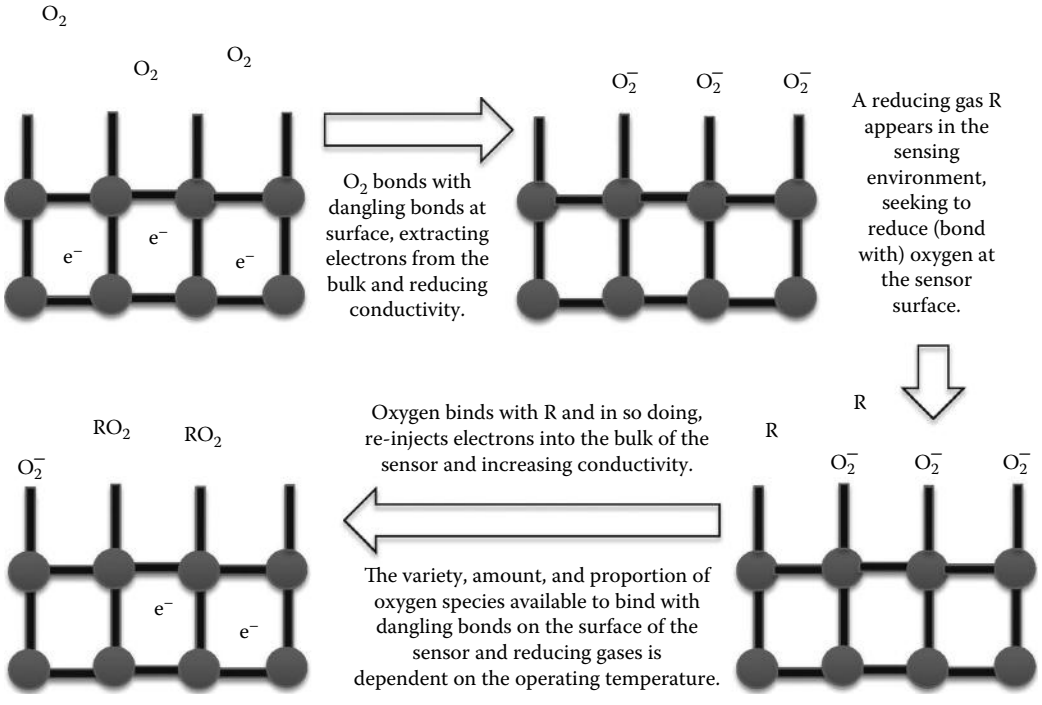


FIGURE 54.2 Basic operation of metal–oxide sensors.

oxygen molecules seek to fill their outer shells to enhance stability, they will extract electrons from the bulk of the semiconductor to fill that outer shell:

$$\left(\frac{m}{2}\right)O_2 + [\text{vacant site}] + e^- \rightarrow (O_{m^-})_{site} \tag{54.1}$$

For each dangling bond, an electron is extracted from the bulk of the semiconductor, decreasing conductivity accordingly. When a reducing gas C is present in the ambient environment, these reducing molecules will bind with the oxygen molecules, completing the outer shell of oxygen molecules which in turn will return electrons back to the bulk of the semiconductor and increase conductivity:

$$C + (O_{m^-})_{site} \rightarrow (CO_{m^-})_{site} + e^- \tag{54.2}$$

The rate K2 at which reaction in (54.1) proceeds relative to the rate K3 at which reaction in (54.2) proceeds influences the proportionality constant r that impacts the input/output relationship of the metal–oxide sensor:

$$\sigma = \mu q(n + [C]^r) \tag{54.3}$$

where

[*C*] is the concentration of the analyte of interest
σ is the conductivity of the sensor
μ is the mobility of electrons in the metal–oxide
q is the electronic charge (1.6×10^{19} C)

Clearly, the metal–oxide sensor responds to any reducing gases that react with oxygen at the surface of the sensor but it does so to varying degrees, as expressed by r in Equation 54.3. Tin oxide is, by far, the most popular material used for broadly selective metal–oxide chemiresistors and similar sensors. However, zinc oxide, tungsten oxide, titanium oxide, and tungsten oxide have also proven popular for a wide variety of gas sensing applications (Table 54.1). Palladium, platinum, zinc, silver, copper oxide, fluorine, lanthanum oxide, and cobalt oxide are popular dopants used to increase the sensitivity of these baseline metal–oxide films (Bochenkov and Sergeev, 2010).

54.3.3 Semiselective Chemical Interactions

Selectivity of many broadly selective, chemically sensitive materials and chemical sensors can be adjusted through a range of parameters. For example, the broadly selective metal–oxide materials highlighted in the previous section can be modified to improve selectivity via the following approaches:

Choice of material: Although metal–oxides react, in general, to gases that reduce the sensor surface, the rates of reaction and the equilibrium established at the sensor surface can vary between materials and certain classes of materials. Common examples of metal–oxides preferred for different classes of reducing chemicals are summarized in Table 54.1.

Operating temperature: The charge of oxygen species that is adsorbed at the metal–oxide sensor's surface is highly dependent on temperature (Iwamoto, 1992). Oxidation rates also increase with temperature and patterns of adsorption, desorption, and diffusion are dependent on temperature (Wlodek et al., 1991a,b). Thus, as a first cut to improving selectivity, the operating temperature of a metal–oxide sensor, regardless of material, can be optimized to an analyte of interest. Further improvements in selectivity can be achieved by modulating operating temperature to obtain a signature or pattern that is indicative

TABLE 54.1 Snapshot of Metal–Oxide Chemical Sensors and Affinity to Specific Analytes

Class of Chemicals	Analyte	Metal–Oxide	Dopant	Approximate LOD (ppm)
Ammonia		Tin oxide	Palladium	100
		Zinc oxide	None	10
		Tungsten oxide	Titanium	5
Alcohols	Ethanol	Tin oxide	Palladium	4000
		Zinc oxide	Aluminum	400
		Zinc oxide	None	30
		Zinc oxide	Indium	2
Carbon oxides	CO	Zinc oxide	None	300
		Tin oxide	Gold	10
		Tin oxide	None	5
Hydrocarbons	Butane	Zinc oxide	None	1000
		Zinc oxide	Gold	1000
	Methane	Zinc oxide	None	5000
		Tin oxide	None	500
Hydrogen		Tin oxide	Palladium	1000
		Zinc oxide	None	600
		Tin oxide	Fluorine	100
Nitrogen oxides	NO_2	Zinc oxide	None	1000
		Zinc oxide	Indium	5
		Tin oxide	Tungsten oxide	500
Volatile organic compounds	Acetone	Tin oxide	Palladium	200
		Zinc oxide	None	10
	Hydrogen sulfide	Tin oxide	Copper oxide	10
		Tin oxide	Silver	1
		Zinc oxide	None	0.01
	Propane	Zinc oxide	None	4000
		Tin oxide	Nickel	600
		Tin oxide	Strontium	10

Source: Bochenkov, V.E. and Sergeev, G.B. *Metal Oxide Nanostructures and Their Applications*, Ahmad Umar and Yoon-bong Hahn, eds., Vol. 3, Chapter 2, American Scientific Publishers, 2010.

LOD, limit of detection.

of particular analytes present in a mixture. Complex features of output patterns generated by varying operating temperature have been used to improve the selectivity of tin oxide sensors in a range of research efforts (Huang et al., 2007; Vergara et al., 2007).

Arrays: Variations in the speed and degree of reaction among different materials and different operating temperatures (for metal–oxide sensors) and similar variables with other classes of materials can be used to create patterns that advance sensors from being broadly selective to being semiselective. A wide variety of arrays have been used to create both electronic noses and electronic tongues capable of distinguishing one mixture from another as well as one analyte from another in both liquid- and gas-phase environments.

Like metal–oxide sensors, many polymers are also inherently sensitive to reduction and oxidation, albeit through the conjugated backbone of the polymer rather than the dangling bonds of the metal–oxide sensor surface. However, conducting polymers have almost no conductivity (in the neutral state) until they are oxidized (p-doping) or reduced (n-doping). The baseline conductivity of the polymer can

be adjusted via chemical doping (similar to semiconductor doping) but this process is not reproducible. Alternatively, electrochemical doping enables fine-tuning of the doping level (and therefore the baseline conductivity of the polymer). Once the baseline is established through appropriate doping, conducting polymers can be further selected or modified to be broadly selective or semiselective.

Used in solution as broadly selective sensors, several conducting polymers make good pH sensors. Polyaniline (PANI), polypyrrole (PPY), polyimide (PI), polycarbazole, and polyazines are readily protonated or deprotonated making them appropriate for pH sensing. Protonation increases conductivity, while deprotonation decreases conductivity. As semiselective sensors, PI and polycarbazoles respond to copper ions in solution, while poly-3-octylthiophene (P3OTH) responds to silver ions. Complementing the sensitivity of these polymers to these inorganic ions are some conducting polymers that respond to classes of organic ions in solution. PANI and PPY respond electrically to classes of organic analytes in solution such as dicarboxylates and amino acids. In solution, however, it is far more common for these conducting polymers to demonstrate measurable responses in optical properties rather than nonoptical or electrical properties.

In air, the situation is different. A wide range of gases can both react chemically and adsorb physically onto conducting polymers, thereby making them semiselective by choice of specific polymers. Chemical reactions with electron acceptors such as NO_2, I_2, O_3, and O_2 partially oxidize many conducting polymers, thereby increasing the doping level and decreasing the resistance of these polymers. On the other end of the spectrum, electron donors such as H_2S, NH_3, and N_2H_4 reduce conducting polymers, thus leading to an increase in resistance. In addition to chemical reactivity, physical adsorption of certain classes of molecules leads to semiselectivity of these polymers to nonreactive organic compounds including chloroform, acetone, aliphatic alcohols, benzene, toluene, and similar (Lange et al., 2008).

Thus, just like metal–oxide chemical sensors, conducting polymers are used to make sensors ranging from broadly selective to semiselective. An excellent review of the use of conducting polymers as nonoptical sensors can be found in Lange et al., 2008.

54.3.4 Catalytic Interactions

Additional selectivity can be incorporated into chemical sensors via the addition of catalysts to the sensor material. Ion-selective electrodes (and electrochemical sensors that use them) are often modified with enzymes for chemical sensing in biological systems. For example, urease has commonly been used as an enzyme catalyst to detect urea. Urease reacts with urea to produce ammonia that is then sensed with an ammonium-selective electrode (Singh et al., 2008). Similar catalysts can be used in nonbiological systems to preferentially increase reaction rates to target chemicals or analytes, thereby increasing selectivity to those analytes. For example, platinum and palladium can be added to tin oxide sensors to detect carbon monoxide in the presence of interfering C_3H_8 and NO_2 as demonstrated in a variety of efforts, but most recently with nanoparticle-based tin oxide sensors (Aruna et al., 2009). These catalysts improve sensor performance via the spillover effect (Sharma and Madou. 2012). The catalyst dissociates a gas molecule, causing the dissociated atoms to spill over onto the surface of the sensor where they then react with the reactive oxygen species on the metal–oxide surface. The better the dispersion of catalyst near the reactive sensor surface, the less catalyst is needed to achieve the desired sensor response (Cukrov et al., 2001; Rothschild and Komem, 2004). The most popular catalyst for metal–oxide materials is palladium Pd (Bochenkov and Sergeev, 2010).

54.3.5 Electrochemical Interactions

Electrochemical sensors are the one of the most popular commercially available chemical sensors. Commonly termed "ion-selective electrodes," a complete electrochemical cell consists of at least two electrodes (Figure 54.3) interacting with ions in solution. This pair of electrodes acts according to similar (but reverse) behavior to a battery; in certain operating modes, a third electrode may be

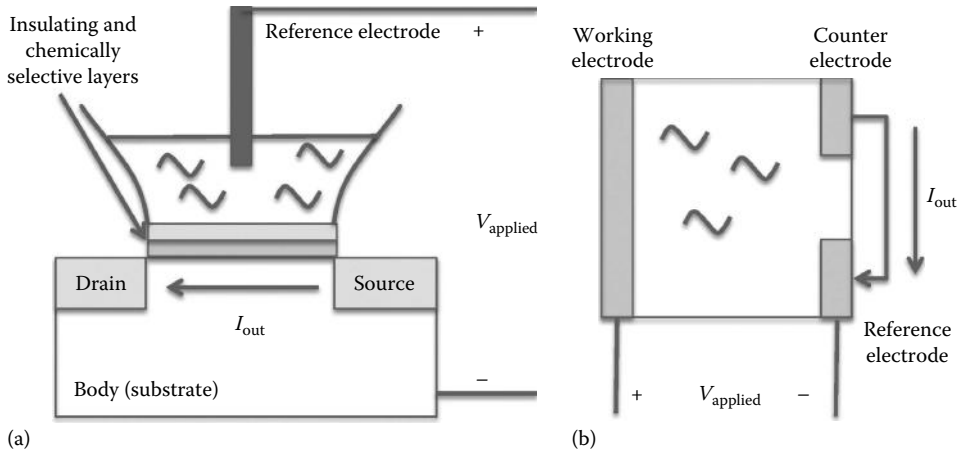


FIGURE 54.3 Electrochemical sensor operation: (a) Potentiometric operation is commonly configured using a field effect transistor (FET). A voltage is applied from the reference electrode to the body of the transistor and the analyte of interest interacts with the insulator layers, creating a charge on those layers. The charge enhances or detracts from the net field effect, either increasing or decreasing the resulting transistor current respectively (I_{out}). No current flows from the solution to the transistor, thereby ensuring Nernstian operation. In (b) amperometric mode, a voltage is applied between the working electrode and reference electrode and the resulting current from working to counter electrode is measured as a function of analyte concentration.

used, but the basic operation of the electrochemical cell remains the same. One of the electrodes in the electrochemical cell has a stable and constant interaction with a known concentration of ion in solution; this electrode is the reference electrode and plays a critical role in the resulting stability, drift, and accuracy of the electrochemical sensor. The other "working" electrode is chosen to interact with the ion of interest in solution. Selectivity of this working electrode can be enhanced in two primary ways.

First, the electrode can be modified to prefer the ion of interest more than other ions that inherently exchange electrons with the working electrode, thereby transitioning the electrode from broadly selective to semiselective. If the voltage (potential) across the two electrodes is then measured as a function of ion concentration without drawing a current, the well-known Nernst equation applies:

$$V_{sensor} = K + \left(\frac{RT}{nF}\right) \ln [C] \tag{54.4}$$

where

V_{sensor} is the voltage across the two electrodes
R, K, and F are constants
n is the number of electrons exchanged between solution and working electrode
$[C]$ is the concentration of the analyte of interest in solution

This Nernst relationship provides the electrochemical sensor the stable, predictable operation inherent to potentiometric operation. Potentiometric mode of operation (Figure 54.3a) is more stable than other modes of operation because electrode material is not typically consumed during sensor operation and it is more predictable than other modes of operation because it is governed by a well-known theory (Nernst behavior) rather than empirical relationships. Potentiometric operation is also largely considered more scalable than other modes of operation, meaning that no penalty in limit of detection or sensitivity is inherent in miniaturizing these sensors. However, sensor selectivity in potentiometric

mode of operation is largely limited by the types of ions the working electrode will exchange electrons with, thereby making these sensors vulnerable to interference from competing ions.

Second, in amperometric operation, the selectivity can be enhanced electrochemically by controlling the voltage applied to the electrode pair (Figure 54.3b). In amperometric mode, if the redox potential (voltage) of a particular molecule and the working electrode is greater than the applied voltage, that molecule will not react with the working electrode. Only molecules whose redox potential is lower than the applied voltage will react with the working electrode, hence improving selectivity well beyond that enabled by the material selectivity of the working electrode. Although amperometric sensors are more selective, have higher signal-to-noise ratio, and lower detection limits than potentiometric sensors, they suffer from the drawbacks of electrode consumption during operation (limited lifetime) and restricted scalability (miniaturization hurts detection limit and signal-to-noise ratio). They are also governed by empirical relationships rather than being grounded in theory:

$$I_{\text{peak}} = (2.68 \times 10^5) n^{3/2} A D^{1/2} [C] v^{1/2} \tag{54.5}$$

where

I_{peak} is the maximum current flowing between the electrodes
v is the sweep rate (of the voltage across the electrodes)
D is the diffusion constant of the analyte in solution
A is the cross-sectional area of the electrode
n is the number of electrons exchanged between solution and working electrode

These input/output relationships derived from data rather than theory can limit the predictability of their behavior (and resulting accuracy) in unknown sensing environments. Nevertheless, the redox potential adds another dimension to selectivity for amperometric sensors that in principle can enhance overall sensor performance. Moreover, advances in nanomaterials have reduced or eliminated the scale penalty of amperometric sensors, further increasing their competitiveness with potentiometric sensors.

54.4 Interference

Any deviation from lock-and-key selectivity is essentially chemical interference. A chemical or analyte that is not of interest in an application (oxygen to a carbon monoxide sensor, for example) is a chemical interferent. However, other sources of interference extend beyond a nondesirable analyte binding with an active chemical surface. Two primary sources of additional interference in chemical sensor operation are (1) physical interference from nonchemical interactions between the sensing environment and the active chemical sensor area and (2) electrical interference where electronic noise underlying the fundamental operation of the transduction mechanism in a chemical sensor limits overall performance. These two forms of interference are discussed next.

54.4.1 Physical Interference

Physical interference is loosely used to describe any (undesirable) interaction between a chemical sensor and the sensing environment that does not involve a strong interaction (covalent bonding) between molecule and sensor material, whether through adsorption or absorption into the sensor surface or bulk, respectively. For example, water vapor may adsorb onto the surface of a metal–oxide sensor, thus blocking reaction sites and reducing chemisorption of oxygen onto the surface. A reduction in oxygen species available for interaction with reducing gases in the environment results in an overall decrease in sensor surface area and resulting sensitivity of the sensor. Likewise, water vapor may absorb into composite polymer structures, thus inducing swelling and generating a misleading

decline in conductivity that can be falsely interpreted as a shift in concentration of the analyte(s) of interest. Similar interference can be caused by fluctuations in pH in liquid solution as fluctuations in humidity in gaseous environments. In addition to water vapor and protons/hydroxide ions, a wide range of analytes can block binding sites in chemical sensors, creating physical interference at multiple levels, both reversible and irreversible.

54.4.2 Electrical Interference

Most electrical interference associated with chemical sensor behavior and performance is noise, loosely defined as "any unwanted disturbance that obscures or interferes with a desired signal" (Motchenbacher and Connelly, 1993). Although noise can come from a variety of sources both internal and external to a sensor, most of the electronic noise that occurs in chemical sensors is directly related to the type of transduction mechanism required to convert the chemical signal to an electrical signal. Common sources of electronic noise (interference) and their impact on sensor performance are briefly discussed here. A full discussion of electronic noise is outside the scope of this chapter but an excellent text on the subject is by Motchenbacher and Connelly (1993).

Resistors: All electronic devices have some resistance associated with them and all resistors operating above temperatures of absolute zero are affected by thermal noise. Thermal or Johnson–Nyquist noise is generated by thermal fluctuations of electrons inside an electrical conductor at equilibrium. Thermal noise is independent of both frequency and applied voltage and can be expressed by an RMS voltage as follows:

$$V_{\mathrm{RMS}} = (4kTR \times \mathrm{Bandwidth})^{0.5} \tag{54.6}$$

where

k is Boltzmann's constant = 1.38×10^{23} m^2 kg s^{-2} K^{-1}
T is the temperature in degrees kelvin
R is the resistance in ohms and Bandwidth is the frequency band over which the resistor is operated (in Hz)

In a simple circuit, where a resistance is being measured directly, the bandwidth is the high frequency limit of operation of the measurement instrument. For example, consider a resistor of value 20,000 Ω, operating at 20 °C (293 K) operating in a circuit whose output voltage is measured by an operational amplifier with a maximum operating frequency of 1 MHz. The RMS noise of the resistor is then

$$V_{\mathrm{RMS}} = (4 \times 1.38 \times 10^{-23}\ \mathrm{m^2\ kg\ s^{-2}\ K^{-1}} \times 293\ \mathrm{K} \times 20{,}000\ \Omega \times 1 \times 10^{6}\ \mathrm{s^{-1}})^{0.5} = 18\ \mu\mathrm{V} \tag{54.7}$$

which simply means that the smallest voltage that can be measured across the 20,000 Ω resistor under these operating conditions is 18 μV. Conductivity in many materials, including metal–oxides and conducting polymers, can be used to directly transduce a chemical signal or analyte concentration into an electrical output (as a resistance). Unfortunately, these devices tend to have high baseline resistances on the order of 10–100 kΩ, which generates significant thermal noise and makes the small changes in resistance generated by chemical concentration even more difficult to measure. One way to reduce electrical noise (interference) in resistors is to use the fact that noise is constant across frequency to an advantage in measurement circuits. Since thermal noise increases as the square root of the bandwidth, reducing the bandwidth over which the measurement of resistance is made can significantly reduce noise. One way to do this is to use a lock-in amplifier, which can measure resistance over a very narrow range of frequencies, thus reducing noise by several orders of magnitude over standard measurement techniques.

Transistors: When chemical sensors are operated in potentiometric mode, a field effect transistor (FET) can be used to convert the electrochemical signal into an electrical output without drawing current, thus guaranteeing Nernstian operation. To do this, the interaction of the analyte of interest with the working electrode of the electrochemical cell generates all or part of the gate voltage of the transistor. As the gate voltage changes, the underlying channel current between the drain and source of the transistor changes, yet no current from gate to channel flows (as is characteristic of the fundamental behavior of a FET). However, unlike chemiresistors or conductivity-based chemical sensors, multiple types of noise are generated in the FET channel. The first is thermal noise, a characteristic of the inherent resistance of the channel. The second component of the noise is the flicker noise, which is shown as the second term in the following equation, after the thermal noise:

$$I_{\mathrm{RMS}} = \left[\left(\frac{4kT}{R_{\mathrm{channel}}} \right) \times \mathrm{Bandwidth} + [I_{\mathrm{channel}} \times K] \times \ln(f_2/f_1) \right]^{0.5} \tag{54.8}$$

where

e is an empirical constant

K is a combination of constants and geometry-dependent terms

Note that the noise in this potentiometric (FET-based) sensor is given in terms of current rather than voltage because the fundamental output of the FET-based transduction mechanism is current rather than voltage; I_{RMS} represents the minimum amount of current that can be detected by the transistor. Flicker ($1/f$) noise clearly dominates at lower frequencies. Thus, one way to reduce noise in chemFET (potentiometric) operation is to operate the sensor at high frequencies and narrow bandwidth thus reducing both $1/f$ noise and thermal noise, respectively.

Other sources of noise exist in electronic devices, yet for standard operation (defined as low to moderate frequencies of operation under room temperature conditions or above), thermal and $1/f$ noise are often the strongest source of electronic interference. Electronic noise is a key player in determining the limit of detection (lowest detectable concentration) in chemical sensors and sensor arrays; hence, careful design of measurement circuits is important to optimizing sensor performance.

54.5 Advances in Selectivity and Interference of Chemical Sensors

An immense amount of research has been dedicated over the last three to four decades to developing materials and functionalizing sensor platforms for specificity to analytes of interest. As important to sensor operation as the material, however, is the transduction mechanism and sensor platform that ultimately determine sensitivity, resolution, and detection limit. Because so many sensing applications, especially in the field, rely on ultralow limits of detection and accuracy (immunity from false alarms) at these limits, we focus this section on two promising advances in chemical sensor technology: (1) developments in nanomaterials that substantially expand usable surface area for chemical sensor response and (2) signal measurement approaches that lower the noise floor of these sensors in the interests of improving limits of detection and resolution at low concentration levels.

54.5.1 Importance of Surface Area

Whether it is of interest to a chemical sensing application or undesirable (an interferent), a chemical analyte impacts the signal produced by an electronic device involved in its transduction via (1) the electronic signal responsible for transducing the chemical signal into a usable, readable output;

(2) the interference associated with the chemical signal (selectivity and physical interference as discussed in previous sections); and (3) the interference associated with the electrical signal (various forms of noise). Reducing the interference involved in creating a chemical signal (binding a target analyte to the active sensor area) is one way to reduce interference (improving selectivity). Reducing electronic interference is another way. However, yet another approach to reducing interference in a chemical sensor is to increase the chemical signal relative to the electrical interference. The most common way of increasing chemical signal while maintaining electrical interference relatively constant is to increase the surface area of the sensor without increasing the overall sensor volume, thus increasing the ratio of chemically active area/electronically active area. In bulk sensors, however, increases in surface area are highly limited because they involve a corresponding increase in overall sensor volume. Nanoscale particles or structures embedded in a chemical sensor, however, can increase active sensor area considerably more than corresponding increases in active electronic volume. When these nanomaterials are integrated into a sensor, they create little microcosms of active sensing area within the larger sensor electronic structure, substantially increasing surface area-to-volume ratio. Nanomaterials cover a wide range of structures from nanoparticles that lack any definite shape to those with distinct shape, including nanospheres, nanowires, and nanotubes. One of the first substantial efforts at incorporating nanomaterials into sensors was by Labeau et al. (1993) involving a simple chemiresistor made of tin oxide nanoparticles, undoped or doped with palladium to increase sensitivity. The sensor delivered an 8 × increase in conductance when exposed to 300 ppm of carbon monoxide.

Since these early efforts in the use of nanomaterials for chemical sensors, the advantage of greater surface area (alongside greater surface area-to-volume ratio) has contributed to a wide range of effort in this area. A brief snapshot of the different kinds of nanomaterials used as chemical sensors and their enhancements over bulk sensor performance is provided in Table 54.2. An excellent review of chemical sensors using gold nanoparticles as the foundation of both nonoptical and optical sensors can be found in Saha et al. (2011) and a similar review for other nanostructures in Huang and Choi (2007). Advances in and theory of operation associated with reducing metal–oxide materials and sensors to nanoscale are explored by Sharma and Madou (2012). Development of carbon nanotubes (CNTs) for electrochemical sensors and the detection of molecules/analytes in biological systems are surveyed by Gao et al. (2012) and Jacobs et al. (2010), respectively. These reviews cover an enormous amount of ground with respect to the development of nanomaterials specifically for chemical sensing. Nanoscale materials for chemical sensing, in addition to providing a lower limit of detection, can also operate faster, use less power, operate at lower temperatures (as in the case of metal–oxide sensors), and sometimes have less need for expensive catalysts. However, nanomaterials may also be susceptible to instability, thermal degradation, and higher drift rates (Korotcenkov and Cho, 2010).

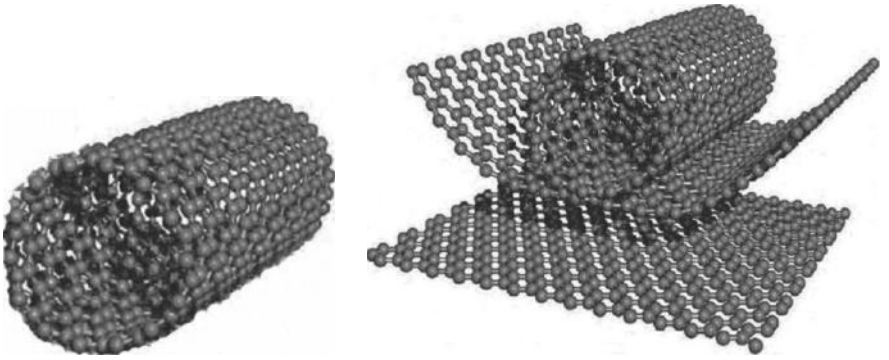
Of the nanomaterials, CNTs (Figure 54.4) have particular promise for use as chemical sensors. CNTs not only have the advantage of high surface area-to-volume ratio but are also chemically and mechanically stable and possess superior electronic characteristics including electronic mobilities up to 100,000 $cm^2/V\ s$ (Fam et al., 2011). Thus, they can offer high sensitivity, lower operating temperatures, and faster response times over their bulk sensor counterparts. CNTs are in a fairly unique position to provide both superior chemical sensor and superior transduction properties for use in complete chemical sensors. CNTs come in two basic forms. Single-walled carbon nanotubes (SWCNTs) are shaped as armchair, zigzag, or chiral depending on how the central axis of the tube is oriented relative to the hexagonal lattice (Figure 54.4a). Multiwall carbon nanotubes (MWCNTs) are typically concentric cylinders of carbon (Figure 54.4b). The presence of a chemical of interest typically changes the Fermi energy of the CNT (similar to doping a semiconductor) or changes the band structure of the tube, leading to a change in conductance. Both SWCNTs and MWCNTs can be used in two fundamentally different ways as chemical sensors. They can be used in their pristine form where a property of the nanotube itself changes directly with chemical concentration or alternatively, as functionalized nanotubes where metals or molecular groups are attached to the tip or sides of the nanotube to react with chemicals of

TABLE 54.2 Snapshot of Nanostructure Use in Chemical Sensors and Systems

			Performance	
Type	Material	Target Analyte	Nanoscale[a]	Bulk[a]
Nanobelts	Tin oxide	NO_2	0.5 ppm	500 ppm
			Comini et al. (2002)	Bochenkov and Sergeev (2010)
	Vanadium oxide	Ethanol	10 ppm	250 ppm
			Liu et al. (2005)	Micocci et al. (1997)
Nanoparticles	Gold, with octanethiol	Toluene; tetrachloroethylene	1 ppm Wolhtjen and Snow (1998)	N/A
	Gold, with hexanethiol	Dichloromethane	10 ppm Raguse et al. (2007)	N/A
Nanorods	Zinc oxide	Ethanol	1 ppm	30 ppm
			Wang et al. (2012)	Bochenkov and Sergeev (2010)
	Zinc oxide	Uric acid	2 μM	1 mM
			Zhang et al. (2004)	Ansari et al. (2009)
	PEDOT	Ammonia	10 ppm	100 ppm
			Jang et al. (2005)	Wongchoosuk et al. (2011)
	Zinc oxide	Formaldehyde	1.8 ppm	32 ppm
			Peng et al. (2009)	Han et al. (2009)
Nanotubes	Tin oxide	Hydrogen	100 ppm	1000 ppm
			Huang et al. (2005)	Bochenkov and Sergeev (2010)
	Titanium oxide	Hydrogen	100 ppm	1000 ppm
			Varghese et al. (2003)	Miyazaki et al. (2005)
	Platinum	Glucose	2 mM	5 mM
			Yuan et al. (2005)	Kasapbasioglu et al. (1992)
Nanowires	Silicon, modified	pH	Range: 2–9	Range: 5–7
			Huang et al. (2001)	Wojas et al. (1998)
	Silicon, modified	DNA	Femtomolar	N/A
			Hahm et al. (2004)	
	Indium oxide	NO_2	0.5 ppm	2 ppm
			Li et al. (2003)	Sbervegileri et al. (1988)
	Indium oxide	Low-density lipoproteins (LDL) Hsiai et al. (2006)		N/A
	Tin oxide: lead doped	Ethanol	10 ppm	4000 ppm
			Wan et al. (2005)	Bochenkov and Sergeev (2010)
	Tin oxide	CO	5 ppm; Hernandez-Ramirez et al. (2007)	5 ppm Bochenkov and Sergeev (2010)
	Zinc oxide	Ethanol	200 ppm	30 ppm
			Wan et al. (2004)	Bochenkov and Sergeev (2010)
	Zinc oxide	Ammonia	40 ppm	10 ppm
			Wang et al. (2006)	Bochenkov and Sergeev (2010)

[a] All performance metrics are LOD unless otherwise noted.

interest (Goldoni et al., 2010). Pristine CNTs have been used to detect oxygen, NO_2, ammonia, sulfur dioxide, and NO (as reviewed by Goldoni et al., 2010). These materials, as they are inherently part of the underlying CNT, make for simpler manufacturing and sensor design yet continue to suffer from the same selectivity limitations as their bulk sensor counterparts. To overcome these limitations inherent in broadly selective sensor materials, the CNT can be functionalized to provide selectivity much closer to the lock-and-key ideal. For example, CNTs functionalized in a wide variety of ways to detect



(a) Unfunctionalized SWCNT (b) Unfunctionalized MWCNT

FIGURE 54.4 CNTs used as chemical sensors: (a) A "pristine" SWCNT sensor consists of a single, unprocessed or bare CNT cylinder structurally configured in armchair, zigzag, or chiral patterns, and (b) single-wall nanotubes wrapped concentrically around each other are MWCNTs. Greater specificity (higher selectivity) can be added to the CNT sensor by adding molecules to the sides or tip of the CNT. (Images courtesy of what-when-how, www.what-when-how.com.)

TABLE 54.3 Snapshot of CNT Advances in Chemical Sensor Performance

			Performance (Nanostructure)	
Type of CNT	Modification	Target	CNT	Non-CNT
SWCNT, pristine	N/A	NO_2	0.1 ppm	0.5 ppm
			Wei et al. (2011)	Li et al. (2003)
	Boron doped	Dopamine	LOD: 1.4 nM	1.55 nM
			Deng et al. (2008)	Atta et al. (2012)
SWCNT, functionalized	Zinc oxide	Formaldehyde	LOD: 0.01 ppm	LOD: 10 ppm
			RT: seconds	RT: 15 min
			Lu et al. (2010)	Peng et al. (2009)
	Ruthenium oxide coated	Insulin	LOD: 14 nM	LOD: 10 fM
			Wang and Musameh (2004)	Regonda et al. (2011)
MWCNT, functionalized	Platinum nanoparticles	Glucose	LOD: 6.18 mM:	LOD: 2 mM
			RT: < 5 s	RT: < 100 s
			Tsai et al. (2009)	Yuan et al. (2005)
	Tin oxide	Ethanol	LOD: 25 ppm	LOD: 10 0 ppm
	Pt/Pd doped		RT: 150 s	RT: 60 s
			Ahmadnia-Feyzabad et al. (2012)	Patil et al. (2010)

LOD, limit of detection; RT, response/recovery time.

chemicals of interest. A snapshot of advances in the chemical sensing capability of CNTs is provided in Table 54.3. More complete reviews of the use of CNTs to detect molecules in biological systems and in natural world environments can be found in Jacobs et al. (2010) and Goldoni et al. (2010)/Bondavalli et al. (2009), respectively.

54.5.2 New Approaches to Signal Measurement

Research and product development for portable chemical and biological sensors tend to favor a single form of signal measurement. A single measurement is often simpler, more compact, and least expensive to implement. However, additions to or adjustments to single dc measurements of chemical sensor

outputs can often be made without significant compromise to cost, size, weight, or power of chemical sensor systems while achieving valuable improvements in performance, especially resolution and limit of detection. We describe one such approach in detail here and its relevance both to conducting polymers where it was originally demonstrated as well as to other chemiresistors. Other creative approaches to signal measurement, transduction, and sensor arrays are also described briefly to convey the largely unrealized contribution that smart design of signal measurement and transduction can make to overall chemical sensor performance.

Multiple simultaneous approaches to capturing conductivity (resistance) changes as a function of a changing concentration of chemical analyte can creatively overcome some of the limitations involved in any single measurement. One such example of the benefits that can be achieved with creative signal measurement schemes without significant compromise to cost, size, or weight of a portable system is demonstrated by Lange et al. (2011). In this heterogeneous signal measurement scheme, the advantages of three ways of extracting conductivity changes resulting from chemical concentration are combined. The two-point measurement scheme used in the traditional chemiresistor consists of two electrodes (Figure 54.5a) over which a voltage (ac or dc) is applied and the resulting current is measured using a simple measurement circuit such as a voltage divider or Wheatstone bridge. Since it involves measuring a current over the same interface at which a voltage is applied, this measurement includes both channel resistance and polymer/electrode contact resistance in the resulting sensor output. Both polymer/electrode contact properties and polymer channel conductivity often change in response to biological or chemical stimulation, in different ways. A four-point measurement (Figure 54.5b) applies a fixed current between one pair of (outer) electrodes and measures the voltage across the remaining (inner) pair of electrodes. This approach, because it does not involve drawing any current at the inner pair of electrodes to make a measurement (of voltage), is insensitive to contact resistance. Using both two- and four-point

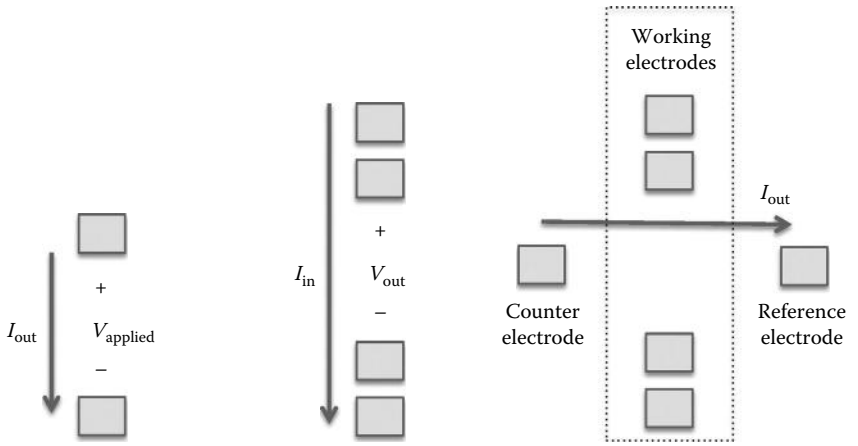


FIGURE 54.5 Multiple measurements of polymer-based chemical sensors: (a) In the standard two-point configuration (Measurement #1), a voltage is applied across two electrodes and the output current measured (usually as a voltage) using a voltage divider or Wheatstone bridge; (b) in the four-point configuration (Measurement #2), current is applied between the two outer electrodes and voltage measured on the inner electrodes (thus eliminating the contact resistance from the output); (c) in the six-point configuration (Measurement #3), the electrochemistry is measured amperometrically by the current between counter and reference electrodes.

measurements simultaneously can allow contact resistance and conductivity changes in the main polymer to be individually extracted from the sensor signal. Bulk measurements (captured by two-point and four-point signals) and contact resistance (captured by only two-point measurements) can change differently with analyte concentration due to delays in the diffusion of analyte through the conducting polymer. Thus, combining two- and four-point measurements provides information about analyte concentration both through bulk resistance and the ratio of bulk to contact resistance. This improves both detection limit and sensitivity of the device by improving signal-to-noise ratio through multiple distinct measurements of the same reaction between sensor and analyte. Successful combination of two- and four-point measurements was reported by Krondak et al. (2006) for the detection of gaseous HCL using polydialkoxybipyrroles (chemiresistor layer) combined with interdigitated gold electrodes.

This approach to signal measurement can be further enhanced by adding at third six-electrode measurement to the overall sensor. A six-electrode measurement (Figure 54.5c) is essentially a chemFET measurement, allowing the redox state of the polymer to be controlled so that only certain analytes will react with the conducting polymer surface. Lange et al. (2011) demonstrated the compact nature and the advantage of this three simultaneous measurement configuration using the conductive polymer polythiophene to detect nitrogen dioxide in a compact footprint, demonstrating faster regeneration time than single measurement configurations. Using this approach with other chemiresistive materials, including metal–oxides, can achieve similar performance benefits.

Multiple measurements of conductivity can improve the limit of detection and other performance outcomes of chemiresistors but are by no means the only signal measurement or transduction approach available for improving chemical sensor performance. Arrays, discussed previously, are another approach to improving performance and detection capability. According to basic statistics, identical sensors in an array (same material, same manufacturing technique) provide output whose resolution and detection limit improves as the square root of the number of sensors. In addition, nonidentical sensors placed in heterogeneous arrays are commonly used to increase selectivity and specificity of chemical performance. As long as the noise added by each sensor in the array does not exceed the information provided, higher density arrays can continue to make contributions to system performance when materials availability, manufacturing technique, and cost/size/weight/power constraints permit.

A wide range of other techniques have also been explored at the signal measurement and transduction level to improve chemical sensor performance. To reduce the size of chemFETs, which are essentially miniaturized potentiometric sensors, light can be used to amplify carrier levels in a much smaller capacitive structure in order to achieve comparable performance to chemFETs; these light-addressable potentiometric sensors (LAPSs) have been successfully used to detect oxygen, pH, and glucose in solution in a much smaller footprint than other electrochemical approaches to chemical sensing (Yotter and Wilson, 2004). LAPSs have been commercialized by molecular devices and have the potential to resolve events at a cell-by-cell level. Modulation of operating temperature for metal–oxide sensors has also proven useful for improving specificity, selectivity, and accuracy of these sensors (Gutierrez-Osuna et al., 2011; Rogers and Semancik, 2012). Careful signal measurement circuit design, including lock-in amplification, can substantially reduce electronic noise (Moe et al., 2004). Likewise, elimination of baseline resistance (McKennoch et al., 2001) from a conductive chemical sensor allows measurement circuit resolution to be focused on capturing chemical signal information rather than irrelevant baseline information. These and additional techniques enable signal measurement, transduction, and sensor interface design to improve overall performance for chemical sensor systems.

54.6 Conclusions

For decades, the primary emphasis on reducing interference and enhancing selectivity and specificity in chemical sensors has been on developing new or modified materials to this purpose. These efforts continue today with a focus on nanomaterials, ranging from nanoparticles to CNTs. Relative to materials development efforts, work that focuses on improving chemical sensor performance by manipulating sensor

arrays, measurement schemes, transduction mechanisms, and interface circuits has been relatively scant. However, these nonmaterial approaches to improving sensor performance can have comparable impact to advances in materials, especially when integrated into a system approach to design. Thus, future efforts in chemical sensors should place a high priority on commercialization and should take a balanced and integrated look at both materials and supporting components that make up a full chemical sensor system.

References

Ahmadnia-Feyzabad, S., A.A. Khodadi, M. Vesali-Naseh, and Y. Mortazavi, Highly sensitive and selective sensors to volatile organic compounds using MWCNTs/SnO_2, *Sensors Actuat. B*, 166–167, 150–155, 2012.

Alizadeh, T., M.R. Ganjali, M. Zare, and P. Norouzi, Development of a voltammetric sensor based on a molecularly imprinted polymer (MIP) for caffeine measurement, *Electrochim. Acta*, 55(5), 1568–1574, 2010.

Ansari, S.G., R. Wahb, Y.-S Kim, Z.A. Ansari, O-B Yang, G. Khang, and H.-S. Shin, Thick film urea sensor based on nanostructured zinc oxide, *Int. J. Nanomanuf.*, 4(1–4), 290–299, 2009.

Aruna, I., F.E. Kruis, S. Kundu, M. Muhler, R. Theissmann, and M. Spasova, CO ppb sensors based on monodispersed SnO_x:Pd mixed nanoparticle layers: Insight into dual conductance response, *J. Appl. Phys.*, 105(6), 064312–064317, 2009.

Atta, N.F., A. Galal, and E.H. El-Ads, Gold nanoparticles-coated poly(3,4-ethylene-dioxythiophene) for the selective determination of sub-nano concentrations of dopamine in presence of sodium dodecyl sulfate, *Electrochim. Acta*, 69, 102–111, 2012.

Benhar I., I. Eshkenazi, T. Neufeld, J. Opatowsky, S. Shaky, and J. Rishpon, Recombinant single chain antibodies in bioelectrochemical sensors, *Talanta*, 55(5), 899–907, 2001.

Bochenkov, V.E. and G.B. Sergeev, Sensitivity, selectivity, and stability of gas-sensitive metal-oxide nanostructures, in *Metal Oxide Nanostructures and Their Applications*, Ahmad Umar and Yoon-bong Hahn, eds., Vol. 3, Chapter 2, American Scientific Publishers, California, 2010.

Bondavalli, P., P. Legagneux, and D. Pribat, Carbon nanotubes based transistors as gas sensors: State of the art and critical review. *Sensors Actuat. B*, 140, 304–318, 2009.

Comini, E., G. Faglia, G. Sberveglieri, Z.W. Pan, and Z.L. Wang, Stable and highly sensitive gas sensors based on semiconducting oxide nanobelts, *Appl. Phys. Lett.*, 81, 1869–1871, 2002.

Cukrov, L.M., P.G. McCormick, K. Galatsis, and W. Wlodarski, Gas sensing properties of nanosized tin oxide synthesized by mechanochemical processing, *Sensors Actuat. B*, 77, 491–495, 2001.

Deng, C., J. Chen, X. Chen, C. Xiao, L. Nie, and S. Yao, Direct electrochemistry of glucose oxidase and biosensing for glucose based on boron-doped carbon nanotubes modified electrode, *Biosens. Bioelectron.*, 23, 1272–1277, 2008.

Ehrentreich-Forster, E., D. Orgel, A. Krause-Griep, B. Cech, V.A. Erdmann, F. Bier, F.W. Scheller, and M. Rimelle, Biosensor-based on-site explosives detection using aptamers as recognition elements, *Anal. Bioanal. Chem.*, 391(5), 1793–1800, 2008.

Fam, D.W.H., A.I. Palaniappan, A.I.Y. Tok, B. Liedberg, and S.M. Moochhala, A review on technological aspects influencing commercialization of carbon nanotube sensors, *Sensors Actuat.*, 157(1), 1–7, 2011.

Fang, C., C. Yi, Y. Wang, Y. Cao, and X. Liu, Electrochemical sensor based on molecular imprinting by photo-sensitive polymers, *Biosens. Bioelectron.*, 24(10), 3164–3169, 2009.

Gao, C., Z. Guo, J.-H. Lui, and X.-J. Huang, The new age of carbon nanotubes: An updated review of functionalized carbon nanotubes in electrochemical sensors, *Nanoscale*, 4, 1948–1963, 2012.

Goldoni, A., L. Petaccia, S. Lizzit, and R. Larciprete, Sensing gases with carbon nanotubes: A review of the actual situation, *J. Phys. Condens. Matter*, 22(1–9), 2010.

Gutierrez-Osuna, R., R. Gosangi, and P. Gouma, Advances in active and adaptive chemical sensing, *AIP Conf. Proc.*, 1362, 11–12, 2011.

Hahm, J. and C.M. Lieber, Direct ultrasensitive electrical detection of DNA and DNA sequence variations using nanowire nanosensors, *Nano Lett.*, 4, 51–54, 2004.

Han, N., Y. Tian, X. Wu, and Y. Chen, Improving humidity selectivity in formaldehyde gas sensing by a two-sensor array made of Ga-doped ZnO_2, *Sensors Actuat. B*, 138(1), 228–235, 2009.

Hernandez-Ramirez, F., A. Tarancon, O. Casals, J. Arbiol, A. Romano-Rodriguez, and J.R. Morante, High response and stability in CO and humidity measures using a single SnO_2 nanowire, *Sensors Actuat. B*, 121(1), 3–17, 2007.

Hsiai, T.K., M. Rouhanizadeh, T. Tao, C. Li, J. Hwang, and C. Zhou, Differentiation of oxidized low density lipoproteins by nanosensors, *Sensors Actuat. B*, 114(2), 788–798, 2006.

Huang, J., N. Matsunaga, K. Shimanoe, N. Yamazoe, and T. Kunitake, Nanotubular SnO_2 templated by cellulose fibers: Synthesis and gas sensing, *Chem. Mater*., 17, 3513–3518, 2005.

Huang, X.-J. and Y.-K Choi, Chemical sensors based on nanostructured materials, *Sensors Actuat. B*, 122, 659–671, 2007.

Huang, Y., X.F. Duan, Q.Q.Wei, and C.M. Lieber, Directed assembly of one dimensional nanostructures into functional networks, *Science*, 291, 630–633, 2001.

Iwamoto, M. Characterization of oxygen adsorbates on semiconductive oxides. In *Chemical Sensor Technology*, Vol. 4, S. Yamauchi, ed. Elsevier Science: London, U.K., 1992.

Jacobs, C.B., M. Jacobs, J. Peairs, and B.J. Venton, Review: Carbon nanotube based electrochemical sensors for biomolecules, *Anal. Chim. Acta*, 662, 105–127, 2010.

Jang, J., M. Chang, and H. Yoon, Chemical sensors based on highly conductive poly(3,4-ethylenedioxythiophene) nanorods, *Adv. Mater*., 17, 1616–1620, 2005.

Kasapbasioglu, B., P. Hesketch, C. Hanly, G.Jordan MacClay, and R. Nowroozi-Esfahani, Impedance based ultra-thin platinum island film glucose sensor, *Sensors Actuat. B*, 14(1–3), 749–751, 1992

Kim, Y.S., Seongbuk-Gu, and Anam-Dong, Specific detection of oxytetracycline using DNA aptamer-immobilized interdigitated array electrode chip, *Anal. Chim. Acta*, 634(2), 250–254, 2009.

Korotcenkov, G. and B. Cho, Grain size effect in structural stability of SnO_2 and In_2O_3 films aimed for gas sensing application. In *Proceedings of the 2nd International Conference on Computer Research and Development,* Kuala Lumpur, Malaysia, pp. 461–464, 2010.

Krondak, M., G. Broncova, S. Anikin, A. Merz, and V.M. Mirsky, Chemosensitive properties of poly-4,4-dialkoxy-2,2-bipyrroles, *J. Solid State Electrochem*., 10(3), 185–191, 2006.

Labeau, M. and B. Gautheron, Synthesis of pure and Pd-doped SnO_2 particles, *Solid State Ionics*, 63–65, 159–163, 1993

Lange, U. and V.M. Mirsky, Chemiresistors based on conducting polymers: A review on measurement techniques, *Anal. Chim. Acta*, 687(2), 105–113, 2008.

Lange, U. and V.M. Mirsky, Integrated electrochemical transistor as a fast recoverable gas sensor, *Anal. Chim. Acta*, 687(1), 7–11, 2011.

Li, C., D.H. Zhang, X.L. Liu, S. Han, T. Tang, J. Han, and C.W. Zhou, In_2O_3 nanowires as chemical sensors, *Appl. Phys. Lett*., 82, 1613–1615, 2003.

Liang, R., R. Zhang, and W. Qin, Potentiometric sensor based on molecularly imprinted polymer for determination of melamine in milk, *Sensors Actuat*., 141(2), 544–550, 2009.

Liu, J.F., X. Wang, Q. Peng, and Y.D. Li, Vanadium pentoxide nanobelts: Highly selective and stable ethanol sensor materials, *Adv. Mat*., 17, 764–767, 2005.

Lu, Yijiang, M. Meyyappan, and J. Li, A carbon-nanotube-based sensor array for formaldehyde detection, *Nanotechnology*, 22(1–4), 2011.

Mazzotta, E., R.A Picca, C. Malitesta, S.A. Piletsky, and E.V. Piletska, Development of a sensor prepared by entrapment of MIP particles in electrosynthesised polymer films for electrochemical detection of ephedrine, *Biosens. Bioelectron*., 23(7), 1152–1156, 2008.

McKennoch, S. and D. Wilson, Autoranging compensation for variable baseline chemical sensors, *SPIE, Photonics East*: Boston, MA, November 2001.

Metilda, P., K. Prasad, R. Kala, J.M. Gladis, T. Rao, and N. Prasada, Ion imprinted polymer based sensor for monitoring toxic uranium in environmental samples, *Anal. Chim. Acta*, 582(1), 147–153, 2007.

Micocci, G., A. Serra, A. Tepore, S. Capone, R. Rella, and P. Siciliano, Properties of vanadium oxide thin films for ethanol sensor, *J. Vacuum Sci. Technol. A (Vacuum, Surfaces Films)*, 15(1), 34–38, 1997.

Miyazaki, H., T. Hyodo, Y. Shimizu, and M. Egashira, Hydrogen-sensing properties of anodically oxidized TiO_2 film sensors, *Sensors Actuat. B*, 108(1–2), 467–472, 2005.

Moe, A.E., S.R. Marx, I. Bhinderwala, and D.M. Wilson, A miniaturized lock-in amplifier design suitable for impedance measurements in cells, *IEEE Sensors 2004*: Vienna, Austria, October 24–27, 2004.

Motchenbacher, C.D. and J.A. Connelly, *Low-Noise Electronic System Design*, John Wiley & Sons, New York, 1993.

Neufeld, T., A.S. Mittelman, V. Buchner, and J. Rishpon, Electrochemical phagemid assay for the specific detection of bacteria using *Escherichia coli* TG-1 and the M13KO7 phagemid in a model system, *Anal. Chem*., 77(2), 652–657, 2005.

Pardieu, E., H. Cheap, C. Vedrine, M. Lazerges, Y. Lattach, F. Garnier, S. Remita, and C. Pernelle, Molecularly imprinted conducting polymer based electrochemical sensor for detection of atrazine, *Anal. Chim. Acta*, 649(20), 236–245, 2009.

Patil, L.A., M.D. Shinde, A.R. Bari, and V.V. Deo, Highly sensitive ethanol sensors based on nanocrystalline SnO_2 thin films, *Curr. Appl. Phys*., 10(5), 1249–1254, 2010.

Peng, L.Q. Zhao, D. Wang, J. Zhai, P. Wang, S. Pang, and T. Xie, Ultraviolet-assisted gas sensing: A potential formaldehyde detection approach at room temperature based on zinc oxide nanorods, *Sensors Actuat. B*, 136, 80–85, 2009.

Raguse, B., E. Chow, C.S. Barton, and L. Wieczorek, Gold nanoparticle chemiresistor sensors: Direct sensing of organics in aqueous electrolyte solution, *Anal. Chem*., 79, 7333–7339, 2007.

Regonda, S., K. Trivedi, R. Tian, L. Spurgin, S. Green, D. Jiahuan, G. Jinming, and W. Hu, Si multi-nanochannel FETs to improve device uniformity/stability and detection of 10 fM insulin in serum, *2011 IEEE 11th International Conference Nanotechnology (IEEE-NANO)*, pp. 621–624, 2011.

Riskin, M., R. Tel-Vered, T. Bourenko, E. Granot, and I. Willmar, Imprinting of molecular recognition sites through electropolymerization of functionalized Au nanoparticles: Development of an electrochemical TNT sensor based on π-donor-acceptor interactions. *J. Am. Chem. Soc*., 130(30), 9726–9733, 2008.

Rogers, P. and S. Semancik, Development of optimization procedures for application-specific chemical sensing, *Sensors Actuat. B*, 163(1), 8–19, 2012.

Rothschild, A. and Y. Komem, The effect of grain size on the sensitivity of nanocrystalline metal-oxide gas sensors, *J. Appl. Phys*., 95, 6374–6380, 2004.

Saha, K., S.S. Agasti, C. Kim, X. Li, and V.M. Rotello, Gold nanoparticles in chemical and biological sensing, *Chem. Rev*., 112, 2739–2779, 2012.

Sberveglieri, G., S. Groppelli, and G. Coccoli, Radio frequency magnetron sputtering growth and characterization of indium-tin oxide (ITO) thin films for NO_2 gas sensors, *Sensors Actuat*., 15(3), 235–242, 1988.

Sergeyeva, T.A., O.A. Slinchenko, L.A. Gorbach, V.F. Matyushov, O.O. Brovko, S.A. Piletsky, L.M. Sergeeva, and G.V. Elska, Catalytic molecularly imprinted polymer membranes: Development of the biomimetic sensor for phenols detection, *Anal. Chim. Acta*, 659(1–2), 274–279, 2010.

Sharma, S. and M. Madou, Review article: A new approach to gas sensing with nanotechnology, *Philos. Trans. Roy. Soc. A Math., Phys. Eng. Sci*., 370(1967), 2448–2473, 2012.

Siddiquee, S., N.A. Yusof, A.B. Salleh, S.G. Tan, and A.F. Bakar, Development of electrochemical DNA biosensor for *Trichoderma harzianum* based on ionic liquid/ZnO nanoparticles/chitosan/gold electrode, *J. Solid State Electrochem*., 16(1), 273–282, 2012.

Singh, M., N. Verma, A.K. Garg, and N. Redhu, Urea biosensors, *Sensors Actuat. B,* 134(1), 345–351, 2008.

Tsai, M.C. and Y.C. Tsai, Adsorption of glucose oxidase at platinum-multiwalled carbon nanotube-alumina-coated silica nanocomposite for amperometric glucose biosensor, *Sensors Actuat. B*, 141, 592–598, 2009.

Van Dorst, B., J. Mehta, K. Bekaert, E. Rouah-Martin, W. De Coen, P. Dubruel, R. Blust, and J. Robbens, Recent advances in recognition elements of food and environmental biosensors: A review, *Biosens. Bioelectron*., 26, 1178–1194, 2010.

Varghese, O.K., D. Gong, M. Paulose, K.G. Ong, E.C. Dickey, and C.A. Grimes, Extreme changes in the electrical resistance of titania nanotubes with hydrogen exposure, *Adv. Mater*., 15(7–8), 624–627, 2003.

Vergara, A., E. Llobet, J. Brezmes, P. Ivanov, C. Cane, I. Gracia, X. Vilanova, and X. Correig, Quantitative gas mixture analysis using temperature-modulated micro-hotplate gas sensors: Selection and validation of the optimal modulating frequencies, *Sensors Actuat. B*, 123(2), 1002–1016, 2007.

Wan, Q., Q.H. Li, Y.J. Chen, T.H.Wang, X.L. He, J.P. Li, and C.L. Lin, Fabrication and ethanol sensing characteristics of ZnO nanowire gas sensors, *Appl. Phys. Lett*., 84, 3654–3656, 2004.

Wan, Q. and T.H. Wang, Single-crystalline Sb-doped SnO_2 nanowires: Synthesis and gas sensor application, *Chem. Commun*., 14(30), 3841–3843, 2005.

Wang, J. and M. Musameh, Electrochemical detection of trace insulin at carbon-nanotube-modified electrodes, *Anal. Chim. Acta*, 511(1), 33–36, 2004.

Wang, L., Y. Kang, X. Liu, S. Zhang, W. Huang, and S. Wang, ZnO nanorod gas sensor for ethanol detection, *Sensors Actuat. B*, 162(1), 237–243, 2012.

Wang, X.H., J. Zhang, and Z.Q. Zhu, Ammonia sensing characteristics of ZnO nanowires studied by quartz crystal microbalance, *Appl. Surface Sci*., 252, 2404–2411, 2006.

Wei, A., L. Wei-Wei, W. Jing-Xia, L. Qing, W. Zhao, X. Li, D. Xiao-Chen, and H. Wei, Single-walled carbon nanotube networked field-effect transistors functionalized with thiolated heme for NO_2 sensing, *Chinese Phys. Lett*., 28(12), 127304–127306, 2011.

Wei, F. and C.-M. Ho, Aptamer-based electrochemical biosensor for Botulinum neurotoxin, *Anal. Bioanal. Chem*., 393(8), 1943–1948, 2009.

Wlodek, S., K. Colbow, and F. Consadori, Signal-shape analysis of a thermally cycled tin-oxide gas sensor, *Sensors Actuat. B*, 63(1), 63–68, 1991a.

Wlodek, S., K. Colbow, and F. Consadori, Kinetic model of thermally cycled tin oxide gas sensor, *Sensors Actuat. B*, 632, 123–127, 1991b.

Wohltjen, H. and A.W. Snow, Colloidal metal-insulator-metal ensemble chemiresistor sensor, *Anal. Chem*., 70, 2856–2859, 1998.

Wojas, P., L. Meixner, and P. Frosti, Slow pH response effects of silicon nitride ISFET sensors, *Sensors Actuat. B*, 48(1–3), 501–504, 1998.

Wongchoosuk, C., P. Jangtawee, S. Lokavee, S. Udomrat, P. Sudkeaw, and T. Kerdsharoe, Novel flexible NH_3 gas sensor prepared by ink-jet printing technique, *Adv. Mater. Res*., 506, 39–42, 2011.

Yotter, R. and D. Wilson, Sensor technologies for monitoring metabolic activity in single cells—Part II: Non-optical methods and applications, *IEEE Sensors J*., 4(4), 412–429, 2004.

Yuan, J.H., K. Wang, and X.H. Xia, Highly ordered platinum-nanotubule arrays for amperometric glucose sensing, *Adv. Funct. Mater*., 15, 803–809, 2005.

Zhang, F.F., X.L.Wang, S.Y. Ai, Z.D. Sun, Q.Wan, Z.Q. Zhu, Y.Z. Xian, L.T. Jin, and K. Yamamoto, Immobilization of uricase on ZnO nanorods for a reagentless uric acid biosensor, *Anal. Chim. Acta*, 519, 155–160, 2004.

55

Electrochemical Composition Measurement

Michael J. Schöning
Aachen University of Applied Sciences

Arshak Poghossian
Aachen University of Applied Sciences

Olaf Glück
Forschungszentrum Julich GmbH

Marion Thust
Forschungszentrum Julich GmbH

Electrochemical analysis in liquid solutions is concerned with the measurement of electrical quantities, such as potential, current, and charge, to gain information about the composition of the solution and the reaction kinetics of its components. The main techniques are based on the quantitative determination of reagents needed to complete a reaction or the reaction products themselves. Four traditional methods of electrochemistry are described here (Figure 55.1): potentiometry, voltammetry, coulometry, and conductometry. Potentiometry implies the measurement of an electrode potential in a system in which the electrode and the solution are in electrochemical equilibrium. Voltammetry is a technique in which the potential is controlled according to some prescribed function while the current is measured. Coulometry involves the measurement of charge needed to completely convert an analyte, and conductometry determines the electrical conductivity of the investigated test solution. The practical applications of these measurement techniques for analytical purposes range from industrial process control and environmental monitoring to food analysis and biomedical diagnostics. The analytical methods and their instrumentation as well as recent trends, such as electrochemical sensors, are discussed.

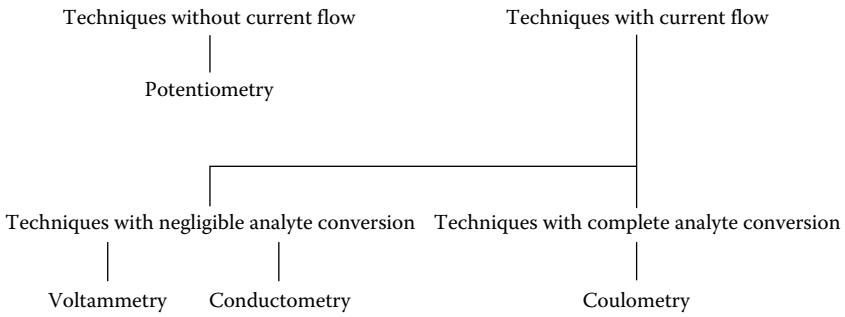


FIGURE 55.1 Electrochemical methods described in this section.

55.1 Basic Concepts and Definitions

55.1.1 Electrodes and the Electrical Double Layer

In electrochemistry, *electrodes* are devices for the detection of charge transfer and charge separation at phase boundaries or for the generation and variation of the charge transfer and separation with an impressed current across the phase boundary. One important feature of electrodes is a potential difference across the electrode/electrolyte phase boundary. At this interface, the conduction mechanism changes since electrode materials conduct the current via electrons, whereas electrolytes conduct via ions. To understand the processes that lead to the formation of the potential difference, it is helpful to consider first an atomistic model, which was given by Helmholtz. It leads to the idea of an *electrical double layer*.

If an electrode is immersed in an electrolyte solution, the bulk regions of the two homogeneous phases—the electrode material and the electrolyte—are in equilibrium. This means that far away from the phase boundary (>1 μm), the sum of the forces on the particles is zero and charges are distributed homogeneously. Since the cohesion forces that bind the individual particles together in the bulk are significantly reduced at the surface of the electrode, particles in this region will have less neighbors or neighbors from the other phase. Thus, close to the phase boundary, the equilibrium conditions are drastically different from the equilibrium conditions in the bulk of the electrolyte. This change in the equilibrium of forces on particles at the interface can lead to an *interfacial tension*. In addition, the surface of a condensed phase usually has different electrical properties than the bulk phase, for example, due to the accumulation of free charge on the surface of an electrically charged solid. Besides, the orientation of dipoles in the surface region and adsorption of ions and dipoles from the electrolyte can lead to a change in the electrical properties. This excess charge from ions, electrons, and dipoles produces an electric field that is accompanied by a potential difference across the phase boundary. The region in which these charges are present is termed the electrical double layer (Hamann et al. 2007). The formation of an electrical double layer at interfaces is a general phenomenon but only the electrode/electrolyte interface will be considered here in more detail.

According to the hypothesis of Helmholtz, the electrical double layer has the character of a plate capacitor, whose plates consist of a homogeneously distributed charge in the metal electrode and ions of opposite charge lying in a parallel plane in the solution at a minimal distance from the surface of the electrode (Rieger 1987). Modern conceptions are based on the assumption that the electron cloud in the metal extends to a certain degree into a thin layer of solvent molecules in the immediate vicinity of the electrode surface. In this layer, the dipoles of the solvent molecules (e.g., H_2O) are oriented to various degrees toward the electrode surface. Ions can accumulate in it due to electrostatic forces or be adsorbed specifically on the electrode through van der Waals and chemical forces. These substances are called *surface-active substances* or *surfactants*. The sum of oriented solvent molecules and surfactants

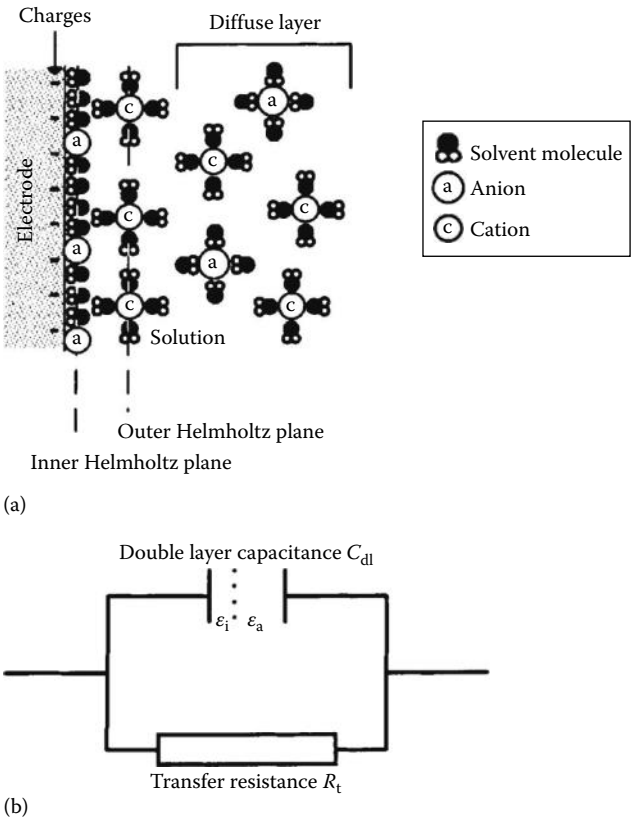


FIGURE 55.2 Helmholtz model describes the electrical double layer as a plate capacitor: (a) capacitor representation with inner and outer planes, and (b) the equivalent circuit representation.

in the immediate vicinity of the electrode is considered as one layer. The plane through the centers of these molecules and ions parallel to the electrode surface is termed *inner Helmholtz plane* (Figure 55.2).

If only electrostatic attraction is taken into account, ions from the solution can approach the surface to a distance given by their primary solvation sheaths. This means that at least a monomolecular solvent layer remains between the electrode and the solvated ion. The plane through the centers of these ions is called *outer Helmholtz plane*, and the solution region between the electrode surface and this outer Helmholtz plane is called *Helmholtz* or *compact layer*. In reality, electrostatic forces cannot retain ions at a minimal distance from the electrode surface. Due to thermal motion, the excess charge is smeared out in the direction of the electrolyte bulk to form a *diffuse layer*, also termed the *Gouy–Chapman layer*. It describes the region between the outer Helmholtz plane and the bulk of the solution. In concentrated electrolyte solutions (approx. 1 mol L^{-1}), the diffuse layer is as thin as the inner Helmholtz plane and may be considered as rigid. In highly dilute solutions, its thickness can be as large as 100 nm. As in the early model of Helmholtz, the double layer acts as a capacitor (Wang 2006). Here, two different dielectric layers with permittivities ε_i and ε_o represent the region between the electrode surface and the inner Helmholtz plane and the region between the inner and the outer Helmholtz plane, respectively (Figure 55.2).

In addition to these ideal electrostatic processes that lead to the formation of the electrical double layer, one has also to consider the transition of charge, ions, and/or electrons from the electrode phase into the electrolyte phase or vice versa. In the equivalent circuit representation, such a charge transport through the double layer is symbolized as a transfer resistance R_t connected in parallel with the capacitor. If any charge transport through the double layer is excluded, the transfer resistance is nearly infinite. According to Ohm's law, any current impressed across the electrode surface leads to a high polarization

voltage determining the electrode as *ideally polarizable*. One example of a polarizable electrode is the dropping mercury electrode (DME), which is frequently used in polarography. In the opposite case with a nearly vanishing transfer resistance, the electrode is termed *ideally unpolarizable*. In the equivalent circuit representation, this corresponds to short circuit of the capacitor. The current flow then does not influence the voltage drop across the phase boundary. *Reference electrodes*, whose voltage has to be constant when immersed in an electrolyte, are nearly unpolarizable electrodes. Since every voltage measurement is accompanied by a small current flow, the difference between polarizable and unpolarizable electrodes is very important in measurement technique.

55.1.2 Nernst Equation

If the electrode phase and the electrolyte phase contain a common ion, the potential difference across the phase boundary is determined by the effective concentration (activity) of this ion in the solution. This fact is described quantitatively by the *Nernst equation* and will be derived in the following. If one mole of ions of a species *i* has to be transferred from a given reference state outside into the bulk of an electrically charged phase, work must be expended to overcome the chemical bonding forces and the electric forces. This work is given by the electrochemical potential $\bar{\mu}_i$. Since the chemical interactions of a species with its environment always possess electric components, generally the electrochemical potential cannot be separated into chemical and electrical parts. Nonetheless, the electrochemical potential is frequently given formally as a sum of the chemical potential μ_i and an electrostatic work $zF\phi$:

$$\bar{\mu}_i = \bar{\mu}_i + zF\phi \tag{55.1}$$

The chemical potential μ_i of an uncharged component of a system is the amount of Gibbs energy G inherent in 1 mol of that component (Koryta and Stulik 2009):

$$\mu_i = \left(\frac{\partial G}{\partial n_i} \right)_{p,T} \tag{55.2}$$

Here, n_i is the number of moles of the given component. In the case of a dilute solution, the chemical potential of a component *i* is

$$\mu_i = \mu_i^0 + RT \ln c_i \tag{55.3}$$

where

$\bar{\mu}_i^0$ denotes the standard chemical potential
c_i denotes the concentration of the species *i*
R is the gas constant
T is the absolute temperature

The values of standard chemical potentials can be found in standard textbooks of thermodynamics and in tables of physicochemical constants under the name standard molar Gibbs energies. μ_i^0 is independent of the concentration c_i. In concentrated electrolytes, the concentration c_i has to be replaced by the respective activity a_i. The activity a_i is given by the relationship $a_i = \gamma c_i$, where γ is the activity coefficient that is a correction factor for nonideal behavior. In the second term of Equation 55.1, z denotes the charge number of the ion *i*, F is the Faraday constant, and ϕ is the *inner electric potential*, which is, in general, the electric work necessary for the transfer of a unit charge: for example, 1 C, from infinity to a given site.

The inner electric potential may consist of two components: an *outer electric potential* ψ and a *surface electric potential* χ. Whereas the outer electric potential of a phase is produced by excess electric charge

supplied from outside, the surface electric potential is an effect of electric forces at the interface that leads to the electrical double layer introduced earlier. The difference of the outer potentials of the electrode (e) and the solution (s)

$$\psi_e - \psi_s = \Delta\psi \tag{55.4}$$

is termed *Volta potential difference* and is the only measurable quantity. Neither the difference of the surface potentials of the appropriate phases $\Delta\chi$ nor the difference of the inner electric potentials

$$\Delta\phi = \Delta\psi + \Delta\chi \tag{55.5}$$

defined as the *Galvani potential difference* can be measured directly. Strictly speaking, even the Volta potential difference between the solution and the electrode is a not measurable quantity since only the Volta potential difference between two electrodes can be measured. To determine the potential of the solution phase, one has to dip an electrode in the solution. This, however, creates a new electrode/solution interphase, and consequently, one measures the sum of two potential differences. This is the reason for the lack of absolute potentials in electrochemistry. Therefore, one uses a reference electrode that has a known potential relative to a standard electrode.

In thermodynamic equilibrium, the electrochemical potentials of the considered species are equal in both phases. For a charged particle *i* that may cross the phase boundary solution/electrode, this means

$$\mu_{i,s}^0 + RT\ln a_{i,s} + z_i F\phi_s = \mu_{i,e}^0 + RT\ln a_{i,e} + z_i F\phi_e \tag{55.6}$$

and therefore, in equilibrium the Galvani potential difference is given by

$$\Delta\phi = \phi_e - \phi_s = \frac{\mu_{i,s}^0 - \mu_{i,e}^0}{z_i F} + \frac{RT}{z_i F}\ln\frac{a_{i,s}}{a_{i,e}} \tag{55.7}$$

Since the chemical standard potentials of the respective phases are constants, the first term in Equation 55.7 can be expressed as a standard Galvani potential difference $\Delta\phi^0$:

$$\Delta\phi = \Delta\phi^0 + \frac{RT}{z_i F}\ln\frac{a_{i,s}}{a_{i,e}} \tag{55.8}$$

For metal electrodes, the activity of the metal atoms M and that of the electrons in the electrode phase equals unity per definition. Thus, for an electrode reaction of type

$$M_s^{z+} + ze^- \leftrightarrow M_e \tag{55.9}$$

Equation 55.8 becomes the *Nernst equation*

$$\Delta\phi = \Delta\phi^0 + \frac{RT}{zF}\ln a_s \tag{55.10}$$

which gives the relation between the activity of the potential determining ion a_s and the Galvani potential difference $\Delta\phi$. Using base 10 logarithms, the Nernst equation is given as

$$\Delta\phi = \Delta\phi^0 + \frac{RT\cdot 2.3}{zF}\log a_s = \Delta\phi^0 + k\cdot\log a_s \tag{55.11}$$

where *k* is called the *Nernst constant*.

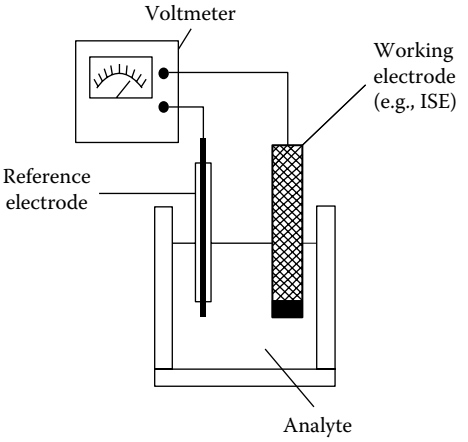


FIGURE 55.3 Schematic of an electrochemical cell with a working electrode and a reference electrode immersed in the test solution (electrolyte).

The classical form of the Nernst equation (Equation 55.10) can be formulated more generally for a redox reaction. If a_{ox} and a_{red} are the activities of the oxidized and reduced form of the considered ion, the Galvani potential difference is given as

$$\Delta\phi = \Delta\phi^0 + \frac{RT}{zF}\ln\frac{a_{ox}}{a_{red}} \tag{55.12}$$

In *potentiometry*, the activity of a certain ion can be determined directly by the measurement of the equilibrium Galvani potential difference of a suitable electrode (*direct potentiometry*). On the other hand, changes of the activity of the detected ion and equivalence points (EPs) can be detected in titration reactions (*potentiometric endpoint titration*).

After this rather theoretical definition of the Galvani potential difference, the question arises on how to measure this potential difference between the bulk of the electrode and the solution. Since a potential difference cannot be measured with only one electrode, a second one must be immersed in the solution. Both are connected to a voltmeter, to complete the *electrochemical cell* (Figure 55.3). An electrochemical cell generally consists of two (or more) electrodes immersed in an analyte. Thus, in some of the old literature, a single electrode is often referred to as a *half-cell* and its potential is called *half-cell potential*. In modern electrochemistry, usually the term *electrode potential* is used. An electrochemical cell is in a current-free state during potentiometric measurements (e.g., with an *ion-selective electrode* [ISE]), but may also supply electric energy (a galvanic cell) or accept electric energy from an external source (an electrolytic cell). Since a second electrode potential arises at the phase boundary second electrode/electrolyte, only the sum of at least two Galvani potential differences can be measured. A separation into the two individual parts is impossible. Hence, the function of the second electrode, named reference electrode, is to act as an electrode of constant potential against which variations in the potential of the measuring electrode in various samples can be measured. In the Nernst equation, the Galvani potential ϕ is then replaced by E, the symbol for measurable voltages.

55.1.3 Classification of Electrodes

Electrodes are termed *reversible electrodes* if they transfer electrons and ions with negligible impedance. Therefore, under current, the electrochemical potential of electrons, ions, and neutral species does not change across the different interfaces that may exist in an electrode. Otherwise, the electrode is not suitable to measure thermodynamic (equilibrium) quantities such as ion activity. Since distribution

equilibrium of charged species is considered here, the electrode and the solution phase must have at least one charged species in common. Depending on the number of equilibria being involved in the forming of the electrode potential, reversible electrodes can be divided into different groups:

1. *Electrodes of the first kind*. These may be cationic or anionic electrodes at which equilibrium is established between the atoms or molecules in the electrode material and the respective cations or anions in the solution. According to the Nernst equation, the equilibrium Galvani potential difference is here determined by the activity of the considered ion in the solution. Examples for electrodes of the first kind are ISEs including metal and amalgam electrodes and the hydrogen electrode.
2. *Electrodes of the second kind*. These electrodes consist of three phases. A metal wire is covered by a layer of its sparingly soluble salt that usually has the character of a solid electrolyte (e.g., Ag and AgCl). This wire is immersed in a solution containing a soluble salt of the anions of this solid electrolyte (e.g., KCl). Here, the equilibrium between the Ag atoms in the metal and the anions in the solution is established through two equilibria: the first one is given between the metal and the cation in its sparingly soluble salt, for example,

$$\mathrm{Ag} \leftrightarrow \mathrm{Ag}^{+} + \mathrm{e}^{-} \tag{55.13}$$

   and the second one between the anion in the sparingly soluble salt and the anion in the solution, for example,

$$\mathrm{AgCl} \leftrightarrow \mathrm{Ag}^{+} + \mathrm{Cl}^{-} \tag{55.14}$$

   The electrode potential of electrodes of the second kind is rather insensitive to small current flows. Thus, they are often used as reference electrodes.
3. *Electrodes of the third kind*. In this electrode, the sparingly soluble salt contains a second cation that also forms a sparingly soluble compound with the common anion but with a higher solubility product than the electrode metal compound (e.g., Ag_2S and PbS). Here, the electrode potential depends on the activity of this cation in the solution.
4. *Oxidation–reduction (redox) electrodes*. They consist of an inert metal such as Pt, Au, or Hg that is immersed in a solution of two soluble oxidation forms of a single substance (e.g., Fe^{3+} and Fe^{2+}). Thus, for the electrode reaction

$$\mathrm{Fe}^{3+} + \mathrm{e}^{-} \leftrightarrow \mathrm{Fe}^{2+} \tag{55.15}$$

   the Nernst equation is

$$E_{\mathrm{Fe}^{3+}/\mathrm{Fe}^{2+}} = E^{0}_{\mathrm{Fe}^{3+}/\mathrm{Fe}^{2+}} + \frac{RT}{F} \ln \frac{a_{\mathrm{Fe}^{3+}}}{a_{\mathrm{Fe}^{2+}}} \tag{55.16}$$

   according to Equation 55.10. Here, E is termed the *electrode potential* and E^0 is designated the *standard electrode* (or *redox*) *potential* of the electrode reaction if it is measured versus the *standard hydrogen electrode* (SHE). The subscripts of E and E^0 denote the redox couple of the considered electrode reaction. The standard redox potential is a measure of the reducing or oxidizing ability of a substance. If one considers, for example, two systems 1 and 2 with their respective standard redox potentials E_1^0 and E_2^0, system 1 is a stronger oxidant than system 2 if $E_1^0 > E_2^0$. This means that in a mixture of the solutions of these two systems where originally the activities of the reduced forms equal that of the oxidized forms ($a^1_{\mathrm{red}} = a^1_{\mathrm{ox}}$ and $a^2_{\mathrm{red}} = a^2_{\mathrm{ox}}$), an equilibrium will be

TABLE 55.1 Some Standard Electrode Potentials and Redox Potentials

Electrode or Half-Cell Reaction	E^0 (V)
$Li^+ + e^- \leftrightarrow Li$	−3.0403
$K^+ + e^- \leftrightarrow K$	−2.931
$Ca^{2+} + 2e^- \leftrightarrow Ca$	−2.868
$Mg^{2+} + 2e^- \leftrightarrow Mg$	−2.372
$Al^{3+} + 3e^- \leftrightarrow Al$	−1.662
$Zn^{2+} 2e^- \leftrightarrow Zn$	−0.762
$Fe^{2+} + 2e^- \leftrightarrow Fe$	−0.447
$Pb^{2+} + 2e^- \leftrightarrow Pb$	−0.1264
$AgCl + e^- \leftrightarrow Ag + Cl^-$	0.22216
$Hg_2Cl_2 + 2e^- \leftrightarrow 2Hg + 2Cl^-$	0.26791
$Cu^{2+} + 2e^- \leftrightarrow Cu$	0.3417
$I_2 + 2e^- \leftrightarrow 2I^-$	0.5353
$Fe^{3+} + e^- \leftrightarrow Fe^{2+}$	0.771
$Ag^+ + e^- \leftrightarrow Ag$	0.7994
$Tl^{3+} + 2e^- \leftrightarrow Tl^+$	1.2152
$2Cl^- \leftrightarrow Cl_2 + 2e^-$	1.35793
$Ce^{4+} + e^- \leftrightarrow Ce^{3+}$	1.610

established with $a^2_{ox} > a^2_{red}$ and $a^1_{red} > a^1_{ox}$. The experimentally determined standard potentials of well-known redox systems are listed in (Haynes 2011). Table 55.1 gives some examples. In redox electrodes, the metal acts as a medium for the electron transfer between the two forms. In contrast to electrodes of the first kind, the solution should not contain ions of the electrode metal in order to avoid an additional Galvani potential difference at the electrode determined by the activity of the electrode metal ions in the solution. This disturbing ion activity is negligible if the standard potential of the electrode metal is a few 100 mV higher than the redox potential to be measured. Thus, mainly platinum electrodes ($E^0_{Pt^{2+}/Pt} = 1.20$ V) and gold electrodes ($E^0_{Au^+/Au} = 1.42$ V) are used as redox electrodes.

55.1.4 Reference Electrodes

The potential of an ISE is always measured with respect to a reference electrode. Ideally, the reference electrode should not cause chemical changes in the sample solution, or vice versa. It should maintain a constant potential relative to the sample solution, regardless of its composition. In practice, any changes of its potential with composition should be at least as small as possible and reproducible. Reference electrodes with liquid junctions, strictly speaking reference electrode *assemblies*, consist of a reference element immersed in a filling solution (often called bridge solution) contained within the electrode. This reference element should possess a fixed activity of the ion defining the potential of the element with respect to the filling solution. The electric contact between the electrode and the sample solution is made by the liquid junction consisting of a porous plug or a flow restriction that permits the filling solution to flow very slowly into the sample.

At the junction between the two electrolyte solutions, ions from both solutions diffuse into each other. Since different ions have different mobilities, they will diffuse at different rates. Thus, a charge separation will occur related in size to the difference in mobilities of the anions and cations in the two solutions. This charge separation produces a potential difference across the junction called the *liquid junction potential* (Morf 1981). In reference electrodes, usually the bridge solution is given a slightly higher pressure than the sample so that the solution, often concentrated potassium chloride, flows out

relatively rapidly into the sample and diffusion of the sample back into the salt bridge is impeded. If the bridge solution is concentrated enough, it is assumed that variations in the liquid junction potential due to the varying composition of the sample are suppressed. This is the basis on which the reference electrode assembly is used. Since the potential of the whole assembly E_{ref} is the sum of the potential of the reference element E_r in the bridge solution and the liquid junction potential E_j

$$E_{ref} = E_r + E_j \tag{55.17}$$

any change in the liquid junction potential appears as a change in the potential of the assembly. An extra liquid junction potential must be included if a *double-junction* reference electrode is considered. When an analysis using a cell with an IES is carried out, standard solutions are used to calibrate the ISE. A change in the liquid junction potential that occurs when the standard solutions are replaced by the sample is termed the *residual liquid junction potential* and constitutes an error in the analytical measurement. The needed constancy of the potential can be approached by a suitable choice of standards and/or sample pretreatment and by the use of a proper bridge solution and the best physical form of the liquid junction.

Several types of liquid junctions exist from which the best ones with regard to stability and reproducibility are complicated to realize in practice and the worst ones are easy to use but much less stable and reproducible. Most of the commercial reference electrodes with adequate properties possess *restrained diffusion junctions* where the most common junctions available are the ceramic plug, the asbestos wick or fiber, two types of ground sleeve junction, and the palladium annulus junction (Figure 55.4). For a very large majority of applications with IESs, a ceramic plug will perform adequately. The flow rate of the bridge solution into the sample solution is sometimes called leak rate and is given in mL per 5 cm head of

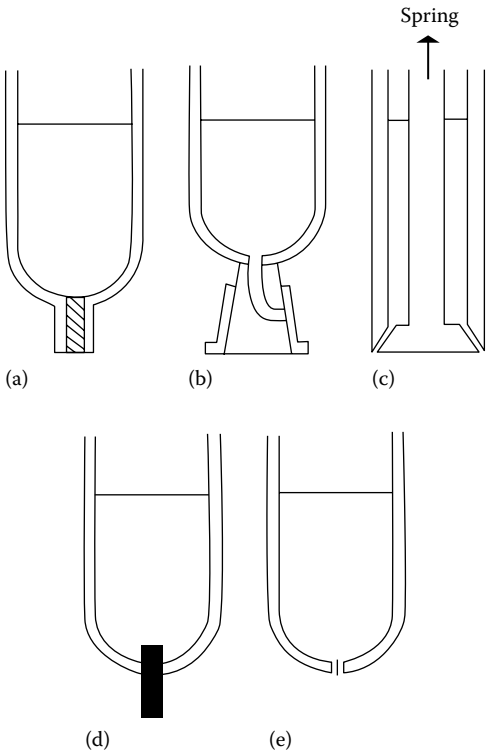


FIGURE 55.4 Different types of liquid junctions: (a) ceramic plug, (b) ground glass sleeve (type 1), (c) ground glass sleeve (type 2), (d) asbestos wick, and (e) palladium annulus.

bridge solution per day. The head of bridge solution is measured as the height of the surface of the bridge solution above the surface of the sample. In order to work satisfactorily, the surface of the bridge solution of all these restricted junction devices has to be at least 1 cm above the sample solution. Otherwise, if the bridge solution falls too low, the junction and the bridge will become contaminated by species diffusing from the sample. The bridge solution has then to be replaced. For the same reason, reference electrodes should be stored, when not in use, with the junction immersed in bridge solution.

Whereas the ceramic plug and the asbestos wick and fiber (Figure 55.4a and d) have relatively slow flow rates of about 0.01–0.1 mL per 5 cm head of bridge solution per day, ground sleeve junctions of type (b) have a flow rate of 1–2 mL. On the other hand, the flow rates of different asbestos wick junctions may vary by a factor up to 100, and the liquid junction potential may have a day-to-day (in) stability of ±2 mV under the favorable conditions of a junction between strong potassium chloride solution and an intermediate pH buffer. Under the same conditions, ground glass sleeve junctions of type (b) and the little-used palladium annulus junction show stabilities of ±0.06 and ±0.2 mV, respectively. It is worth mentioning that palladium annulus junctions may partly respond as a redox electrode in strong oxidants (e.g., 0.2 M $KMnO_4$ in 0.05 M H_2SO_4) and mild or strong reductants (e.g., 0.5 M $SnCl_2$ in 1 M HCl). In such samples, reference electrodes with palladium or platinum annulus junctions should not be used. Although ground glass sleeve junctions have inconveniently high flow rates and the bridge solution needs to be replenished frequently, these junction types have found particular use in applications where the junction has the tendency to clog, such as measurements in protein solutions. However, the stability of the liquid junction potential appears to be relatively poor in fast-flowing sample solutions and may be very sensitive to sample flow rate. Asbestos wick junctions are particularly liable to blockage and should consequently be used in clear solutions only.

In *double-junction reference electrodes*, the filling solution in which the reference element is immersed (reference solution) makes contact with another solution, the bridge solution, by means of a liquid junction. A second liquid junction enables contact to be made between the bridge solution and the sample. Such electrodes are useful when it is essential that contamination of the sample by the inner filling solution must be kept at a very low level. The outer bridge solution can be selected to be compatible with the sample. In order to minimize the liquid junction potentials that can drift and cause instability, the bridge solution should be equitransferent; that is, the transport numbers of its anion and cation should be nearly equal. However, the complication of a second liquid junction in the cell should be avoided if possible.

55.1.4.1 Standard Hydrogen Electrode

Aqueous solutions are of major concern in electrochemistry because of their hydrogen ion content. Thus, it is advantageous to use a reference electrode where a reaction occurs that involves the participation of hydrogen ions. One of these reactions is

$$\frac{1}{2}H_2 + H_2O \leftrightarrow H_3O^+ + e^- \tag{55.18}$$

Figure 55.5 shows a hydrogen electrode. A hydrogen electrode usually consists of a platinum sheet covered by a thin layer of spongelike structured platinum, so-called platinum black that has a high specific surface area. This electrode is rinsed with pure gaseous hydrogen in order to form a complete layer of adsorbed H_2 molecules at the surface. If this electrode is immersed in an electrolyte, it acts as an electrode consisting of hydrogen at which the gaseous hydrogen is oxidized to hydronium ions or the hydronium ions are reduced to hydrogen, respectively, according to Equation 55.18. The real mechanism of this electrode process is rather complicated because the platinum electrode is in contact with the hydronium ions in the solution as well as with the gaseous hydrogen that is bubbled through it. Thus, the final equilibrium between the gaseous hydrogen, the dissolved hydronium ions, and the electrode phase consists of several successive equilibrium steps that can be found in Koryta (1991).

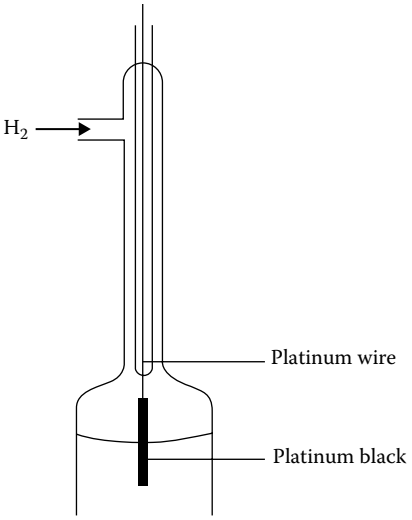


FIGURE 55.5 Schematic of a hydrogen electrode.

To calculate the potential of a hydrogen electrode, which is strictly speaking the difference between the potential of the electrode and that of the solution, one has to consider the electrochemical potentials of the respective phases. The chemical potential of gases is usually expressed in terms of the pressure p instead of the molar concentration c. Due to the elementary relationship $pV = nRT$ for ideal gases, where V is the volume of the gas and n is the amount of moles the pressure, p is proportional to the molar concentration $c = n/V$. Thus, according to Equation 55.3

$$\mu_{H_2} = \mu^0_{H_2} + RT \ln \frac{P_{H_2}}{P^0_{H_2}} \tag{55.19}$$

where μ_{H_2} and $\mu^0_{H_2}$ are the pressure and standard pressure of hydrogen, respectively. In the case of moderate ion concentrations, the chemical potential of the solvent water is equal to its standard chemical potential. Hence, the potential difference between the electrode and the solution is, according to Equation 55.7,

$$\Delta\phi = \frac{\mu^0_{H_3O^+} - \frac{1}{2}\mu^0_{H_2} - \mu^0_{H_2O}}{F} + \frac{RT}{F} \ln c_{H_3O^+} - \frac{RT}{2F} \ln \frac{P_{H_2}}{P^0_{H_2}} \tag{55.20}$$

This equation is generally valid for hydrogen electrodes. The electrode is called SHE if the molar concentration is such that the activity of the hydronium ions is unity ($a_{H_3O} = 1$) and the pressure of hydrogen is equal to its standard pressure. Hence, for an SHE, the second and third terms in Equation 55.20 vanish. The combination of standard chemical potentials in the first term of Equation 55.20 is defined as zero. Consequently, the total potential difference across the interface SHE/electrolyte is equal to zero *by definition* at any temperature. Since SHEs are very difficult to prepare, they are not used as reference electrodes in practice. However, electrode potentials are usually standardized with respect to the SHE, and their values are thus called "on the hydrogen scale."

55.1.4.2 Calomel Electrode

The calomel electrode is the most common of all reference electrodes. It consists of a pool of mercury that is covered by a layer of mercurous chloride (calomel, Hg_2Cl_2). The calomel is in contact with a

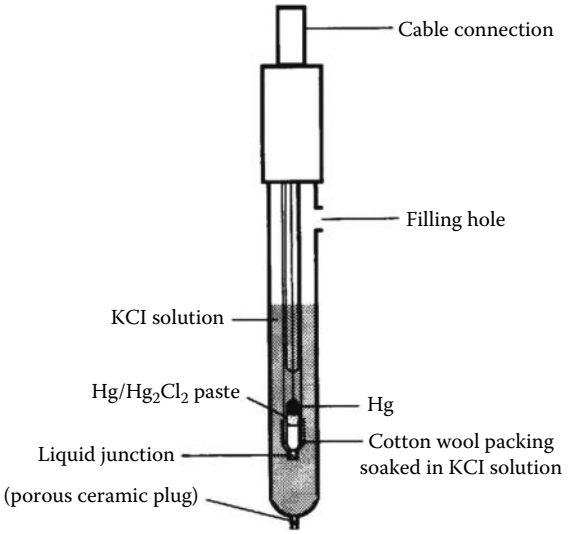


FIGURE 55.6 Schematic of a calomel reference electrode.

reference solution that is nearly always a solution of potassium chloride, saturated with mercurous chloride. Thus, the calomel electrode is a typical electrode of the second kind. Figure 55.6 shows a typical arrangement of a commercial calomel electrode assembly where the electrode is inverted, with the mercury uppermost, and packed into a narrow tube. Depending on the strength of the potassium chloride solution used, the electrode is called saturated calomel electrode (SCE), 3.8 or 3.5 M calomel electrode, respectively. Potassium chloride is used as reference solution because it gives rise to a small liquid junction potential at the outer liquid junction of the electrode, that is, the liquid junction with the sample. Hence, potassium chloride is a suitable reference solution as well as a good bridge solution. Furthermore, mercurous chloride has a very low solubility in potassium chloride solutions, regardless of concentration. The electrode reaction of a calomel electrode is

$$Hg + Cl^- \leftrightarrow \frac{1}{2}Hg_2Cl_2 + e^- \tag{55.21}$$

Its standard potential, including the liquid junction, is 0.2444 V versus SHE at 25 °C for the SCE and 0.2501 V for the 3.5 M calomel electrode (Bailey 1980; Koryta et al. 1993).

The components of a calomel electrode are chemically stable except for the mercurous chloride, which significantly disproportionates at temperatures above 70 °C according to the following equation:

$$Hg_2Cl_2 \leftrightarrow Hg + HgCl_2 \tag{55.22}$$

Hence, potential drift occurs and lifetime decreases with increasing working temperature. On the other hand, calomel electrodes can be used at temperatures down to −30 °C if 50% glycerol is added to the potassium chloride solution.

Impurities in the potassium chloride solution, such as bromide and sulfide ions as well as redox agents and complexants, cause a small shift in the electrode potential. Nevertheless, the measurement of potential differences is not affected. However, the most unsatisfactory feature of the performance of the calomel electrode is its thermal hysteresis that occurs if the electrode filling material is not in thermal equilibrium or if the electrode and the sample have different temperatures. Thus, temperature stability

during the storage and measurements is very important. In any cases where the temperature of the reference electrode or the sample has to be varied, it is thus usually better to use a silver/silver chloride electrode instead of a calomel electrode.

55.1.4.3 Silver/Silver Chloride Electrode

The silver/silver chloride electrode consists of a silver wire or plate that is coated with silver chloride. For the same reasons as with the calomel electrode, this phase is in contact with a strong potassium chloride solution, here saturated with silver chloride. Figure 55.7 shows the diagram of a typical Ag/AgCl reference electrode. Since this kind of reference electrode is the simplest and for many applications the most satisfactory one, it is commonly used as internal reference electrode of pH electrodes and other IESs. Besides, Ag/AgCl electrodes can be easily prepared in the laboratory. In contrast to mercury-based electrodes, the Ag/AgCl electrode does not contain toxic chemicals and is therefore recommendable for measurements in food.

The major problem with the Ag/AgCl electrode is the considerably high solubility of AgCl in concentrated potassium chloride solution. Thus, especially for the use at high temperatures, a sufficient excess of solid silver chloride must be present in the reference solution. This can be achieved, for example, through the addition of a few drops of diluted silver nitrate solution. Otherwise, silver chloride will dissolve off the electrode until saturation is reached. As a consequence, the electrode potential will drift and the lifetime of the electrode will be shortened. However, in contrast to the calomel electrode, the Ag/AgCl electrode can be used successfully up to 125 °C. Its electrode potential is very stable in the long term in pure potassium chloride solutions, but is affected by impurities like redox reagents and species that react with the silver chloride, as with the calomel electrode. Unlike the calomel electrode, in the Ag/AgCl electrode, the concentration of the electrode coating in the bridge solution is rather high. Thus, a greater amount of reaction products (e.g., solid silver sulfide) may arise in the reference solution and block the liquid junction causing drift and instability of the electrode potential. In contrast to the calomel electrode, the silver/silver chloride electrode shows only very small thermal hysteresis effects that are usually negligible. Hence, this kind of electrode is suitable for measurements in samples with varying temperatures. Ag/AgCl electrodes are relatively insensitive to polarization. The standard potentials, including the liquid junction potentials of saturated and 3.5 M silver/silver chloride electrodes, at 25°C are 0.1989 and 0.2046 V, respectively (Bailey 1980). As with the calomel electrode, the nomenclature of the electrodes is derived from the potassium chloride concentration of the respective reference solution.

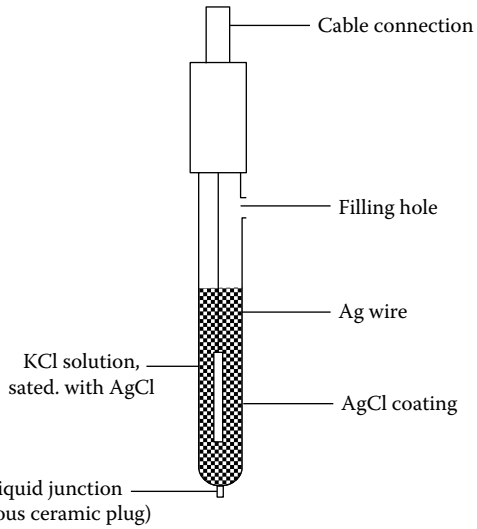


FIGURE 55.7 Schematic of a Ag/AgCl reference electrode.

55.2 Volumetry

The basic concept of *voltammetry* is the measurement of the current *i* at a redox electrode as a function of the electrode potential *E* (Bard and Faulkner 2001). During the experiment, the electrode is immersed in a solution that contains an electroactive species, that is, a species that can undergo an electrode reaction (standard redox potential E^0). The electrode potential is changed from a value $E_1 < E^0$ to a value $E_2 > E^0$ or vice versa in a manner that is predetermined by the operator. Thus, during the measurement, the electrochemical equilibrium shifts from the oxidized (reduced) form of the analyte to the reduced (oxidized) form. The resulting charge transfer across the interface electrode/solution can be observed as a current flow, which is termed *faradaic*.

55.2.1 Instrumentation

Voltammetric measurements are usually performed with a cell arrangement of three electrodes (Figure 55.8). The redox electrode at which the electrode processes occurs is called *working electrode*. Its potential is measured against a suitable reference electrode, often Ag/AgCl or calomel. To adjust the potential difference between the working and the reference electrode to a certain value, a current is forced through the working electrode. Because the current and the electrode potential are related functionally, this current is unique. However, the current through the reference electrode must be kept as small as possible. Therefore, a third electrode called *auxiliary electrode* or *counter electrode* is usually employed to close the current circuit. It should be emphasized that there are two circuits: one in which the current flows and which contains the working and the auxiliary electrode and another, a current-free one, in which the potential difference between the working and the reference electrode is measured. Since almost no current flows through the reference electrode, its potential can be regarded as constant and the measured change in potential equals the potential change of the working electrode. The current through the working electrode, and thus its potential, can be adjusted by controlling the voltage between the working and the auxiliary electrode. This task is performed by an instrument called a *potentiostat*, which basically consists of a voltage source and a high-impedance feedback loop. With a function generator that may be integrated into the potentiostat, the potential time course can be predetermined. Modern potentiostats are controlled by a PC and offer the possibility to program many different potential time courses. Thus, they allow the performance of several voltammetric techniques, as are discussed later. The measured current can be displayed as a function of the electrode potential or of time using a strip-chart or *xy* recorder or a PC.

There are two possibilities to operate an electrochemical cell: in so-called *batch cells*, the electrolyte solution rests stationary during the measurement, whereas in *flow-through cells*, it flows across the electrode. Between two measurements with different solutions, the cell must be cleaned in order to remove

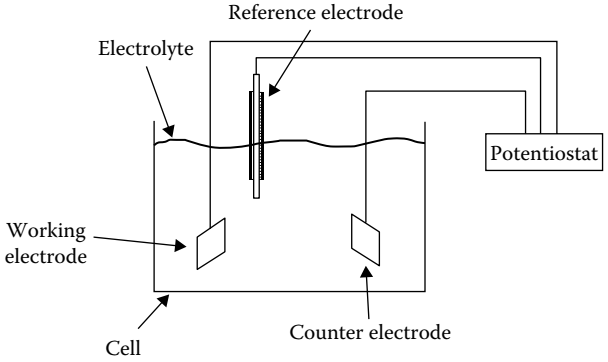


FIGURE 55.8 For voltammetric measurements a three-electrode arrangement is usually employed.

residues of the preceding measurement's solution that could disturb the new measurement. The electrochemical cell is usually built of glass or Teflon because of these materials' chemical inertness.

The chemical inertness is also important for the choice of the working electrode because the electrode must not change during the measurement. Common materials are gold, platinum, and mercury. Several kinds of carbon electrodes (e.g., glassy carbon) are also used but are often covered with gold or mercury. An advantage of the solid-state electrodes is their easy handling. They can be employed as planar or as wire electrodes. Further, with the noble metal electrodes, substances having a more positive redox potential than mercury can be investigated. However, the use of mercury electrodes offers distinct advantages and the voltammetric techniques using mercury electrodes are extremely well developed. These techniques play a major role in electroanalytical methods and are summarized under the term *polarography*.

In polarography, mercury is used either as a *thin mercury film electrode* (TMFE) or as a *hanging mercury drop electrode* (HMDE). The HMDE can be a *stationary mercury drop electrode* (SMDE) or a DME. The drop is produced from a thin capillary with an inner diameter that can range from several ten to a few hundred micrometers. The size of an SMDE is held constant during the measurement, whereas a DME constantly grows during its lifetime until it falls from the capillary due to its weight.

The main advantages of mercury drop electrodes are their good reproducibility and their high *overpotential* for the hydrogen evolution, that is, the fact that hydrogen evolution is inhibited and thus occurs at much higher potentials than would be expected from the standard potential. The good reproducibility is achieved because a new drop can easily and rapidly be produced from the capillary for each measurement. Hence, the contamination of the electrode with substances from a preceding measurement and from impurities in the solution is near zero. However, a drawback of HMDEs is their relative mechanical instability, which can be a problem in flow-through cells, in field measurements, and if the solution is stirred.

Stirring of the solution is often applied during the measurement if the supply of reactive species at the electrode should be enhanced. However, this forced convection affects the electrode current. Moreover, the electrolyte is often stirred and bubbled with an inert gas like nitrogen or argon before voltammetric measurements are carried out to remove dissolved oxygen. This is usually necessary to reduce background currents from oxygen reduction and to prevent undesirable oxidation or precipitation of solution components. Because the electrode currents, especially in trace and ultratrace analysis, can be quite small, it is common to place the cell in a faradaic cage to shield it from electromagnetic stray fields. Coaxial cables are then used for the electric connections from the cell to the instruments.

55.2.2 Principles of Voltammetry

Actually, the electrode current measured in voltammetry is a sum of two currents that arise due to different processes. Besides the faradaic current i_f, a capacitive current i_c results from changes in the *double-layer charging*. Although the faradaic current is a direct measure for the rate of the electrode reaction, several effects usually occur that have to be considered.

55.2.2.1 Diffusion Limitation of the Faradaic Current

The decrease of the analyte concentration at the electrode surface due to an electrode reaction must be balanced by the diffusion of species from the bulk solution. In most measurements, the consumption of reactive species is faster than the supply by diffusion, and the effect of *diffusion limitation* of the faradaic electrode current is observed. To understand this important point, the time-dependent concentration profile of the analyte has to be calculated using *Fick's laws*. The electrode current can then be derived as a function of time. According to Fick's first law, the flux j of the analyte at the point r and at the time f is proportional to the gradient of the analyte concentration c:

$$j(r,t) = -D\nabla c(r,t) \tag{55.23}$$

The proportionality factor D is called the *diffusion coefficient*. At the electrode surface, the flux must be equal to the number of moles N converted per unit of time and surface area by the electrode reaction:

$$j(0,t) = \frac{\mathrm{d}N}{\mathrm{d}t} \tag{55.24}$$

The faradaic current i_f is related to dN/dt according to

$$i_f = \frac{nFA\mathrm{d}N}{\mathrm{d}t} \tag{55.25}$$

where

n is the number of electrons involved in the reaction of a single analyte particle
F is the Faraday constant and
A is the surface area of the working electrode

The *Nernst diffusion layer* model assumes that within a layer of thickness δ, the analyte concentration depends linearly on the distance from the electrode surface until it reaches the bulk concentration c_0. For simplicity, the diffusion problem is often considered to be 1D as it is the case for a planar working electrode in a cylindrical cell. The combination of Equations 55.23 through 55.25 then gives

$$i_f = nFAD\left(\frac{(c_0 - c_e)}{\delta}\right) \tag{55.26}$$

where c_e is the concentration at the electrode surface. For a sufficiently large difference between the applied potential and the standard potential of the analyte's redox couple, all species reaching the electrode surface by diffusion are immediately converted, and the faradaic current reaches a maximum. In this case, the analyte concentration c_e at the electrode surface can be regarded as zero.

The diffusion profile and thus the dependence of δ from time can be obtained by solving the differential equation that is known as Fick's second law:

$$\frac{\partial c(r,t)}{\partial t} = D\nabla^2 c(r,t) \tag{55.27}$$

where ∇^2 is the Laplacian operator. For *linear diffusion*—that is, 1D diffusion as it was considered in Equation 55.26—the solution of Equation 55.27 with the appropriate boundary conditions ($c_e(t = 0) = c_0$; $c_e\ (t > 0) = 0$; $c(x > \delta = c_0)$) yields

$$\delta = \mathrm{sqrt}(\pi Dt) \tag{55.28}$$

Combination with Equation 55.26 leads to the *Cottrell equation*:

$$i_f(t) = nFAc_0\mathrm{sqrt}\left(\frac{D}{\pi t}\right) \tag{55.29}$$

After reaching a maximum value, the current decreases with $t^{-1/2}$ and is proportional to c_0, whereas the diffusion layer thickness increases with $t^{1/2}$ (Figure 55.9).

For a spherical electrode of radius r_0, as it is the case for HMDEs, one has to change to spherical coordinates and Fick's second law becomes

$$\frac{\mathrm{d}c(r,t)}{\mathrm{d}t} = D\left[\frac{\mathrm{d}^2 c(r,t)}{\mathrm{d}r^2} + \frac{r\ \mathrm{d}c(r,t)}{\mathrm{d}r}\right] \tag{55.30}$$

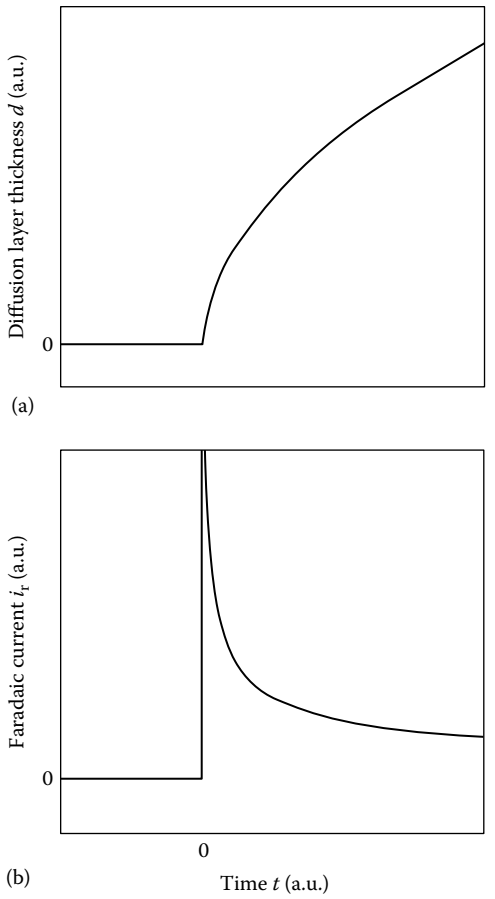


FIGURE 55.9 At a planar electrode, the diffusion layer thickness increases with $t^{1/2}$ (a), whereas the diffusion-limited current decreases with $t^{-1/2}$ (b).

where $r > r_0$ is the radial distance from the electrode center. The solution of Equation 55.30 with the appropriate boundary conditions $c(r, 0) = c_0$, $\lim(r \to \infty)\, c(r, t) = c_0$, $c(r_0, t > 0) = 0$ yields the current–time relation

$$i_f(t) = n\mathrm{FAD}c_0\left[\frac{1}{(\pi Dt)^{1/2}} + \frac{1}{r_0}\right] \tag{55.31}$$

The first term in brackets equals that for the linear case; the second, constant term reflects the fact that the surface of the spherical diffusion layer grows and thus can draw an increasing number of reactive species.

The situation is even more complicated for DMEs because in addition to the surface of the diffusion layer, the surface and the radius of the drop are growing during the drop's lifetime. At any time, the growing electrode surface forces the depletion layer to stretch over a still larger sphere, which makes the layer thinner than it otherwise would be. A rigorous mathematical approach to this is rather difficult because the relative convective movement between the solution and the electrode during drop growth must be considered (Bard and Faulkner 2001). However, a simplified approach that is valid when the second term in Equation 55.31 is negligible and the diffusion problem can be regarded as linear yields the *Ilkovic equation*

$$i_f(t) = 708nDcm^{2/3}t^{1/6} \tag{55.32}$$

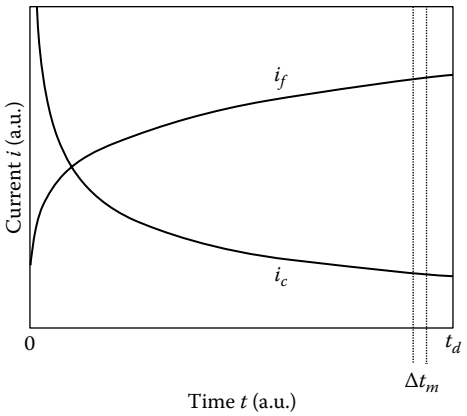


FIGURE 55.10 At a DME the measurement is performed during a time interval Δt_m at the end of the drop's lifetime when the ratio i_f/i_c is very large.

where m is the mercury flow rate (mass/time) from the capillary. Consequently, the current increases during the lifetime t_d of the drop (*drop time*), whereas it decreases with time in the other arrangements that have been described. Figure 55.10 depicts this current–time relation of a DME with the characteristic current plateau at the end of the drop's lifetime.

In the considerations that have been made previously, analyte transport by convection and migration in the electric field have been neglected. Convection can be regarded as absent if the solution is unstirred and if the working electrode rests motionless. However, in longer-lasting measurements, convective mass transport can play a role due to arising inhomogeneities in the density of the solution. Furthermore, if a DME is employed, the growth of the drop may cause a considerable convection of the solution. When the drop falls off, it stirs the surrounding solution and the depletion effect almost vanishes. Consequently, every drop is born in an almost homogeneous environment. The migration of electrically charged analyte particles due to the electric field in the solution can easily be suppressed using an inert supporting electrolyte with a concentration that is much larger than the analyte concentration. Since all charged species contribute to the migration current, the migration of the analyte species can then be neglected.

55.2.2.2 Double-Layer Charging Current

A process that affects all kinds of voltammetric measurements is the flow of *capacitive current*. The accumulation of charge on one side of the electrode/solution interface causes the necessity of a mirror charge on the other side. Hence, a change of the electrode potential (i.e., in the electrode charging) causes a corresponding flux of charged particles between the double layer and the bulk solution. Therefore, the interface has a certain capacitance that is called the double-layer capacitance. The resulting *double-layer charging current* i_c is superimposed on the faradaic current and often perturbs its measurement. In analytical techniques, one is often concerned with the reduction of the capacitive/faradaic current ratio. However, the actual measurement of the double-layer capacitance is demanding and requires the technique of *impedance spectroscopy*, as described, for example, in Gileadi (1993) and Bard and Faulkner (2001).

55.2.2.3 Irreversible Electrode Processes

Another assumption that has been made implicitly is that the rate of the electrode reaction is very fast in comparison to the supply of analyte by diffusion (*reversible electrode process*). Under this condition, all analyte species reaching the electrode are immediately converted. However, if the reaction rate is too slow, the consumption of reactive species is compensated by the diffusion of the analyte (*irreversible electrode process*), and thus, the concentration at the electrode surface never drops to zero. The electrode current is then determined by the reaction rate, and the previous calculations do not hold. In practice,

the situation is sometimes complicated if so-called *quasi-reversible* electrode processes with intermediate reaction rates occur. Although this concept of electrochemical *reversibility* is a simplification, it is a suitable working basis and can be summarized in the following statement: in a given electrochemical experiment, an electrode process that follows the *Nernst equation* at *any time* is called *reversible*.

55.2.2.4 Influence of Adsorption, Catalysts, and Chemical Reactions

Besides the diffusion and reaction rate, some other processes can influence the electrode current. *Adsorption* of the analyte or its reaction product on the electrode changes the double-layer capacitance or can passivate the electrode surface and thus lower the current. Moreover, if a species serves as a *catalyst*, it may shift the equilibrium potential. In the case that the catalyst returns the product of the electrode reaction back into the initial form of the analyte, the analyte concentration at the electrode surface will always be large and thus increases the limiting current and shifts the equilibrium. All these *catalytic currents* are subject to analytical studies. Besides adsorption and catalysis, complicated scenarios occur if the electrode reaction is followed by a chemical reaction whose product itself undergoes an electrode reaction within the observed potential range.

55.2.3 Techniques

The several voltammetric (i.e., *potential-controlled*) techniques differ just in the manner in which the electrode potential is varied with time. The potential can be changed in distinct steps, in a continuous sweep, or it can be pulsed or superimposed with an ac signal. In addition, the rate of potential change can be varied. The characteristics, advantages, and drawbacks of the most important techniques will be discussed in the following sections. Special attention will be given to polarography due to its practical importance in electroanalysis. Besides, the emphasis will be on reversible electrode processes because only they allow the realization of *analytical* investigations, on which this chapter is focused.

55.2.3.1 Amperometry

If in a *potential step* experiment the working electrode potential is abruptly changed from a constant value E_1 where faradaic processes do not occur to another constant value E_2 where the electrochemical equilibrium is on the side of the oxidized or reduced form of the analyte, then a faradaic current begins to flow (Figure 55.11a). In the case that the difference between the applied potential and the standard potential E^0 of the analyte's redox couple is sufficiently large, the effect of *diffusion limitation* sets in and a further increase of the potential difference yields no increase in the electrode current. The current is then called *limiting current*. The current–time relationship follows the Cottrell equation (Equation 55.29), with the current decreasing while the diffusion layer thickness increases.

If the diffusion layer thickness could be held constant, then from Equation 55.26, it follows that the current would not decrease with time but remain at a constant value. This can be accomplished if the solution is stirred or flows across the electrode in a proper way. According to Equation 55.26 (with $c_e = 0$), the current then is proportional to the analyte concentration in the solution.

The described method corresponds to the electroanalytical technique called *amperometry* (Oehme 1991), with the exception that in this the potential step is omitted and the electrode current is measured at a fixed potential E at which the analyte undergoes an electrode reaction and the faradaic current is in the limited region. The solution usually crosses the electrode in a laminar flow, keeping the diffusion layer thickness constant.

Because the electrode current is proportional to the concentration of the analyte, only two measurements are needed for calibration. The base current is measured in an analyte-free solution, and a second measurement is performed at a known analyte concentration. It should be mentioned that amperometry cannot only be used to determine liquid and ionic components of a solution but also to measure the amount of dissolved gas in a liquid. Moreover, with modified electrochemical cells, even gas analysis can be accomplished.

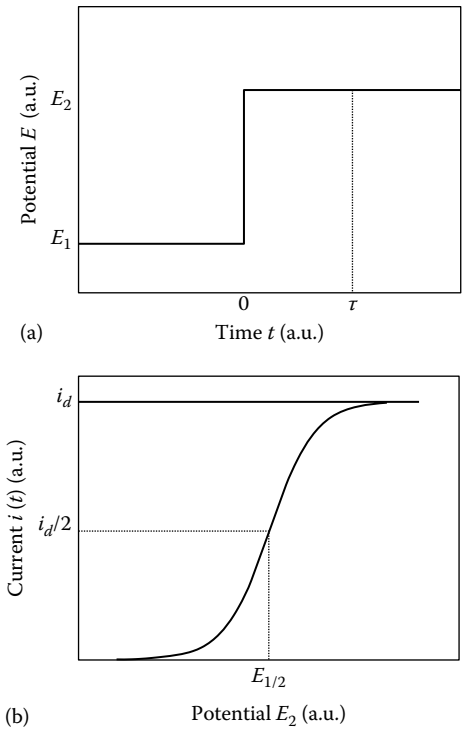


FIGURE 55.11 In potential step techniques, the current is measured a fixed time τ after the potential step (a). The measurement of $i(\tau)$ for different potential steps ΔE yields a wave-shaped current–potential relation with a half-wave potential $E_{1/2} \approx E^0$ (b). The maximum current is proportional to the analyte's bulk concentration c_0.

The main disadvantage of amperometry is its poor selectivity. Given a certain analyte and operating at a higher potential than the corresponding standard potential E^0, all components of the solution with a standard potential smaller than E also contribute to the faradaic current. Operating at a potential $E < E^0$, the same problem occurs if substances with a standard potential larger than E are present in the solution. For this reason, amperometry is preferably carried out in solutions containing only one electroactive substance or, if possible, at a potential at which only one substance is involved in an electrode reaction. If this is impossible, the selectivity can often be enhanced by covering the working electrode with a membrane that, in comparison to the diffusion rate of the analyte through the membrane, is virtually impermeable for the interfering substances.

In addition to analytical purposes, amperometric methods can also be used to investigate reaction constants of chemical reactions. In *reversed potential step techniques*, the first potential step is followed by a second one in the opposite direction, often back to the initial value. The reaction product B of the first step is then reconverted into the original analyte A. However, if the first electrode reaction is followed by an additional chemical reaction, a certain part of B is converted into a product C before the reversed step is applied. Therefore, the current during the reversed step is reduced. The ratio of the electrode currents during the forward and reversed steps depends on the reaction constant of the chemical reaction. Because the reconversion of B into A is required, batch arrangements without convection of the electrolyte are used for reversed step methods. Otherwise, a large part of B would be flushed away from the electrode surface and could not be reconverted.

55.2.3.2 Amperometric Titration

In *amperometric titration techniques* (Willard et al. 1988; Settle et al. 2011), a titrant that reacts with the analyte is added to the analyte solution. During the titration, the limiting current is measured

as a function of the volume of titrant added. The titrant has to be chosen such that the reaction product is not reducible or oxidizable at the applied potential and, hence, does not contribute to the current.

If the analyte and the titrant are electroactive at the applied potential, then the current flow will be large at the beginning of the measurement and decreases linearly with the volume of the titrant added, because both, the analyte and the titrant, are consumed by the reaction. The concentration of electroactive species then diminishes until the analyte is totally consumed. Further addition of titrant leads to a linearly increasing current because the titrant is no longer consumed. In the plot of the current versus the volume of titrant added, the point where the slope changes is called *endpoint* of the titration. From the corresponding amount of titrant added and the stoichiometry of the reaction, the original volume of analyte can be computed. If only the analyte is electroactive, then from the endpoint the current will not increase but remain zero. If only the titrant undergoes an electrode reaction, the current will be zero until all analyte is consumed and then will linearly increase from the endpoint. In practical operation, the slope of the current does not change abruptly due to background currents, and the endpoint has to be determined by extrapolation of the two linear regions.

In contrast to the majority of other electrochemical techniques, amperometric titration offers the advantage that even analytes that are not reducible or oxidizable can be determined using the oxidation–reduction characteristics of the titrant. Moreover, it is possible to analyze systems that have no measurable standard potential but can be electrolyzed.

55.2.3.3 Sampled-Current Voltammetry

Consider a *potential step* experiment like the one in the section next to the previous one. If the potential difference between E_2 and E^0 is too small, the electrode reaction is not so efficient that the analyte concentration at the electrode surface becomes zero (i.e., $c_e > 0$ in Equation 55.26). Within this region, the current depends on the applied potential. However, even in this situation, a depletion effect occurs so that the current always decreases with time. Recording the current i for different values of E_2 at a fixed time τ after switching the potential (*sampled-current voltammetry*), a sigmoidal (wave-shaped) curve is obtained (Figure 55.11b).

The shape of this curve can also be calculated by exactly solving the diffusion problem. A wave rising from a baseline to the diffusion-limited current i_d is obtained. The diffusion coefficients of the analyte and its redox partner are nearly equal to the *half-wave potential* $E_{1/2}$, where $i = i_d/2$ is almost identical with the standard potential E^0. Therefore, $E_{1/2}$ is often used in qualitative analysis to determine the analyte. Quantitative information about the analyte concentration is obtained from the maximum current (Cottrell current), which according to Equation 55.29 is proportional to c_0.

The influence of the double-layer charging current has been neglected thus far, but is worth considering. It obeys the equation:

$$i_c = \frac{\Delta E}{R_s} \exp\left(\frac{-t}{(R_s C_{dl})}\right) \tag{55.33}$$

where

ΔE is the potential step width
R_s the solution resistance
C_{dl} is the double-layer capacitance

Although the measurement of R_s and C_{dl} is not trivial, one can obtain qualitative information from this formula. Comparison of Equations 55.33 and 55.29 yields that the capacitive current decreases exponentially, while the faradaic current decreases according to $t^{-1/2}$. Consequently, the electrode current is measured a sufficiently long time after the potential step when the capacitive current has largely decayed,

whereas the faradaic current is still significant. In polarography with DMEs, the growth of the electrode surface alters the temporal decrease of the double-layer charging current according to

$$i_c \sim m^{2/3} t^{-1/3} \tag{55.34}$$

whereas the faradaic current increases according to $t^{1/6}$ (Equation 55.32). The current is measured shortly before the drop falls off (Figure 55.10).

The lower detection limit (LDL) amounts to 10^{-5}–10^{-6} mol L^{-1} for the determination of organic and inorganic analytes. The half-wave potential of different substances should be at least 100 mV apart for a simultaneous determination.

55.2.3.4 Linear Sweep and Cyclic Voltammetry

In *linear sweep voltammetry* (LSV), the electrode potential is changed *continuously* from an initial to a final value at a constant rate $\nu = dE/dt$, such that $E(t) = E_1 \pm \nu t$. Starting at a potential E_1 where no faradaic process occurs, a current begins to flow when the electrode potential comes into the vicinity of E^0. The current rises to a maximum and then decreases due to the depletion effect (Figure 55.12). The solution of the diffusion equations, which yields the shape of the *i*–*E* wave, can only be found numerically. For the electrode process to always follow the Nernst equation and thus be reversible, the sweep rate must not be too high (e.g., $\nu < 100$ mV s^{-1}). The peak potential E_p can then be calculated to be

$$E_p = E_{1/2} \pm 1.1 \cdot \left(\frac{RT}{nF}\right) = E_{1/2} \pm \left(\frac{28.0}{n}\right) \text{mV (at 25 °C)} \tag{55.35}$$

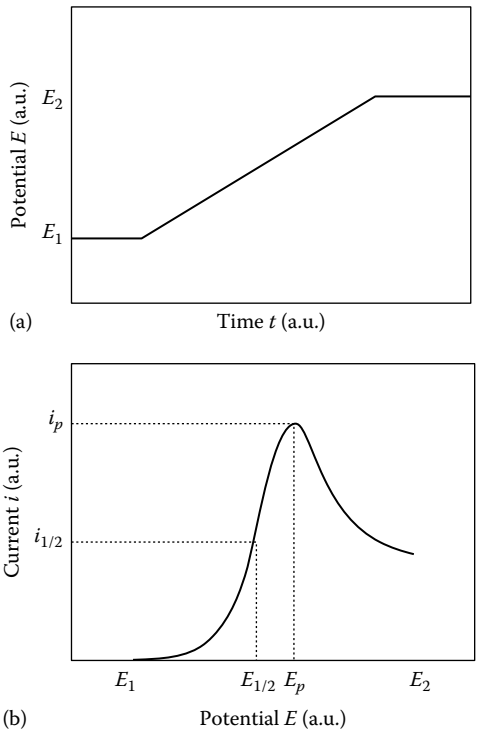


FIGURE 55.12 In LSV the potential varies linearly with time (a). The current–potential relation yields a peak-shaped curve with a half-wave potential $E_{1/2} \approx E^0$ (b). The peak current is proportional to the analyte's bulk concentration c_0.

The positive sign in Equation 55.35 applies to an anodic sweep (from negative to positive potential with $v > 0$) and the negative sign to a cathodic one (from positive to negative potentials with $v < 0$). The peak current is given by

$$i_p = 0.446nFA\left(\frac{nF}{RT}\right)^{1/2} D^{1/2} c_0 v^{1/2} \tag{55.36}$$

Thus, the peak current is proportional to the bulk concentration c_0 of the analyte and depends on the sweep rate according to $v^{1/2}$.

Another contribution to the measured current is the capacitive double-layer-charging current i_c, which always flows in LSV due to the continuous change of potential. It can be calculated using the following equation:

$$i_c = C\left(\frac{dE}{dt}\right) = Cv \tag{55.37}$$

which yields a proportionality to v while the faradaic peak current is proportional to $v^{1/2}$. Thus, for the faradaic current to dominate the measurement, the sweep rate should not be chosen too large. A sweep rate of 100 mV s^{-1} can be regarded as an upper limit. Moreover, the surface area of the working electrode must be taken into consideration. Rough electrodes have a much larger active than geometric surface area and thus a very large capacitance. Therefore, small, very smooth electrodes should be chosen.

A variation of LSV is a technique called *cyclic voltammetry* (CV). Here, the electrode potential is swept forth and back between two potentials E_1 and E_2 (Figure 55.13a). Although the bulk concentration of the reaction product is essentially zero, its concentration at the electrode surface after the first sweep is quite large. In the backward sweep, the reaction product of the analyte is converted into the analyte again. The current flows in the opposite direction and using an *xy* recorder, an *i*–*E* curve is obtained (Figure 55.13b). From Equation 55.35, it follows that for reversible processes, the peak potentials of the forward and backward sweep have a distance of (56/*n*) mV at room temperature. Therefore, CV is a favorable method for the investigation of the reversibility of a system. If the electrode current totally decays in the forward sweep, the analyte concentration has dropped to zero and the product concentration at the electrode surface is about c_0. Ideally, the peak current during the reverse scan should be equal (with reversed sign) to the peak current of the forward sweep.

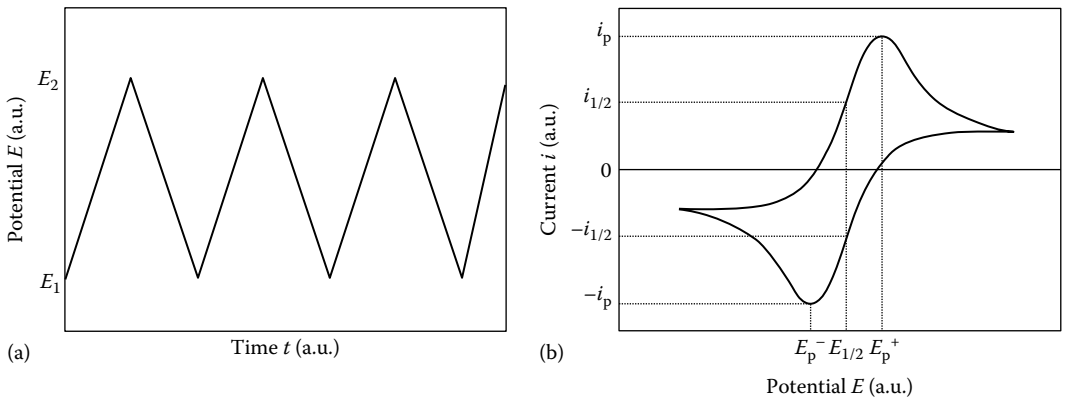


FIGURE 55.13 In CV, the potential is swept forth and back between two fixed values (a). The current–potential relation yields a peak-shaped curve with a half-wave potential $E_{1/2} \approx E^0$ (b). The peak current is proportional to the analyte's bulk concentration c_0. For totally reversible systems, the peak currents of the forward and the backward sweep are equal in magnitude but of opposite sign.

Although the theory of LSV and CV measurements is very promising, the methods have several practical limitations. One is the frequently insufficient stability of the *i*–*E* characteristic during the first cycles in CV. However, after 5–10 cycles, it tends to become highly reproducible. Yet, one must be careful deriving quantitative information from these later cycles because the initial and boundary conditions of the diffusion problem have changed and convective mass transport may already play a role. Thus, the equations developed for LSV cannot be used. Another problem that concerns both LSV and CV is the potential drop that occurs in the solution between the working and the reference electrode and that leads to a distortion of the shape of the *i*–*E* wave. This error increases with increasing current flow. Thus, the rate of change v of the electrode potential is not really constant, as has been assumed in the boundary conditions for solving the diffusion equations. Furthermore, the quantitative information is usually obtained from the position E_p and the height i_p of the current peak where the error is maximum. Finally, the determination of the peak height itself is sometimes problematic due to difficulties in the extrapolation of the baseline. For all these reasons, it may be advisable to verify the results of quantitative analysis with additional methods. Nevertheless, on easy terms, the LDL of LSV and CV in quantitative analysis can amount to 10^{-7} mol L^{-1} with a resolution of about 50 mV.

Besides the analysis of faradaic processes, LSV and CV are favorable techniques for the investigation of the adsorption of species on the electrode surface (Bard and Faulkner 2001). In such adsorption processes, the current is called *pseudocapacitive current*. Although it is a charge transfer across the interface, it exhibits many of the properties of a pure capacitive current. The current–potential wave has a very similar shape as for faradaic processes. If Θ denotes the coverage ($0 \leq \Theta \leq 1$) and q_1 the charge that is required to form a monolayer of a species, the pseudocapacitive current i_a can be expressed as

$$i_a = q_1\left(\frac{d\Theta}{dt}\right) = q_1\left(\frac{d\Theta}{dE}\right)\left(\frac{dE}{dt}\right) = C_a v \tag{55.38}$$

where C_a is called the *adsorption pseudocapacitance*. The calculation of C_a yields that the pseudocapacitance does not depend on v. Therefore, at any potential the current is proportional to the sweep rate ($i \sim v$). The peak potential gives information about the adsorption kinetics. In contrast to faradaic CV, it has the same value for the forward and the backward sweep.

55.2.3.5 Pulse Techniques

Voltammetric pulse techniques are derived from potential step experiments to suppress the capacitive currents during the measurement. A potential step that can vary in amplitude and sign is periodically repeated and superimposed with a potential ramp. The current is measured at the end of the step when the double-layer charging current has largely decayed.

Normal Pulse Voltammetry: In *normal pulse techniques*, periodic voltage pulses with an increasing amplitude from pulse to pulse are superimposed on a constant potential. Typical pulse duration is about 50 ms and the current is measured during a time interval Δt_m of about 10–15 ms at the end of each pulse. Between two pulses there is a waiting period of a few seconds (Figure 55.14). In polarography with a DME, each drop is dislodged directly after the pulse and thus used for just one measurement.

Because normal pulse voltammetry equals a series of potential step measurements with increasing step widths, the current obeys to Equation 55.29 and the evaluation of the measured current values can be carried out using the sampled-current method. In comparison with the step technique, the LDL is enhanced for one to two orders of magnitude up to 10^{-6} and 10^{-7} mol L^{-1} (Henze 2008). The peak resolution is about 100 mV.

Square-Wave Voltammetry: In *square-wave* techniques, a periodic rectangular voltage is superimposed on a linearly rising potential ramp. The measuring interval lies at the end of a pulse when the capacitive current can be neglected (Figure 55.15a). Typical pulses have frequencies between 200 and 250 Hz and an amplitude of ΔE_p = 5–30 mV (Henze 2008). The capacitive current is suppressed even more effectively if the pulse is tilted to decrease during the pulse period. No pulse tilt is required

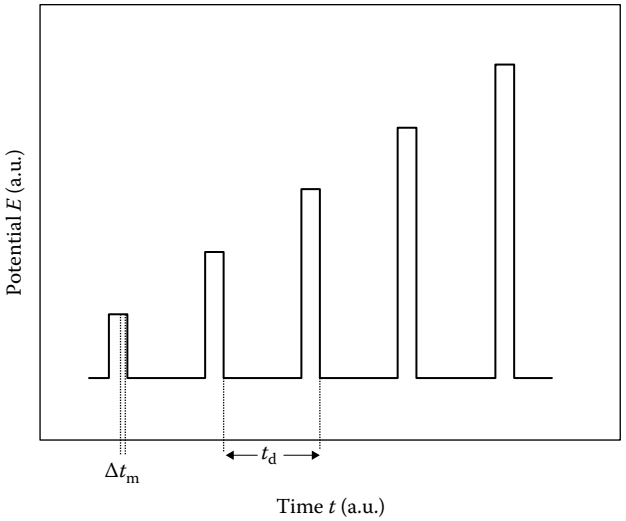


FIGURE 55.14 Normal pulse voltammetry equals a series of potential step measurements with increasing step widths. The current is measured during a time interval t_m near the end of the pulse. In polarography with a DME, the drop is dislodged after each measurement. The drop's lifetime is denoted by t_d.

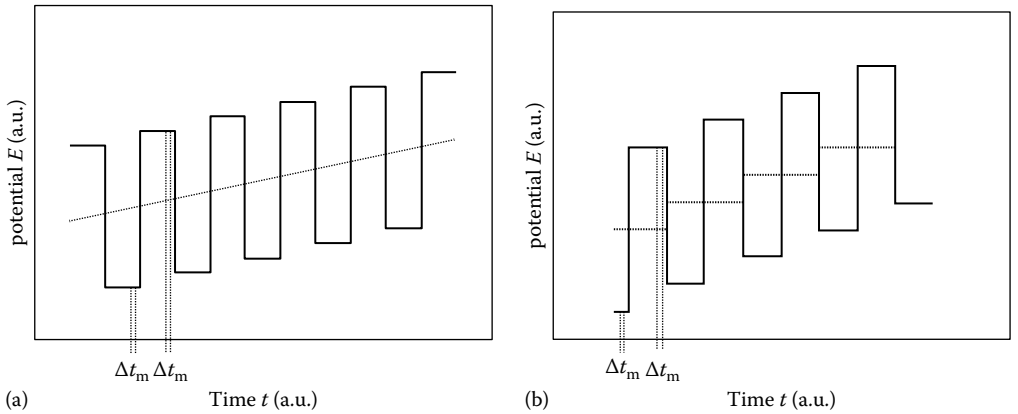


FIGURE 55.15 In square-wave voltammetry, a periodic rectangular voltage pulse is superimposed on (a) a linearly changing potential ramp (dotted line) or (b) on a stepped ramp (dashed curve). The current is measured during a time interval Δt_m at the end of each pulse.

if the potential ramp is stepped (staircase ramp) instead of a linear ramp. The voltage pulse is then applied on the plateau of the stepped ramp (Figure 55.15b).

After rectification of the measured current values, one obtains peak-shaped i–E curves. The peak potential corresponds to the half-wave potential of LSV and, thus, to the standard potential of the analyte's redox couple. The peak current i_p depends on the frequency and amplitude ΔE of the voltage pulses and obeys

$$i_p \sim n^2 D \Delta E c_0 \tag{55.39}$$

where the frequency dependence is included in the proportionality constant. The LDL is in the range of 10^{-8} mol L^{-1} and the peak resolution amounts to 40–50 mV.

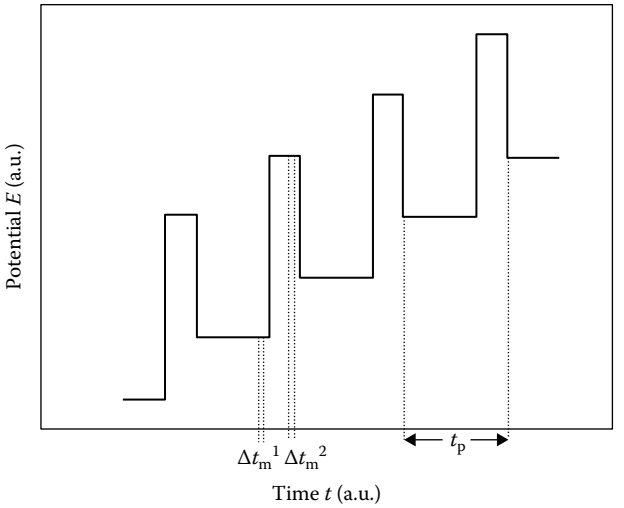


FIGURE 55.16 In differential pulse voltammetry, periodic rectangular pulses are superimposed on a stepped potential ramp. The difference between the current measured in a time interval Δt_m^1 directly before each pulse and during a time interval Δt_m^2 at the end of each pulse is plotted against the base potential. In polarography with a DME, the drop is dislodged after each pulse. The drop's lifetime is denoted by t_d.

Shorter analysis times are achieved if very short and relatively large rectangular pulses with a duration t_p = 5–10 ms and an amplitude of ΔE = 50 mV are superimposed on a stepped potential ramp with the same duration but smaller potential steps of about 10 mV. The potential can then be scanned at extremely high rates of up to 1200 mV s^{-1}. However, the sensitivity decreases because the ratio of faradaic to capacitive currents is lowered by the short pulse times.

Differential Pulse Voltammetry: Differential pulse methods are the most important ones in analytical voltammetry. Periodically repeated rectangular voltage pulses with a constant amplitude ΔE of several 10 mV are superimposed on a stepped potential ramp (Figure 55.16). The pulse duration Δt_p is about 5–100 ms (Henze 2008). Between two pulses, the potential is held constant for a few seconds. The current is measured in a short time interval ($\Delta t_m \approx$ 1–20 ms) directly before a pulse is applied and for the same duration near the pulse end. If a DME is used, the drop is knocked off mechanically between two pulses and each drop serves for just one measurement.

For the evaluation, the difference between the two measured current values Δi that corresponds to one pulse is recorded as a function of the base potential. A peak-shaped curve is obtained with a maximum very close to the half-wave potential $E_{1/2}$. The peak height is proportional to the analyte concentration in the bulk:

$$\Delta i_p \sim nFA\left(\frac{D}{\pi t_p}\right)^{1/2} c_0 \tag{55.40}$$

With differential pulse measurements, a LDL of 10^{-8} mol L^{-1} and a resolution of 50–100 mV can be achieved.

55.2.3.6 Alternating Current Voltammetry

Alternating current techniques are similar to differential pulse methods. A linear potential ramp is modulated with a low-frequency ($f \sim$ 50 Hz) sinusoidal alternating voltage of small amplitude ($\Delta E \sim$ 50 mV) (Henze 2008). The amplitude of the resulting alternating current is plotted against

the base potential. A peak-shaped curve is obtained with a maximum that is proportional to the bulk concentration of the analyte:

$$i_p \sim \Delta F f^{1/2} c_0 \tag{55.41}$$

The LDL is 10^{-5} mol L^{-1} due to the large capacitive currents. It can be enhanced by phase-selective rectification because the capacitive and the faradaic currents have a phase shift of 90° and 45°, respectively. The peak resolution amounts to 50–100 mV.

55.2.3.7 Stripping Voltammetry

Stripping techniques can be performed with analytes whose reaction products adsorb on the electrode surface. For accumulation, the electrode potential is held at a value at which the electrochemical equilibrium is on the product's side. Accumulation times usually amount up to several minutes. During this period, the solution is stirred to prevent the depletion of the analyte at the electrode surface. The accumulation is followed by a rest period of 2–30 s, during which the solution remains unstirred and the current falls to a small residual value. In the subsequent *stripping step*, the electrode potential is shifted to a value at which the adsorbed product is reconverted into the analyte by oxidation or reduction. Depending on whether an oxidation or reduction process occurs, the method is called *anodic stripping voltammetry* (ASV) or *cathodic stripping voltammetry* (CSV), respectively. The stripping step can be performed in various manners (Wang 1998) of which the linear sweep method shall be exemplary discussed here. It yields a peak-shaped *i*–*E* curve with a maximum at

$$E_p = E_{1/2} - \frac{1.1\,RT}{nF} \tag{55.42}$$

and a peak height that is proportional to the bulk analyte concentration according to

$$i_p \sim n^{3/2} v^{1/2} c_0 \tag{55.43}$$

Different substances can be determined in successive experiments with an adequate choice of the accumulation potentials. For the first measurement, the accumulation potential is chosen to allow adsorption of only one species; in the next experiment, the first and one further analyte adsorb, and so on. For the simultaneous determination of two or more substances, their peak potentials should be at least 150 mV apart.

A special case of the stripping techniques is *adsorptive stripping voltammetry* (AdSV). Here, the analyte is deposited in the form of metal chelates or organic molecules. For the formation of metal chelates, a complexing agent is added to the electrolyte or the surface of a solid-state electrode is modified with it. The stripping current is then due to the oxidation or reduction of the central atom or the ligand of the metal chelate complex. With this method, organic and organometallic compounds can be determined in the ultratrace range.

A crucial point in stripping analysis is the reproducibility. All experimental parameters have to be selected very carefully. In particular, the electrode surface must not be changed significantly by the adsorption and dissolution processes. Therefore, HMDEs are frequently employed for stripping analysis. A new drop is produced for each measurement. Another advantage of mercury electrodes is the fact that not only their surface but rather the hole bulk is used for the accumulation of analyte species. Consequently, more material can be collected. This leads to an enhanced lower determination limit that can be below 10^{-8} mol L^{-1}. Comprehensive monographs about stripping techniques are given in Brainina et al. (1993) and Wang (1998).

55.2.4 Applications

Analytical applications of voltammetry concern the determination of (heavy) metal cations, typical anions (halides, pseudohalides), organometallic, and organic compounds in the 10^{-4}–10^{-9} mol L^{-1} concentration range. Therefore, they are established in several fields like environmental, medical, food, and water analysis. A disadvantage is the usually labor-intensive sample preparation necessary, for example, to disintegrate ions from complexes, to adjust the pH of the solution, or to remove interfering species like oxygen and organic molecules. Principally, the preparation of the electrode (surface) is also crucial. However, commercially available equipment is well developed not only to enhance determination limits, sensitivity, selectivity, and reproducibility but also to reduce the expense for electrode and cell preparation. Moreover, sample and electrode preparation can be automated to a certain degree by devices that pump different solutions for cleaning, conditioning, and analysis through the cell setup. A further improvement of the instrumentation is the use of *microelectrodes* with dimensions of 1–100 μm. Because their dimensions are small in comparison with the diffusion length of the analyte, even for planar microelectrodes the diffusion is rather hemispherical than linear. Therefore, the depletion effect is less strong and the faradaic current is increased. Moreover, planar microelectrodes can be rotated (*rotating disk electrode* (RDE)) to intensify convection, and the solution can be stirred with ultrasound. Another advantage of microelectrodes is the possibility to realize several electrodes in a close neighborhood, so-called *electrode arrays*. They serve as *one* electrode if they are held at *one* potential and exhibit an improved signal-to-noise ratio due to the better diffusion conditions. In contrast, if different potentials are applied at different electrodes, the simultaneous determination of different species is possible. These techniques have just become commercially available as electrochemical detectors, for example, for high-performance liquid chromatography (HPLC). In this arrangement, the different species in the solution are separated by the HPLC and flow through the detector cell one after the other. Thus, interference between different analytes is minimized. The selectivity can often be further improved by the use of membrane-covered microelectrodes. The well-known *Clark oxygen sensor* and different biochemical sensors represent promising examples of this application in amperometry. Moreover, it opens up new possibilities for the creation of microelectrode arrays.

Due to the high analytical potential and the relatively low costs of voltammetric methods in comparison with spectroscopic techniques, all aspects of voltammetry are still subject of intense research. Current efforts concern the *miniaturization* of the whole cell, including microchannels, microvalves, micropumps, and microelectrodes by means of precision mechanics and *micromachining techniques* (Mastrangelo and Tang 1994). They employ fabrication methods of silicon planar technology and Lithographie, Galvanoformung, Abformung (LIGA) technique. Thin-film techniques, like physical and chemical vapor deposition (PVD, CVD), allow the fabrication of electrodes with a thickness in the submicrometer range and with lateral dimensions from the micrometer to the nanometer range. One goal is the realization of a *microsystem* with the sensitive components (i.e., the electrodes) and microelectronics integrated on a single chip. Currently, there are several commercially available systems. Two companies dominate the market of microelectrode arrays: Multichannel Systems (Reutlingen, Germany) and Panasonic (Tokyo, Japan). Modern microelectrode systems comprise the microelectrode arrays, electronic circuitry, and computer softwares for signal amplification and processing, respectively (Ryynänen et al. 2011).

55.3 Potentiometry

Potentiometry implies the measurement of an electrode potential in a system in which the electrode and the solution are in electrochemical equilibrium. Thus, the potential becomes the dependent variable, for example, as a function of time. In potentiometry, the current is attempted to be kept as small as possible; ideally, it should be zero. Potentiometry implies known fluxes (i.e., concentration gradients at the electrode surface) and thus information on the composition of the sample. In this section, potentiometry is

related to the measurement of potentials, where the voltage source is a form of a galvanic cell, consisting of a measuring electrode and a reference electrode (in general, electrodes of the second kind). The principles of *direct potentiometric measurements* as well as *potentiometric titrations* will be described.

55.3.1 Ion-Selective Electrodes

The equipment required for potentiometric analysis includes a measuring electrode, also called an ISE or *indicator electrode*, and a reference electrode. In addition to the sensitivity, the most important characteristic of the ISE is given by its *selectivity*. Depending on the type of membrane, ISEs can be classified into four different groups: *glass electrodes, solid-state electrodes, liquid-membrane electrodes*, and miscellaneous *combined electrodes* (Pungor 1998; Koryta and Stulik 2009). For all ISEs, the validity of the Nernst equation could be proved.

55.3.1.1 Glass Electrodes

The most common *glass electrode* is the *pH electrode*, widely used for hydrogen ion determination. The pH-glass electrode consists of a thin, pH-sensitive *glass membrane* sealed to the bottom of an ordinary glass tube. The tube is filled with a solution of hydrochloric acid (e.g., 0.1 M HCl) that is saturated with silver chloride. A silver wire, connected to an external potential-measuring device, is immersed in this solution. Note that the internal HCl concentration is constant and, thus, the internal potential (inner surface of glass membrane) of the pH electrode is fixed. Only the potential that occurs between the outer surface of the glass bulb and the test solution responds to pH changes. To measure the hydrogen ion concentration of the test solution, the glass electrode (indicator electrode) must be combined with an external reference electrode, which is required for all kinds of ISE determination. Often, pH-glass electrodes are available as a combination of the indicator electrode and an internal reference electrode (e.g., Ag/AgCl in saturated KCl solution) as schematically shown in Figure 55.17.

The composition of the glass membrane clearly influences the sensitivity of the pH electrode. Usually, three-component systems of, for example, $SiO_2/Na_2O/CaO$, are employed (Covington 1979). The pH

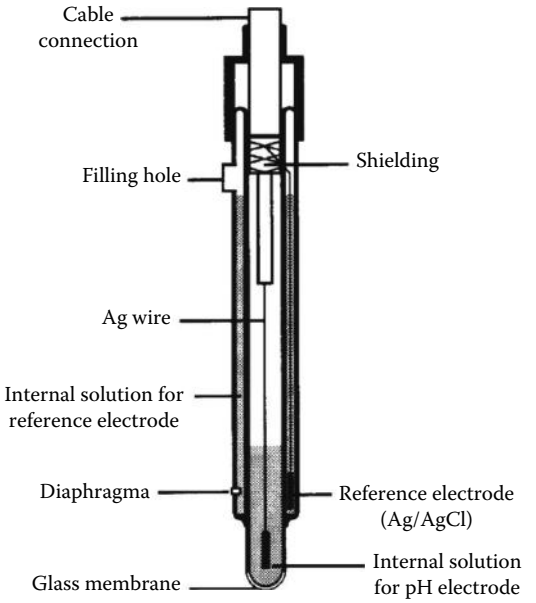


FIGURE 55.17 Combination of pH-glass electrode with an integrated Ag/AgCl reference electrode.

dependence can be expressed by the Nernst equation (Equation 55.11). At room temperature ($T = 25$ °C), Equation 55.11 can be simplified by

$$E = E^0 + 59.1 \text{ mV pH} \tag{55.44}$$

where E^0 is the standard Galvani potential with respect to the SHE. Thus, the measured potential is a linear function of pH within an extremely wide range (10–14 decades). The selective pH response of the pH ISE is due to the ion exchange process, in particular, due to the replacement of sodium ions in the glass membrane (m) by protons in the solution (s), and vice versa:

$$H^+_{(s)} + Na^+_{(m)} + \Leftrightarrow H^+_{(m)} + Na^+_{(s)} \tag{55.45}$$

The sodium ion exchange is also responsible for the *alkaline error* of pH electrodes in solution with pH greater than 10. In spite of the high resistance of the glass membrane against chemical attack, one has to deal with deviations (alkaline error) from the linear pH dependence. This error (i.e., the sensitivity toward alkali-metal ions) can be greatly reduced if Na_2O is replaced by LiO_2. Because pH-glass electrodes can be used in the presence of substances that interfere with other electrodes (e.g., proteins, oxidants, reductants, and viscous media), they have a wide range of applications. Typical fields are the clinical and food analysis, environmental monitoring (e.g., industrial waste, acidity of rain), and process control (e.g., fermentation, boiler water, galvanization, and precipitation).

The employment of glass membranes prepared with different glass compositions allows an electrode response sensitive to cations. For example, sodium-, potassium-, and ammonium-selective glasses consist of a mixture of Na_2O, Al_2O_3, and SiO_2 in various proportions (aluminosilicate glasses). Using specific compositions and mixtures of chalcogenides, ion-selective *chalcogenide glass electrodes* with sensitivities toward monovalent ions (e.g., Ag^+, Tl^+, F^-, Cl^-, Br^-, I^-) and double-charged species (e.g., Cu^{2+}, Pb^{2+}, Cd^{2+}, Hg^{2+}, S^{2-}) can be prepared (Vlasov et al. 1994). However, in all cases, some sensitivity to charged species (e.g., H^+ ions) remains. The electrode potential under these conditions is described by the *Nikolsky–Eisenman* equation:

$$E = E^0 \pm \frac{RT}{zF} \ln(a_i + K_{ij} a_j^{z_i/z_j}) \tag{55.46}$$

where

z_i, z_j and a_i, a_j are the ionic charge and activity of the primary or determined (i) and the interfering (j) ion

K_{ij} is the *selectivity coefficient*

It is a measure of the ISE ability to discriminate against the interfering ion. A small value of K_{ij} indicates an ISE with a poor selectivity.

55.3.1.2 Solid-State Electrodes

The glass membrane of an ISE can be replaced by a single or a mixed crystal or a polycrystalline (pressed) pellet (Figure 55.18a). With respect to their membrane composition, *solid-state electrodes* are divided into *homogeneous* and *heterogeneous membrane electrodes*.

A typical single-crystal electrode (homogeneous membrane electrode) is the fluoride-sensitive ISE, which contains a LaF_3 crystal doped with Eu^{2+}. The crystal with a thickness of about 2 mm is sealed into the bottom of a plastic tube. The internal solution (0.1 M of NaF and NaCl) controls the potential at the crystal inner side by means of an Ag/AgCl wire as reference electrode. In contact with the test solution at the crystal outer side, an electrochemical equilibrium is established, proportional to the fluoride ion activity. This is due to an ion exchange process at the phase boundary membrane/electrolyte. In particular, fluoride ions from the membrane are replaced by fluoride ions from the solution and vice versa, where the fluoride ions can migrate from one lattice defect to another inside the crystalline membrane.

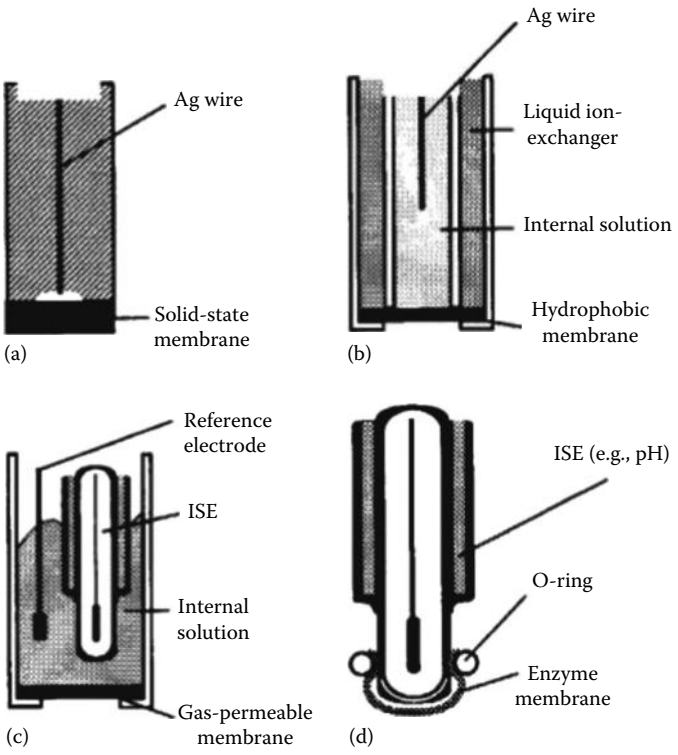


FIGURE 55.18 Typical membrane electrode types: solid-state electrode (a), liquid-membrane electrode (b), gas-sensing electrode (c), and enzyme-based electrode (d).

Further homogeneous membrane electrodes are silver halide electrodes, where the respective silver halide (AgCl, AgI, AgBr, Ag_2S) is pressed into a pellet, placed in a tube, and contacted via a silver wire. In these substances, silver ions are accordingly able to migrate. Such electrodes have been successfully used for the selective determination of chloride, bromide, iodide, silver, and sulfide ions. Likewise, if the pellets contain Ag_2S together with the silver halides or mixtures of PbS, CdS, and CuS, solid-state electrodes sensitive toward Pb^{2+}, Cd^{2+}, Cu^{2+}, and SCN^- can be realized. Moreover, the general problem of light sensitivity and high membrane resistance can be reduced by the additional use of Ag_2S.

Instead of the pressed pellets, the ion-selective material can be incorporated into an organic polymer matrix, like silicon rubber, carbon paste, or paraffin. In heterogeneous membrane electrode preparation, a mixture of the precipitate (e.g., AgI/Ag_2S) and polysiloxane is homogenized, and the polymerization is carried out. The resulting disks are fixed on the end of a tube and the internal solution (e.g., 0.1 M Kl) is contacted via a Ag/AgCl wire. *Coated-wire electrodes* represent another possibility. They can be manufactured by coating an appropriate polymeric membrane onto a conducting wire. Often, the conductor (Pt, Ag, Cu, or graphite) is dipped in a solution of polymer (e.g., polyvinylbenzylchloride (PVC) or polyacrylic acid) and the active substance. These electrodes allow the determination of K^+, Na^+, amino acids, and some drugs (e.g., cocaine). In addition to their simple miniaturization, the preparation is easy and inexpensive. However, further work is necessary to improve their analytical performance with regard to reproducibility and long-term stability.

55.3.1.3 Liquid-Membrane Electrodes

Liquid-membrane electrodes are based on two different membrane-active components: *solid ion-exchanger* and *complex-forming neutral-charged carriers*. They permit the determination of several polyvalent cations as well as certain anions. The sensor membrane (10–100 μm thickness) is usually prepared

of a plasticized PVC containing the organic sensor-active component that is insoluble in water; a Ag/AgCl wire is immersed into the internal reference solution. The liquid-membrane electrode differs from the glass electrode only in that the test solution is separated from the solution with the known target ion activity by a hydrophobic membrane, instead of the glass layer (Figure 55.18b). As membrane materials besides PVC, Teflon, sintered glass, filtering textile, or disks can be employed to hold the organic layer.

Liquid-membrane electrodes with ion exchangers have been realized for the determination of, for example, Ca^{2+}, K^+, BF_4^-, ClO_4^-, IO_4^-, SCN^-, I^-, Br^-, Cl^-, HCO_3^-, $H_2PO_4^-$, and NO_3^-. On the other hand, the synthesis of compounds containing individual cavities of molecule-sized dimensions results in complex-forming neutral-charged carriers. These *ionophores* (e.g., crown ethers like cyclic polyether, depsipeptides like valinomycin, and macrotetrolides like nonactin and monactin) are capable of enveloping various target ions reversibly in their pockets. For example, valinomycin membranes show a high K^+ selectivity. Many cyclic and monocyclic carriers with remarkable ion selectivities have been successfully developed for the determination of Li^+, Cs^{2+}, Ca^{2+}, Na^+, NH_3^+, Mg^{2+}, Ag^+, Hg^{2+}, SCN^-, and $H_2PO_4^-$. For all kind of membranes, a high molecular weight (i.e., a slight overpressure) prevents the quick intrusion of the test solution inside. Hence, the electrode's lifetime is limited as a consequence of diffusion of the sensor-active component into the analyte (*leaching out*).

55.3.1.4 Combined Electrodes

Two different types of *combined electrodes* will be presented here: *gas-sensing electrodes* and *enzyme-based electrodes*. Gas-sensing electrodes can be used to determine solutions of gases. They consist of an inner sensing element, normally a suitable ISE with an electrolyte solution (0.1 M), surrounded by a gas-permeable membrane (Figure 55.18c). On immersion of this ISE, the gas-permeable membrane contacts the liquid of the gas that diffuses through it, and the resultant internal solution will be examined with the ISE. The partial pressure of the gas attains equilibrium between the test solution/membrane and the membrane/ISE phase boundary. For example, the determination of carbon dioxide, which diffuses through the semipermeable membrane, lowers the pH values of the inner solution:

$$CO_2 + H_2O \Leftrightarrow HCO_3^- + H^+ \tag{55.47}$$

Such pH changes are detected by the ISE, in this case by a pH-sensitive glass electrode. Semipermeable membrane materials are polytetrafluoroethylene, polypropylene, or silicone rubber. The internal solution contains sodium chloride and an electrolyte with the corresponding ion that is determined. Gas-sensing electrodes have been realized for gases dissolved in solution, such as NH_3, NH_4Cl, CO_2, H_2CO_3, $NaHCO_3$, NO_2, $NaNO_2$, SO_2, H_2SO_3, $K_2S_2O_5$, CN, SCN, Cl_2, Br_2, I_2, and H_2S.

Enzyme electrodes are based on the coupling of an enzymatic membrane with any type of appropriate ISE. The enzyme converts (*catalyzes*) the analyte (*substrate*) to be determined extremely selective into an ionic product. The latter can be detected by the known ISE (Figure 55.18d). The coupling of the enzyme can be carried out by several *immobilization procedures*, such as entrapping in a gauze or gel, adsorptive or covalent binding, and cross-linking. A typical example for the operation of an enzyme electrode is given by the urea electrode. The enzyme urease hydrolyzes urea in order to liberate ammonium ions:

$$CO(NH_2)_2 + H_2O + 2H^+ \xrightarrow{\text{urease}} 2NH_4^+ + CO_2 \tag{55.48}$$

Either the alteration of the pH by a pH ISE or the variation of the NH_4^+ concentration by an ammonium-sensitive gas electrode can be detected. Likewise, penicillin, glucose, lactate, phenol, creatinine, cholesterol, salicylate, or ethanol will be catalyzed by means of the respective enzyme. Using different biological components (enzymes, cells, tissues, antibodies, receptors, or nucleic acids), a wide variety of analytically important substances for clinical, environmental, and food analysis can be determined. However, disadvantages of this type of electrode are its slow response time (several minutes) and the insufficient stability in the long term.

55.3.2 Instrumentation and Measurement

For potentiometric measurements, one uses an indicator electrode (ISE) versus a reference electrode and a *potentiometer*, also called *pH meter* or *ion meter*. Owing to the high resistance of the ISE membranes (e.g., 5–500 MΩ for the glass membrane), a potentiometer with a high input resistance is required. Modern potentiometers consist of an electronic digital voltmeter with a suitable operational amplifier, scaled directly to pH units or mV, with a resolution of better than ±0.002 pH and ±0.1 mV. They may range from simple handheld instruments for field applications to more convenient laboratory models. Frequently, potentiometers include a bias control that can be adjusted to correspond to the temperature of the test solution (*automatic temperature compensation*).

Direct Potentiometry: Direct potentiometric measurements can be performed for the determination of ionic species for which an appropriate ISE is available. A schematic measuring setup for direct potentiometry is shown in Figure 55.3. The measuring technique is quite simple: comparing the potential of the ISE in the test solution with its potential in a known *standard solution*. That means, before the determination, the ISE must be calibrated in solutions of known concentration of the chosen ionic species. Thus, for the ion determination to be made, at least two to three reference solutions are necessary that differ by two to five concentration decades. Typical resulting *calibration curves* for anions and cations are plotted in Figure 55.19. The curves can be separated into three distinct regions: (1) the straight part corresponding to the Nernstian slope (i.e., the sensitivity of the ISE), (2) the curve portion, and (3) the horizontal part below the LDL, where almost no sensitivity exists. The LDL of the ISE is defined as the concentration at which the extrapolated horizontal portion of the graph intersects the extrapolated Nernstian portion of the graph.

For practical applications, there are two aspects to be dealt with: often a *total ionic strength adjuster buffer* (TISAB) is added to both the standard solutions and the test solution (same temperature) to achieve comparable ionic strengths. Then, the potential difference can be assigned to the equivalent concentration of the calibration curve. Various methods for calibration calculations are described by, for example, Gran's plot or the standard addition method (Mendham et al. 2000). Because all measurements take place in dilute solutions (≤0.1 M), ion concentrations can be used in the Nernst equation instead of ion activities.

Potentiometric Titration: Potentiometric titration can be applied in the fields of *acid–base, precipitation, complex formation*, and *redox reactions*. Therefore, the ISE is used in combination with a reference electrode in order to establish the *EP* in a titration curve. A typical S-shaped potentiometric titration curve,

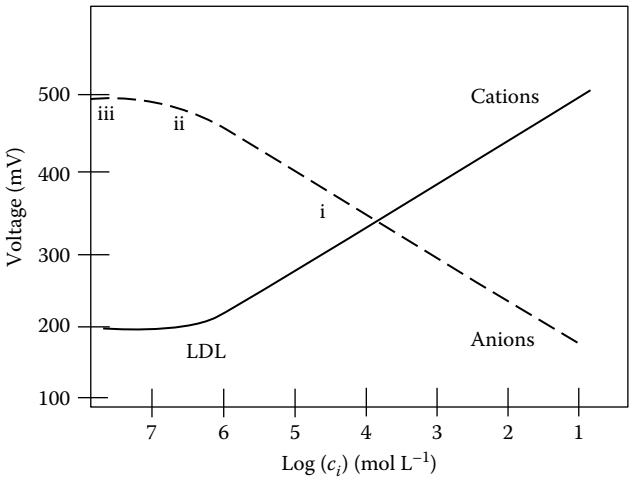


FIGURE 55.19 Schematic calibration curves for anions and cations (LDL).

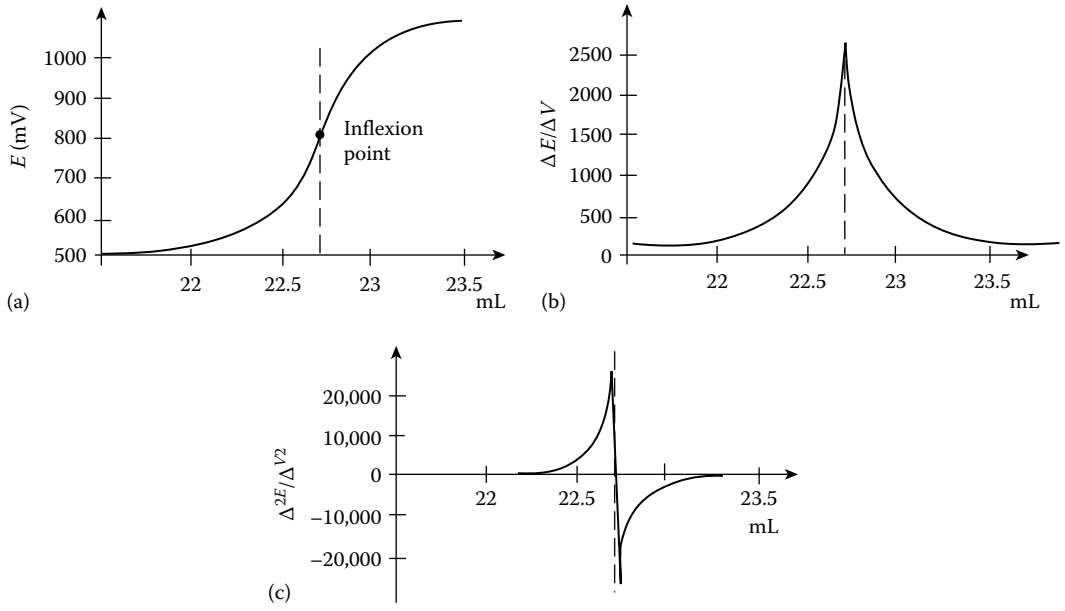


FIGURE 55.20 Characteristic potentiometric titration curve (a), first-derivative curve (b), and second-derivative curve (c).

where the electrode potential is plotted versus the reagent volume (*titrant*), is given in Figure 55.20a. The titrant is added to the initial solution that is stirred, and the ISE records the potential value at equilibrium. The EP (*endpoint*) of the reaction is reached when a sudden change in the potential of the ISE occurs. The midpoint in the curve (i.e., the steeply rising portion) is termed *endpoint* or *inflection point*. It can be evaluated by analytical methods, namely, the first- and second-derivative curves (Figure 55.20b and c). The first-derivative curve gives the potential change per unit change in volume of reagent and depicts the endpoint at the maximum of the inflection point. The second-derivative curve is zero where $\Delta E/\Delta V$ reaches its maximum. The greater the slope at the endpoint, the smaller should be the volume increment in order to reduce titration errors.

For practical applications, modern microprocessor-controlled titrators are commercially available (*auto-titrator*), coupled to a chart recorder to produce the titration curve directly. Such instruments also allow to evaluate the first- and second-derivative curves and provide Gran's plot. Acid–base (neutralization) titrations are performed with a glass/calomel electrode system and can be used to titrate a mixture of acids that differ greatly in their strengths (e.g., acetic (ethanoic) and hydrochloric acids). For precipitation titrations, the ISE consists of an electrode (e.g., silver or a platinum wire) that quickly reaches equilibrium with the ions to be precipitated. A typical precipitate reagent represents silver nitrate for the determination of halogens, halogenides, mercaptans, sulfides, arsenates, phosphates, and oxalates. For complex-formation titrations, membrane electrodes can be used that involve the formation of soluble complexes, like ethylenediaminetetraacetic acid (EDTA) or silver cyanide ($Ag(CN_2)^-$). Oxidation–reduction titrations are performed by a platinum indicator electrode to any redox couples where the potential depends on the concentration ratio of the reactants. Some experimental details for potentiometric titration are described in Mendham et al. (2000).

As an alternative principle, *chronopotentiometry* is based on the observation of the change in potential of a working electrode as a function of time during electrolysis. Usually, this electrolysis is performed with a constant current, whereas the time is measured that is necessary for the potential to go from one level to another. Since chronopotentiometry is disappointing at concentrations below 10^{-4} mol L^{-1}, it is only a powerful tool for studying electrode processes at higher concentrations. Consequently, this method is not very important for practical applications.

55.4 Semiconductor Field-Effect Chemical and Biological Sensors

The integration of thin ion-selective membranes with solid-state devices leads to miniaturized *chemically sensitive solid-state devices*. Among the variety of concepts and transducers principles proposed for solid-state chemical and biological sensors, the integration of chemically or biologically sensitive materials (recognition elements) with semiconductor *field-effect* devices based on the *electrolyte–insulator–semiconductor* (*EIS*) system is one of the most attractive approaches (Bergveld and Sibbald 1988; Madou and Morrison 1989; Grattarola and Massobrio 1998; Bergveld 2003). Representative examples are *chemically sensitive field-effect transistors* (*ChemFETs*), *chemically sensitive capacitive* EIS sensors, and *light-addressable potentiometric sensors* (*LAPS*), which are currently being one of the basic structural elements of various chemical and biological microsensors (Poghossian and Schöning 2006, 2007). Although these three kinds of (bio-)chemical sensors have different device configurations and measurement setups, the transducer principle of using an electric field to create excess charge regions in a semiconductor substrate is common to all of them. These field-effect devices are based on the technology used for manufacturing microelectronic chips and thus offer the possibility of mass production. (Bio-)chemical sensors based on other types of field-effect devices, like *silicon-nanowire transistors* (Timko et al. 2010), *silicon thin-film resistors* (Neff et al. 2006), or *carbon-nanotube transistors* (Lee et al. 2009a), are still in the initial state of research and development.

55.4.1 Chemically Sensitive Field-Effect Transistors

The ChemFETs can react sensitive to some ions (*ion-sensitive FET [ISFET]*), biomolecules (*biologically sensitive FET [BioFET]*), or gases (*gas-sensitive FET [GasFET]*) in aqueous media, or they can be insensitive (*reference FET [ReFET]*). They incorporate the sensitive membrane directly on the gate area of a field-effect transistor (FET). A schematic of an *n*-channel ISFET, mounted in a measuring cell and contacted via a reference electrode (e.g., Ag/AgCl electrode), is given in Figure 55.21a.

It consists of a *p*-type silicon substrate with two *n*-doped regions, source (*S*) and drain (*D*), separated by a short channel that is covered by the gate insulator (typically, a thin SiO_2 layer) and the ion-sensitive membrane. For operating an ISFET, the gate voltage, V_G, is applied by a reference electrode. When the sensor membrane is placed into contact with the test solution containing the ions or analytes to be detected, an additional potential (ΔV) is generated at the membrane–electrolyte interface. If a sufficiently positive bias potential is applied to the gate (with respect to the bulk silicon substrate), an n-type *inversion layer* (usually less than 10 nm) is formed in the channel and the current, I_D, starts to flow between source and drain.

The *drain current* and *threshold voltage* V_T (the gate voltage at which a conductive channel is formed between source and drain) of the ISFET can be deduced from that of its electronic analogue *metal–oxide–semiconductor FET* (*MOSFET*) by simply adding the potential drops at the additional interfaces (Bergveld and Sibbald 1988; Poghossian and Schöning 2006):

$$I_d = \frac{C_i \mu b}{L}\left[(V_G - V_T)V_D - \frac{V_D^2}{2}\right] \tag{55.49}$$

for the *nonsaturated region* ($V_D < V_G - V_T$), and

$$I_d = \frac{C_i \mu b}{2L}(V_G - V_T)^2 \tag{55.50}$$

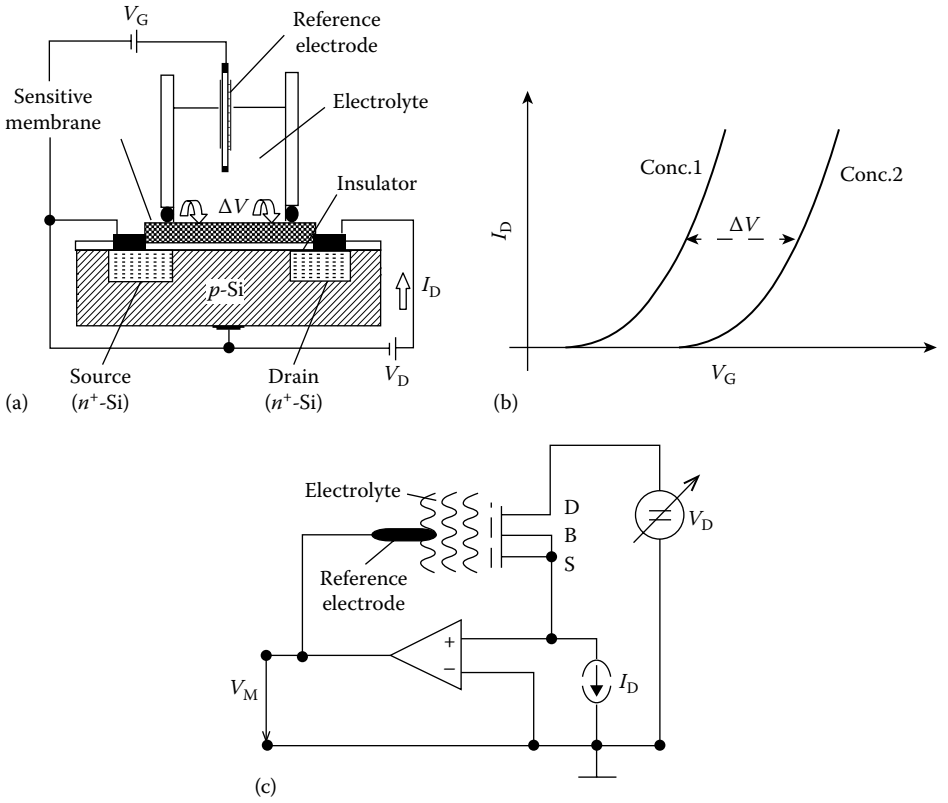


FIGURE 55.21 ISFET configuration (a), transfer characteristic (b), and schematic circuit of CCM (c). The metallic gate from a MOSFET is replaced by the arrangement sensitive membrane/test solution/reference electrode (V_G, gate-source voltage; V_D, drain-source voltage).

for the *saturated region* ($V_D > V_G - V_T$), with

$$V_T = E_{ref} - \Delta V + \chi - \frac{W_S}{q} - \frac{Q_i + Q_{ss}}{C_i} - \frac{Q_{sc}}{C_i} + 2\phi_B \tag{55.51}$$

where,

μ is the mobility of the electrons in the channel between source and drain
b is the width
L is the length of the channel
C_i represents the capacitance of the gate insulator per unit area
E_{ref} is the potential of the reference electrode
χ is the surface-dipole potential of the solution
W_S is the silicon electron work functions
q is the elementary charge (1.6×10^{-19} C)
Q_{sc}, Q_i, and Q_{ss} are the charges per unit area in the space-charge region as well as located in the oxide and the surface/interface states, respectively
ϕ_B is the potential difference between the Fermi level in the bulk semiconductor and the intrinsic Fermi level

If the applied gate (V_G) and source-drain (V_D) voltages are fixed, the only variable term is the analyte concentration-dependent interfacial potential ΔV. Changes in the chemical composition will

induce changes in the potential drop at the electrolyte–membrane interface that consequently will modulate the drain current of the ISFET. After calibration of the ISFET with standard solutions of known ion activity, the variation of I_D can be used to determine the ion concentration in the test solution (Figure 55.21b). Often, the ISFET is operated in a feedback loop (e.g., the *constant charge mode*, Figure 55.21c), and the voltage V_M needed to maintain I_D at a fixed value represents the sensor response. The sensor response can be described by the same Nernst and Nikolsky equations that characterize conventional IESs.

Like with the ISEs, the most attention is gained to pH-sensitive ISFETs, built up of SiO_2 and an additional thin film (about 30–100 nm) of other oxides and nitrides. The additional layer is necessary because the pH response of SiO_2, initially used as pH-sensitive dielectric, is poor (20–40 mV/pH) and the material was indeed found to be unstable and to suffer from considerable drift of the sensor signal. Therefore, different materials, like Si_3N_4, Al_2O_3, Ta_2O_5, IrO_x, ZrO_2, TiO_2, SnO_2, WO_3, HfO_2, AlN, TiN, $PbTiO_3$, barium strontium titanate, diamond-like carbon, and nanocrystalline diamond, have been investigated with respect to their pH sensitivity and stability (see, e.g., Poghossian and Schöning 2006). For example, Si_3N_4 shows a sensitivity of about 45–55 mV/pH. The sensitivity can be improved by using Al_2O_3 or Ta_2O_5 with 53–57 and 55–59 mV/pH, respectively. Also, the reported drift values are less than 1 mV h^{-1}. The pH sensitivity of these gate-insulating materials can be explained by the site-binding theory, which is exemplary discussed for SiO_2 in Madou and Morrison (1989). Other more exotic materials such as AlN, HfO_2, PbTiO3, and WO_3 sometimes show nearly Nernstian sensitivity, but have only rarely been studied. Methods that have been used to deposit pH-sensitive materials include electron-beam evaporation, thermal evaporation, CVD, sputtering, thermal oxidation, pulsed laser deposition (PLD), and solgel technique.

Being in its basic structure a pH-sensitive device, the selectivity of an ISFET toward other ions can be achieved by means of a modification of the gate surface or a deposition of subsequent organic or inorganic ion-sensitive membranes on top of the gate insulator. For instance, by implantation of high doses of B, Al, Ga, In, Ti, Li, or Na, potassium- and sodium-sensitive ISFETs were achieved. Also, the deposition of thin layers of modified chalcogenide glasses offers the determination of heavy-metal ions for biological investigations and industrial applications (Vlasov 1994; Schöning and Kloock 2007). By means of vacuum evaporation, ion-sensitive films of LaF_3, Ag_2S, or AgX (X = Cl, Br, I) for the determination of F^-, Cl^-, Br^-, I^-, Ag^+, and S^{2-} can be prepared. A chemical surface modification of the original gate insulator (e.g., the covalent linking of hydrophilic layers that contain the sensing molecule) leads to organic gate materials for the determination of different ions such as Ca^{2+}, NH_4^+, K^+, Cl^-, NO_3^-, Na^+, and Ag^+ (Rheinhoudt 1992). Similar results were obtained for homogeneous polymeric membranes, containing solid ion-exchanger or neutral-charged carriers (see section on ISE). In order to achieve well-defined and highly ordered sensor membranes, the gate can be coated with ultrathin Langmuir–Blodgett films (Schöning et al. 1995).

BioFETs are constructed from an ISFET by modifying the gate or coupling it with *biological recognition elements*. The biomolecules of various complexity (e.g., enzyme, antibody, protein, deoxyribonucleic acid [DNA]) or even living biological systems (e.g., cell, tissue slice, or whole organism) can be used as recognition element. Based on the hierarchy of biological complexity, the BioFETs can be subdivided as follows: *enzyme-modified FET* (*EnFET*), *DNA-modified FET* (*DNA-FET*), *cell-based FET*, and *beetle/chip FET*. A review of basic concepts of different kinds of BioFETs is given in Schöning and Poghossian and Schoning (2006), Lee et al. (2009b), and Poghossian et al. (2009).

EnFETs are usually constructed by immobilizing an enzyme or multienzyme system onto the gate insulator of an ISFET using different immobilization techniques (e.g., physical adsorption, enzyme entrapment within polymeric matrices, covalent binding, cross-linking). The EnFET directly corresponds to the enzyme ISE and detects the potentiometric response to either the concentration change in one of the products or reactants catalyzed by the enzyme. A multitude of EnFETs have been reported for the detection of numerous analytes such as glucose, penicillin, urea, lactose, sucrose, lactate, organophosphorus pesticides, creatinine, and glycoalkaloids (see, e.g., Poghossian and Schöning 2006 and

references there). Frequently, EnFETs are built up of pH ISFETs, in which hydrogen ions are produced or consumed by the enzymatic reactions. For an exact measurement, an on-chip integrated pH-ISFET/EnFET differential arrangement is often employed, where the pH ISFET contains a blank enzyme-free membrane and serves as a reference device. The demand of compatibility with integrated circuit technology provides enzymatic membranes that can be photolithographically patterned (e.g., photo-crosslinkable materials). Intensive researches during the last years were focused on the improvement of EnFET characteristics and to circumvent problems, which actually prevent the successful commercialization of EnFETs (e.g., dependence of the sensor response on buffer capacity, ionic strength, and pH of the test sample; restricted dynamic range and nonlinearity; relatively slow response and recovery times; operating and storage stability; reproducibility; dependence on enzyme immobilization method; the incompatibility of most used enzyme-containing layer deposition and patterning methods with silicon technology).

DNA-FETs: Generally, field-effect devices are surface-charge-measuring devices, and therefore, they are principally able to detect charged biomolecules (including DNA molecules) by their intrinsic molecular charge. Due to their small size and compatibility with advanced micro- and nano-fabrication technologies, these devices could offer a next generation of DNA chips with direct electronic readout for a label-free, fast, simple, and real-time analysis of nucleic acid samples. The core principle behind DNA chips is the hybridization event between two single-stranded DNA (ssDNA), where an unknown target ssDNA is identified by its complementary ssDNA molecule. As a result, a double-stranded DNA (dsDNA) helix structure with the two complementary strands is formed. Since DNA molecules are polyanions with negative charges at their phosphate backbone, the charge associated with the target molecule can effectively alter the charge applied to the gate, resulting in a modulation of the threshold voltage and the drain current of the DNA-FET. A critical review of different types of DNA-FETs is given in Poghossian et al. (2007) and Schöning and Poghossian (2008). The major disadvantage of electrostatic DNA detection by its intrinsic charge is the screening of the DNA charge by mobile ions in the surrounding solution. This could significantly reduce the expected sensor signal, especially in high-ionic-strength solutions, where the screening length (Debye length) is very short (~1 nm for a 0.1 M solution of monovalent 1:1 salt). To enhance the sensor signal, the DNA-FET must be operated in a very low-ionic-strength solution (<1 mM) and/or with a tightly packed probe ssDNA ($>10^{13}$ ssDNA cm^{-2}). To overcome the described problems, an alternative approach based on the detection of the DNA hybridization-induced redistribution of the ion concentration within the intermolecular spaces and/or the alteration of the ion sensitivity of the field-effect device has been proposed in (Poghossian and Schöning 2006).

Cell-based FETs or cell/transistor hybrids are obtained by direct coupling of single living cells (the smallest self-sustaining biological entity) or cell systems to the gate insulator of FETs (Wang 2005; Poghossian et al. 2009). Cell-based FETs can be used for both the monitoring of cellular metabolism (the extracellular acidification rate, concentration of ions, oxygen consumption, CO_2 production, and other metabolic products caused by different external stimuli) and the measurement of action potentials of electrogenic cells such as neuronal and muscle cells. For example, with the commercially available Bionas® 2500, analyzing system metabolically relevant data, including oxygen consumption, acidification rate, and adhesion (cell impedance) of cells, can be noninvasively measured in parallel and over a long period of time (Thedinga et al. 2007). Moreover, a cell-monitoring system, which includes both FETs for detecting the action potential and ISFETs coupled with different ion-sensitive membranes for the measurement of concentration of extracellular ions, like Na^+, K^+, and Ca^{2+}, has been realized (Wang et al. 2005). Those devices are of great interest for a large variety of applications including the detection of pharmaceutical agents, toxic substances, or pollutants and the monitoring of electrical communication within neuronal networks or transmission paths of ionic channels. Moreover, direct electrical interfacing of semiconductor and nerve cells is the physical basis for the development of hybrid neuroelectronic devices.

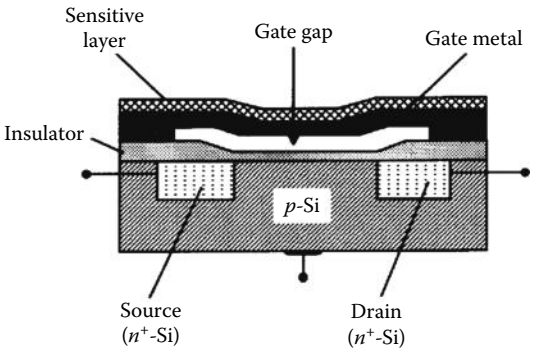


FIGURE 55.22 Schematic of an SGFET.

ReFETs consist of a sensor surface that is as insensitive as possible to all kinds of substances in the test solution. Thus, a differential pair of an ISFET and a ReFET eliminates disturbances, like influence of temperature or light. Appropriate materials to cover the ISFET surface with an insensitive layer are blocking materials, such as Teflon or different polymers (e.g., parylene, polyacrylate, PVC). However, not well-defined potential processes as well as some ion exchange will result in nonideal behavior. Alternative concepts use nonblocking polymer membranes with a fixed membrane potential or quasi-ReFETs with a delayed pH response. The most promising approaches are the application of an inert metallic layer or wire in a differential ISFET setup as a *quasi-reference electrode* and the miniaturization of conventional reference electrodes. For example, by means of PVD methods, Ag/AgCl electrodes were miniaturized on silicon chips inside anisotropically etched cavities (Madou and Morrison 1989).

The basic mechanism of *GasFETs* is due to the chemical modification of the *electron work function* of a metal–insulator–semiconductor (MIS) field-effect structure, for example, of a *suspended gate FET* (*SGFET*) as schematically shown in Figure 55.22.

The SGFET contains an additional insulator, the "*gap*" within the gate structure, which consists of a vacuum, a gas, or a nonconducting liquid. As gate metal, usually a platinum layer or mesh is used. The chemically sensitive layer on top of this structure, for example, palladium, exhibits sensitivity toward hydrogen. The hydrogen molecules adsorb and dissociate atoms (H_a) on the metal surface (Pd), depending on their partial pressure, as well as desorb from the metal surface by recombination into H_2 and reacting with oxygen to form water:

$$H_2 \Leftrightarrow 2H_a \quad \text{and} \quad 4\,H_a + O_2 \Leftrightarrow 2H_2O \tag{55.52}$$

The adsorbed atoms diffuse rapidly to the inner surface gap/insulator where they become polarized and form an interface dipole layer, resulting in a potential drop. For example, SGFETs with Pd, operated at 100 °C–140 °C, are sensitive to H_2, CO, and H_2S in the ppm range, whereas an increased operating temperature up to 240 °C allows the detection of alcohols (methanol, ethanol, propanol, butanol). To achieve selectivity, the surface of the suspended gate can be modified by inorganic or organic layers. Ammonia sensitivity can be achieved by catalytic metals such as Pt, Ir, Ru, or Rh. By the deposition of organic layers, like polypyrrole, sensitivities to alcohols and aromatic hydrocarbons are achieved. Several related devices based on SGFETs are explained in Josowicz and Janata (1988).

55.4.2 (Bio)Chemically Sensitive Capacitors

Sensors on the basis of capacitive field-effect EIS structures are much simpler to fabricate than ISFETs, and consequently they are favorable for laboratory use. Such *structures* correspond to MIS *capacitors*

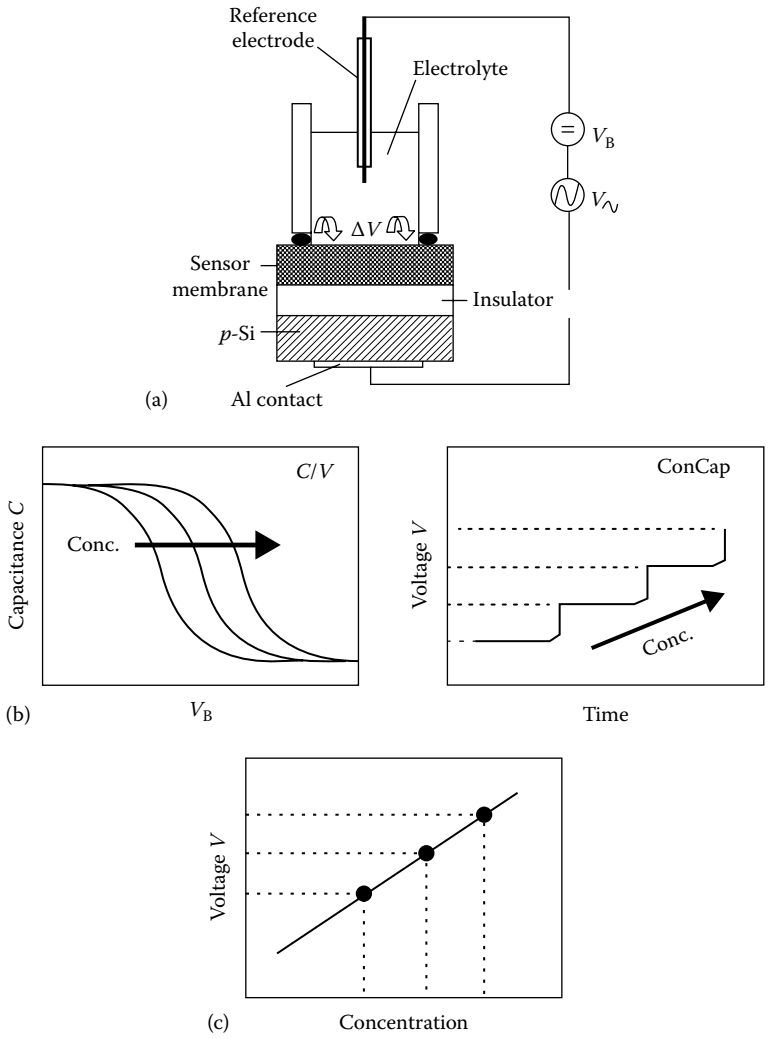


FIGURE 55.23 Schematic of an EIS structure (a), measurement of the EIS sensor in the *C*/*V* (capacitance/voltage) and the ConCap mode (b), and resulting calibration curve (c).

and their operation principle can be derived from the fundamental MIS devices (Sze and Ng 2007). A schematic buildup of an EIS structure and the measuring principle is given in Figure 55.23a.

The sensor consists of a *p*- or *n*-type semiconductor (silicon) covered by a thermally grown SiO_2 insulating layer (<100 nm) and the sensor membrane that is directly exposed to the test solution. For operation, a dc polarization voltage V_B is applied via the reference electrode to set the working point, and a small superimposed ac voltage (10–50 mV) is applied to the system in order to measure the capacitance of the sensor.

The functioning principle and physical properties of a capacitive EIS sensor can be explained by the charge carrier distribution at the insulator/semiconductor interface, which is controlled by both an external dc voltage (V_B) and an electrochemical interaction between the test solution and the sensor membrane (ΔV) (Poghossian and Schöning 2006). For a *p*-Si substrate, a negative potential ($V_B < 0$) on the reference electrode accumulates positively charged mobile carriers (i.e., holes) at the Si/SiO_2 interface (*accumulation regime*). When V_B becomes positive ($V_B > 0$), the holes are displaced from the interface, resulting in a space-charge region depleted of mobile carriers (*depletion regime*). If the potential

gets more positive ($V_B >> 0$), an inversion layer of *n*-type Si is created although the substrate is *p*-type (*inversion regime*). The electric behavior is given by the small-signal capacitance of the EIS structure. The whole capacitance (C) can be described as a series connection of the membrane capacitance (C_M), insulator capacitance (C_i), and space-charge capacitance (C_{SC}) of the semiconductor:

$$\frac{1}{C} = \frac{1}{C_M} + \frac{1}{C_i} + \frac{1}{C_{SC}} \tag{55.53}$$

with

$$\frac{1}{C_i} = \frac{\varepsilon_i}{d} A \tag{55.54}$$

where

A is the contact area of the gate with an electrolyte solution
d is the thickness
ε_i is the permittivity of the gate insulator

C_{SC} is, among others, a function of the gate voltage applied to the EIS structure and interfacial potential at the membrane–electrolyte interface. Typical C/V (*capacitance/voltage*) curves of a *p*-type EIS sensor are exemplary given in Figure 55.23b (left) for different ion concentrations. Due to the electrochemical interaction (ΔV), a horizontal shift of the C/V curve is provided, depending on the change of the ion concentration in the test solution. As resulting measuring signal (calibration curve), the potential shift can be evaluated at a fixed capacitance value within the linear region of the C/V curves (e.g., 60% of the maximum capacitance, Figure 55.23c). Using a feedback circuit, the measured capacitance can be adjusted at a fixed value in the *constant capacitance* (*ConCap*) mode (Figure 55.23b, right). Thus, potential shifts can be recorded directly.

Chemical and biological sensing EIS structures with different organic and inorganic sensor membranes have been developed within the last years (see, e.g., Poghossian and Schöning 2006 and references there). They consist of nearly identical sensor membrane materials and compositions as ISFETs, ranging from inorganic pH-sensitive layers (e.g., Si_3N_4, Al_2O_3, Ta_2O_5) or crystalline films (e.g., LaF_3, silver halides) over organic Langmuir–Blodgett films to enzymatic layers (e.g., urease, penicillinase). Much effort has been done in order to improve the limiting long-term stability that is often disclosed by FET devices in permanent contact with the analyte. Novel approaches pursue a further optimization with regard to the preparation (e.g., due to specific immobilization procedures) or the deposition of the sensor membrane in order to raise the sensor performance. For example, an extremely long-term stable pH sensor was developed by the suggestion of the PLD process as the thin-film preparation method. The EIS structure consists of a layer sequence of Al/p-Si/SiO_2/Al_2O_3, where no degradation of the pH sensitivity during a measurement period of 2 years was found (Schöning et al. 1996). In addition, a highly sensitive and corrosion-resistant non-glass unbreakable pH sensor has been developed using a Ta_2O_5 layer as pH-sensitive material (Schöning et al. 2005). The capacitive EIS sensors have been integrated into a flow-through microfluidic channel (fabricated by combining Si and SU-8 technologies) for pH and penicillin measurements (Poghossian and Schöning 2007). Moreover, EIS sensors have been applied for the detection of charged macromolecules (e.g., polyelectrolytes, DNA, dendrimers) and nano-objects (e.g., gold nanoparticles, carbon nanotubes) (Siqueira et al. 2009; Abouzar et al. 2011; Gun et al. 2011) Recently, a highly sensitive DNA hybridization/denaturation sensor has been realized using an array of individually addressable field-effect nanoplate *silicon-on-insulator* (*SOI*) capacitors modified with gold nanoparticles (5–8 nm) (Abouzar et al. 2012). Finally, the feasibility of EIS sensors for the creation of *molecular "**AND**" and "**OR**" logic gates* has been demonstrated in Poghossian et al. (2011).

Like GasFETs, MIS capacitors and *MIS Schottky diodes* are also available as gas-sensitive devices. For the MIS capacitor, a concentration-dependent dipole layer is detected as a shift of the C/V curve.

To reduce the drift of these devices, additional insulating layers, such as Al_2O_3, Si_3N_4, or Ta_2O_5, can be deposited between the metal layer and the SiO_2 insulator. Experimental results of $Pd/Al_2O_3/SiO_2/Si$ structures show sensitivities of 25 mV ppm^{-1} around 1 ppm (Armgarth and Nylander 1991). Schottky barrier diodes consist of a thin insulating layer (e.g., 2 nm SiO_2) between the metallic gate (e.g., Pd) and the semiconductor, in order to allow the current to pass through it. By variation of the metallic gate films, different sensitivities can be achieved, comparable to those of the SGFETs.

55.4.3 Light-Addressable Potentiometric Sensors

A schematic structure (a), a simplified equivalent circuit (b), and typical *photocurrent* response (c) of the LAPS are presented in Figure 55.24. The LAPS structure is similar to that of an EIS structure. Therefore, in the absence of illumination, the LAPS behaves like an EIS capacitor. As in the case of the capacitive EIS sensor, a dc bias voltage (V_{Bias}) is applied between the reference electrode and the semiconductor substrate in order to induce a space-charge region at the gate insulator–semiconductor interface. However, to detect the variation of the capacitance, the LAPS is, in contrast to EIS-based measurements, illuminated with a intensity-modulated light, which induces an ac photocurrent to be measured as the sensor signal (Wagner and Schöning 2007; Poghossian et al. 2009; Yu et al. 2010). The excitation light can be directed to the semiconductor either from front- or backside of the LAPS structure by placing

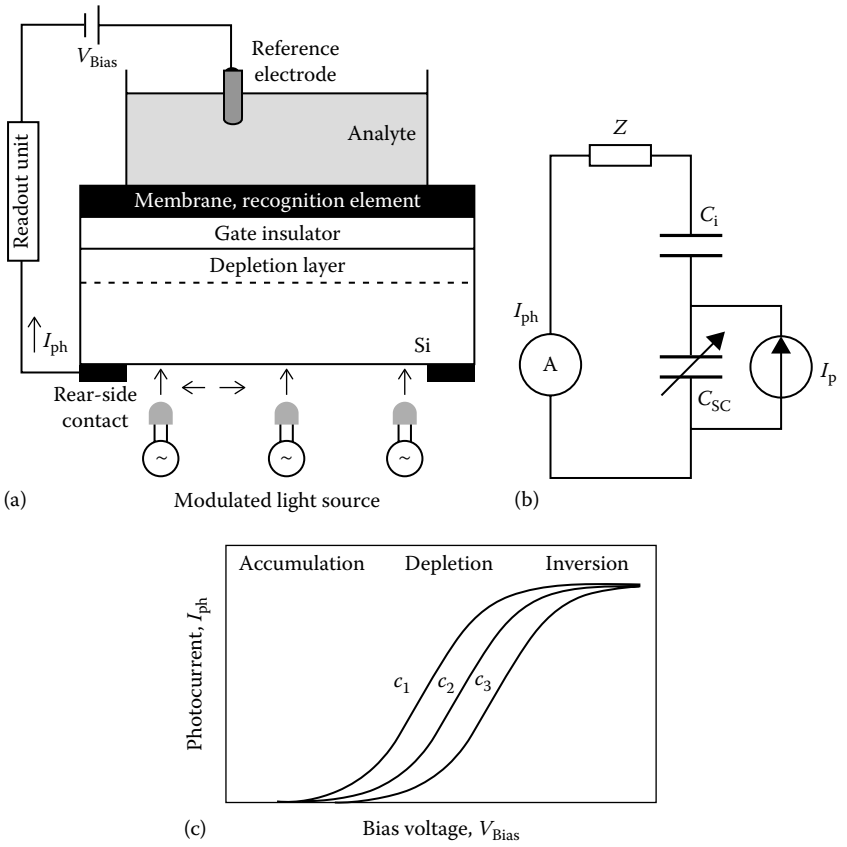


FIGURE 55.24 Schematic structure (a), simplified equivalent circuit (b), and typical photocurrent response by different analyte concentrations (c) of the LAPS. I_{ph}, measured photocurrent; I_p, generated photocurrent; Z, impedance of a reference electrode and a solution; C_i, gate capacitance; C_{SC}, space-charge capacitance; c_1, c_2, and c_3, analyte concentrations.

one or multiple light pointers (light-emitting diode [LED] or laser beam) just below the semiconductor, as shown in Figure 55.24a. When silicon is illuminated with infrared light, there is a generation of electron–hole pairs in the semiconductor. The electron–hole pairs that have diffused from the bulk semiconductor into the depletion region or that are photogenerated within this region are separated in the electric field. Modulating the intensity of an excitation light will modulate the depletion capacitance in the illuminated region of the semiconductor, thus, generating a photocurrent (I_{ph}) in the external circuit. Under the condition of strong depletion, the photocurrent I_{ph} of the LAPS can be described by the following simplified equation (Grattarola and Massobrio 1998):

$$I_{ph} = I_p \frac{C_i}{C_i + C_{SC}} \tag{55.55}$$

where I_p is the alternating component of the photogeneration of electron–hole pairs. Because the capacitance of the space-charge region, C_{SC}, depends on the applied dc voltage, the measured photocurrent is a function of the bias voltage applied to the LAPS structure.

Figure 55.24c depicts typical photocurrent–voltage characteristics of a *p*-type LAPS. By sweeping the applied voltage, the semiconductor surface potential can be driven from accumulation to inversion condition. The amplitude of photocurrent saturates as the bias voltage is increased toward inversion, and the photocurrent disappears when the device is biased into accumulation, where the depletion region and therefore the electric field no longer exist. The LAPS uses this dependence of the photocurrent on bias voltage to sense the interfacial potential changes at the gate/solution interface. For different ion or analyte concentrations, the photocurrent–voltage curve shifts along the voltage axis (Figure 55.24c). By measuring this shift, the analyte concentration can be quantitatively determined. The photocurrent is most sensitive to changes in the chemical composition of the solution in the depletion range (a sharp transition region between a maximum [in inversion mode] and a virtually zero value [in accumulation mode]). Therefore, the shift of the photocurrent–voltage curve is often quantified by tracking the potential of the inflection point of the curve. For chemical imaging applications (a typical chemical image consists of an enormous number of measured spots or pixels), however, the measurement of the complete current–voltage curve is too time-consuming, and a faster measuring method is required. One possible measuring mode is the constant-voltage mode, which is often used for the LAPS-based chemical imaging sensor (Wagner and Schöning 2007). In this mode, the applied bias voltage is simply fixed at a constant value (usually near the inflection point of the current–voltage curve, where the linearity of the curve is high), and the analyte concentration-dependent variation of the photocurrent is recorded. Another mode of operation is the constant-photocurrent mode, in which a feedback system controls the applied bias voltage so that the photocurrent is maintained at a constant value. This mode is most suitable for applications, where measurement of dynamic changes of the sensor signal is necessary.

Thus, not only the sensor structure but also the measurement principle of the LAPS is very similar to that of the EIS capacitance sensor, in which the current signal is induced by the application of an ac voltage. However, in contrast to the capacitive EIS sensor, where the measured value of the ion or analyte concentration is an average over the whole sensing surface in contact with the analyte, the LAPS measurement is spatially resolved. An attractive feature of the LAPS, compared to the EIS sensor and ISFET, is its light addressability. The measurement spot on the sensing surface is defined by the illumination area, where the ac photocurrent to be measured is generated. The main disadvantages of the LAPS are the light sensitivity and the dependence of the photocurrent on the electrolyte conductivity.

Due to the structural similarity between LAPS, capacitive EIS sensor, and ISFET, many of the sensitive materials, deposition or immobilization techniques, and surface modification strategies, which have already been developed for ISFETs and EIS sensors, are also applicable to LAPS devices. Traditionally, the LAPS is mostly employed for pH recording (see, e.g., Poghossian and Schöning 2006, Wagner and Schöning 2007, and references there). Si_3N_4 is the most frequently used pH-sensitive material in LAPS

devices, although several alternative materials such as Ta_2O_5 and Al_2O_3 have also been proven as alternative pH-sensitive materials. The first and successfully commercialized system using the pH-sensitive LAPS for the determination of extracellular acidification of living cells is the cytometer microphysiometer system, realized in 1990s by the company Molecular Devices Corporation (United States) (Hafner 2000). More recently, various portable miniaturized LAPS devices (e.g., LAPS-card sensor or a pen-shaped LAPS) with integrated signal processing unit have been realized (Wagner and Schöning 2007).

In contrast to ISFETs or capacitive EIS sensors, ion-sensitive and enzyme-modified LAPS are studied in less detail. For example, the application of LAPS as an ion sensor for the detection of different cations (Li^+, K^+, Cs^+, Ca^{2+}, and Mg^{2+}) and anions (NO_3^- and SO_4^-) has been demonstrated. A LAPS for heavy-metal detection using a chalcogenide glass membrane was developed. An enzyme-modified LAPS has been realized for the detection of penicillin, urea, glucose, and butyrylcholine. LAPS devices became popular in many chemical and biological applications such as the detection of bacterial growth, the measurement of cell metabolism, and the study of mechanisms of drug action on cell physiology. For example, with the aim to study the influence of drugs on the metabolism of cells, a microphysiometer, based on the multi-light LAPS concept, has been developed for the simultaneous measurement of several extracellular ion concentrations (H^+, Na^+, K^+, and Ca^{2+}) (Yicong et al. 2001). More details about different types of LAPS can be found in Poghossian and Schöning (2007) and Wagner and Schöning (2007).

In order to measure a pH- or ion-concentration distribution along the LAPS sensor surface with a spatial resolution, either the light pointer (e.g., focused laser beam) can be scanned along the surface or multiple light pointers can be used. In this way, a map of a 2D distribution of the pH value or the ion concentration can be visualized (chemical imaging). The detection of the metabolic activity of bacteria immobilized on a LAPS surface and the potentiometric imaging of a solution inside the fluidic microchannel are two examples of a possible application of LAPS as chemical imaging sensor. The spatial resolution of the LAPS is the most important parameter for both the multi-sensor application and the chemical imaging application. It limits both the smallest size of structures that can be visualized by the chemical imaging sensor and the density of measuring points on the sensor surface. Because of lateral diffusion of photogenerated carriers, the spatial resolution is limited by the diffusion length of the minority carriers or the thickness of the Si substrate. The best values of lateral resolution reported for bulk silicon have been about 20 μm. The high spatial resolution down to several micrometers or less can be achieved using SOI substrates with ultrathin Si (<0.5 μm) or poly-Si films as well as a LAPS based on amorphous silicon films.

For a detailed understanding of physicochemical phenomena at semiconductor device/cell interfaces, electrically excitable cells have to be connected individually, that is, each cell has to be interfaced with a separate potential-sensitive device (e.g., with a FET). However, because of the restriction of microelectronic fabrication, the number of FETs and therefore the number of active measuring sites are limited. Instead of placing single cells on the separate gates of the FET arrays, cells grown on a LAPS surface can be individually addressed by scanning the light pointer and illuminating the LAPS surface below the single cell of interest. In this way, it should in principle be possible to record the metabolic or electrical activity of a single cell, although there are many cells cultured on the chip surface. Therefore, several attempts have been made to record action potentials of single cells by means of a LAPS; however, the measured signals were small (about 10 μV). In addition, a hybrid device, a so-called scanning probe potentiometer based on cantilever-type micro-LAPS, capable for the measurement of the pH distribution in microvolume solution with the spatial resolution of 10 μm has been developed (Manalis et al. 2000). It is expected that such a scanning probe potentiometer could be used to image pH gradients produced by individual cells. The main technical problem preventing measurements of extracellular acidification rate of single cells with LAPS devices is the problem of confining the produced protons in a very small and defined volume.

55.4.4 Practical Applications and Limitations

Possible practical applications of field-effect chemical and biological sensors reach from medicine, biotechnology, and environmental monitoring over food and drug industries up to defense

and security purposes (see, e.g., Bergveld and Sibbald 1988, Wang et al. 2005, Poghossian and Schöning 2006, Lee et al. 2009a, and Jimenez-Jorquera et al. 2010). For example, ISFETs have been utilized for pH- and ion-concentration measurements in whole blood, plasma, and urine; in vivo pH monitoring in the stomach; acid-rain monitoring; marine monitoring; soil analysis; monitoring of nutrient solutions in greenhouses; online detection of microorganisms in water; determination of pH in meat and Ca^{2+} in milk; and online process control of pH, K^{+}, Ca^{2+} in wine industry. Moreover, EnFETs were applied for the determination of glucose in blood serum and urine, urea in blood serum and in hemodialysis fluids, and creatinine in hemodialysis solutions and in serum of renal failure patients as well as for transcutaneous blood glucose monitoring of diabetic patients. However, despite the intensive research and practical realization of different field-effect (bio-) chemical devices, generally, it could be concluded that their transfer from scientific laboratories to real life remains rather slow. Only very few of field-effect (bio-)chemical sensors, namely, pH-ISFET electrodes and cell-monitoring systems based on ISFET (Bionas 2500 analyzing system) and LAPS (cytometer microphysiometer, Molecular Devices Corporation), have successfully been commercialized so far. Commercially available pH ISFETs are exceptionally stable, fully temperature-compensated, rugged, reliable devices and exhibit performances that are comparable to those of pH-glass electrodes. Resistance to breakage is the most obvious feature of the pH ISFETs compared to the pH-glass electrode. Therefore, nowadays in many in-line process-monitoring systems in biotechnology, food, and drug industries, the breakable pH-glass electrode is gradually replaced by non-glass, unbreakable pH sensors based on ISFETs.

The study of the current state of BioFETs reveals that some BioFETs, like EnFETs or cell-based FETs, are at a well-developed stage, whereas other BioFETs (e.g., DNA-FETs) are still in the experimental stage or starting phase. Although, many improvements have been made in the last few years, there are, however, still a number of fundamental and technological problems that must be overcome before the first reliable BioFET-based bioanalytical microsystem will appear on to the market. The same can be stated for biologically sensitive capacitive EIS sensors and LAPS.

In general, ChemFETs possess significant advantages over classical ISEs, such as high input impedance, a fast response time, and a small weight. The small sensor area includes the possibility of multiple sensor applications (*sensor arrays*) on a single chip. Moreover, temperature compensation is possible. However, most of these sensors are exposed to a chemically very reactive environment, and therefore, a highly long-term stable protection (encapsulation) of the electronics from the analyte is required. The instability of the materials used induces a sensor drift. In some cases, attachment and fixation of the sensor membranes must be improved. To take the advantage of miniaturized FET devices, there is also the necessity of a small reference electrode. For ChemFETs, there exist two approaches for successful commercialization: dealing with small sample volumes for biomedical use and the high-volume fabrication for a low-price market (e.g., environmental and process monitoring, agriculture and food analysis, leak detectors). The employment of capacitive EIS and MIS sensors offers besides the easier manufacturing technique distinct advantages concerning the improved mechanical and electrochemical stability and sensor lifetime.

55.5 Conductometry

In addition to potentiometry, *conductometric analysis* represents the most important nonfaradaic method. *Conductometry* is based on the measurement of the electrical conductance of an electrolyte solution, which directly depends on the number of positively and negatively charged species in the solution. This analysis method is limited due to its nonselective nature, because all ions in the solution will contribute to the total conductance. Nevertheless, *direct conductance measurements* play an important role in the analysis of binary water/electrolyte mixtures, for example, in chemical water monitoring. The technique can also be applied to ascertain the endpoint detection in *conductometric titrations* for the determination of numerous substances.

55.5.1 Measurement of Conductance and Instrumentation

The *conductance G* of a solution is the reciprocal of the electrical resistance *R* and has the units of siemens (S) that correspond to ohm^{-1} (Ω^{-1}). The conductance of a uniform sample with the length *l* and cross-sectional area *A* is given by

$$G = \kappa \frac{A}{l} \tag{55.56}$$

where the proportionality constant $\kappa = 1/\rho$ (ρ: resistivity) describes the *conductivity* (*specific conductance*) of the solution, expressed in units of S cm^{-1}. The *equivalent conductivity* Λ (*molar conductivity*) of a solution is defined as the conductivity due to one mole, measured between two electrodes that are spaced 1 cm apart and is

$$\Lambda = \frac{1000\ \kappa}{c} \tag{55.57}$$

where *c* corresponds to the concentration of the solution in mol L^{-1}. The units of Λ are S cm^{-1} mol^{-1}. Equation 55.56 permits the calculation of the molar conductivity for a solution of known concentration by considering the experimental values of κ. The molar conductivity Λ, that is, the mobility of ions in solution, is mainly influenced by interionic effects for strong electrolytes and the degree of dissociation for weak solutions. For strong electrolytes, the molar conductivity increases as the dilution is increased. By linear graphical extrapolation for diluted solutions of strong electrolytes, a limiting value is defined as *molar conductivity at infinite dilution* Λ_0. At infinite dilution, the interionic attraction is nil, the ions are independent of each other, and the total conductivity is

$$\Lambda_0 = \lambda_+^0 + \lambda_-^0 \tag{55.58}$$

where λ_+^0 and λ_-^0 are the ionic molar conductivities of the cations and anions, respectively, at infinite dilution. For weak electrolytes, due to the nonlinear relationship between Λ and *c*, a graphical extrapolation cannot be made. Typical values for the limiting molar conductivities for various species in water are listed in Table 55.2.

TABLE 55.2 Molar Conductivity at Infinite Dilution Λ_0 (Ω^{-1} cm^2 mol^{-1})

Cations λ_+^0		Anions λ_-^0	
H^+	349.8	OH^-	198.3
Na^+	50.1	F^-	55.4
K^+	73.5	Cl^-	76.3
Li^+	38.7	Br^-	78.1
NH_4^+	73.4	I^-	76.8
Ag^+	61.9	NO_3^-	71.5
$N(CH_3)_4^+$	44.9	ClO_4^-	67.3
Ca^{2+}	119.0	$C_2H_3O_2^-$	40.9
Mg^{2+}	106.2	HCO_3^-	44.5
Cu^{2+}	107.2	AcO^-	40.9
Zn^{2+}	105.6	SO_4^{2-}	160.0
Ba^{2+}	127.7	CO_3^{2-}	138.6
Pb^{2+}	139.0	$C_2O_4^{2-}$	148.4
Fe^{3+}	204.0	PO_4^{3-}	240.0
La^{3+}	208.8	$Fe(CN)_6^{4-}$	442.0

The equipment needed for measuring the conductivity includes an electric power source, a cell containing the solution, and a suitable measuring bridge. The electric power source consists of an alternating current source that produces signals of about 1 kHz in order to eliminate effects of faradaic current. The measurement is performed by a Wheatstone bridge arrangement. Modern *conductivity meters* supply the alternating current and allow the measurement in a wide range of conductivities (0.001 μS cm^{-1} to 1300 mS cm^{-1}). Additional electronics eliminate disturbing capacitance effects and offer automatic range switching. An integrated temperature sensor corrects automatically conductivities to their value at 25°C. The conductivity cell consists of a pair of electrodes placed in a defined geometry to each other. Usually, the electrodes are platinized to increase their effective surface (high capacitance). Thus, disturbing faradaic currents are minimized. For accurate conductivity determination, the precise area of the electrodes A and their distance apart d, the *cell constant* K, must known exactly. Therefore, the cell constant ($K = A/d$) must be evaluated by calibration with a solution of accurately known conductivity (e.g., a standard KCl solution). Details of calibration standards and concepts of conductivity cells are given in Oehme (1991).

55.5.2 Applications of Conductometry

55.5.2.1 Direct Conductometric Measurement

In spite of the insufficient selectivity of *direct conductometric measurements*, the high sensitivity of this procedure makes it an important analytical tool for certain applications. The specific conductivity of pure water (distilled or deionized) is about 5×10^{-8} S cm^{-1}, and the smallest trace of ionic impurity leads to a large increase in conductivity by an order of magnitude and more. Therefore, conductometric monitoring is employed where a high purity of water is required (e.g., laboratories, semiconductor processing, steam-generating power plants, ion exchanger). Conductometric measurements are widely used to control pollution of rivers and lakes and in oceanography to control the salinity of seawater.

55.5.2.2 Conductometric Titration

In *conductometric titration*, the reaction is followed by means of conductometry and is used for locating endpoints (i.e., the EP in acid–base titrations [neutralization titration]). To define the titration curve, at least three or four measurements before and after the EP are required. The obtained data of the conductivity are plotted as a function of the titrant volume, and the EP is given as the intersection of the two linear extrapolated fractions. A characteristic titration curve of a strong acid (hydrochloric acid) with a strong base (sodium hydroxide) is depicted in Figure 55.25. The solid line represents the resulting

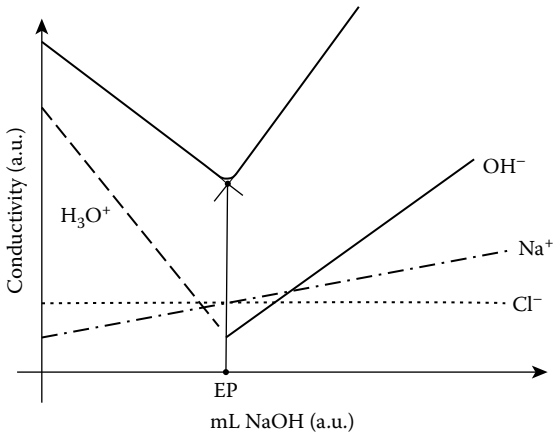


FIGURE 55.25 Conductometric titration of a strong acid (HCl) with a strong base (NaOH). The EP is represented by EP.

titration curve, whereas the broken lines indicate the contribution of the individual species. By adding NaOH to the solution, the hydrogen ions are replaced by the equivalent number of less mobile sodium ions (and $H^+ + OH^- \rightarrow H_2O$). As a result, the conductivity decreases to lower values. The solution exhibits its lowest conductivity at the EP, where the concentrations of hydrogen and hydroxide ions are at the minimum. Further addition of NaOH reverses the slope of the titration curve, since both the sodium ion concentration and hydroxide ion concentration increase.

Due to the high linearity between the conductance and the volume of the added species, this method possesses a high accuracy and can be employed in dilute as well as in more concentrated solutions. In contrast to potentiometric titration methods, the immediate EP region has no strong significance. Thus, very weak acids, such as basic acid and phenol, can be titrated. Moreover, mixtures of hydrochloric acid or another strong acid and acetic (ethanoic) acid or any other weak acid can be titrated with a weak base (e.g., aqueous ammonia, acetate) or with a strong base (e.g., sodium hydroxide). Moreover, precipitation and complex-formation titrations of, for example, sodium chloride with silver nitrate are possible. For practical applications, the volume of the solution should not change appreciably during the titration. Therefore, the titrating reagent may be 20–100 times more concentrated than the solution being titrated, whereas the latter should be as diluted as practicable. For additional examples of analytical procedures and results of conductometric titrations, see Skoog et al. (2007).

55.5.2.3 Oscillometry

In order to investigate electrolyte solutions with high resistivities and dielectric constants, *high-frequency titration* (*oscillometry*) can be performed at 10^5–10^7 Hz. For that, a specific measuring cell is required, where the metal electrodes encircle the outside of a glass container. In this arrangement, the electrodes are not in contact with the test solution, which is advantageous for dealing with corrosive materials. Oscillometric measurements can be employed for the determination of binary mixtures of nonionic species, where the dielectric behavior predominates (e.g., ethanol/nitrobenzene, benzene/chlorobenzene, and alcohol/water). Further practical examples are EDTA titrations and the determination of thorium Th^{4+}) with sodium carbonate, beryllium (Be^{2+}) with sodium hydroxide, and hydrocarbons (e.g., benzene). However, the instrumentation and the interrelations are more complicated than for the classical conductivity method. Thus, oscillometry gets only significance for specific applications, where the presence of the electrodes interferes.

55.5.2.4 Conductometric Sensors

Depending on the demanded size and geometry, miniaturized cells with two or more electrodes (e.g., a four-electrode conductivity meter) as well as contactless cells are commercially available as *conductometric sensors*. The contactless methods use *capacitive* and *inductive conductivity cells*, which are advantageous to circumvent electrochemically caused electrode reactions. Conductivity cells can be coupled as detectors to ion chromatographic systems for measuring ionic concentration in the eluate. For this, special *micro-conductivity cells* with a volume of about 1.5 µL have been developed.

Within the last 10 years, two aspects of conductometric applications became important: *conductometric gas sensors* and the use of *conductometric chemiresistors* as sensors. In the former, a phase change that transfers the gaseous component into a solution is necessary (e.g., by a bubbler nebulizer). All methods deal with acidic gases such as HCl or SO_2 or with alkaline gases, like NH_3. Also, organic halogens can be detected after their conversion into HCl or HF. By means of integrated circuit technology, thin metal films can be photolithographically patterned as interdigital electrodes onto semiconductor substrates with insulating dielectric layers of SiO_2 or Si_3N_4. Both the thin metal films and additionally deposited organic layers on top of the metallic films can lead to a change of the total resistance by variation of the ionic composition of the reacting solution. For chemiresistors, the organic layer usually consists of an ion-selective polymer layer or a Langmuir–Blodgett membrane; for biosensors enzymatic layers are used (see the ISE section). Such sensors allow the determination of different gaseous components, such as CO, NO_2, H_2S, SO_2, or NH_3, as well as the detection of

biologically relevant species like urea, glucose, penicillin, and choline chlorides. Although several companies offer such gas analyzer systems, conductometric sensors and chemiresistors are still in the state of research and development.

55.6 Coulometry

Coulometry represents an electroanalytical method, where the analyte is specifically and completely converted due to direct or indirect electrolysis. The quantity of electricity (in coulombs) consumed by this reaction, the charge, is measured. A fundamental requirement of coulometry is that the species in the solution interact with 100% current efficiency; that is, the reaction corresponds to the Faraday law. According to this condition, there exist two alternatives: the analyte participates in the electrode reaction (*primary* or *direct coulometric analysis*) and the analyte reacts with a reagent, generated by an electrode reaction (*secondary* or *indirect coulometric analysis*). Two general techniques—*controlled-potential coulometry* and *coulometric titration* (*controlled-current coulometry*)—are used for coulometric analysis.

55.6.1 Controlled-Potential Coulometry

In this method, the potential of the working electrode is held at a constant value compared to a reference electrode. The resulting current is adjusted continuously to maintain the desired potential. The substance being determined reacts without involvement of other components in the sample. The reaction is completed when the current has practically decreased to zero. To measure the charge, a potentiostat, an instrument for measuring the time-dependent current, and a current–time integrating device are used. Modern potentiostats have a built-in electronic coulometer and allow extremely accurate determinations. Otherwise, one can use free-standing coulometers.

Controlled-potential coulometry has been widely employed for the determination of various metal ions, such as Cu, Bi, Cd, Zn, Ni, Co, Pu, and U. To apply this method, current–voltage diagrams must be available for the oxidation–reduction system to be measured as well as for any reaction system at the working electrode. Current–voltage diagrams can be obtained by plotting the measured current versus the cathode-reference electrode potential. To fulfill the requirement of the 100% current efficiency in generation, it is necessary to control the potential of the working electrode. With regard to their determination, the metals are deposited at controlled potentials with a mercury cathode as working electrode and a silver wire or a platinum cylinder as anode. Typical applications are the electrolytic determination and synthesis of organic compounds like acetic acid and picric acid. Further, controlled-potential coulometry is frequently used for monitoring the concentration of constituents in gas or liquid streams, typically small oxygen contents. Here, the reduction of oxygen takes place within the pores of a porous silver cathode:

$$O_2(g) + 2H_2O + 4e^- \Leftrightarrow 4OH^- \tag{55.59}$$

Using a cadmium sheet (m) as anode, the electrode reaction in solution (s) is

$$Cd(m) + 2OH^- \Leftrightarrow Cd(OH)_2(s) + 2e^- \tag{55.60}$$

The quantity of the electricity (current) is passed through a standard resistor and converted to a voltage signal. Hence, the oxygen concentration is proportional to the recorded potential drop. Controlled-potential coulometry needs relatively long electrolysis times, although it proceeds virtually unattended with automatic coulometers. With a multimeter, changes in the range from 1 ppm to 1% can be dissoluted. Thus, controlled-potential coulometry permits analysis with an accuracy of a few tenths of a percent.

55.6.2 Coulometric Titration (Controlled-Current Coulometry)

Controlled-current coulometry maintains a constant current throughout the reaction period. Here, an excess of a redox buffer substance must be added in such a way that the potential does not cause any undesirable reaction. That means the product of the electrolysis of the redox buffer must react quantitatively with the unknown substance to be determined. *Coulometric titrations* need an electrolytically generated titrant that reacts stoichiometrically with the analyte to be determined. As in controlled-potential coulometry, 100% current efficiency is required. The current is accurately fixed at a constant value and the quantity of electricity can be calculated by the product of the current (in A) and the time (in s) using endpoint detection. In principle, any endpoint detection system that fits chemically can be used, for example, chemical indicators (color change) and potentiometric, amperometric, or conductometric procedures. For coulometric titrations the instrumentation consists of a titrator (constant-current source, integrator) and a cell. As the constant-current source, an electronically controlled amperostat is preferably used. The integrator measures the product of current and time (i.e., the number of coulombs). The electrolysis cell, filled with the solution from which the titrant will be generated electrolytically and the solution to be titrated, is schematically shown in Figure 55.26. The generator electrode, at which the reagent is formed, possesses a large surface area (e.g., a rectangular strip of platinum). The auxiliary electrode (e.g., a platinum wire) is in contact with an appropriate electrolyte of higher concentration than the solution to be titrated. It is isolated from the analyte by a sintered disk or some other porous media. This is required to avoid the interference of additional products generated at the second electrode. To circumvent these limitations of internal generation, an external generator cell is often used.

Typical applications of coulometric titrations are neutralization titrations, precipitation and complex-formation titrations, and oxidation–reduction titrations. Neutralization titrations can be employed for both weak and strong acids and bases. The former can be performed with hydroxide ions generated at a platinum anode by the following reaction:

$$2H_2O + 2e^- \Leftrightarrow 2OH^- + H_2(g) \tag{55.61}$$

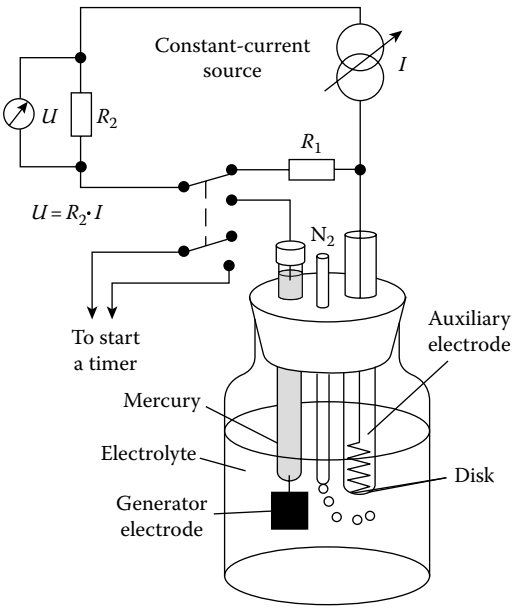


FIGURE 55.26 Coulometric titration cell with working electrode and auxiliary electrode and equivalent circuit diagram (schematically).

the latter one with hydrogen ions by the following reaction:

$$H_2O \Leftrightarrow \tfrac{1}{2}O_2(g) + 2H^+ + 2e^- \tag{55.62}$$

A working (*generator*) electrode of silver as anode offers the determination of Cl^-, Br^-, I^-, and mercaptans in solution(s). For bromide, the reaction becomes

$$Ag + Br^-(s) \Leftrightarrow AgBr(s) + e^- \tag{55.63}$$

Similar precipitation and complex-formation titrations as well as oxidation–reduction titrations are described in Skoog et al. (2007).

Coulometric titrations possess some practical advantages: no standard solutions are required and unstable reagents can be generated or consumed immediately, small amounts of titrants can be electrically quantified with high accuracy, pretitration is possible, and the method can be readily adapted to automatic remote control. Thus, with respect to controlled-potential coulometry, a wider field of practical applications exists. Often, automatic titrators for multipurpose and single analysis employ potentiometric endpoint detection. Examples are sulfur dioxide monitors and water titrators (*Karl Fischer*). For more detailed information concerning applications of coulometry, see Dahmen (1986).

References

Abouzar, M.H., Poghossian, A., Cherstvy, A.G., Pedraza, A.M., Ingebrandt, S., and Schöning, M.J. 2012. Label-free electrical detection of DNA by means of field-effect nanoplate capacitors: Experiments and modeling. *Physica Status Solidi A* 209: 925–934.

Abouzar, M.H., Poghossian, A., Pedraza, A.M. et al. 2011. An array of field-effect nanoplate capacitors for (bio-)chemical sensing. *Biosensors Bioelectronics* 26: 3023–3028.

Armgarth, M. and Nylander, C.I. 1991. Field-effect gas sensors. In *Sensors: A Comprehensive Survey*, Vol. 2, eds. W. Göpel, J. Hesse, and J. N. Zemel, pp. 509–512. Weinheim, Germany: Wiley-VCH.

Bailey, P.L. 1980. *Analysis with Ion-Selective Electrodes*. London, U.K.: Heyden & Son.

Bard, A.J. and Faulkner, L.R. 2001. *Electrochemical Methods: Fundamentals and Applications*. New York: Wiley-VCH.

Bergveld, P. 2003. Thirty years of ISFETOLOGY: What happened in the past 30 years and what may happen in the next 30 years. *Sensors and Actuators B* 88: 1–20.

Bergveld, P. and Sibbald, A. 1988. Analytical and biomedical applications of ion-selective field-effect transistors. Amsterdam, the Netherlands: Elsevier.

Brainina, Kh. Z. and Neyman, E. 1993. *Electroanalytical Stripping Methods*. New York: Wiley-VCH.

Covington, A.K. 1979. Glass electrodes. In *Ion-Selective Electrode Methodology*, ed. A.K. Covington, pp. 77–84. Boca Raton, FL: CRC Press.

Dahmen, E.A.M.F. 1986. *Electroanalysis*. Amsterdam, the Netherlands: Elsevier.

Gileadi, E. 1993. *Electrode Kinetics for Chemists, Chemical Engineers, and Materials Scientists*. New York: Wiley-VCH.

Grattarola, M. and Massobrio, G. 1998. *Bioelectronics Handbook: MOSFETs, Biosensors and Neurons*. New York: McGraw-Hill.

Gun, J., Gutkin, V., Lev, O. et al. 2011. Tracing gold nanoparticle charge by electrolyte-insulator-semiconductor devices. *The Journal of Physical Chemistry C* 115: 4439–4445.

Hafner, F. 2000. Cytosensor microphysiometer: Technology and recent applications. *Biosensors Bioelectronics* 15: 149–158.

Hamann, C.H., Hamnett, A., and Vielstich, W. 2007. *Electrochemistry*. Weinheim, Germany: Wiley-VCH.

Haynes, W.M. 2011. *CRC Handbook of Chemistry and Physics*. Boca Raton, FL: CRC Press.

Henze, G. 2008. Analytical voltammetry and polarography. In *Ullmann's Encyclopedia of Industrial Chemistry*, pp. 409–450. Weinheim, Germany: Wiley-VCH.

Jimenez-Jorquera, C., Orozco J., and Baldi, A. 2010. ISFET based microsensors for environmental monitoring. *Sensors* 10: 61–83.

Josowicz, M. and Janata, J. 1988. Suspended gate field-effect transistor. In *Chemical Sensor Technology*, Vol. 1, ed. T. Seiyama, pp. 167–175. Amsterdam, the Netherlands: Elsevier.

Koryta, J. 1991. *Ions, Electrodes and Membranes*. Chichester, U.K.: Wiley-VCH.

Koryta, J., Dvorák, J., and Kavan, L. 1993. *Principles of Electrochemistry*. Chichester, U.K.: Wiley-VCH.

Koryta, J. and Stulik, K. 2009. *Ion-Selective Electrodes*. Cambridge, U.K.: Cambridge University Press.

Lee, C.-S., Kim, S.K., and Kim, M. 2009a. Ion-sensitive field-effect transistor for biological sensing. *Sensors* 9: 7111–7131.

Lee, M., Baik, K.Y., Noch, M., Kwon, Y.-K., Lee, J.-O., and Hong, S. 2009b. Nanowire and nanotube transistors for lab-on-a-chip applications, *Lab on a Chip* 9: 2267–2280.

Madou, M.J. and Morrison, S.R. 1989. *Chemical Sensing with Solid State Devices*. San Diego, CA: Academic Press.

Manalis, S.R., Cooper, E.B., Indermuhle, P.F. et al. 2000. Microvolume field-effect pH sensor for the scanning probe microscope. *Applied Physics Letters* 76: 1072–1074.

Mastrangelo, C.H. and Tang, W.C. 1994. Semiconductor sensor technologies. In *Semiconductor Sensors*, ed. S. M. Sze, pp. 17–95. New York: Wiley-VCH.

Mendham, J., Denney, R.C., Barnes, J.D., and Thomas, M.J.K. 2000. *Vogel's Textbook of Quantitative Chemical Analysis*. Harlow, U.K. Prentice Hall.

Morf, W.E. 1981. *The Principles of Ion-Selective Electrodes and Membrane Transport*. Amsterdam, the Netherlands: Elsevier.

Neff, P.A., Wunderlich, B.K., Lud, S.Q., and Bausch, A.R. 2006. Silicon-on-insulator based thin film resistors for quantitative biosensing. *Physica Status Solidi A* 203: 3417–3423.

Oehme, F. 1991. Liquid electrolyte sensors: Potentiometry, amperometry and conductometry. In *Sensors: A Comprehensive Survey*, Vol. 2, eds. W. Göpel, J. Hesse, and J.N. Zemel, 302–312. Weinheim, Germany: Wiley.

Poghossian, A., Ingebrandt, S., Abouzar, M.H., and Schöning, M.J. 2007. Label-free detection of charged macromolecules by using a field-effect-based sensor platform: Experiments and possible mechanisms of signal generation. *Journal of Applied Physics A* 87: 517–524.

Poghossian, A., Ingebrandt, S., Offenhäusser, A., and Schöning, M.J. 2009. Field-effect devices for detecting cellular signals. *Seminars Cell Developmental Biology* 20: 41–48.

Poghossian, A., Malzahn, K., Abouzar, M.H., Mehndiratta, P., Katz, E., and Schöning, M.J. 2011. Integration of biomolecular logic gates with field-effect transducers. *Electrochimica Acta* 56: 9661–9665.

Poghossian, A. and Schöning, M.J. 2006. Silicon-based chemical and biological field-effect sensors. In *Encyclopedia of Sensors* Vol. 9, eds. C.A. Grimes, E.C. Dickey, and M.V. Pishko, pp. 463–533. Stevenson Ranch, CA: American Scientific Publishers.

Poghossian, A. and Schöning M.J. 2007. Chemical and biological field-effect sensors for liquids—A status report. In *Handbook of Biosensors and Biochips*, eds. R.S. Marks, D.C. Cullen, I. Karube, C.R. Lowe, and H.H. Weetall, Chapter 24, pp. 1–17. Weinheim, Germany: Wiley-VCH.

Pungor, E. 1998. The theory of ion-selective electrodes. *Analytical Sciences* 14: 249–256.

Rheinhoudt, D.N. 1992. Application of supramolecular chemistry in the development of ion-selective ChemFETs. *Sensors Actuators B* 6: 179–185.

Rieger, P.H. 1987. *Electrochemistry*. Englewood Cliffs, NJ: Prentice-Hall.

Ryynänen, T., Kujala, V., Ylä-Outinen, L. et al. 2011. All titanium microelectrode array for field potential measurements from neurons and cardiomyocytes—A feasibility study. *Micromachines* 3: 394–409.

Schöning, M.J., Brinkman, D., Rolka, D., Demuth, C., and Poghossian, A. 2005. CIP (cleaning-in-place) suitable "nonglass" pH sensor based on a Ta_2O_5-gate EIS structure. *Sensors Actuators B* 111–112: 423–429.

Schöning, M.J. and Kloock, J.P. 2007. About 20 years of silicon-based thin-film sensors with chalcogenide glass materials for heavy metal analysis: Technological aspects of fabrication and miniaturization. *Electroanalysis* 19: 2029–2038.

Schöning, M.J. and Poghossian, A. 2006. BioFEDs (field-effect devices): State-of-the-art and new directions. *Electroanalysis* 18: 1893–1900.

Schöning, M.J. and Poghossian, A. 2008. Detection of charged macromolecules by means of field effect devices (FEDs): Possibilities and limitations. In *Electrochemical Sensors, Biosensors and Their Biomedical Applications,* eds. X. Zhang, H. Ju, and J. Wang, pp. 187–212. San Diego, CA: Academic Press.

Schöning, M.J., Sauke, M., Steffen, A. et al. 1995. Ion-sensitive field-effect transistors with ultrathin Langmuir-Blodgett membranes. *Sensors Actuators B* 26–27: 325–328.

Schöning, M.J., Tsarouchas, D., Schaub, A. et al. 1996. A highly long-term stable silicon-based pH sensor fabricated by pulsed laser deposition technique. *Sensors Actuators B* 35: 228–233.

Settle, F.A., Lamp, B.D., McCurdy, D.L. Vitha M.F., Gregory B.W., and Ma, Y. 2011. *Instrumental Methods of Analysis*. Oxford, U.K. Wiley-Interscience.

Siqueira, J.R.Jr., Abouzar, M.H., Poghossian, A., Zucolotto, V., Oliveira, O.N.Jr., and Schöning M.J. 2009. Penicillin biosensor based on a capacitive field-effect structure functionalized with a dendrimer/carbon nanotube multilayer. *Biosensors Bioelectronics* 25: 497–501.

Skoog, D.A., Holler, F.J., and Crouch, S.R. 2007. *Principles of Instrumental Analysis*. Belmont, CA: Thomson Brooks/Cole Publishing.

Sze, S.M. and Ng, K.K. 2007. *Physics of Semiconductor Devices*. Hoboken, NJ: Wiley-Interscience.

Thedinga, E., Kob, A., Holst, H. et al. 2007. Online monitoring of cell metabolism for studying pharmacodynamic effects. *Toxicology and Applied Pharmacology* 220: 33–44.

Timko, B.P., Cohen-Karni, T., Qing, Q., Tian, B., and Lieber, C.M. 2010. Design and implementation of functional nanoelectronic interfaces with biomolecules, cells, and tissue using nanowire device arrays. *IEEE Transactions on Nanotechnology* 9: 269–280.

Vlasov, Y.G., Bychkov, E.A., and Bratov, A.V. 1994. Ion-selective field-effect transistor and chalcogenide glass ion-selective electrode systems for biological investigations and industrial applications. *Analyst* 119: 449–454.

Wagner, T. and Schöning M.J. 2007. Light-addressable potentiometric sensors (LAPS): Recent trends and applications. In *Electrochemical Sensor Analysis*, eds. S. Alegret and A. Markoci, pp. 87–128. Amsterdam, the Netherlands: Elsevier.

Wang, J. 1998. *Stripping Analysis*. Weinheim, Germany: Wiley-VCH.

Wang, J. 2006. *Analytical Electrochemistry*. Hoboken, NJ: Wiley-VCH.

Wang, P., Xu, G., Qin, L. et al. 2005. Cell-based biosensors and its application in biomedicine. *Sensors Actuators B* 108: 576–584.

Willard, H.H., Dean, J.A., Settle, F.A., and Merritt, L.L. 1988. *Instrumental Methods of Analysis*. Belmont, CA: Wadsworth.

Yicong, W., Ping, W., Xuesong, Y. et al. 2001. A novel microphysiometer based on MLAPS for drugs screening. *Biosensors Bioelectronics* 16: 277–286.

Yu, H., Liu, Q.J., Wang, P. et al. 2010. Light-addressable potentiometric sensor (LAPS) as cell-based biosensors. In *Cell-Based Biosensors: Principles and Applications*, eds. P. Wang and Q. Liu, pp. 119–149. Norwood, MA: Artech House.

Further Information

Bakker E. and Pretsch E. 2007. Modern potentiometry. *Angewandte Chemie International Edition* 46: 5660–5668.

Blackburn, G.F. 1987. Chemically sensitive field-effect transistors. In *Biosensors: Fundamentals and Applications*, eds. A.P.F. Turner, I. Karube, and G.S. Wilson, pp. 481–530. Oxford, U.K.: Oxford University Press.

Bockris J. and Khan, S. 1993. *Surface Electrochemistry: A Molecular Level Approach*. New York: Plenum Press.

Brett, C.M.A. and Brett, A M.O. 1998. *Electrochemistry Principles, Methods and Applications*. Oxford, U.K.: Oxford University Press.

Durst, R.A. 2012. Ion-selective electrodes—The early year. *Electroanalysis* 24: 15–22.

Gellings P.J. and Bouwmeester, H.J.M. 1997. *The CRC Handbook of Solid State Electrochemistry*. Boca Raton, FL: CRC Press.

Kissinger P.T. and Heinemann, W.R. 1996. *Laboratory Techniques in Electroanalytical Chemistry*. New York: Marcel Dekker.

Nakazato K. 2009. An integrated ISFET sensor array. *Sensors* 9: 8831–8851.

Schöning, M.J. and Poghossian, A. 2002. Recent advances in biologically sensitive field-effect transistors (BioFETs). *Analyst* 127: 1137–1151.

Tani, Y. and Umezawa, Y. 2005. Ion-selective electrodes. *Sensor Letters* 3: 99–107.

56

Thermal Composition Measurement

Mushtaq Ali
The National Grid Company

Behrooz Pahlavanpour
Nynas AB

Maria Eklund
Nynas AB

56.1 Introduction

Thermal analysis is the measurement of a physical parameter as a function of temperature. The area comprises several techniques where thermogravimetry (TG), thermometric titrimetry, thermomechanical analysis (TMA), differential thermal analysis, differential scanning calorimetry, and some specialized techniques are among those described in this section. Applications are found in the characterization of organic materials, both solids and liquids. Materials of interest can be polymers, mineral and synthetic oils, lubricants, greases, paper (cellulose), and pharmaceuticals. Material to be analyzed (typically 10 mg) must be isolated and subjected to thermal treatment; hence, the technique is destructive. The obvious advantage is that the thermal profile or structure of a large specimen can be investigated. The disadvantage is that the small sample size can give rise to excessive statistical errors. However, recent advances in microcalorimetry techniques to look at slow degradation of pharmaceuticals allow thermal analysis to be performed at room temperature on samples up to a few grams without destruction of the sample. A general schematic of thermal analysis apparatus is shown in Figure 56.1.

The history of the development of thermal analysis methods from the sixteenth century is the subject of a number of excellent papers by Mackenzie [1–3], Wendlandt [4], and Keattch [5]. Lavoisier and Laplace [6] were pioneers in the development of thermal analysis by their practical approach.

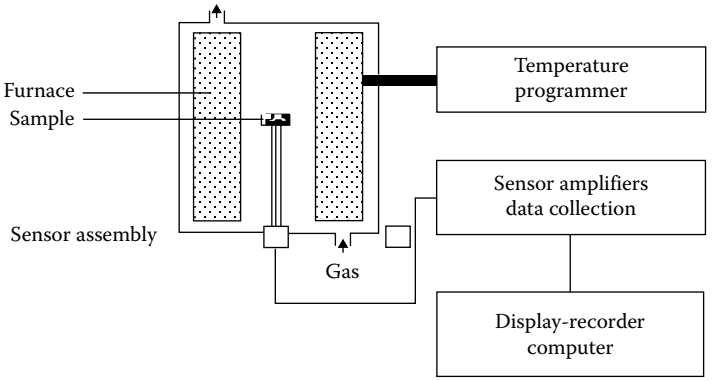


FIGURE 56.1 Schematic thermal analysis apparatus. The figure shows the essential components of a generalized thermal analysis apparatus.

TABLE 56.1 Factors Affecting Thermal Analysis

Factor	Details
Sample	History of sample and preparative technique used can affect the curve, and the presence of trace impurities (in some cases) may catalyze decompositions. Particle size can alter shape of curve (e.g., by surface reaction).
Crucible	The crucible (or sample holder) material should be such that it does not react with the sample or catalyze a reaction. The geometry of the sample holder may affect the results [12].
Rate of heating	Thermal lag: rate of heat transfer between furnace and all parts of the sample is not instantaneous. Therefore, care should be taken when working at different heating. Corrections can be applied [12].
Atmosphere	Various effects [12], including dissociation of sample.
Mass of sample	Size and packing density.

Note: Details the five main areas that would affect analysis of a sample via thermal experiments. The effects can be marked and would certainly affect repeatability also.

The International Confederation for Thermal Analysis and Calorimetry (ICTAC) has produced definitive guidelines regarding nomenclature and calibration [7–10].

56.1.1 Factors Affecting Results

The five factors affecting thermal analysis can be remembered by the acronym SCRAM [11]. This refers to the Sample, Crucible, Rate of heating, Atmosphere, and Mass. See Table 56.1.

56.2 Thermogravimetry

TG or thermogravimetric analysis (TGA) [8] is a technique in which the mass of the sample is monitored against time or temperature while the temperature of the sample, in a specified atmosphere, is programmed.

Derivative thermogravimetry (DTG) shows the change in mass per unit time as a function of temperature.

56.2.1 Apparatus

The apparatus is referred to as a thermobalance or thermogravimetric analyzer. There are a number of configurations: horizontal, simultaneous (TGA–differential thermal analysis [DTA]), and vertical. The vertical design provides better sensitivity and weight capacity. The thermobalance consists of five essential components: furnace, temperature regulator, weighing mechanism, atmosphere controller, and recording system.

56.2.2 Calibration

Small furnaces can be calibrated by a method [8] using Curie points of a range of metals and alloys. The Curie point is the temperature at which a ferromagnetic material loses its ferromagnetism. At the Curie point, the magnetic force is reduced to zero and an apparent mass change is observed.

The study of the reactions can be divided into the following stages: (1) intermediates and (2) products of reaction, (3) energetics of reaction, and (4) the reaction kinetics. Stages (1) and (2) can be readily studied by TG and differential scanning calorimetry (DSC).

For example, the decomposition of calcium oxalate monohydrate shows three distinct steps. The first around 200°C with a loss of 12.4% corresponds to dehydration, while those at 500°C and 800°C match with a loss of CO and CO_2. These are confirmed by analysis of residues.

56.2.2.1 Kinetics of Reaction Including Measurement of α and dα/dt

The use of TG as a means for the elucidation of the reaction kinetics is attractive. The nature of solid–solid interactions is quite complex [13] and will not be discussed in this section.
Consider an endothermic solid-state reaction:

$$\text{A(solid)} \rightarrow \text{B(solid)} + \text{C(gas)} \tag{56.1}$$

During the course of the reaction, there is a mass loss, combined with the loss of gas. Heat absorption also occurs. This process can be modeled. However, it should be noted that the equation (although generally applicable) is not valid for all cases. Methods and mathematical treatment of results are given in the papers by Šatava, Šesták, and Škvára [14–17].

Static (isothermal) and dynamic methods can be used in a kinetic study of the weight change. The former is based on the determination of the degree of transformation at constant temperature as a function of time. The latter is the determination of the degree of transformation as a function of time during a linear increase of temperature. The static method is probably better suited for obtaining information about the slowest process, the reaction order, and reaction mechanism. The dynamic method is better if data on the kinetics of the reaction from a single curve for the whole temperature range are required. Comparisons between both methods have shown comparable results with respect to precision [18].

The extent of a reaction ξ may be defined [19] by Equation 56.2:

$$n_B = n_{B,0} + \nu_B \xi \tag{56.2}$$

where

n_B is the amount of substance B
$n_{B,0}$ is the amount of substance B at $t = 0$
ν_B is the stoichiometric number of B (positive number if B is a product and negative if B is a reactant)

For solid-state reactions, the changes in the portion reacted α are followed with respect to time. Therefore, the rate of reaction can be defined by Equation 56.3:

$$\text{Rate} = \frac{d\alpha}{dt} \tag{56.3}$$

For solution reactions (referring to Equation 56.1), the change in concentration C_B of B is followed.

Since the rate of reaction varies with time (even at constant temperature), at a value of α, Equation 56.2 is derived:

$$\text{Rate} = \frac{d\alpha}{dt} = k_T, f(\alpha) \tag{56.4}$$

where

k_T is the rate constant at temperature T

$f(\alpha)$ is a mathematical expression in α

It should be noted that the form of $f(\alpha)$ sometimes alters partway through a reaction.

If hyphenated group-specific techniques are employed to study a reaction simultaneously (e.g., as is the case in TGA and FTIR), the IR-active species may not contribute the greatest mass loss, and therefore, the values of α will not be the same.

There are many equations relating the rate of solid-state reactions to α and they have been summarized by Sestak and Berggren [20].

A general *integrated* kinetic equation is given in Equation 56.5:

$$g(\alpha) = k_T t \tag{56.5}$$

where

$$g(\alpha) = \int \frac{d\alpha}{f(\alpha)}$$

The rate constant k_T can be calculated from the Arrhenius equation given in Equation 56.6:

$$k_T = A \exp\left(-\frac{E_A}{RT}\right) \tag{56.6}$$

where

E_A is the activation energy (J mol^{-1})

A is the preexponential factor

R is the molar gas constant, 8.314 J (K mol)$^{-1}$

56.2.3 Measurement of α and dα/dt

Consider a thermogravimetric curve consisting of one step. α at a particular time can be found using

$$\alpha = \left(\frac{m_i}{m_i - m_f}\right) - \left(\frac{m_t}{m_i - m_f}\right) \tag{56.7}$$

The differential is hence,

$$\frac{d\alpha}{dt} = -\left[\frac{dm_t/dt}{(m_i - m_f)}\right] \tag{56.8}$$

This states that the rate of reaction can be measured from the slope of the mass–time curve. Since dm_t/dt is already measured by the DTG curve, $d\alpha/dt$ can be found directly from the curve.

Combination of a number of the equations discussed [21] gives

$$\ln \frac{d\alpha}{dt} - \ln(f(\alpha)) = \ln\left(\frac{A}{\beta}\right) - \frac{E_A}{RT} \tag{56.9}$$

where $\beta = dT/dt$.

56.3 Thermometric Titrimetry

Thermometric titration is the measurement of the temperature change in a system as a function of time or volume of titrant. The technique consists of the measurement of the change in temperature as the titrant is added to it, under near-adiabatic or more commonly referred to as isoperibol conditions. The experiments are typically carried out in a small dewar flask submerged in a well-controlled constant-temperature bath. The method can be used to study oxidation–reduction, complexation, precipitation, and neutralization reactions in aqueous solvents. Publications by Zenchelsky [22] and Jordan [23] review the technique in detail.

The basic principle is that a free energy change occurs in the system [24] and is based on the measurement of the free energy-dependent term

$$\Delta G^{\Theta} = -RT \ln K \tag{56.10}$$

where

ΔG^{Θ} is the change in free energy under standard conditions
R is the molar gas constant
T is the temperature in kelvin
K is the equilibrium constant for the system at the temperature T

A calorimetric method (entropy titration) for the determination of ΔG, ΔH, and ΔS from one thermometric titration has been described by Christensen et al. [24].

56.4 Thermomechanical Analysis

TMA relates to techniques where deformation is measured as a change in either volume or length. The deformation is plotted against temperature when a sample is heated under a controlled temperature program. Thermodilatometry measures the dimensional changes as a function of time under negligible loads. TMA is similar to thermodilatometry but also provides information regarding penetration, extension, and flexure using various types of loads on the test specimen. In dynamic mechanical analysis (DMA), the test specimen is subjected to a sinusoidally modulated stress under specified temperature. The viscoelastic response of a material is then monitored under tensile, compressive, shear, or torsional load [25].

56.4.1 Apparatus

A typical instrument for thermal mechanical analysis is called a dilatometer and is equipped with a linear variable differential transformer (LVDT). The displacement of the sample is transferred to the LVDT via a rod (probe) that is unaffected by heat and dimensional changes. A zero weight is accomplished for thermodilatometry by a float system so that a minimum of a load is subjected to the sample. The sample is placed on a sample holder in an oven. A force is applied through the probe in TMA and DMA. The sample cylinder and the probe are independently connected to the measuring device. The top of the probe is also connected to a balance arm. Probe movement and sample length changes are detected. The recorded signals are time, temperature, dimensional changes, and load. Various probes are available, depending on the analysis needs. Expansion, compression, and penetration probes are standard. Tension, three-point bending, and cubical expansion probes are available. Measuring temperature from –150 °C to 600 °C or even up to 1500 °C is possible, depending on the instrument.

56.4.2 Calibration of Probe

The temperature is usually the measured quantity in TMA. Therefore, calibrating the temperature axis is important. Thermomechanical analyzers can be temperature calibrated according to ASTM standard test method E 1363 [26]. An equation is developed for a linear correlation of the experimentally obtained

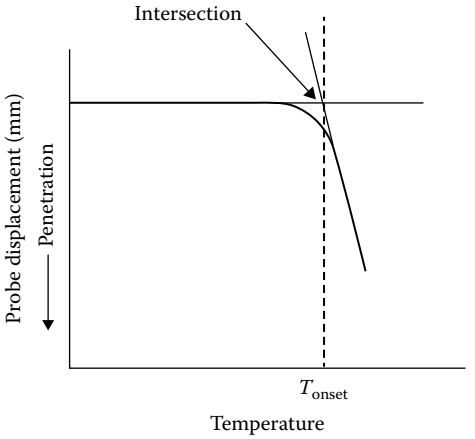


FIGURE 56.2 The calibration of a TMA instrument is a two-point method. There is an assumption that there is a relationship $T_t = (T_o \times S) + I$ between the actual specimen temperature and the onset temperature (Equation 56.11).

program temperature and the actual melting temperature for known melting standards (i.e., mercury, water, tin, benzoic acid). A penetration probe is used to obtain the onset temperatures for two melting standards. The two-point calibration assumes the relationship Equation 56.11 between the actual specimen temperature (T_t) and the observed extrapolated onset temperature (T_0). S and I are the slopes and intercept, respectively, in the TMA thermal curve (Figure 56.2):

$$T_t = (T_0 \times S) + I \tag{56.11}$$

Thermomechanical methods are generally applied on solid-shaped samples like polymeric products. Special clamps are used for testing of soft samples made of rubbers, adhesives, fats, etc. Films and fibers can be tested using clamps. Liquid polymers are tested on support. DMA is used for detecting α, β, and γ transitions in cured epoxy systems [25]. TMA is used for measuring the volume change of bitumen. Scratching and crack propagation at low temperatures is simulated. This is useful when investigating asphalt paving materials [27]. The thermal expansion coefficient of linear expansion is calculated from the slope of the expansion–temperature curve. This is obtained under zero load in thermodilatometry mode. Thermodilatometry can also provide information on phase changes, sintering, and chemical reactions. Softening temperatures are measured using small-diameter tips on the probe under a load (TMA). This sensitive technique is also used for the measurement of heat distortion temperatures and glass transition temperatures of polymers [28].

56.5 Differential Thermal Analysis and Differential Scanning Calorimetry

DTA is the detection of the temperature difference between the sample holder and the reference holder using the electromotive force of the thermocouples, which are attached to the holders. The sample and reference are subjected to a controlled temperature program. The differential is output as the DTA signal. DSC is similar to a DTA in construction, but the DSC measures the difference in heat flow rate to the sample and the reference. Consequently, more information is received on the thermodynamic behavior of the material using DSC. Quantitative DTA is also addressed as a DSC. This definition results in that the major part of all differential thermal analyses performed today uses DSC. An application is found for combined TGA/DTA analysis for kinetic evaluation of petroleum products [29].

56.5.1 Apparatus

The DTA apparatus has a sample and a reference cell subjected to the same temperature program. The measuring device consisting of a thermocouple or any temperature measurement device placed in each cell, measuring the difference in temperature. Operating temperature range is ambient to 1000°C or higher depending on the construction of the instrument and sample pan material. Differential scanning calorimetry is originally defined as individually heated cells. Equal temperature is maintained in the cells, giving an electrical signal proportional to the power needed. The DSC curves represent the rate of energy absorption. Today, most DTA units that can be calibrated to give calorimetric response are called DSC [25]. For qualitative applications, both classical DTA and DSC are equally good. In quantitative work, the DSC is claimed to be better at low heating rates [30].

56.5.2 Calibration and Reference Materials

The dynamic nature of thermal analysis requires a calibration and standard compound to be able to relate results obtained by different instruments. Temperature calibrations can be done using a range of selected materials. Different materials are chosen depending on the temperature range. Common standards are 1,2-dichloroethane, indium, silver sulfate, and quartz. Other organic compounds, metals, inorganic nitrates, sulfates, or chromates are also used [25].

56.5.3 Theory of DTA and DSC

The measured quantity is ΔT, the difference between the temperature of the sample and the reference material. In Equation 56.12, T_S is the temperature of the sample and T_R is the temperature of the reference material:

$$T_S - T_R = \Delta T \tag{56.12}$$

The result is presented as a plot of ΔT against T under a stated temperature program, the differential thermal curve. An endothermic process is then shown as a negative signal. A quantity of material decomposition or the enthalpy of the process is obtained from the area of the peak. It is then, in fact, a calorimetric analysis and the technique is referred to as differential scanning calorimetry.

In DTA, heat transfer to a sample and reference causes a difference in temperature ΔT, which can be related to the energy of any transition of the sample:

$$\Delta H = K \text{ (peak area)} \tag{56.13}$$

For heat flux DSC, a similar process occurs, whereas in power-compensated DSC, electric heating is supplied to the sample and reference to keep their temperatures as close as possible. For best calorimetric accuracy, the constant K should vary little with T.

56.6 Specialized Techniques

56.6.1 Thermoelectrometry

Electrical properties such as resistance/conductance and capacitance can be measured as a function of temperature. A variation that can measure the generated EMF is called thermovoltaic detection [8].

56.6.2 Modulated DSC

In Modulated DSC (MDSC), the heating rate is modulated. This is performed using a small alternating power supply in combination with the standard programmed heating. The heating program is given by the equation

$$T = T_0 + \beta t + B\sin(\omega t) \tag{56.14}$$

and the heat flow is given by

$$\frac{dq}{dt} = C_p[\beta\omega\cos(\omega t) + f(t,T) + C\sin(\omega t)] \tag{56.15}$$

where

T_0 is the initiation temperature
B is the amplitude of temperature modulation
ω (angular frequency) = $(2\pi f)$
C_p is the heat capacity
$f(t,T)$ is the kinetic response (average)
C is the amplitude of response to sine-wave modulation

56.6.3 Simultaneous Techniques

Each of the techniques discussed earlier provides information about the sample. However, a synergistic effect exists, in that the total amount of information obtained (by using techniques simultaneously) regarding the sample is greater than the sum of the information from the individual techniques.

56.6.3.1 Evolved Gas Analysis

This allows the identification of gases evolved during thermal analysis and is performed by replacing the detector with a mass spectrometer or FTIR. An alternative technique is to precede the detector by passing gases evolved during the thermal analysis through a gas chromatograph.

56.6.3.2 Thermomicroscopy

This can be incorporated under thermoptometry (a family of techniques that measure changes of an optical property with temperature change). Thermomicroscopy uses observations under a microscope.

56.7 Applications (Including the Analysis of Electrical Insulating Materials)

56.7.1 Oxidative Stability of Oils and Greases and Polymers

Oxidative degradation of oils upon heating can be monitored using a DSC apparatus. The detected onset *time* or *temperature* of the exotherm can be taken as a measure of the thermal/oxidative stability of the oil. The detected onsets are a strong function of the sample size, instrument sensitivity, kinetics, and scan rate. This enables DSC to be used in an oxidation test. Isothermal high-pressure DSC (PDSC) has been used to characterize the oxidative stability and the oxidation mechanisms of lubricants [34,35]. A PDSC works at pressures up to 3.5 MPa of a selected gas, using a wide temperature range. The technique is useful in the development of new lubricants with improved thermal and oxidative properties. The influence of metal catalysis on oil oxidation can be determined using PDSC. The volatile degradation products have been determined using combined PDSC–gas chromatography–mass spectrometry (GC/MS) [36]. PDSC gives information about relative oxidation stability used for comparing the lifetime of oils [37,38]. It has been a good technique for evaluating the thermal and oxidative stability of lubricating oils [34,39]. PDSC has also been used for evaluating deposit-forming tendencies of liquid lubricants [39].

Volatilization occurs when low-boiling oil is heated, especially at high temperatures. This leads to uncontrolled changes in composition. It also affects the size and shape of the DSC exotherm, causing imprecise determination of the oxidation onset [34]. Use of high pressure in the DSC cell reduces

volatility and evaporation interference with it. Added to this, the onset value is shifted to lower temperatures [34]. The onset becomes better defined and the peak size increases [40].

DSC is a fast technique for oxidation stability testing. This is a great advantage in the quality control of electrical insulating oils. Experimental evaluations of transformer insulating oils have shown ranking to be possible. The remaining lifetime of inhibited oils may correlate to the oxidation induction time [41]. Important parameters in PDSC are sample weight, pressure, and temperature program and have to be carefully considered before applying the technique. The sample pans must not be overlooked. Results are significantly influenced by variations in metallurgy, due to the catalytic and inhibiting effects of various metals. Oxidation induction time of lubricating greases can be determined by ASTM method D5483-93.

56.7.1.1 Predicting the Lifetime of a Product

Estimating the lifetime of a product typically uses some form of accelerated testing. TGA decomposition kinetics can be used to arrive at aging stability information and lifetime predictions in relatively short timescales (hours compared months in conventional oven aging). The sample (e.g., insulating paper) can be heated through its decomposition at several heating rates and the weight loss as a function of temperature recorded. The activation energy is calculated from a plot of log heating rate versus the reciprocal of the temperature for a constant decomposition level. The activation energy is subsequently used to calculate kinetic parameters such as specific rate constant (k) or half-life times, as well as to estimate the lifetime of the material at a given temperature. DTA and DSC have been employed in the electric industry to study polymeric insulation and for the determination of dielectric stability and lifetime prediction.

56.7.1.2 Thermal Analysis and Stability of Materials

TGA is widely employed in the determination of thermal stability of materials and analysis of their composition. The thermal history of electric cable insulation has been determined using DSC [42]. Thermal analysis techniques have greatly improved the quality control and inspection of electric cables. Hyphenated techniques have been employed in the analysis of trace components in electrical insulation [42–44]. ASTM method D3386-84 standardizes measurement of coefficient of linear thermal expansion of electrical insulating materials, while D3850-84 refers to the rapid determination of thermal degradation of solid electrical insulating materials by thermogravimetric methods.

56.7.1.3 Mechanical Stress Determinations

Longitudinal mechanical stresses, frozen into electric cable insulation during the fabrication process, can produce "shrink back." This causes the insulation to shrink away from freshly cut cable ends to varying degrees. TMA can be used to determine these stresses and has been found more versatile than the traditional BS6469 shrinkage measurement [45,46].

56.7.1.4 Evolved Gas Detection and Evolved Gas Analysis

The main use of EGD is to distinguish between phase transitions and endothermic decompositions (e.g., coordination chemistry). It has been used for the analysis of effluents [47]. TGA coupled with FTIR has been used to establish the failure mechanisms of electrical insulating materials [48]. EGA is also used for assessing the thermal endurance of polymeric materials and is of particular value in thermosetting polymers used in the electric industry [49].

56.7.1.5 Investigation of Polymeric Systems

TG can be applied to the study of polymer processes (pyrolysis, oxidative degradation, volatilization, absorption, adsorption, and polymerization) in which a change in weight occurs. The degree of crystallinity provides information regarding the thermal history of a polymer and can be measured by DSC. Physical and mechanical properties of polymers are related to the degree of crystallinity [50]. Thermophysical

property measurements and analysis of additives in polymers can also be performed using thermal analysis techniques [51]. ASTM method D4000-89 can be used for the identification of plastic materials.

56.7.1.6 Pharmaceutical Applications

Calorimetric purity determinations are used in the pharmaceutical industry. The concentration of the impurity is regarded as inversely proportional to its melting point. Therefore, an increase in the sample's impurity content decreases the melting point and broadens the melting range. DTA can also be used but DSC is preferred since it also gives the ΔH_f (heat of fusion) of the melt [52]. The DSC method is based on the van't Hoff equation. A compound may exist in various crystal forms; DSC and TG are used to characterize polymorphs and assess the stability of the compounds. DSC has been used for investigating the effect of inhibitors with model membranes [53]. Drug incompatibility is defined as "an interaction between two or more components to produce changes in the chemical, physical, microbiological, or therapeutic properties of the preparation" [54]. DTA and DSC are used to record reactions as a function of temperature and investigate drug compatibility [55]. Recent advances in microcalorimetry have allowed nondestructive analysis at room temperature [56,57]. The technique is gaining popularity in the pharmaceutical industry and also in the study of ballistics.

56.7.1.7 Characterization of Greases and Lubricants

Greases and lubricants are, in application, exposed to high temperatures in both inert and in oxidizing atmosphere. Material losses due to evaporation and loss or alteration due to thermal cracking or oxidation of the molecular structure are possible. The various aging reactions are usually inhibited by additives. TG, DTA, and DSC are used as test instruments, but the overall difficulty is to find methods that correlate with real thermal aging of the greases and lubricants.

The peak onset and peak maximum temperatures from DTG, DTA, and DSC curves are used or the peak onset from TGA curve. The evaporation behavior of greases is the most used parameter, but wax content, glass temperature, and cloud point are other characteristics of greases that are studied using thermoanalytical techniques [27].

In the Noack test of evaporative loss (DIN 51 581), the sample is held at 250°C for 60 min in an airflow. The sample is weighed before and after treatment. The cause of the weight loss is not clear, whether it is evaporation of parts of the original sample or evaporation of oxidative degradation products. The question has arisen whether isothermal TG could replace the Noack test. This would provide continuous loss information during the thermal exposure. It has been shown that there is a higher weight loss in the thermobalance than in the Noack test at equal test conditions. The deviation is caused by the difference in surface: volume ratio between the two methods [27].

Oxidation studies of low-boiling lubricant or lubricating oils do not give representative results. This is due to the evaporation of oil and existence of low-boiling oxidative degradation products. A TGA curve of a lubricant produced in an air atmosphere does not always represent the oxidation reaction. The use of an elevated oxygen or air pressure in DSC has been shown to reduce sample evaporation due to increased evaporation temperature and increase the rate of the oxidation. Several papers deal with this technique, which has found application in the characterization of lubricants [27,34–41,58].

56.7.1.8 Insulation Paper/Cellulose

The rate of weight loss on pyrolysis of cellulosic materials has applications to engineering problems in many industries. On heating, cellulose undergoes a number of linked physical and chemical changes [59]. Properties such as weight, strength, crystallinity, and enthalpy are affected.

TGA can be used to perform a collective measurement of the weight loss due to the production of H_2O, CO, and CO_2 during degradation. Of course, the measurement will include evaporation of other pyrolysis products. The enthalpy changes can also be measured by DSC. These methods are very useful in determining the temperature range at which physical and chemical processes occur. The rate of these processes can be determined by using DTG.

Acknowledgments

The authors would like to thank P. Haines, Chapman & Hall, and TA Instruments for their advice and permission.

Defining Terms

Derivative: Techniques where a measurement or calculation of the first derivative is performed.

Differential: Techniques where a difference in a property is measured.

Thermal analysis: A group of techniques in which a property of the sample is monitored against time or programmed temperature (in a specified atmosphere).

References

1. R.C. Mackenzie, A history of thermal analysis. *Thermochim. Acta*, 73, 249, 1984.
2. R.C. Mackenzie, Origin of thermal analysis. *Isr. J. Chem*., 22, 203–205, 1982.
3. R.C. Mackenzie, Early thermometry and differential thermometry. *Thermochim. Acta*, 148, 57–62, 1989.
4. W.W. Wendlandt, The development of thermal analysis instrumentation 1955–1985. *Thermochim. Acta*, 100, 1–22, 1986.
5. C.J. Keattch and D. Dollimore, *Introduction to Thermogravimetry*, Heyden, London, U.K., 1975.
6. A.L. Lavoisier and P.S. de Laplace, "Mémoire sur la. Chaleur", *Mém. Acad. R. Sci*., 355–408, 1780 (1784).
7. R.C. Mackenzie, C.J. Keattch, D. Dollimore, J.A. Forrester, A. Hodgson, and J.P. Redfern, Nomenclature in thermal analysis II. *Talanta*, 19, 1079–1081, 1972.
8. J.O. Hill, *For Better Thermal Analysis and Calorimetry III*. ICTA, 1991.
9. R.C. Mackenzie, Nomenclature in thermal analysis III. *J. Therm. Anal*., 8(1), 197–199, 1975; *Thermochim. Acta*, 28, 197, 1975.
10. R.C. Mackenzie, Nomenclature in thermal analysis, in *Treatise on Analytical Chemistry,* Part 1, Vol. 12, P.J. Elving, (Ed.), John Wiley & Sons, New York, pp. 1–16, 1983.
11. P.J. Haines, *Thermal Methods of Analysis*, Blackie/Chapman & Hall, London, U.K., 1995.
12. E.L. Charsley, J.P. Davies, E. Gloeggler, N. Hawkins, G.W.H. Hoehne, T. Lever, K. Peters, M.J. Richardson, I. Rothemund, and A. Stegmayer, *J. Therm. Anal*., 40, 1405–1414, 1993.
13. G.M. Lukaszewski, Accuracy in thermogravimetric analysis. *Nature*, 194, 959, 1962.
14. W.E. Garner, *The Chemistry of the Solid State*, Butterworths, London, U.K., 1955.
15. V. Šatava, Pouziti Terografickych Metod ke Studiu Reakcni Kinetiky. *Silikáty*, 5(1), 68, 1961. (The Thermographic merod [sic] of Determination of Kinetic Data).
16. V. Šatava and J. Šesták, Kinetika Analysa Termogravimetrickych Da. *Silikáty*, 8(2), 134, 1964. (Kinetic Analysis of Thermogravimetric Measurements), Source: un-numbered English-language contents page preceding p. 93 in *Silikaty*, 8(2), 1964.
17. V. Šatava and F. Škvára, Mechanism and kinetics of the decomposition of solids by a thermogravimetric method. *J. Amer. Ceram. Soc*., 52, 591–595, 1969.
18. J. Šesták, A review of methods for the mathematical evaluation kinetic data from nonisothermal and isothermal thermogravimetric measurements. *Silikáty*, 11, 153–190, 1967.
19. J. Šesták, Errors of kinetic data obtained from thermogravimetric curves at increasing temperature. *Talanta*, 13, 567, 1966.
20. M.L. McGashan, *Physicochemical Quantities and Units*, Vol. 39, RSC, London, U.K., 1968.
21. J. Šesták and G. Berggren, Kinetics of the mechanism of solid state reactions at increasing temperatures. *Thermochim. Acta*, 3, 1–12, 1971.
22. S.T. Zenchelsky, Thermometric titration. *Anal. Chem*., 32, 289R, 1960.

23. J. Jordan, *Handbook of Analytical Chemistry*, L. Meites (Ed.), McGraw-Hill, New York, Sec. 8–3, 1963.
24. L.S. Bark and S.M. Bark, *Thermometric Titrimetry*, International Series of Monographs in Analytical Chemistry, Vol. 33, Pergamon Press, 1969; J.J. Christensen, R.M. Izatt, L.D. Hansen, and J.A. Partridge, Entropy titration. A calorimetric method for the determination of ΔG, ΔH, ΔS from a single thermometric titration. *J. Phys. Chem.*, 70(6), 1966 and 40(1), 1968.
25. D. Dollimore, Thermoanalytical instrumentation, in *Analytical Instrumentation Handbook*, G.W. Ewing (Ed.), Marcel Dekker, New York, 1990.
26. ASTM Designation, E 1363-90, Standard test method for temperature calibration of thermomechanical analyzers. 1990.
27. H. Kopsch, *Thermal Methods in Petroleum Analysis*, VCH, Verlagsgesellschaft mbH, Weinheim, Germany, 1995.
28. H.H. Willard, L.L. Merritt Jr., J.A. Dean, and F.A. Settle Jr., *Instrumental Methods of Analysis,* 6th edn., Wadsworth Publishing Company, Belmont, CA, 1981.
29. ASTM Designation, E 698-79, Standard test method for Arrhenius kinetic constants for thermally unstable materials. 1979.
30. M.I. Pope and M.D. Judd, *Differential Thermal Analysis*, Heyden & Sons Ltd., London, U.K.,1980.
31. W.W. Wendlandt, Thermoelectrometry—A review of recent thermal analysis applications. *Thermochim. Acta*, 73, 89–100, 1984.
32. P.D. Garn and G.D. Anthony, Repetitive gas chromatographic analysis of thermal decomposition products. *Anal. Chem.*, 39,1445–1448, 1967.
33. M.R. Holdiness and R. Mack, Evolved gas analysis by mass spectrometry: A review. *Thermochim. Acta*, 75, 361–399, 1984.
34. S.M. Hsu, A.L. Cummings, and D.B. Clark, Society of Automotive Engineers SAE, Technical Paper 821252, 1982.
35. R. Schumacher, Practical thermoanalysis in tribology. *Tribol. Int.*, 25(4), 259–270, 1992.
36. A. Zeman, DSC cell—A versatile tool to study thermooxidation of aviation lubricants. *J. Synthetic Lubricants*, 5, 133–148, 1988.
37. E. Gimzewski, A multi-sample high pressure DTA for measuring oxidation induction times. *Thermochim. Acta*, 170, 97–105, 1990.
38. R.E. Kauffman and W.E. Rhine, *J. Soc. Tribol. Lubr. Eng.*, 44, 154–161, 1988.
39. Y. Zhang, P. Pei, J.M. Perez, and S.M. Hsu, *J. Soc. Tribol. Lubr. Eng.*, 48, 189–195, 1992.
40. R.L. Blaine, Thermal-analytical characterization of oils and lubricants. *American Laboratory*, Reprint, 6(1), 18–20, January 22, 1974.
41. M. Eklund, Literature review of DSC oxidation tests on petroleum products. Report to International Electrotechnical Commission Technical Committee, April 10, 1996.
42. J.W. Billing, Thermal history of cable insulation revealed by DSC examination. *IEEE DMMA Conference*, 289, 309–312, June 1988.
43. R.L. Hutchinson, Thermal analysis to spectroscopy, an overview of analytical instrumentation for electrical insulating materials. *Proceedings of the 17th Electrical/Electronics Insulation Conference*, Boston, MA, 1985.
44. M.T. Baker, S. O'Connor, and J.F. Johnson, Hyphenated analysis for trace components in electrical insulations. *Proceedings of the 17th Electrical/Electronics Insulation Conference*, Boston, MA, 1985.
45. J.W. Billing and D.J. Groves, Treeing in mechanically strained h.v. cable polymers using conducting polymer electrodes. *Proc. Inst. Elec. Eng.*, 121, 1451–1456, 1974.
46. BS6469 1992 Insulating and sheathing materials of electric cable Part 1 section 1.3 (equivalent to IEC 811–1–3. 1985 + Al: 1990).
47. W. Lodding (Ed.), *Gas Effluent Analysis*, Edward Arnold, London, U.K., 1967.
48. M. Ali, J.M. Cooper, S.J. Fitton, and S.P. McCann, The development of techniques for the analysis of materials. *7th INSUCON 1994, BEAMA Int. Elect. Insul. Conf.*, pp. 131–135.

49. M. Ali, J.M. Cooper, S.G. Swingler, and S.P. Waters, Simultaneous thermal and infrared analysis of insulating resins. *IEE 6th Int. Conf. Dielect. Mater. Meas. Appl*., Vol. 363, Manchester, U.K., pp. 77–80, 1992.
50. E.L. Charsley and S.B. Warrington (Eds.), *Thermal Analysis—Techniques and Applications*, RSC Special Publication No. 117, 1992.
51. E.A. Turi (Ed.), *Thermal Characterization of Polymeric Materials*, Academic Press, New York, 1981.
52. L. Kofler and A. Kofler, *Thermomikromethoden zur Kennzeidung Organisher Stoffe und Stoffgemische*, Verlag Chemie, Weinheim, Germany, 1954.
53. D.R. Reid, L.K. MacLachlan, R.C. Mitchell, M.J. Graham, M.J. Raw, and P.A. Smith, Spectroscopic and physicochemical studies on the interactions of reversible hydrogen ion. *Biochim. Biophys. Acta*, 1029, 24–32, 1990.
54. A. Wade (Ed.), *Pharmaceutical Handbook*, Vol. 28, 19th edn., Pharmaceutical Press, London, U.K., 1980.
55. J.L. Ford and P. Timmins, *Pharmaceutical Thermal Analysis*, Academic Press, New York, 1989.
56. R.J. Willson, A.E. Beezer, J.C. Mitchell, and W. Loh, Determination of thermodynamic and kinetic parameters from isothermal heat conduction microcalorimetry: Applications to long-term-reaction studies. *J. Phys. Chem*., 99, 7108–7113, 1995.
57. D.L. Hansen, Instrument selection for calorimetric drug stability studies. *Pharm. Technol*., 20(4), 64–65, 68, 70, 72, 74, 1996.
58. ASTM Designation D 5483-93, Standard test method for the oxidation induction time of lubricating greases by pressure differential scanning calorimetry.
59. D. Dollimore and J.M. Hoath, The preparation and examination of partially combusted cellulose chars. *Theromchim. Acta*, 45, 103–113, 1981.

57

Kinetic Composition Measurement

E.E. Uzgiris
General Electric Company

J.Y. Gui
General Electric Company Schenectady

57.1 Introduction

Kinetic methods involve the measurement of chemical reactions or processes in a time-dependent manner. Rates of dynamic processes are measured rather than the properties of a system at equilibrium. Of course, this approach is a central one for the study of chemical reactions and reaction mechanisms; however, it has much value in analytic chemistry; that is, in the determination of the composition of materials. This fact has been recognized for some time, but in recent years, there has been resurgence in interest in the use of kinetic methods in analytic chemistry. There have been several world congresses on this subject and numerous monographs [1,2], and the number of papers on kinetic methods has dramatically increased in the last decade [3].

Why is there such an interest? After all, there are many analytic procedures that are quite general and sensitive. As a specific example, consider the analysis for various metals in environmental samples. Metal ions can be detected by numerous means such as by ion-selective electrodes, atomic flame spectroscopy, or ion-coupled plasma spectroscopy, yet there is abundant literature on metal detection by catalyzed reactions in a kinetic manner [1–3]. In this case, the method of choice is dictated by cost of analysis, speed, sensitivity, and convenience. Furthermore, certain molecular species may be difficult to discriminate from others in conventional analysis. In this case, with a proper reaction, the kinetic approach is a powerful tool in detecting such constituents. Finally, the kinetic approach is the only method capable of elucidating the nature of binding sites in molecular binding because the determination of an equilibrium association constant alone is insufficient to elucidate mixed binding sites [4]. It is also the principal means of identifying short-lived intermediate species in a reaction [5].

Thus, kinetic methods comprise an important group of methods available for the analysis of substances. In some cases, kinetic methods offer unique advantages as in the study of mixed binding sites,

TABLE 57.1 Classification of Kinetic Methods Based on System and Catalyst

System	Catalyst	Reaction Examples
Homogeneous	Enzymatic	Hydrolysis electron transfer
	Nonenzymatic	Redox
		Complexation
		Chemiluminescence
	No catalyst	Redox
		Chemiluminescence
Heterogeneous	Enzymatic	Immunoenzymatic
		Electrode reactions
		Electrocatalysis
		Fluorescence
	Nonenzymatic	Electrode reactions
		Electrocatalysis
		Fluorescence
	No catalyst	Fluorescence
		Radioimmunoassay

in the delineation of competing species and in the determination of short-lived intermediates. In other cases, kinetic methods offer speed and convenience, and low cost, as, for example, in such applications as clinical analysis and environmental field analysis.

In a broader sense, time-dependent changes in chemical, physical, and biological processes are universal. Because equilibrium may not be achieved in certain processes, time-dependent effects must be considered and accounted for in a satisfactory manner for analytic determinations to be accurate and reproducible. In some instances, for reasons of speed of analysis, kinetic rates are measured rather than equilibrium values. The range of time dependencies can range from picoseconds, studied with mode-locked lasers, to seconds or minutes, studied with batch mixing procedures. Kinetic methods encompass a broad range of processes and time domains. The methods of simple chemical reactions can often be applied to complicated biological processes. This is possible because often one reaction in a group of coupled reactions controls the overall rate of the process.

Kinetic methods have been classified according to different criteria. The most common classification is based on whether the method involves a catalyst. This is so because reactions are frequently quite slow. In such cases, a catalyst must be added to speed up the reactions and make rate determinations practical. In other instances, the catalyst is the analyte itself. There are two major groups of catalysts: enzymatic and nonenzymatic. Another common classification of kinetic methods is based on whether the reaction proceeds in a homogeneous or heterogeneous system. Most of the discussion will be focused on homogeneous liquid and heterogeneous liquid–solid systems because these comprise the majority of kinetic analytical methods that have been developed. Presented in Table 57.1 are classifications based on the aforementioned criteria along with example reactions.

57.2 Theoretical Aspects

A reaction involving species A and B proceeds to a product with a rate constant, k', such that the rate of change of species A is given by

$$-\frac{\mathrm{d[A]}}{\mathrm{d}t} = k'[\mathrm{A}][\mathrm{B}] \tag{57.1}$$

where the brackets denote concentration. If the species A is of interest, then the reactant B can be in excess, in which case changes in [B] can be ignored. A pseudo-first-order reaction can be written as

$$-\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = k[\mathrm{A}] \tag{57.2}$$

where $k = k'$ [B] and the time evolution of [A] is just

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-kt} \tag{57.3}$$

The product, P, which is the species that is usually detected, evolves as

$$[\mathrm{P}] = [\mathrm{A}]_0 (1 - e^{-kt}) \tag{57.4}$$

The species A can be expressed in terms of product by

$$[\mathrm{A}] = [\mathrm{P}]_\infty - [\mathrm{P}] \tag{57.5}$$

where $[\mathrm{P}]_\infty = [\mathrm{A}]_0$. By measuring [P] as a function of time, the initial concentration of A can be deduced from a plot of

$$\ln\,[\mathrm{A}] = \ln\,[\mathrm{A}]_0 - kt \tag{57.6}$$

In this way, a calibration curve can be generated against which an unknown sample can be measured for the content of species A.

In case the reaction is of a different order, the time-dependent plots for determining $[\mathrm{A}]_0$ take on a different form. For example, in a second-order reaction, the rate of change of [A] is given by

$$-\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = k[\mathrm{A}]\,[\mathrm{A}] \tag{57.7}$$

where, as before, the reactant [B] is considered in excess and its time dependence can be assumed to be negligible. The calibration curve is now

$$\frac{1}{[\mathrm{A}]} = \frac{1}{[\mathrm{A}]_0} + kt \tag{57.8}$$

Clearly, the order of the reaction under study must be known for a correct analysis. There are straightforward ways to determine the order by varying the initial concentration of [A] and noting the initial velocity of the reaction. A plot of the initial velocities versus initial [A] will reveal the order of the reaction [1].

One of the strong points of kinetic methods is that closely related species that may be difficult to resolve by other means can be resolved by kinetic measurements. This is particularly true when enzyme reactions are employed. Enzymatic reactions are extremely sensitive to molecular structure and closely related structural analogs may have significantly different kinetics. For example, consider species A and B going through a reaction to a product but each having a different rate constant, k_a and k_b. The detected product is given as a sum of the two components by

$$[\mathrm{P}]_\infty - [\mathrm{P}] = [\mathrm{A}]_0 \exp(-k_\mathrm{a} t) + [\mathrm{B}]_0 \exp(-k_\mathrm{b} t) \tag{57.9}$$

Then, by computer fitting or graphical analysis of a semilog plot of

$$\ln \{[P]_\infty - [P]\} = \ln [A]_0 - k_a t + \ln [B]_0 - k_b t \tag{57.10}$$

one can extrapolate to $t = 0$ and determine $[A]_0$ and $[B]_0$.

The important case of catalyzed reactions must be considered separately as there are important differences from the case of uncatalyzed reactions considered earlier. The catalyst is usually the species to be determined as it often is a metal ion or nonorganic ion of interest. Usually, the catalyst combines with the reactant species [B] in a very fast reaction with a given equilibrium constant to give

$$[C] + [B] \leftrightarrows [CB] + [Y] \tag{57.11}$$

Here, CB, the reactant B bound to C, reacts with A with a much faster rate than if B is unbound. This develops because of the reduction of the activation energy provided by the catalyst C in combination with B as discussed later. Thus,

$$[CB] + [A] \leftrightarrows [P] + [Y] \tag{57.12}$$

This more complex kinetics simplifies to pseudo-first-order if one considers only the initial rates of the reaction. The initial velocity of the indicator product, P, takes the simple form

$$V_0 = \frac{d[P]}{dt} = K'[C]_0 + K'' \tag{57.13}$$

where K' are K'' constants. A calibration curve for C can thereby be generated through initial velocity measurements.

57.2.1 Enzyme Reactions

Enzymes are a class of proteins that catalyze reactions with exquisite specificity. The activity of certain enzymes is in itself of great importance in clinical diagnosis, but enzymes can be useful in determining substrate concentration—also very important for clinical applications and for environmental analysis. The rates of enzyme reactions are directly proportional to enzyme concentration; however, there is a saturation of reaction rates with increasing substrate concentration. This saturation effect must be considered when analyzing such reactions. The essential feature of enzyme reactions involves the enzyme, the substrate, the enzyme–substrate complex, and the product. It is the formation of the enzyme–substrate complex that leads to the saturation kinetics [5]. The reaction can be represented as follows:

$$[E] + [S] \overset{k_2}{\underset{k_1}{\rightleftarrows}} [ES] \overset{}{\underset{k_3}{\rightarrow}} [P] + [E] \tag{57.14}$$

where the reaction to form the enzyme–substrate complex is reversible as indicated by the arrows and k_1 is the forward rate and k_2 is the backward dissociation rate, and there is no reversion of product to substrate in the initial stages of reaction.

With the condition that initially $P \sim 0$, and setting $d[ES]/dt = 0$, it is easy to show that

$$\frac{[ES]}{[E][S]} = \frac{1}{K_M} \tag{57.15}$$

where K_M is the Michaelis–Menten constant. Now, since the velocity of the reaction (and here, one considers the initial velocity only) is given by

$$V_i = k_3[ES] \tag{57.16}$$

and one can define a maximum velocity such that

$$V_{imax} = k_3[E]_{tot} \tag{57.17}$$

The velocity is maximum when all of the enzyme binding sites are filled with substrate. Solving for [ES] and using $[E] = [E]_{tot} - [ES]$, one obtains

$$V_i = \frac{k_3[E]_0[S]}{(K_M + [S])} \tag{57.18}$$

This is the functional form that expresses saturation kinetics with respect to substrate concentration. Generally, for determination of activities, enzyme reactions are performed in a fully saturated regime (i.e., $[S] \gg K_M$); otherwise, the kinetic rates need corrections and the Michaelis–Menten constant must be known or needs to be determined. For determination of substrate concentration, the analysis must account for the nonlinearity of V_i with respect to [S].

Enzyme activity is defined in terms of units, rate of formation of product under given conditions, since the protein content in the enzyme preparation can be misleading—not all of the enzymes in a preparation need be active. Because enzymes are proteins, and in some cases rather delicate ones, great care must be exercised in handling and storing. The activities of enzymes are very sensitive to pH, salinity, and temperature. All of these factors must be precisely controlled for reliable kinetic determinations.

57.2.2 Temperature Dependence

Rate constants obey the Arrhenius relation

$$k = A \exp\left\{\frac{-E_a}{RT}\right\} \tag{57.19}$$

where

E_a is the activation energy
R is the gas constant
T is the absolute temperature
A is a prefactor term

Knowledge of the activation energy allows for the extrapolation of a kinetic rate to any temperature. It is the lowering of this activation energy that is at the heart of catalysis and enzymatic reactions. Because of the exponential dependence, a reduction of the activation energy can lead to a rate constant increase of many orders of magnitude.

The prefactor A is determined by some collision frequency. However, in general, reactions proceed slower than the collision theory would predict. This is because, in addition to collisional frequency, there are also configurational and entropic terms that play a role in determining A. Nevertheless, it is useful to consider the concept of diffusion-controlled reactions. Here, it is the collisional frequency that

dominates the reaction. In that situation, it is possible to utilize the diffusion theory of random motion in a medium to derive A such that

$$A_{\text{diff}} = \frac{4\pi(r_{ij})(D_i + D_j)N_0}{1000} \tag{57.20}$$

where

N_0 is Avogadro's number
D_i and D_j are the diffusion constants for species i and species j
r_{ij} is the encounter distance

For D of the order of 1.5×10^{-5} cm^2 s^{-1}, which is a value appropriate for small molecules, A_{diff} is 10^{10} M^{-1} s^{-1}. Reactions involving protonation or the OH^- ion proceed at this rate, but only a few enzyme–substrate complex-formation reactions approach the diffusion limited rate [6].

57.3 Experimental

The kinetic methods can be further classified according to experimental approaches as presented in Table 57.2.

The principal instrumental elements of a kinetic apparatus are the mixing chamber, timing device and control of data acquisition, and detector. Automation and computer controls have allowed kinetic measurements to be done routinely and with great accuracy for even very fast reactions. We consider those aspects of instrumentation unique to the problem of mixing and proper fast sampling—the essential issues of the experimental method. The other components of instrumentation are beyond the scope of this chapter; the readers may refer to the monographs for more details on those topics [1,2].

Although the nature of kinetic measurements does not require absolute quantitation of a product, it does require care in accurate timing and fast mixing of reactants. For slow reactions, the mixing chambers can be closed systems without any need for elaborate devices or techniques to initiate the reaction

TABLE 57.2 Classification of Kinetic Methods Based on Mixing Technique or Equilibrium Perturbation

Technique	Methods
Batch mixing (for slow reactions)	Stirring in cuvette or flask
Flow mixing (fast reactions)	Continuous flow
	Accelerated flow
	Pulsed flow
	Stopped flow
Thermodynamic jump	Temperature jump
	Pressure jump
	Electric current jump
	Concentration jump
Periodic relaxation	Cyclic voltammetry
	Dielectric relaxation
Pulse relaxation	Time-resolved fluorescence
	Time-resolved phosphorescence
	Flash photolysis
	Pulse NMR
	Pulse EPR

of interest. The so-called batch mixing can be done in ordinary optical cuvettes with a suitable magnetic stirring rod or mixing plunger. These straightforward experimental techniques are not discussed here; rather, the time domain for which kinetic methods require specialized equipment will be considered. This domain is in the region of 1 ms to 1 s, for example. Reactions with time constants in this domain are very common in current applications of kinetic methods.

57.3.1 Mixing Methods

In the so-called open systems, there are three approaches to initiating and monitoring reactions: (1) continuous flow, (2) pulse and accelerated flow, and (3) stopped flow.

In method (1), the reactants are brought together into a capillary under fast-flow conditions and the product is monitored (e.g., by a photodiode) along the length of the capillary, thus tracing out the kinetics in so far as the time dependence of the reaction is transformed into distance along the capillary by

$$t = \frac{d}{v} \tag{57.21}$$

where

v is the flow velocity

d is the distance along the capillary after the junction in which the reactants are introduced

A high flow rate ensures a high Reynolds number condition and the achievement of turbulent flow and good mixing in the capillary. This method has the disadvantage of requiring rather high molar extinction coefficient for the product to achieve sensitivity and the high consumption of sample and reactant. In addition, multiple measurements along the tube are required to trace out the kinetics.

Method (2), pulsed and accelerated flow, was devised to address these deficiencies. By accelerating the flow, it is possible to do a single-point measurement: the kinetics can be deconvoluted from the known change of flow as a function of time. In addition, integrated detection can be used in which the light path of the detector and source look down the flow tube, thus affording much greater sensitivity by virtue of a long absorption path length. Rather small quantities of analyte and reactant are consumed by this method because the flow is not continuous and a single-point measurement is sufficient for the measurement of the kinetic parameters.

Method (3), the stopped-flow method has all the advantages of method (2), is simpler analytically, and can measure even faster kinetics. In this method, reactant and analyte are combined from two syringes driven simultaneously by a push block as shown schematically in Figure 57.1. As the stop syringe plunger hits a precalibrated stop position, the flow is halted. Data are accumulated after the flow is stopped, free from effects of flow turbulence and other time-dependent interferences. Dead times (i.e., the time between inception of mixing and start of measurements) can be as short as 0.5 ms. The steps involved in the measurements can be automated for multisample, high throughput applications. A particularly simple stopped-flow system has been described by Harvey [7]. The drive syringes are standard 10 mL syringes that are manually pushed by a plunger. The mixing chamber is at the bottom of a 3 cm^2 observation cell. As the mixed solution enters the cell, it pushes up a float past the level of the light beam by which the reaction is monitored. After the measurement, the spent solution is displaced by pushing down on the float. This very simple approach is adequate for reactions slower than some 100 ms or so.

Reactions fast as 1 ms can be measured by stopped-flow techniques. Fully automated sampling and data acquisition systems have been implemented [8]. Computer-controlled, three-wave valves are used to charge up the drive syringes and to flush them clean between measurements. Very fast reactions such as those involved in the folding of proteins have been studied in this way [9].

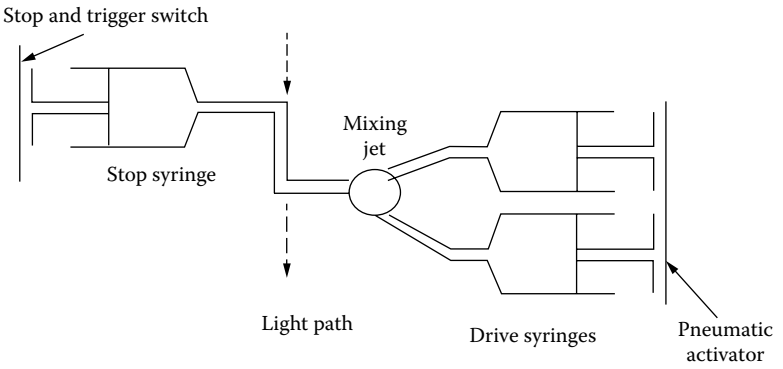


FIGURE 57.1 Schematic diagram of stopped-flow system. The reactant solutions are taken up into the two drive syringes as shown. The charging of the syringes is accomplished by valves and reservoirs not shown in the schematic for simplicity. As the activator plunger pushes the solutions through the mixing chamber, which is designed for efficient and fast mixing through tangential injection and turbulent flow (as in the Dionex Corp. system, for example), the old spent solution in the observation chamber is forced out into the stop syringe. The stop syringe plunger hits a stop, which causes immediate cessation of flow and activates the data acquisition system, which may be an oscilloscope, strip chart, or computer. The observation chamber shown here is oriented parallel to the light beam path for maximum path length and maximum sensitivity to absorption changes. The dead time of such a system can be as low as 0.5 ms and the quantity of solutions required can be as low as 100–500 μL. At the end of a measurement, the stop syringe is purged and the drive syringes are recharged for another measurement cycle.

57.3.2 Relaxation Methods

An entirely different approach to kinetics is to probe the reactions of two reacting species that are in equilibrium by perturbing the equilibrium by a sudden change of temperature or pressure. These methods, known as relaxation methods, have as their virtue the ability to resolve kinetics in the very fast time regime much shorter than 1 ms [10]. If the equilibrium is disturbed, the relaxation to the new equilibrium state will proceed with a time constant τ given by

$$\frac{1}{\tau} = k_1 + k_{-1} \tag{57.22}$$

where k_1 and k_{-1} are the forward and back reaction rate constants between the two species, respectively. The magnitude of the response depends on the enthalpy change with temperature or volume change with pressure of the particular reaction under study.

In the temperature jump method, a pulse of energy is supplied to the sample, either by a current pulse if the solution is conducting or by a light pulse if the solution is absorptive at a suitable wavelength. Light pulses can be made extremely short with a suitable laser source—nanoseconds to picoseconds—and this approach lends itself to the examination of the very fast molecular processes such as the intermediate states in photoreception [11].

These methods are well suited for the study of fast reactions but less useful for compositional analysis. However, a type of relaxation that is well known (i.e., fluorescence and phosphorescence) has become very valuable analytical tools. In such methods, a light pulse populates an excited state of molecules under study, and the rate of decay of that state provides an identifying signature of the species, even in a background of other emissions at the same wavelength. This holds as long as the signal of interest decays with a different time constant from the background signals.

The steady-state aspect of fluorescence and phosphorescence spectroscopy is a well-established "equilibrium" analytical technique that relies on the spectral differences for identification and intensity

differences for quantitation. Its application, however, becomes invalid when an analytical sample contains multiple species that have indistinguishable luminescence spectra. However, it is frequently the case that the different species have different luminescence lifetimes. Thus, time-resolved spectra may produce a resolution of the species. For example, some tetracyclines have overlapping phosphorescence spectra that prevent characterization of each individual tetracycline. By using time-resolved, room-temperature phosphorescence, simultaneous determination of these tetracyclines was achieved based on their decay times in a continuous-flow system [12]. Even if two species have similar lifetimes, one can attach luminescence groups with different lifetimes to differentiate them. For example, different antigens tagged with different dyes with different lifetimes [13] were used to allow simultaneous detection of the antigens.

There are many benchtop fluorescence instruments capable of measuring lifetimes as short as nanoseconds. However, most of them are capable of monitoring only one specific wavelength as a function of time. Recently, fast optical spectrometers have been developed that have nanosecond time resolution over the entire visible spectrum [3].

57.4 Catalytic Reactions

Catalytic methods are based on the kinetic determination of catalyzed reactions. Such reactions can be extremely sensitive when the catalyst is the analyte. For example, chemiluminescence reactions of the oxidation of luminol by hydrogen peroxide catalyzed by metal ions provide extremely low detection limits for Co(II), Cu(II), Ni(II), Cr(III), and Mn(II). It should be pointed out that the term "catalyst" is loosely defined here as a substance that modifies the rate of a reaction without altering its equilibrium. Thus, the term "catalyst" includes the notion of promotion, inhibition, and, of course, true catalysis in which the catalyst remains chemically unchanged at the end of the reaction. Catalysts are usually categorized into two groups: enzymatic and nonenzymatic. Discussed later are overviews of catalytic-based kinetic methods applied in both homogeneous and heterogeneous systems.

57.4.1 Homogeneous Systems

Most applications of homogeneous kinetic methods are based on rate determination of catalyzed indicator (or substrate) reactions. Most frequently, the catalyst is the analyte to be determined, although, occasionally, it may serve simply as a reagent. Enzymes are one special type of catalyst. They are proteins possessing a very high degree of specificity. For example, certain enzymes can only exert catalytic actions on particular chemical bonds or steric isomers. Homogeneous enzymatic methods are widely used in clinic diagnoses to determine enzyme activity as well as enzyme–substrate concentrations. The theoretical aspects of enzyme kinetics have been discussed in the previous section. Analytical applications for both enzyme activity determination and enzyme–substrate detection can be found in the literature [3,14].

Homogeneous nonenzymatic catalytic methods are mainly applied for detection of metal ions and other simple inorganic and organic species [3,15]. There are three major types of indicator reactions: redox, chemiluminescence, and complexation. One popular redox indicator reaction is the reduction of hydrogen peroxide by iodide catalyzed by metal ions (Fe, Mo, W, and Zr) that are also the analytes. The most common chemiluminescence indicator reaction is the decomposition of luminol (5-amino-2,3-dihydrophthalazine-1,4-dione) accompanied by the generation of luminescence at 425 nm. This decomposition is achieved through the oxidation of the doubly charged anion by the oxidant in this reaction (metal ions in most cases). Although the oxidant in this case is consumed during the reaction, it is often termed a "catalyst" in the literature because its consumption is negligible in the time frame of the initial rate measurement, principally because of the ultrasensitivity of the chemiluminescence measurement. For complexation reactions, there are two main groups: ligand-exchange and complex-formation reactions. They are less studied compared with the aforementioned two indicator reactions

but have promising application in the determination of nontransition metals. For example, a reaction involving ligand exchange can be used to detect 0.4 ppm Ca. The most widely used detection technique for the complexation indicator reaction is UV/VIS absorbance.

57.4.2 Heterogeneous Systems

Many kinetic methods depend on the application of different heterogeneous catalysis processes where the catalytic reaction takes place at the interface between two immiscible phases, usually between the liquid–solid phases. The discussion here focuses on two main areas of heterogeneous catalysis that are important in chemical analysis. The first encompasses immobilized enzymes in which the labeled enzymes are either physically or chemically attached onto a solid surface. The measurement of surface enzyme activity is then related to the analyte concentration. The second is the area of electrocatalysis, in which chemical reactions occur at the interface of an electrode and an electrolyte solution. The catalyst in this case is the charged electrode surface in either the intrinsic state or in a chemically modified state. The analyte concentration in the solution is determined by the electrode dynamic current.

The most widely used format for immobilized immunoenzymatic techniques is known as enzyme-linked immunosorbent assay (ELISA). This type of assay combines the great selectivity provided by specific antibody–antigen recognition, the high sensitivity provided by enzymatic amplification, and general applicability provided by the use of common detection methods. It has proven to be a very powerful technique for simple, rapid, and cost-effective trace analysis and is widely used today in clinical diagnosis [16], drug screening [17], food safety inspection [18], and environmental analysis [19,20].

ELISA can be operated in several different modes, depending on the nature of analyte, sample environment, and requirements on speed, cost, and detection limits. Different assays are usually classified according to their operating procedure (competitive or noncompetitive), to the signal detection technique used (calorimetric, luminescent, electrochemical, or radioactive), or to the physical arrangement of the antibody–antigen binding structure (single layer or sandwich layers). For a more detailed description, the reader is referred to several references [21,22].

A typical immunoassay procedure involves three steps: (1) immobilization of antibodies onto a solid surface, (2) competitive binding of analytes and enzyme-tagged conjugates to the antibody sites, and (3) rate measurement of a substrate reaction catalyzed by the enzyme. In most cases, only the latter two steps operate in a kinetic mode. Illustrated in Figure 57.2 is the chemiluminescence ELISA developed for rapid field analysis for polychlorobiphenyls (PCBs) in which the kinetic response of the enzymatic reaction enables the quantitative determination of PCB concentration [19]. First, a solid support of specified material and format is chosen based on the analysis requirement. The support surface is then treated with protein A, a procedure to allow for the immobilization of antibodies in the proper orientation as shown in Figure 57.2. The third step is to immobilize the antibodies onto the protein A–coated surface. Then an enzyme–antigen conjugate (specifically the bromobiphenyl–alkaline phosphatase) is introduced to the well so as to saturate all of the antibody binding sites. After thorough rinse with pH 7 buffer solution, these conjugate treated well plates are ready for use in analysis of samples. The analysis of PCB-containing samples proceeds simply by adding the PCB-containing solution into the well for a fixed time to allow the PCBs to displace the previously bound enzyme conjugates. The higher the PCB concentration in solution, the higher will be the displacement of the enzyme conjugates in a given amount of time. After a fixed time, the well is then thoroughly rinsed and a chemiluminescence substrate is added. Under the catalysis of alkaline phosphatase, the substrate is transformed into a luminescent species that is then detected. The initial luminescence generation rate or the total intensity within a fixed time is proportional to the surface alkaline phosphatase and thus inversely related to the PCB concentration, as shown by the results in Figure 57.3.

Electrocatalytic reactions have been widely used for measuring chemical variables for electroactive species. However, not all electroactive species can be measured by electrochemical methods because

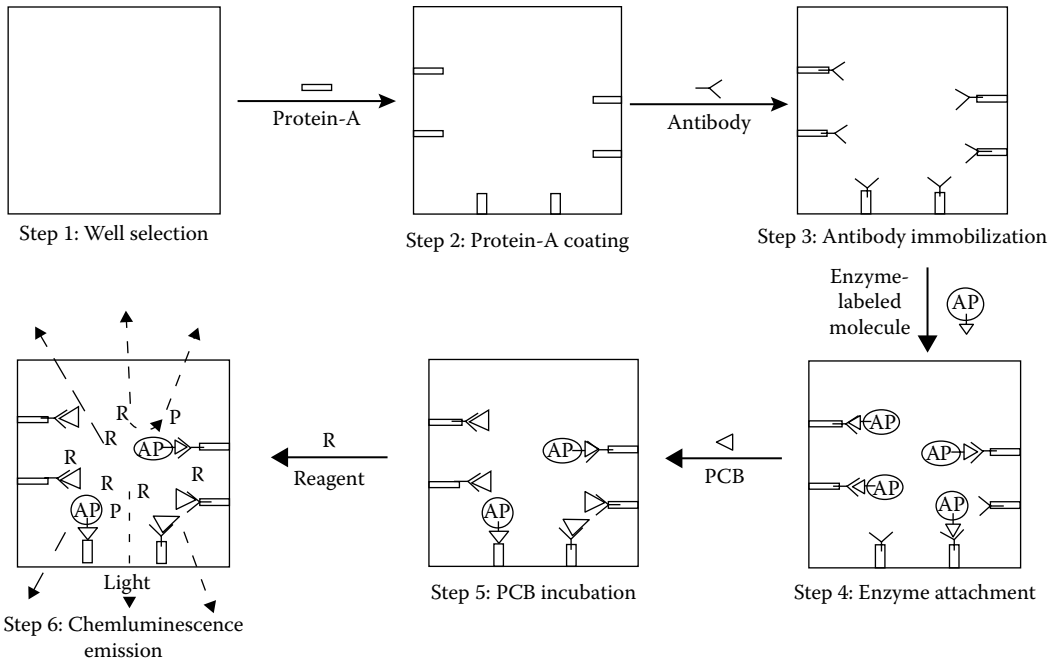


FIGURE 57.2 Pictorial presentation of chemiluminescence immunoassay.

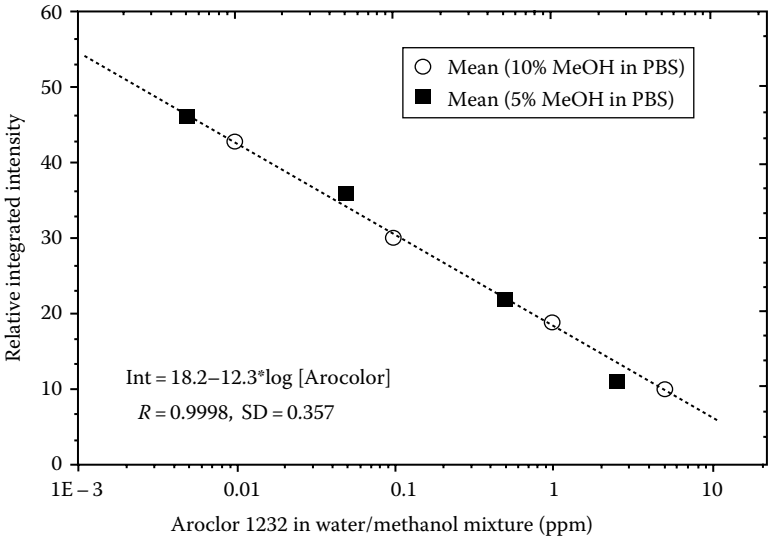


FIGURE 57.3 Dependence of chemiluminescence relative intensity on PCB Aroclor concentration. Plotted are chemiluminescence signals integrated during the first minute of enzymatic reaction (adamantyl dioxetane decomposition catalyzed by alkaline phosphatase in pH 10 buffer.) Samples contain various amount of Aroclor 1232 in pH 7 PBS buffer solution containing 5% (open circles) or 10% (solid squares) methanol.

for some species, the electrode reaction kinetics may be very slow. A simple example is the reduction of molecular oxygen (O_2) at bare Pt electrodes in an aqueous solution. Oxygen cannot be reduced at the thermodynamic potential of the electrode. In this case, one can apply a large overpotential to drive the O_2 reduction. Unfortunately, in many cases, a large overpotential cannot be used because of the limited available potential window or because of the interference from other electroactive species.

Thus, to overcome this problem, electrochemists have chemically modified electrode surfaces in order to accelerate electron transfer rates at the electrode–solution interface [23].

Chemical modification is produced by coating a monolayer of atoms, molecules, or thin layers of polymers onto the electrode surface. These surface-attached molecules may or may not be electrochemically active, but they can accelerate electrode kinetics for the target analyte. When the surface species is electrochemically active, it is termed a mediator; when inactive, it is called a promoter. For example, cytochrome c, like many other large macromolecules, has a large electron transfer rate in a homogenous solution phase, but it exhibits extremely slow electron transfer kinetics at many metal electrode surfaces. Eddowes and Hill, as well as Gui and Kuwana [24], have successfully demonstrated that by adsorbing a monolayer of heteroaromatic molecules such as 4, 4′-bipyridyl and *trans*-1,2-bis (4-pyridine) ethylene onto Au or Pt electrode surfaces, electron transfer kinetics of cytochrome c is significantly promoted. Recently, Dong, Cotton, and coworkers [25] have used a halide-modified Au electrode to study cytochrome c electrode kinetics. They adsorb different halides onto the Au electrode and find that they all can accelerate the electron transfer rate for cytochrome c and the promoting effort is of the order of $F^- < Cl^- < Br^- < I^-$. Various theories for the aforementioned phenomenon have been proposed. One possible explanation is that cytochrome c and related electron transfer molecules can adsorb onto bare electrode surfaces in undesired orientations. Besides the aforementioned atomic and molecular modified electrodes, lipid-modified electrodes [26] have also shown some promoting effect for cytochrome c electron transfer. Direct immobilization of cytochrome oxidase in a lipid bilayer at an Au electrode has resulted in electrochemical reactivity of cytochrome c in solution [27].

Electrochemical methods can also be applied to analyze electrochemically inactive species. The electrochemical immunoassay is a typical example. It combines the great selectivity provided by specific antibody–antigen recognition, the sensitivity provided by catalytic amplification, and the simplicity of electrochemical detection. It has proven to be a useful technique for measuring chemical variables for biological, clinical, and environmental samples. There are many forms of electrochemical assays: homogeneous versus heterogeneous, competitive versus noncompetitive, enzymatic versus nonenzymatic, and simple versus sandwich. Details can be found in Refs. [21,28,29].

The great advantage of the electrochemical immunoassay compared with enzyme-modified electrode methods is that it is a universal method and can be configured to analyze wide range of analytes, regardless of their electrochemical reactivity. For example, Heineman and coworkers have used this technique to detect dioxin, with a detection limit of one attomole using alkaline phosphatase as enzyme to convert 4-aminophenyl phosphate to the electroactive species 4-aminophenol [28]. They also used multiple metal labels rather than enzyme labels for simultaneous detection of multiple analytes [29].

57.5 Noncatalytic Reactions

As stated earlier, most kinetic-based analytical methods are catalytic systems. Noncatalytic systems have more limited applications because equilibrium methods are usually adequate in providing the necessary accuracy and sensitivity, and the noncatalytic kinetic methods do not provide any advantages of sensitivity. However, kinetic methods have been found to be more valuable or even the only choice in some special cases, as illustrated by the following examples:

1. When a sample contains hard-to-separate interference species that demand laborious and time-consuming separation before final measurement with a classic equilibrium method. A kinetic method may provide a simpler and faster determination by not requiring a prior separation.
2. When the equilibrium method is based on a very slow reaction or a reaction cannot proceed to completion due to side reactions. In this case, an initial rate measurement is much preferred and may be the only method of analysis.
3. When the species of interest has an extremely short lifetime, such as in the case of a reaction intermediate.

TABLE 57.3 Companies Providing Kinetic Instrumentation or Kinetic Assay Materials

Instrument	Company
ELISA apparatus	Dynatech Laboratories 14340 Sullyfield Circle Chantilly, VA 22021
Immunoenzymatic assays	Becton Dickinson Microbiology Systems P.O. Box 243 Cockeysville, MD 21030
Enzyme assay kits, ELISA kits	Pierce Chemical Co. 3747 N Meridian Rd. P.O. Box 117 Rockford, IL 61105
Electroanalytic instruments	EG&G Princeton Applied Research P.O. Box 2565 Princeton, NJ 08543
Stopped-flow apparatus	Dionex Corp. 1228 Titian Way Sunnyvale, CA 44088
Spectrometer with stopped-flow attachment	On-Line Instruments, Inc. 130 Conway Drive Bogart, GA 30622
Time-resolved spectrometers	Perkin-Elmer Corp. 761 Main Ave. Norwalk, CT 06859

Note: As examples, one company is listed for each category. For more complete listings, the reader is referred to the latest buyer's guide of *Analytic Chemistry*.

Table 57.3 lists some vendors of appropriate apparatus and assay kits for performing kinetic determinations.

Defining Terms

Antibody: One of a class of immunoglobulins produced by an animal's immune response to antigens (i.e., substances foreign to the body). Antibodies bind to molecular determinants of the antigen with great specificity.

Catalyst: A substance that accelerates a chemical reaction but is not itself consumed by the reaction.

Enzyme: A protein molecule that catalyzes reactions with great specificity.

Substrate: That which is being transformed by an enzyme-mediated reaction.

References

1. D. Perez-Bendito and M. Silva, *Kinetic Methods in Analytic Chemistry*, Chichester, England, U.K.: Ellis Horwood, 1988.
2. A. Mottola, *Kinetic Aspects of Analytical Chemistry*, New York: John Wiley & Sons, 1988.
3. (a) H. A. Mottola and D. Perez-Bendito, Kinetic determinations and some kinetic aspects of analytical chemistry, *Anal. Chem*., 66, 131R–162R, 1994; (b) Kinetic determinations and some kinetic aspects of analytical chemistry, *ibid*., 68, 257R–289R, 1996.

4. S. F. Feldman, E. E. Uzgiris, C. M. Penny, J. Y. Gui, E. Y. Shu, and E. B. Stokes, Evanescent wave immunoprobe with high bivalent antibody activity, *Biosens. Bioelectron.*, 10, 423–434, 1995.
5. R. J. H. Clark and R. E. Hester, *Time Resolved Spectroscopy*, New York: John Wiley & Sons, 1989.
6. I. Tinoco, Jr., K. Sauer, and J. C. Wang, *Physical Chemistry, Principles and Applications in Biological Sciences*, 2nd edn., Englewood Cliffs, NJ: Prentice-Hall, 1985.
7. R. A. Harvey, A simple stopped-flow photometer, *Anal. Biochem.*, 29, 58, 1969.
8. S. R. Crouch, F. J. Holler, P. K. Notz, and P. M. Beckwith, Automated stopped-flow systems for fast reaction-rate methods, *Appl. Spectrosc. Rev.*, 1, 165, 1977.
9. M. S. Briggs and H. Roder, Early hydrogen-binding events in folding reaction of ubiquitin, *Proc. Natl. Acad. Sci. USA*, 89, 2017–2021, 1992.
10. P. Fasella and G. G. Hammes, A temperature jump study of aspartate aminotransferase, *Biochemistry*, 6, 1798–1804, 1967.
11. H. Shichi (Ed.), *Biochemistry of Vision*, New York: Academic Press, 1983.
12. F. Alava-Moreno, Y.-M. Liu, M. E. Diaz-Garcia, and A. Sanz-Medel, Kalman filtering-aided time-resolved solid-surface room temperature phosphorimetry for simultaneous determination of tetracyclines in solution, *Mikrochim. Acta*, 112, 47–54, 1993.
13. J. Choo, E. Cortez, J. Laane, R. Majors, R. Verastegui, and J. R. Villarreal, Far-infrared spectra and ring-puckering potential energy functions of two oxygen-containing ring molecules with unusual bonding interactions, *Proc. SPIE-Int. Soc. Opt. Eng.*, 2089, 538–539, 1993.
14. A. Mottola, *Kinetic Aspects of Analytical Chemistry*, New York: John Wiley & Sons, 1988, Chapter 3.
15. (a) D. Perez-Bendito and M. Silva, *Kinetic Methods in Analytic Chemistry*, Chichester, England, U.K.: Ellis Horwood, 1988, Chapter 2; (b) G. G. Guilbault, in *Treatise on Analytical Chemistry*, I. M. Kolthoff and P. Elving (Eds.), 2nd edn., Part I, Vol. 1, Chapter 11, New York: John Wiley & Sons, 1978.
16. D. S. Hage, Immunoassays, *Anal. Chem.*, 65, 420R–422R, 1993.
17. T. A. Brettell and R. Saferstein, Forensic science, *Anal. Chem.*, 65, 293R–310R, 1993.
18. S. K. C. Chang, P. Rayas-Duarte, E. Holm, and C. McDonald, Food, *Anal. Chem.*, 65, 334R–363R, 1993.
19. (a) J. Y. Gui, S. F. Feldman, E. Y. Shu, D. R. Berdahl, and E. B. Stokes, Chemiluminescence immunoassay for rapid PCB analysis, *Real-Time Analysis*, 1, 45–55, 1995; (b) J. Y. Gui, D. R. Berdahl, E. Y. Shu, J. J. Salvo, S. F. Feldman, and E. B. Stokes, *Chemiluminescence Immunoassay for PCB Detection*. U.S. Patent No. 5,580,741, Dec. 3, 1996.
20. J. M. Van Emon and R. O. Mumma (Eds.), *Immunochemical Methods for Environmental Analysis: 198th National Meeting of the American Chemical Society, ACS Symposium Series*, Miami Beach, FL, September 10–15, 1989.
21. (a) C. P. Price and D. J. Newman, *Principles and Practice of Immunoassay*, New York: Stockton Press, 1991; (b) T. T. Ngo (Ed.) *Electrochemical Sensors in Immunological Analysis*, New York: Plenum Press, 1987.
22. (a) E. Harlow and D. Lane, *Antibodies—A Laboratory Manual*, New York: Cold Spring Harbor Laboratory, 1988; (b) A. L. Ghindilis, P. Atanasov, and E. Wilkins, Enzyme-catalyzed direct electron transfer: Fundamentals and analytical applications, *Electroanalysis*, 9, 661–674, 1997; (c) B. Liedberg, C. Nylander, and I. Lundstrom, Biosensing with surface plasmon resonance—How it all started, *Biosensors Bioelectronics*, 10, i–ix, 1995.
23. M. D. Ryan, E. F. Bowden, and J. Q. Chambers, Dynamic electrochemistry: methodology and application, *Anal. Chem.*, 66, 360R–427R, 1994.
24. (a) M. J. Eddowes and H. A. O. Hill, Novel method for the investigation of the electrochemistry of metalloproteins: Cytochrome c, *J. Chem. Soc. Chem. Commun.*, 771, 1977; (b) Y. Gui and T. K. Kuwana, Electrochemistry and spectroelectrochemistry of cytochrome c at platinum, *J. Electroanal. Chem.*, 226, 199–209, 1987.

25. (a) T. Lu, X. Yu, S. Dong, C. Zhou, S. Ye, and T. M. Cotton, Direct electrochemical reactions of cytochrome c at iodine-modified electrodes, *J. Electroanal. Chem*., 369, 79–86, 1994; (b) X. Qu, J. Chou, T. Lu, S. Dong, and C. Zhou, T. M. Cotton, Promoter effect of halogen anions on the direct electrochemical reaction of cytochrome c at gold electrodes, *J. Electroanal. Chem*., 381, 81–85, 1995.
26. (a) Z. Salamon and G. Tollin, Chlorophyll-photosensitized electron transfer between cytochrome c and a lipid-modified transparent indium oxide electrode, *Photochem. Photobiol*., 58, 730–736, 1993; (b) P. Bianco, and J. Haladjian, Control of the electron transfer reactions between c-type cytochromes and lipid-modified electrodes, *J. Electrochim. Acta*, 39, 911–916, 1994.
27. J. K Cullison, F. M. Hawkridge, N. Nakashima, and S. Yoshikawa, A study of cytochrome c oxidase in lipid bilayer membranes on electrode surfaces, *Langmuir*, 10, 877–882, 1994.
28. N. Kaneki, Y. Xu, A. Kumari, H. B. Halsall, W. R. Heineman, and P. T. Kissinger, Electrochemical enzyme immunoassay using sequential saturation technique in a 20 ml capillary: Dioxin as a model analyte, *Anal. Chim. Acta*, 287, 253–258, 1994.
29. (a) W. R. Heineman, H. B. Halsall, K. R. Wehmeyer, M. J. Doyle, and D. S. Wright, Immunoassay with electrochemical detection in methods of biochemical analysis, *Method. Biochem. Anal.*, 32, 345–393, 1987; (b) M. J. Doyle, H. B. Halsall, and W. R. Heineman, Heterogeneous immunoassay for serum proteins by differential pulse anodic stripping voltammetry, *Anal. Chem*., 54, 2318–2322, 1982.

58

Chromatography Composition Measurement

Behrooz Pahlavanpour
Nynas AB

Mushtaq Ali
The National Grid Company

C.K. Laird
Kingston University

58.1 Introduction

During the early development of modern analytical chemistry, the study of natural materials was a primary concern of organic chemists and biologists. A major problem facing these scientists was the formulation of methods to separate and analyze the complex mixtures encountered in biological research.

Chromatography (literally "color writing") is a physical or physicochemical technique for separation of mixtures into their components on the basis of their molecular distribution between two immiscible phases. One phase is stationary and is in a finely divided state to provide a large surface area relative to volume. The second phase is mobile and is caused to move in a fixed direction relative to the stationary phase. The mixture is transported in the mobile phase, but interaction with the stationary phase causes the components to move at different rates.

Origination of the technique, early in the twentieth century, is generally attributed to Tswett, who separated plant chlorophylls by allowing solutions in petroleum ether to percolate through a vertical glass tube or column, packed with calcium carbonate. The separated components formed colored bands

that were later isolated. Chromatography was adapted for qualitative or quantitative analysis of mixtures by inclusion of a suitable detector at the downstream end of the column and allowing the separated components to pass completely (elute) through the column and detector.

To analyze a sample, a suitable volume is injected into the stream of mobile phase or onto the upstream end of the column and the output of the detector is continuously monitored. The composition of the stream (eluent) passing through the detector then alternates between the pure mobile phase and mixtures with each of the components of the sample. The output record of the detector (chromatogram), plotted as a graph of response versus time, shows a series of deflections or peaks, spaced in time and each related to a component of the mixture. For a given column, mobile phase, and set of operating conditions, the time for a component to pass through the column (retention time) is characteristic and can be used to identify the component. The peak area is proportional to the concentration of the component in the mobile phase.

In modern instrumental applications of chromatography, the stationary phase is either a solid or a liquid, and the mobile phase either a liquid or a gas. The various types of chromatography are classified according to the particular mobile and stationary phases employed. The solid stationary phase may be a granular solid packed in a tube (column) or coated as a thin layer on a suitable supporting plate (thin-layer chromatography, TLC). Liquid stationary phases may be coated onto granular solids or bonded as a thin film to the inner wall of a capillary tube. In gas chromatography (GC), the mobile phase is a gas (carrier gas), and the stationary phase is either a high-boiling liquid (gas–liquid chromatography, GLC) or a solid (gas–solid chromatography, GSC). In liquid chromatography (LC), the mobile phase is a liquid and the stationary phase is either a solid (liquid–solid chromatography) or a second liquid, immiscible with the mobile phase, coated on a granular solid (liquid–LC).

58.2 Principles

Chromatographic theory is given in general textbooks [1,2] and also in specialized texts on different types of chromatography [3,4]. Chromatographic separation involves continuous interchange of solute molecules between the mobile and stationary phases. Four principal processes are involved: adsorption, liquid–liquid partition, ion exchange, and size exclusion. In GC, the predominant processes are adsorption, while LC may involve all four processes. Where liquid–liquid partition is the predominant separation mechanism, the sample components are eluted in order of increasing boiling points.

Since the analyte is transported in the mobile phase, chromatography is limited to solutes that are distributed between the two phases. In practice, this means that GC is limited to substances that are thermally stable in the vapor phase and are volatile at temperatures up to the maximum operating temperature of the GC column (about 350 °C–400 °C for most columns and packings, although some metal columns can be operated at higher temperatures). LC can be used for analysis of thermally labile and high molecular weight materials such as polymeric materials and proteins, at temperatures below their boiling point and that of the eluent.

The separating power of a chromatographic column is described by analogy with distillation–separation processes. It is given as the number of theoretical separation plate (either per meter of column length or in total for the column) and depends on its length, internal diameter, and the stationary phase employed. The height equivalent to a theoretical plate (HETP) value may also be quoted. Separating power is enhanced by use of long, narrow-bore columns with packings of the finest possible mesh size to allow intimate contact between the mobile and stationary phases, but these columns require higher operating pressures to overcome the column resistance, and analysis times increase with column length.

The chromatographic separation process is highly dependent on the temperature of the column, and temperature effects can be related to the temperature dependence of the distribution or adsorption equilibria of the solute between the stationary and mobile phases. However, in practice, the choice of column operating temperature involves a compromise between resolution and speed of analysis. LC is commonly carried out with the column at ambient temperature, although applications requiring

operation at temperatures up to 100 °C are becoming more common. In GC, the column temperature has a major influence on the speed of elution and the separation of sample components. GC can be carried out at a single controlled temperature (isothermal) or the oven temperature can be increased during the analysis in one or more linear ramps (temperature programming). Temperature programming speeds the elution of later components relative to early ones and enables mixtures containing a range of components to be separated more quickly than would be possible with isothermal operation. The broadening of later peaks due to diffusion in passing through the column is also minimized by temperature programming.

58.3 Gas Chromatography

A block diagram of a gas chromatograph is shown in Figure 58.1. The essential components are the column or columns; the carrier gas supply and flow and pressure controllers to enable carrier gas to be delivered to the column at a constant, controlled, and known rate; and the detector or detectors and associated electronics and data recording and processing system. An injector or facility for introducing suitable volumes of sample must be provided at the upstream end of the column. The column must be contained in an environment whose temperature can be held at a constant known value or heated or cooled at known rates. Temperature control in the range 25 °C–400 °C (–0.1 °C) and heating and cooling rates of 0.1 °C–40 °C min^{-1} are typically required. Subambient operation at temperatures down to about –30 °C may be required for separation of some volatile materials or for certain specialized eluents such as liquid carbon dioxide. Both injectors and detectors must be temperature controlled to allow rapid volatilization of the sample in the injector and to prevent condensation in the detector.

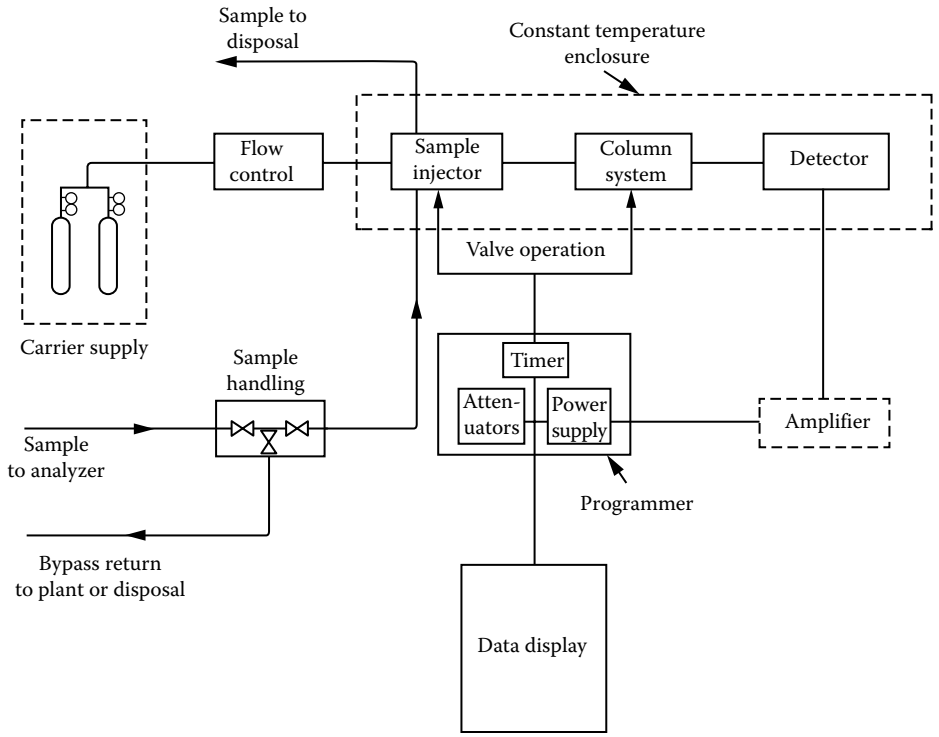


FIGURE 58.1 Functional diagram of process GC. The system consists of gas flow control, sample injection, separation of the components through the column, and a detector. Carrier gas at high pressure is used to move the sample through the column.

58.3.1 Columns

In GC, the processes involved in separation are predominantly adsorption and liquid–liquid partition when the eluent is liquid CO_2. Separation is almost entirely dependent on the nature of the stationary phase, with the gas phase acting mainly as an inert carrier. Separations of permanent gases are carried out by GSC using adsorbents such as silica gel, alumina, or synthetic zeolites (molecular sieves, particularly MS 5A and MS 13X) as the stationary phase. Proprietary porous polymers, such as Porapak (styrene–divinylbenzene copolymer), Chromosorb, and Tenax (polymer of 2,6-diphenyl-*p*-phenylene oxide), and various "carbon molecular sieves" are also used, and some of these materials are also used in separation of liquid samples.

In GLC, the stationary phase is a high-boiling liquid, coated to a few percent by weight on an inert granular support such as silica, firebrick, diatomaceous earth, or Teflon. A wide range of liquids, gums, and waxes have been employed that provide stationary phases that are usable over different temperature ranges and with different polarities. Examples include silicone oils and gums, hydrocarbons, polyphenyl ethers, and high molecular weight polymeric alcohols.

Originally, in both GLC and GSC, the stationary phase or coated support was packed into a column, typically a glass or stainless steel tube, 1–3 m long, and coiled to fit the chromatograph oven. Developments in column technology have led to the gradual, but not yet complete, replacement of the packed GC columns by capillary columns. These columns are usually formed from drawn silica tubing, typically 10–100 m long, 0.2–0.5 mm o.d., with an outer protective coating of polyimide or, for operation above about 350 °C, aluminum. The stationary phase is often silicone oil—for example, polydimethylsiloxane, which instead of being coated on a granular support, is present as a film, 0.1–5 μm thick, chemically bonded to the inner wall of the column (wall-coated, open-tubular or WCOT column). Such open-tubular columns operate at lower carrier gas pressures (typically 34–69 kPa [5–10 psig] instead of 138–340 kPa [20–50 psig] and carrier gas flow rates [1 mL min^{-1} or less instead of 20–30 mL min^{-1}]) than packed columns. Capillary GC columns can typically have 3,000 theoretical plates per meter (50,000 plates per column), compared with 1,000 plates per meter or 2,000 plates per column for a typical packed column. The capacity of the column (i.e., the size of sample that can be separated) depends on the thickness of the film of stationary phase but is of course smaller than for a packed column and is typically in the microgram to nanogram range. Columns with thicker films have higher capacity, but lower resolving power, than thin film columns. Chemical bonding reduces the loss of stationary phase (column bleed), especially during temperature programming, and bonded columns are almost essential for critical applications such as coupled gas chromatography–mass spectrometry (GC–MS). Although low-or medium-polarity general-purpose capillary columns have high performance and can be used with a variety of samples, customized columns are available with stationary phases developed and optimized for particular analyses.

As an alternative to coating the inner wall of the capillary column with liquid stationary phase, the column wall can be coated with finely divided support, which is itself coated with stationary phase (support-coated open-tubular or SCOT column). SCOT columns are one category of the more general group of porous-layer, open-tubular (PLOT) capillary columns where the inner wall of the capillary is coated or bonded with the stationary phase. PLOT columns are available with a range of solid adsorbents, including Porapak, molecular sieve, carbon molecular sieve, and alumina suitable for separation of mixtures of permanent gases and gaseous hydrocarbons, and bring the separating power of capillary columns to GSC. However, the difficulty of reproducibly injecting gas samples into these columns has meant that packed columns are often still favored for separation of gaseous samples.

58.3.2 Carrier Gas

The theory of the influence of carrier gas on the separation process was given by van Deemter, Zuidweg, and Klinkenberg [5]. The van Deemter equation combines rate theory and plate theory

and gives the relation between carrier gas velocity, u, and HETP, H, for a given carrier gas and column. The equation has the form

$$H = A + \frac{B}{u} + Cu \tag{58.1}$$

where A, B, and C are constants. A depends on the particle diameter and irregularity of column packing, B depends on the tortuosity of the channels and the first power of the diffusion coefficient of solute molecules in the gas phase, and C depends on the distribution coefficient of the solute, on the ratio of stationary-phase and gas-phase volumes, on the effective film thickness, and inversely on the diffusion coefficient of the solute in the gas phase.

Plots of H versus u show a minimum value for H corresponding to an optimum carrier gas flow rate, where

$$u_{\text{opt}} = \left(\frac{B}{C}\right)^{1/2} \tag{58.2}$$

The van Deemter equation shows that for a given column, a carrier gas of higher molecular weight can give a more efficient separation (lower value of H_{min}) than one of lower molecular weight, and H_{min} occurs at higher gas velocities for carrier gases of lower molecular weight. However, the equation refers to a single solute, and since chromatography involves separation of several solutes, the optimum carrier gas velocity is necessarily a compromise. The van Deemter equation also shows that for low molecular weight carrier gases, particularly hydrogen and helium, the minimum is less pronounced; that is, that carrier gas flow is less critical to column performance, and these two gases are the preferred choice, especially for capillary chromatography. Where hydrogen or helium is not suitable for the detector in use, a separate "makeup" gas supply is provided at the downstream end of the column; for example, nitrogen makeup is necessary for capillary operation of an electron capture detector (ECD).

The H versus u curve is not symmetrical, carrier gas flow rates being more critical at values below u_{opt} than above. Thus, analyses may be speeded by increasing carrier gas flow rates above the optimum without much deterioration in column performance, but operating at flow rates that are too low leads to a relatively rapid loss of separating power.

58.3.3 Detectors

Detectors for GC should ideally have high sensitivity, rapid and reproducible response, and a wide range of linear response to concentration. Early detectors had universal or near universal response to all solute molecules; more recently, the emphasis has been on the development of detectors with some selectivity to particular groups of compounds.

58.3.3.1 Thermal Conductivity Detector

The thermal conductivity detector (TCD) or katharometer was one of the earliest GC detectors and utilizes the change in thermal conductivity of a gas mixture with composition. The detector consists of either two or four electrically heated filaments or, for highest sensitivity especially at low temperatures, thermistors. The filaments or thermistors are connected in a Wheatstone bridge circuit, with external resistors to complete the bridge if there are only two sensing elements. The filaments or thermistors are mounted in a metal block to provide thermal stability, and provided with channels to allow the effluent from the GC column, and a separate, controlled "reference" flow of pure carrier gas to pass over the sensors or pairs of sensors. The loss of heat from the filaments depends on the filament temperature and on the conductivity of the surrounding gas. The katharometer can be operated under constant current or

constant voltage conditions, or feedback circuitry can be used to maintain the filament resistance constant, but in each case, changes in gas composition lead to an out-of-balance voltage in the Wheatstone bridge circuit.

The TCD is a universal detector; however, it is less sensitive than other detectors such as the FID and is principally used for detection and measurement of permanent gases such as oxygen, argon, nitrogen, carbon monoxide, and carbon dioxide, which either cannot be measured by the FID or require special pretreatment of the effluent gas (see succeeding text). It can be shown that the sensitivity is greatest when the filaments are operated at the maximum possible current and when the difference in conductivity between the carrier gas and sample components is greatest. Both these conditions are fulfilled by use of helium or, better, hydrogen as the carrier gas, as these two gases have higher thermal conductivities than other common gases.

58.3.3.2 Flame Ionization Detector

The flame ionization detector (FID) is one of a group of gas detectors in which changes in the ionization current inside a chamber are measured. The ionization process occurs when a particle of high energy collides with a target particle that is thus ionized. The collision produces positive ions and secondary electrons that can be moved toward electrodes by application of an electric field, giving a measurable current, known as the ionization current, in the external circuit.

The FID utilizes the fact that while a hydrogen–oxygen flame contains relatively few ions, it does contain highly energetic atoms. When trace amounts of organic components are added to the flame, the number of ions increases and a measurable ionization current is produced. The effluent from the GC column is fed into a hydrogen–air flame. The flame jet serves as one electrode, and a second collector electrode is placed above the flame. A potential is applied between the electrodes. When sample molecules enter the flame, ionization occurs—yielding a current that can be amplified and recorded.

The main reaction in the flame is

$$>\mathrm{CH} + \frac{1}{2}\mathrm{O_2} \rightarrow -\mathrm{CHO} + \mathrm{e^-} \tag{58.3}$$

However, the FID also gives a small response to substances not containing hydrogen, such as CCl_4 and CS_2. It is probable that the reaction earlier is preceded by hydrogenation to form CH_4 or CH_3 in the reducing part of the flame. In addition to the ionization reactions, recombination also occurs, and the response of the FID is determined by the net overall ionization process.

The FID is a mass-sensitive detector; that is, the response is proportional to the amount of organic material entering the detector per unit time. For many substances, the response is effectively proportional to the number of carbon atoms present in the flame, and the detector sensitivity can be expressed as the mass of carbon per second required giving a detectable signal. A typical figure is 10^{-11} g C s^{-1}.

The FID responds to practically all organic molecules. It is robust; has high sensitivity, good stability, and a wide range of linear response; and is widely used.

The FID is insensitive to inorganic molecules and water. However, it can be used for measurement of carbon oxides (CO and CO_2) by mixing the effluent from the GC column with a controlled stream of hydrogen. The mixed gas is passed over a heated catalyst to convert the CO or CO_2 to methane (methanation), followed by FID measurement of the methane generated. This allows GC determination of these gases at lower concentrations than can be detected by the TCD.

58.3.3.3 Photoionization Detector

The photoionization detector (PID) has some similarities to the FID, and like the FID, it responds to a wide range of organic and also to some inorganic molecules. An interchangeable sealed lamp produces monochromatic radiation in the ultraviolet (UV) region. Molecules having ionization potentials less

than the energy of the radiation can be ionized on passing through the beam. In practice, molecules with ionization potentials just above the photon energy may also be ionized, due to a proportion being in excited vibrational states. The ions formed are driven to a collector electrode by an electric field, and the ion current is measured by an electrometer amplifier.

The flame in the FID is a high-energy ionization source and produces highly fragmented ions from the molecules detected. The UV lamp in the PID is of lower energy, leading to the predominant formation of molecular ions. The response of the PID is therefore determined mainly by the ionization potential of the molecule, rather than the number of carbon atoms it contains. In addition, the ionization energy in the PID can be selected by choice of the wavelength of the UV source, and the detector can be made selective in its response. Commonly available UV lamps for the PID have energies of 11.7, 10.2, and 9.5 eV. The ionization potentials of N_2, He, CH_3CN, CO, and CO_2 are above the energy of all the lamps, and the PID does not respond to these gases. The 10.2 eV lamp is particularly useful as it allows ionization, and thus detection of alkynes and alkenes (except ethene), but not alkanes.

The PID is highly sensitive, typically to picogram levels or about one order of magnitude more sensitive than an FID, and has a wide linear range. Any of the commonly used carrier gases is suitable, although some gases (e.g., CO_2) absorb UV radiation and their presence may reduce the sensitivity of the detector. The main disadvantage of the detector is the fragility of the UV lamp, the need for periodic cleaning of the UV window, and the difficulty in cleaning the window if the detector becomes heavily contaminated.

58.3.3.4 Electron Capture Detector

The ECD is an ionization chamber in which molecules of electronegative species are allowed to attach to or "capture" electrons that have been slowed to thermal velocities by collision with inert gas molecules. The detector consists of a cell containing an emitting radioactive source (usually ^{63}Ni) and purged with inert gas. Electrons emitted from the source are slowed to thermal velocities (thermalized) by collision with the gas molecules and are eventually collected by a suitable electrode, giving rise to a standing current in the cell. If molecules with greater electron affinity are introduced into the cell, some of the electrons are "captured," forming negative ions that are more massive with less mobile from free electrons, and the current in the cell is reduced. This effect is the basis of the ECD.

Originally, the ECD was operated under dc conditions, potentials up to 5 V being used, but under some conditions, space charge effects produced anomalous results. Modern detectors operate under constant current conditions and use a pulsed supply, typically 25–50 V. The pulse width and/or frequency are varied by feedback circuitry to maintain the ionization current in the cell at a constant level. This extends the linear range of the detector response and ensures optimum response for a range of molecules. The ECD must be used with a suitable inert gas, usually either an argon–methane mixture or nitrogen, either as the chromatograph carrier gas or (more usually) as an auxiliary gas supply.

The ECD is extremely sensitive to electronegative species, particularly halogenated molecules. It is widely used in the analysis of pesticides and some trace atmospheric components such as halocarbons, halogenated solvents, and nitrous oxide. The selectivity and extreme sensitivity is valuable, but the use of a radioactive source is a disadvantage. In certain cases, the detector response is highly sensitive to the cell temperature. The cell may be contaminated by "dirty" samples and cleaning can be difficult or impossible.

58.3.3.5 Flame Photometric Detector

In the flame photometric detector (FPD), the column effluent is passed through a fuel-rich hydrogen–air or hydrogen–oxygen flame, where the sample molecules are broken into simple molecular species and excited to higher electronic states. Under these conditions, most organic and other volatile compounds containing sulfur or phosphorus produce chemiluminescent species. The excited species return to their ground state, emitting characteristic molecular band spectra. The emission is monitored by a photomultiplier through a suitable filter, thus making the detector selective to either sulfur or phosphorus.

The FPD is most commonly used as a detector for sulfur-containing species. In this application, the response is due to the formation of excited S_2 molecules, S_2^*, and their subsequent chemiluminescent emission. The original sulfur-containing molecules are decomposed in the hot inner zone of the flame, and sulfur atoms are formed, which combine to form S_2^*, in the cooler outer zone. As the S_2^* revert to their ground state, they emit light in a series of bands in the range 300–450 nm, with the most intense bands at 384.0 and 394.1 nm. The 394 nm band is monitored.

The FPD is highly sensitive (typically 10^{-11} g S s^{-1} or 10^{-12} g P s^{-1}), selective, and relatively simple. However, the response is nonlinear, given by

$$I_S = I_0[S]^n \tag{58.4}$$

where

I_S is the observed emission intensity (photomultiplier tube output)
$[S]$ is the mass-flow rate of sulfur atoms
n is a constant (value of 1.5–2, depending on flame conditions)

Some systems incorporate circuitry to produce a linear output over 2 or 3 orders of magnitude concentration range.

58.3.3.6 Other Detectors

A number of other GC detectors are available but are less frequently used than those listed earlier. Examples include the nitrogen phosphorus detector, which is selectively sensitive to analytes containing those elements and is used in pesticide analysis, and the helium ionization detector, in which the ionization process is due to highly energetic metastable helium atoms. The helium ionization detector is the only GC detector that permits analysis of the permanent gases at ppb levels. However, it is difficult to use, as the response is highly susceptible to the presence of trace impurities in the helium carrier gas. Additionally, gas chromatographic separation can be combined with detector systems specifically developed to measure a particular analyte. For example, trace levels of carbon monoxide in air can be measured by passing the effluent of the GC column through a heated bed of mercuric oxide. Carbon monoxide reduces the mercuric oxide, liberating mercury vapor, which is detected with high sensitivity by UV atomic absorption spectrometry.

58.3.4 Sample Injection

The purpose of the injection system introduces defined and reproducible aliquots of sample into the chromatograph column. To minimize loss of resolution by diffusion, the sample must be injected as quickly as possible and as a sharply defined slug. For analysis of liquids using packed columns, the injector is usually a zone heated to a temperature to ensure rapid volatilization of the sample, swept by carrier gas and fitted with a silicone rubber septum cap. Liquid injections are made by microsyringe through the septum cap. The syringe can be manual or controlled by an autosampler. Gas injections can also be made by syringe, but this is unsatisfactory for quantitative work as the compressive effect of the column head pressure in the injector makes syringe injections of gas unreproducible, and valve injection is preferable.

The injection valve is a six-port changeover valve that allows a fixed volume of gas, defined by a length of tubing (the sample loop), to be connected in either one of two gas streams with only momentary interruption of either stream. The valve is connected in the carrier gas stream just upstream of the column. The sample loop is filled with sample gas and the valve is operated to connect the loop into the carrier gas stream.

Injection into capillary columns is more difficult than into packed columns due to their smaller sample capacity. An internal standard (i.e., a compound similar to the analyte but separated from it by the

chromatograph) is often added to the sample to enable correction to be made for random differences in injection volume. The most widely used injection technique is the so-called split–splitless injector. In the "split" mode, an aliquot of liquid sample is injected into a heated zone at the head of the column. The injector is swept by carrier gas at constant pressure. A valve allows a variable but known proportion of the carrier gas to flow to waste. Carrier gas flow rates through the column are typically 1–2 mL min^{-1} and the "split" flow is typically 30–50 mL min^{-1}. Thus, only a few percent of the aliquot of sample injected actually passes through the column. In the "splitless" mode, the valve controlling the waste stream of carrier gas is momentarily closed for a fixed time after injection, thus increasing the amount of sample transferred to the column for increased sensitivity.

Split/splitless injectors are relatively simple and can be used with conventional syringes. However, the injector must be carefully designed to obtain reproducible results, and there may be some discrimination (particularly loss of high-boiling components) in samples containing components with a wide range of boiling points. Techniques where the sample is injected directly onto the column, without splitting, can give less discrimination and better sensitivity and reproducibility but are more difficult to use. Examples include cold on-column or programmed temperature injectors. In this technique, a length of column at the head is cooled to trap the sample. The injector is subsequently flash-heated to a high temperature to release the sample.

Alternative injection techniques include headspace analysis, where the vapor in equilibrium with a volatile liquid is sampled, and purge-and-trap techniques, where volatile components are purged from a liquid sample by a stream of inert gas. The purged components are trapped in a cooled zone at the head of the GC column and subsequently released by rapid heating for chromatographic separation and determination.

58.4 Liquid Chromatography

In early applications of LC, the adsorbent was contained in a vertical column through which the liquid phase was passed under gravity. The performance of such systems is limited by the tortuosity of the passage of the liquid phase through the column and by the efficiency and speed of solute exchange between the mobile and stationary phases. Column performance is enhanced by use of microparticulate packings, but to achieve reasonably rapid separations in such systems, the mobile phase must be pumped through the column under pressure. The technique is known as high-performance (or sometimes high-pressure) liquid chromatography (HPLC).

58.4.1 High-Performance Liquid Chromatography

Figure 58.2 provides an illustration of process HPLC.

58.4.1.1 Columns

HPLC columns are formed from precision-bore stainless steel tubing, 30–3 cm long, with 25 and 12.5 cm being the most common lengths. Standard columns have bores in the range 3.0–4.6 mm. Narrow-bore (2 mm) or microbore (1 mm) columns give increased sensitivity and higher chromatographic performance, together with reduced consumption of mobile phase, but require higher operating pressures and specially designed detector systems. The columns are packed under controlled conditions with granular packing material of controlled size range, particle shape, and porosity. The "standard" particle diameter for packing materials is 5 μm, as it allows good column performance at moderate operating pressures, but columns packed with 10 or 3–4.5 μm particles can also be used.

A wide range of general-purpose and specialist packing materials have been developed for HPLC. The most widely used base material is silica, as it has high efficiency and physical rigidity and good solvent compatibility and can be bonded with organosilanes for reversed-phase chromatography (see succeeding text). Other base materials include ceramics such as alumina, polymers, and graphitic carbon.

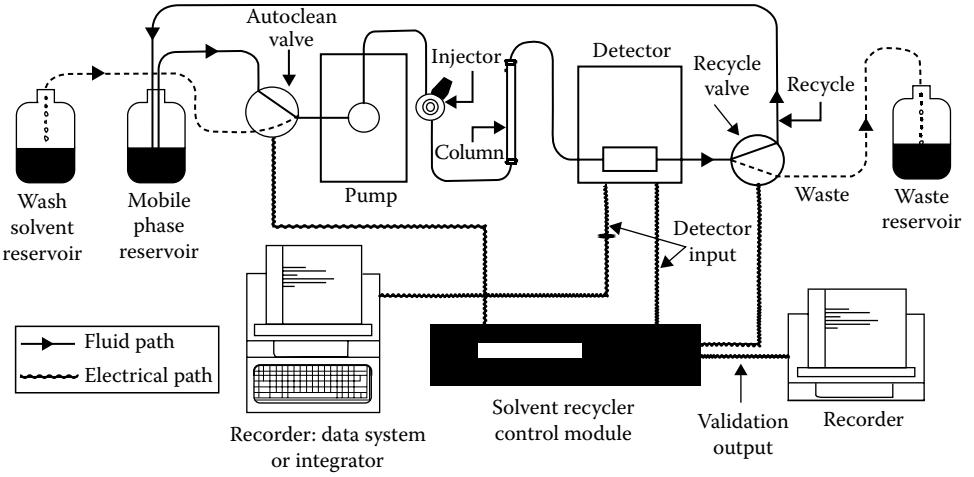


FIGURE 58.2 Functional diagram of process high-performance chromatography. The system consists of pump, sample injector, separation column, and detector. Mobile phase is liquid and wash solvent is used for cleanup. The used solvent can be recovered.

Alumina has better pH stability than silica, but cannot be bonded with organosilanes. Polymers have limited organic solvent compatibility but tolerate strong alkali; they are less robust than silica and cannot be used at high operating pressures.

In liquid–solid chromatography, the samples are retained by adsorption on the support surface. The support may be coated with a liquid phase (liquid–LC). In bonded-phase HPLC, the support (usually silica, although alumina has also been used) is derivatized with a functional group covalently attached to the surface. In both coated and bonded phases, the separation is predominantly by partition, but like their counterparts in GC, bonded phases are inherently more stable than coated phases and can be used with a range of solvent and buffer systems.

In "normal" or traditional LC, the stationary phase is polar and the mobile phase relatively nonpolar. The polar surface may be silica or may be modified by chemical bonding of a suitable functional group. In reversed-phase chromatography, the stationary phase is nonpolar and the mobile phase is relatively polar. Reversed-phase chromatography is almost always carried out on modified silica columns, with octadecylsilyl (ODS or C18) or C8 groups being most commonly used, and the development of these systems has been largely responsible for the present popularity of HPLC.

58.4.1.2 Mobile Phase

In contrast to GC, the mobile phase in HPLC plays a vital role in the separation process, with the rate and order of elution being determined by the relative polarities of the mobile and stationary phases and by the nature of the sample components. In normal-phase chromatography, the eluting power of the mobile phase increases with polarity, while in reversed-phase chromatography, the reverse situation applies, and eluting power decreases with increasing solvent polarity. Most of the common solvents can be used as mobile phase, but *n*-hexane is a common nonpolar solvent, while polar solvents in reversed-phase chromatography are often mixtures of water and methanol or acetonitrile. When a single solvent or mixture of fixed composition is used throughout a separation, the process is referred to as isocratic. In gradient elution procedures, the composition of the solvent is changed continuously during the separation process. The change in solvent composition can be either linear with time or according to a predetermined profile. Gradient elution has some similarities in its effects to temperature programming in GC.

In HPLC, the mobile phase is pumped through the column. Flow rates for mobile phase are 0.5–10 mL min^{-1} for conventional columns and 50 μL–5 mL min^{-1} for narrow-bore columns at

pressures up to 48,263 kPa (7000 psi). The flow rate must be precisely controlled and the outlet stream must be pulse-free. Metered mixing of up to four solvent streams may be required for gradient elution and column washing. HPLC pumps must be capable of delivering a pulse-free stream of mobile phase, at pressures up to 48,263 kPa (7000 psi). Variations in mobile-phase flow rate lead to irreproducibility in chromatographic retention times, and pulses in the flow give noise in the detector baseline.

To ensure optimum pump performance and to remove dissolved oxygen, the mobile phase must be degassed. Degassing can be by vacuum, but periodic purging with helium is more usual. Dissolved oxygen causes noise and drift in UV detectors, quenching in fluorescence detectors, and high background currents and noise in electrochemical detectors.

58.4.1.3 Detectors

Performance requirements for HPLC detectors are similar to those in GC, and the detector system should be sensitive, have a rapid and reproducible response, and have a wide linear range of response. The following types are most commonly used:

1. *Refractive index*: Detectors based on measurement of the refractive index of the mobile phase are applicable to a wide range of solutes. However, they are temperature sensitive, are generally less sensitive than UV or fluorescence detectors, and cannot easily be used with gradient elution systems.
2. *UV–VIS*: UV or visible detection is the most common detection technique. Photometers or spectrophotometers measure the absorption of UV or visible radiation in the range 190–700 nm by the solute molecules. Detectors can be either fixed or variable wavelength. Fixed-wavelength detection at 254 nm is suitable for many solutes. The response is linear according to the Beer–Lambert law and detection limits are subnanogram in favorable cases. "HPLC grade" solvents that have been specially purified to remove UV-absorbing impurities may be required.
3. *Diode array*: The diode array, or photodiode array (PDA), detector allows the UV or UV–VIS spectrum of the mobile phase to be repeatedly scanned during the elution. The resulting spectra are stored by the data system and are useful for checking or monitoring peak purity during elution and may provide some information on the identity of the analyte. Diode array detectors are typically two or three times less sensitive than conventional, single-wavelength UV or visible detectors.
4. *Fluorescence*: Detection of suitable molecules by fluorescence is typically one or two orders of magnitude more sensitive than UV–VIS detection, allowing detection down to low picogram levels in favorable cases. The fluorescence cell volume can be made as small as 5 μL, making fluorescence detection suitable for use with narrow-bore columns.
5. *Electrochemical*: Electroactive (oxidizable or reducible) substances can be detected by electrochemical techniques. The flow-cell volume may be as low as 1 μL, making the detectors suitable for narrow-bore columns. By choice of working conditions, the detector can be made specific to particular compounds or groups of compounds. Examples of suitable compounds for electrochemical detection include aromatic amines, phenols, and chlorinated phenols.

58.4.1.4 Injectors

The injection process is critical in obtaining good performance in HPLC. Injection may be made directly by syringe, but for optimum results, valve injection is essential and is now almost universal. The injection valve is a six-port changeover valve, which allows a defined volume of liquid sample, contained in a sample loop, to be injected into the mobile phase. The sample loop typically has a volume in the range of 10–100 μL for conventional HPLC columns and is loaded by a syringe.

58.4.2 Ion Chromatography

Ion chromatography (IC) is a variant of HPLC in which inorganic and some organic cations and anions are separated on columns packed with high-efficiency pellicular ion-exchange resins. Anion separator

columns are resin based with positively charged fixed ionic sites, usually quaternary amines. The eluent is commonly dilute aqueous sodium hydroxide or sodium carbonate/sodium bicarbonate mixture. Cation separator columns have negatively charged fixed ionic sites, usually sulfonic acid groups, with methane sulfonic acid as eluent. Although conductimetric, amperometric (polarographic), UV–VIS photometric, and fluorescence detectors can all be used in IC, conductimetric detection is the most commonly used technique. When conductimetric detection is used, the sensitivity of the measurement is increased by reduction of the conductivity of the mobile phase (suppression) before detection. The conductivity is suppressed by conversion of the eluent to the corresponding acid or base form in a suppressor column, or membrane suppressor, downstream of the separation column. Thus, for anion analysis, the suppressor is a cation exchanger that replaces sodium ions in the eluent with hydrogen ions. IC allows detection and measurement of anions and cations in solution, typically down to ppb ($\mu g\ L^{-1}$) levels. The common inorganic anions can be determined in a single sample aliquot. The technique is particularly useful for routine analysis of water and environmental samples.

58.4.3 Gel Permeation and Size-Exclusion Chromatography

Gel permeation chromatography (GPC) and size-exclusion chromatography (SEC) separate sample molecules on the basis of their effective molecular size in solution in aqueous or nonaqueous media. Column packings are porous materials with pores in controlled size ranges and can be either resin (e.g., styrene–divinylbenzene copolymers) or silica based. The solute molecules interact with the column packing. Small molecules are "trapped" in the pores and pass through the column more slowly than larger molecules that do not interact so strongly. GPC was developed for polymer chemistry and is useful for the determination of the molecular weight distribution of polymers, but the techniques are also finding other applications, for example, size separation of small organics and petrochemicals and in sequential analysis and sample cleanup of environmental samples.

58.5 Hyphenated Techniques

Chromatographic methods give powerful techniques for separation of mixtures. However, the commonly used detectors give little information about the identity of the separated components, and identification from retention times may be incomplete or ambiguous. The so-called hyphenated techniques, in which chromatographic separation is combined with another, usually spectrometric, technique, have been developed to combine chromatographic separating power with spectrometric identification.

58.5.1 Gas Chromatography–Mass Spectrometry

Developments in column, mass spectrometer, and computer technology have made GC–MS the most widely used of the hyphenated techniques. Modern systems commonly incorporate the following features:

1. Capillary column gas chromatograph, with the column effluent fed directly to the ion source of the mass spectrometer through a heated transfer line.
2. Miniaturized quadrupole or quadrupole-type mass spectrometer, optimized for use as a GC detector and limited to this application. The mass spectrometer may have a limited mass range (maximum mass 650 or 700), corresponding to the maximum molecular weight of compounds that can commonly be analyzed by GC.
3. Control of the mass spectrometer and gas chromatograph, data recording, and mass spectral library searching by dedicated computer.

Such so-called "benchtop" GC–MS systems typically have sensitivities in the nanogram to picogram range and unit mass spectral resolution. They give a valuable means of identification and quantitative

determination of a variety of analytes, particularly in complex matrices such as biological or environmental samples. However, for applications requiring the highest sensitivity or mass spectrometric resolution, gas chromatographs coupled to high-resolution magnetic sector mass spectrometers must still be used. If it is necessary to use a packed chromatograph column, the carrier gas flow rate is likely to be too high for direct coupling to the mass spectrometer, and an interface, which allows selective removal of carrier gas, will be required.

58.5.2 Gas Chromatography–Infrared Spectrometry

Coupling a gas chromatograph with an infrared (GC–IR) spectrometer is an alternative to GC–MS. The effluent from the GC column is fed to a miniaturized flow-through absorption cell (light pipe) where the IR spectra are measured. Fourier transform infrared (FTIR) spectrometry is used to achieve the combination of high spectral scan rates and resolution required to measure spectra from peaks eluting from capillary columns. Computer control and data processing are used, together with computer matching with stored libraries of IR spectra. The IR spectra are vapor phase and differ in some respects from liquid- or solid-phase spectra. GC–IR is to some extent complementary to GC–MS in that some molecular properties, particularly those relating to overall structure or shape, which may be lost in the fragmentation process in the mass spectrometer, may be identified in the IR spectra. However, the technique is some three or more orders of magnitude less sensitive than GC–MS and has received less attention. GC–IR–MS can be used for particularly complex mixtures.

58.5.3 Liquid Chromatography–Mass Spectrometry

Coupling a mass spectrometer to a liquid chromatograph, in principle, offers the advantages of GC–MS to the greater range of materials that can be analyzed by liquid, compared to gas, chromatography. However, interfacing a liquid chromatograph to a mass spectrometer is more difficult than a gas chromatograph, and while GC–MS is a well-established technique, liquid chromatography–mass spectrometry (LC–MS) is only now becoming a useful routine method.

The principal difficulty in interfacing LC and MS lies in the very different physical conditions required for operation of the two techniques. LC uses relatively large quantities of liquid mobile phase, which may include inorganic buffers, while the mass spectrometer operates under vacuum. An interface must be used to selectively remove mobile phase before sample can be introduced to the mass spectrometer. Direct liquid injection and moving belt interfaces have been used but are unreliable. Three main types of interface are currently in use:

1. *Thermospray*. The LC effluent is passed through a probe and heated to 350 °C–400 °C, in an evacuated region just outside the source of the mass spectrometer. The mobile phase, which often includes a volatile buffer such as ammonium ethanoate, is vaporized, and the sample molecules are ionized by a chemical ionization (CI) process. A series of lenses focuses a proportion of the ions into the mass spectrometer, while the solvent is pumped away. Under these conditions, the ionization is soft, that is, there is little fragmentation in the mass spectrometer and the principal peak in the mass spectra is the molecular ion, M^+, MH^+, or MNH_4^+. This can be useful for molecular weight determination, but fragmentation may be increased and, in some cases, sensitivity enhanced, by including a filament in the ion source to give electron impact (EI) mass spectra. The thermospray interface is suitable for ionic and nonvolatile compounds. However, it requires the use of volatile buffers and the spectra are dependent on the solvent matrix.
2. *Particle beam interface*. The particle beam interface utilizes the principle of momentum to separate the solvent from the heavier solute molecules. The column effluent, mixed with helium in some designs, is passed through a series of chambers under pressure, exiting through a nozzle. As the effluent emerges, the solvent is vented, while the solute molecules continue on their original

trajectory and pass into the ion source of the mass spectrometer. The interface allows the use of a standard ion source in either El or CI modes. It can therefore be used to produce standard library-searchable mass spectra and is the only commonly used interface where this facility is routinely available. The interface requires some sample volatility, and some compounds (such as complex sugars) do not give satisfactory spectra.

3. *Electrospray.* In the electrospray interface, the column effluent is mixed with a nebulizing gas and passed through a jet nebulizer into a high-voltage electric field. Drop formation and ionization, by a CI process, occur. The ions enter the mass spectrometer through a capillary tube charged to a different voltage from the remainder of the interface, while the solvent is pumped away. Electrospray interfaces have, thus far, mainly been used for very high molecular weight analytes, although systems for lower molecular weights are being developed. Atmospheric pressure chemical ionization (APCI) is somewhat similar to electrospray except that the ionization process takes place at atmospheric pressure. Both systems can be used for polar compounds, molecular weights up to 100,000 Da, and are highly sensitive. However, ionization and separation involve a complex series of mechanisms, and setup and operation of the system, and interpretation of the spectra produced, may be difficult.

58.6 Applications in the Electricity Industry

58.6.1 Dissolved Gas Analysis

GC analysis of gases dissolved in transformer oil has been used for condition monitoring since the early 1970s [6–8]. The large volumes of gas often generated during a transformer fault have been used to trip mechanical relay for some 60 years [9]. It was later realized that if gases are evolved from the oil in sufficient quantities to operate a Buchholz relay, then slowly developing faults would also produce decomposition gases that would be dissolved in the oil. They only appear in the Buchholz at the end of a complicated system of interchange between the gases contained in bubbles rising to the surface and the less soluble atmospheric gases dissolved in the oil. It should therefore be possible to detect any incipient faults that may be present in the transformer early by analysis of the gases dissolved in the oil, using a gas chromatograph. Thermal and electric faults in a transformer produce various characteristic gases that are, to some extent, soluble in the oil. Extraction and GC analysis of dissolved gases can be used for monitoring of transformer condition. Dissolved gas analysis (DGA) has been accepted as an important and vital condition monitoring technique for power transformers [10–12].

Oil samples can be collected from the equipment using syringes, bottles, or other sampling techniques, as described in IEC 567 [13]. The analysis requires extraction of the dissolved gases from the oil and then injection into a GC. The details of extraction of the gases from the oil are given in IEC 567. Generally, gases are extracted under vacuum using a mercury Toepler pump and the total volume of the extracted gas is measured by bringing to atmospheric pressure. The gases are then separated and determined by GC. An automated mercury-free instrument can also be used for extraction of dissolved gases [14] and other techniques have been used for extraction of the gases from the oil [15]. A static headspace sampling technique has been combined with capillary GC to allow dissolved gases and furan-related compounds to be determined in power transformer oils in a single GC run [16].

Regardless of the technique used for the extraction of the gases from the oil, GC is used for analysis of the gases. Alternative techniques such as MS, although very sensitive, have not been used for routine analysis. IR spectrometry has been used for the detection of gaseous hydrocarbons and carbon oxides [17]. The technique is rapid and accurate, and the detector is very stable. However, it cannot detect hydrogen and atmospheric gases. Hydrogen is a very important incipient gas for transformer condition monitoring. The ratio of oxygen to nitrogen dissolved in the oil is also used as an indication of oil or paper degradation in the transformer (the oxygen is used by reaction with the cellulose).

The gas chromatograph used for analysis of the gases is usually dual channel with FID and TCD detectors. A Porapak column is used for separation of hydrocarbons, and a methanizer is used for converting carbon oxides to methane followed by FID detector, while hydrogen, oxygen, and nitrogen are separated on a molecular sieve column and measured by TCD. Other arrangements such as column switching–backflushing should be used if a single detector is going to be used. In such cases, a TCD is usually used as a detector. IR detection of hydrocarbons followed by TCD is an alternative arrangement. A combination of Porapak and molecular sieve column is used for separation of hydrogen, oxygen, and nitrogen, followed by TCD detection. This arrangement requires a flash backflush system to prevent carbon dioxide from entering the molecular sieve column—where it would be so strongly adsorbed that it would require prolonged heating at a high temperature to remove it. The system is capable of detecting 1 ppm hydrocarbons and carbon oxides and 5 ppm hydrogen in the oil. Oxygen and nitrogen in the oil are usually present at high concentration, and therefore, their detection does not present any problem. High concentrations of acetylene gas in the oil may present some problems, such as poisoning the methanizer catalyst, and it may stay in the column for a long time. In such cases, a longer isothermal time and higher oven temperature for cleaning of the column is the recommended technique.

Water has been determined in transformer oils with accuracy better than 3%, precision better than 4% at the 10 ppm level, and detection limit of 0.3 ppm, by headspace sampling and capillary chromatography with TCD detection. The technique could be automated using the headspace GC system proposed for DGA [18].

58.6.2 Furfuraldehyde Analysis

Under normal operating conditions, the insulation system of transformers gradually deteriorates and produces various degradation by-products. Thermal degradation, or aging of the paper insulation, is one of the most important factors in limiting the lifetime of a transformer. Detection and analysis of the degradation by-products have been widely used to evaluate and monitor the degradation state of the insulation. The aging process of the paper is accompanied by the production of several by-products—mainly carbon monoxide, carbon dioxide, and furfurals. Carbon oxides can be monitored by DGA, but their production is not specific to paper degradation. The measurement of furfurals could provide an early indication of paper degradation, and their analysis by HPLC has been used as a tool for the monitoring of transformer performance [19–22]. A spectrophotometric method has also been used for analysis of furfuraldehyde [23]. The method is only capable of measuring furfuraldehyde and not other furanic compounds present in the oil.

The oil is dissolved in cyclohexane for HPLC analysis and passed through a solid-phase silica cartridge, where the furfurals, phenol, and *m*-cresol (which are products of degradation of phenol-formaldehyde resins in the transformer) are retained. The remaining oil in the cartridge is removed by washing with the solvent. The furfurals, phenol, and *m*-cresol are extracted from the cartridge with water/acetonitrile. The collected extracts are analyzed by HPLC. For separation of furfuraldehyde and other compounds, a C18 column is recommended. A UV detector at 276 nm is usually employed for detecting these compounds. The use of a UV PDA detector gives improved discrimination of products in what is often a complex chromatogram at the expense of some slight loss of sensitivity.

58.6.3 Analysis of Antioxidants in Oil

The presence of antioxidants is a key factor in controlling the oxidation of insulating oil. Their use results in substantial savings by prolonging the oil service life and slowing down the transformer aging process. A large number of antioxidant additives are used, and HPLC is a useful technique for quantitative determination of their concentrations in insulating oil [24,25]. TLC has also been used for quantitative determination of antioxidant in insulating oil [26].

HPLC is also used for evaluation of the quality of mineral insulating oil [27,28]. The presence of certain characteristic chemical compounds is of considerable importance in the electric industry. This technique is also used to identify the presence of some by-products of oil under electric stress, such as x-wax and fluorescence materials. Polar compounds, such as acids and aldehydes, are also products of oxidation of oil and they can also be determined by HPLC. Early detection of such products is important in transformer condition monitoring to provide prior warning of a developing fault and to enable appropriate corrective action to be taken.

HPLC is also used for health and safety monitoring of the polyaromatic hydrocarbon (PAH) content of transformer oils [29]. The toxic and carcinogenic nature of PAH compounds is well established.

58.6.4 Molecular Weight Distribution of Insulating Paper

GPC has been used to measure the change in molecular weight distribution of insulating paper during aging [30]. This technique has been used for measurement of the degree of polymerization of the paper but requires a sampling of the paper in the transformer, which is not generally practical. Measurement of paper degradation by analysis of products dissolved in the oil, such as furfuraldehyde analysis (FFA), although much easier, is indirect and therefore dependent on a knowledge of the history of the transformer and its components. GPC can provide direct information of the state of the paper and the average molecular weight of the cellulosic chains and currently is mostly used as a forensic analysis tool for the investigation of failures.

Two techniques are available for getting the paper into solution for analysis. Direct dissolution into dimethylacetamide/8% lithium chloride is possible for most forms of cellulose, followed by dilution to 1% lithium chloride for analysis. However, high levels of lignin (3%–4%) in the paper interfere with the dissolution process. The alternative is to derivatize the cellulose to the tricarbanilate in pyridine. The product can then be analyzed in solution in tetrahydrofuran, but there is evidence in the literature that the derivatization process itself degrades the paper.

58.6.5 Polychlorinated Biphenyls

Polychlorinated biphenyls (PCBs) are synthetic materials with dielectric and chemical properties that in the past made them attractive alternatives to petroleum-based products. However, they have since been identified as environmental pollutants and possible health hazards. PCBs are not used today, but their use in the past has led to widespread contamination of mineral oil with PCBs. Current legislation requires that oil in service shall have a PCB content of less than 50 ppm. Therefore, the PCB content of insulating oil in the transformer must be measured quantitatively and capillary GC is the technique currently used [31].

The sample is diluted in hexane and deoxygenated. A small volume of the resulting solution is injected into a narrow-bore capillary gas chromatographic column. The capillary column separates the PCBs into individual or small groups of overlapping congeners. Their presence in the effluent is measured by an ECD.

58.6.6 Feedwater and Boiler Water Analysis

IC is well suited to the analysis of highly pure water such as boiler feedwater and often gives better precision, sensitivity, and speed of analysis than established techniques. For example, IC has been used for the determination of carbonic acid in steam–condensate cycles [32].

Morpholine is added to the thermal cycle of some CANDU reactors and has been determined, together with its amine breakdown products, by reversed-phase HPLC on a C18 column with visible detection at 456 nm [33].

58.6.7 Other Applications

The size and scope of the electricity supply industry imply that it has a major effect on the environment. Chromatographic techniques are widely applied in environmental monitoring and research, and the electricity industry is both a major user of standard techniques and sponsor of research.

Defining Terms

Adsorption: The noncovalent attachment of one substance to the surface of another.

Analyte: The substance that is being analyzed.

Baseline resolution: Separation of components at the peak base (no overlap of any peak area).

Chromatography: The physicochemical technique for separation of mixtures into their components.

Column: A steel, glass, or plastic tube containing the stationary phase.

Detector: A device for monitoring the separated compounds from the chromatography by sensing chemical or physical properties of the sample.

Eluent: The moving solvent in a chromatographic column.

Elute: To travel through and emerge from the column.

Gel permeation chromatography (GPC): A mode of LC in which samples are separated according to molecular size.

HPLC: High-performance liquid chromatography.

IC: Ion chromatography.

Mobile phase: The following solvent.

Stationary phase: The material that is contained in the column and does not move during the chromatographic process.

Acknowledgments

The authors would like to thank PerkinElmer and Phenomenex for permission to reproduce diagrams.

References

1. C.F. Poole and S.A. Schuette, *Contemporary Practice of Chromatography*, Elsevier Science BV, Amsterdam, the Netherlands, 1984.
2. R.L. Grob (ed.), *Modern Practice of Gas Chromatography*, 3rd edn., John Wiley & Sons, New York, 1995.
3. G. Guiochon and C.L. Guillemin, Quantitative gas chromatography for laboratory analyses and on-line process control, *Journal of Chromatography Library*, Vol. 42, Elsevier Science, Amsterdam, the Netherlands, 1988.
4. J. Weiss, *Ion Chromatography*, 2nd edn., VCH Publishers, Cambridge, U.K., 1994.
5. J.J. van Deemter, F.J. Zuiderweg, and A. Klinkenberg, Longitudinal diffusion and resistance to mass transfer as causes of non-ideality in chromatography, *Chem. Eng. Sci*., 5, 271, 1956.
6. P.S. Pugh and H.H. Wagner, Detection of incipient faults by gas analysis, *TAIEE*, 80, 189–195, 1961.
7. L.C. Aicher and J.P. Vora, Gas analysis—A transformer diagnostic tool, *Allis Chalmers Electrical Rev*., 28, 22–24, 1963.
8. E. Dornenburg and W. Strittmatter, Monitoring oil-cooled transformers by gas analysis, *Brown Boveri Rev*., 61(5), 238, 1970.
9. M. Buchholz, The Buchholz protection system and its application in practice, *E. T.Z*., 49, 1257–1260, 1928.

10. R. Muller, K. Potthoff, and K. Soldner, The analysis of gases dissolved in the oil as a means of monitoring transformers and detecting incipient faults, *CIGRE*, Paper 12–02, Paris, France, 11 pp., 1970.
11. E. Dornenburg and W. Strittmatter, Monitoring oil-cooled transformers by gas analysis, *Brown Boveri Review*, 61, 238–247, 1974.
12. S.R. Barbosa, O.M. Almeida,, A.P.S Braga, C.M. Tavares, M.A.B. Amora, F.A.P. Aragoa, P.R.O. Braga, and S. Sergio, Artificial neural network application in estimation of dissolved gases in insulating mineral oil from physical-chemical data for incipient fault diagnosis, *5th International Conference on Intelligent System Applications to Power Systems, ISAP 09*, Curtibia, Brazil, 5 pp., 2009.
13. BS EN 60567:2011. Oil-filled electrical equipment. Sampling of gases and analysis of free and dissolved gases. Guidance 2012, http://shop.bsigroup.com/
14. B. Pahlavanpour, Mercury-free extraction of dissolved gases in transformer oil, *AVO International Technical Conference*, Dallas, TX, 1996.
15. J. Jalbert and R. Gilbert, Comparison between headspace and vacuum gas extraction techniques for the gas chromotographic determination of dissolved gases from transformer insulating oil, *IEEE International Conference on Electric Insulation*, Pittsburgh, PA, pp. 123–129, 1994.
16. Y. Leblanc, R. Gilbert, J. Jalbert, M. Duval, and J. Hubert, Determination of dissolved gases and furan-related compounds in a single chromatographic run by headspace/capillary gas chromatography, *Journal of Chromatography A*, 657, 111–118, 1993.
17. J. Jalbert, S. Charbonneau, and R. Gilbert, A new analytical method for the determination of moisture in transformer oil samples, *Sixty Third Annual Conference Doble Clients*, Boston, MA, March 25–29, 1996.
18. B. Pahlavanpour, S. McCann, and M. Ali, Preliminary investigation into the use of Fourier transform infrared analysis of continuous monitoring of dissolved gases in transformer oil, *National Grid Technical Report*, TR(T)38, 1993.
19. P.J. Burton, J. Graham, A.C. Hall, J.A. Laver, and A.J. Oliver, Recent developments by CEGB to improve the prediction and monitoring of transformer performance, *CIGRE*, paper 12–09, Paris, France, pp. 12.9/1-10, 1984.
20. X. Chendong, Monitoring paper insulation aging by measuring furfural contents in oil, *Proceedings of 7th International Symposium on High Voltage Engineering*, Dresden, Germany, Vol. 7 p. 139–142, 1991.
21. M. Duval and A. DePablo, Interpretation of gas-in-oil analysis using new IEC publication 60599 in the IEC TC 10 database, *IEEE Electrical Insulation Magazine*, 17, 31–41, 2001.
22. J. Unsworth and F. Mitchell, Degradation of electrical insulating paper monitoring with high performance liquid chromatography, *IEEE Trans. Elec. Insul.*, 25, 737, 1990.
23. B. Pahlavanpour and G. Duffy, Development of a rapid spectrophotometry method for analysis of furfuraldehyde in transformer oil as an indication of paper aging, *CEIDP Annual Report*, Pocono Manor, PA, pp. 493–498, 1993.
24. C. Lamarre and A. Gendron, Analyse quantitative par chromatographie liquide haute performance du di-tert-butyl-2,6-para-cresol dans les huiles de transformateurs, *J. Chromatogr.*, 464, 448–452, 1989.
25. M. Duval, S. Lamotte, D. Gauchon, C. Lamarre, and Y. Giguette, Determination of low-improved additives in new and aged insulating oil by gel permeation chromatography, *J. Chromatogr.*, 244, 169–173, 1982.
26. IEC 60666 ed2.0 (2010-04), TC/SC 10, Detection and determination of specified additives in mineral insulating oils, http://webstore.iec.ch/
27. M. Duval and C. Lamarre, The characterization of electrical insulating oils by high-performance liquid chromatography, *IEEE Trans. Electr. Insul.*, E12, 340–348, 1977.
28. M. Duval, C. Lamarre, and Y. Giguere, Reversed-phase high-performance liquid chromatographie analysis of polar oxidation products in transformer oils, *J. Chromatogr.*, 284, 237–280, 1984.

29. A.N. Gachanja, Analysis of polycyclic aromatic hydrocarbons by liquid chromatography, *Chromatography Analysis*, February–March, 5–7, 1993.
30. D.J. Hill, T.T. Le, M. Darveniza, and T. Saha, A study of degradation of cellulosic insulation materials in a power transformers. Part 1. Molecular weight study of cellulose insulation paper, *Polym. Degrad. Stab*, 48, 79–87, 1995.
31. P. Wiklund, A. Biverstal, B. Pahlavanpour, and J. Nunes, Study of formation of low molecular acids from aging of cellulose and insulating oil, *44th International Conference on Large High Voltage Electric Systems*, Paris, France, 8 pp., 2012. Analysis polychlorinated biphenyls in transformer oils by gas chromatography, *National Grid Publication*, RDC/0045/R91, 1991.
32. S. Charbonneau, R. Gilbert, and L. Lepine, Determination of carbonic acid in steam-condensate cycle samples using non-suppressed ion chromatography, *Anal. Chem*., 67, 1204–1209, 1995.
33. C. Lamarre, R. Gilbert, and A. Gendron, Liquid chromatographic determination of morpholine and its thermal breakdown products in steam-water cycles at nuclear power plants, *J. Chromatogr*., 467, 249–258, 1989.

59

pH Measurement

Norman F. Sheppard, Jr.
Gamera Bioscience Corporation

Anthony Guiseppi-Elie
Virginia Commonwealth University and Abtech Scientific, Inc.

The measurement of pH is arguably the most widely performed test in the chemical laboratory, reflecting the importance of water as a ubiquitous solvent and reactant. In the 90 years since the first use of an electrode to determine hydrogen ion concentration, the glass electrode and its variants have matured into routine tools of analytical and process chemists. Yet, there continue to be developments that promise to broaden the scope and reach of these measurements. Among recent developments are miniature pH-sensitive field-effect transistors (pHFETs) being incorporated into pocket-sized pH "pens," metal/metal oxide pH sensors for measurements at high temperatures and pressures, and flexible fiber-optic pH sensors for measuring pH within the body. This chapter discusses electrochemical and optical methods for pH measurement and is by necessity limited in scope; readers interested in a more comprehensive treatment may wish to refer to the recent text by Galster [1].

59.1 Definition of pH

In its most common interpretation, pH is used to specify the degree of acidity or basicity of an aqueous solution. Historically, pH was first defined as the negative logarithm of the hydrogen ion concentration (*Pondus hydrogenii*, literally hydrogen exponent), to simplify the handling of the very small concentrations (on the order of 10^{-7} mol L^{-1}) encountered most commonly in nature. This definition of pH is expressed as Equation 59.1, where $[H^+]$ is the molar concentration of solvated protons in units of moles per liter:

$$pH = -\log[H^+] \tag{59.1}$$

As a greater understanding of the behavior of ionic solutes in solution developed, chemists recognized that the measurement techniques used to determine hydrogen ion concentration were in fact

TABLE 59.1 pH of Common Substances

Substance	pH
Soft drinks	2.0–4.0
Lemon juice	2.3
Vinegar	2.4–3.4
Wine	2.8–3.8
Beer	4.0–5.0
Cow's milk	6.3–6.6
Pure water	7.0
Blood	7.3–7.5
Seawater	8.3

Source: Lide D.R. (Ed.), *CRC Handbook of Chemistry and Physics*, 74th edn., CRC Press, Boca Raton, FL, 1993.

measuring the hydrogen ion activity, often referred to as the "effective concentration." This led to the adoption of the more rigorous definition of pH as the negative logarithm of the hydrogen ion activity in solution:

$$\mathrm{pH} = -\log a_{\mathrm{H}^+} = -\log \gamma[\mathrm{H}^+] \tag{59.2}$$

where

a_{H^+} is the hydrogen ion activity
γ is the activity coefficient

The two definitions expressed in Equations 59.1 and 59.2 are equivalent in dilute solution where concentration approximates activity, a_{H^+}. In practice, the routine measurement of pH is not accomplished by the direct determination of the hydrogen ion activity. Rather, pH is determined relative to one or more standard solutions of known pH.

The hydrogen ion activity of common substances ranges over many orders of magnitude, as evidenced from the tabulation presented in Table 59.1. Water, in the absence of other chemicals that might alter its equilibrium, contains 10^{-7} molar hydronium ions [H^+] and 10^{-7} molar hydroxonium ions [OH^-] at 25 °C. Under these idealized conditions, water has a pH of 7.0 and is said to be neutral:

$$2\mathrm{H_2O} \xleftrightarrow{K_W} \mathrm{H_3O^+} + \mathrm{OH^-} \quad K_W = [\mathrm{H_3O^+}][\mathrm{OH^-}] = 10^{-14} \tag{59.3}$$

The addition of acids or bases to pure water increases or decreases the hydrogen ion activity, respectively. The resulting pH depends on a number of factors, such as the concentration of the added acid or base and the strength as quantified by its dissociation constant [2].

59.2 Electrochemical Methods of pH Measurement

Electrochemical measurement of pH utilizes devices that transduce the chemical activity of the hydrogen ion into an electronic signal, such as an electrical potential difference or a change in electrical conductance. The following sections review electrochemical pH measurement, with emphasis on the glass membrane electrode. Also discussed is the operation of hydrogen ion-selective pHFETs, metal/metal oxide electrodes, and other approaches used in specialized applications. Table 59.2 provides a listing of sources, features, and key properties of selected, commercially available pH electrodes.

TABLE 59.2 Sources, Features, and Key Properties of Selected pH Electrodes

Manufacturer	Reference	Range	Accuracy pH Units	Cost ($US, 1997)	Key Features
Glass combination electrodes					
ABTECH Scientific, Inc.[a]	Ag/AgCl	0–14	±0.1	285	Handheld
Miniature Combination					Refillable bulb
pH electrode, CPE 905-X					Min vol. = 10 µL
ATI-Orion, Inc.[a]	Proprietary	0–14	±0.03	245	0 °C–100 °C
Combination electrode, 8102BN					
Brinkman[a]	Ag/AgCl	0–14	±0.05	132	Benchtop
futura plus combination					Handheld
Electrode, 39539					Refillable bulb
corning[a]	Ag/AgCl	0–14	±0.05	225	Increased sensitivity
High-performance	"Scavenger"				TRIS compatible
combination electrode, 476146					Rapid response
Hanna[b]	Ag/AgCl	0–14	±0.1	52	Handheld, waterproof
pHep 3 Pocket	Fabric junction				Auto 2 pt. cal.
pH tester					Auto temp. comp.
Hanna[c]	Ag/AgCl	0–12	±0.1	115	Handheld
combination electrode					Field portable
research grade HI1270					Sealed rugged epoxy body
Hach[c]	Ag/AgCl	0–14	±0.05	185	Benchtop, handheld
Hach one electrode with temperature sensor, 48600-00					Field portable, free-flowing reference junction 0 °C–100 °C
Microelectrodes, Inc.[a]	Ag/AgCl	0–14	±0.1	185	Handheld
Micro-combination pH					Min vol. = 5 µL
probe MI-410					
Radiometer[a]	Ag/AgCl	0–12	±0.1	185	General purpose
PHC2401					
Sensorex[a]	Ag/AgCl	0–14	±0.05	65	Epoxy body
Combination electrode					0–100 psig
450CUS					Flat tip for use on moist surfaces 0 °C–100 °C
ISFET electrodes					
ATI-Orion[a,c]	Ag wire	0–14	±0.02	326[d]	Built-in temperature probe
pHuture Sureflow				393[e]	
electrode, 616500					
Bioanalytical Systems, Inc.[b]	Ag/AgCl	2–10	±0.1	180	Handheld, stores dry
pH Boy, MF8960					Min vol. = 50 µL
Corning[c]	Ag/AgCl	0–14	±0.02	320	Handheld, stores dry
ISFET electrode					Min vol. = 20 µL
476395					Auto temp. comp. 0 °C–60 °C
IQ Scientific Instruments	Ag/AgCl	0–14	±0.01	229[d]	Handheld, sterilizable
IQ200	KC1 gel			695[e]	Auto. temp. comp. −5 °C to +105 °C
					Automatic two-point calibration
					Min vol. = 50 µL

(*continued*)

TABLE 59.2 (continued) Sources, Features, and Key Properties of Selected pH Electrodes

Manufacturer	Reference	Range	Accuracy pH Units	Cost ($US, 1997)	Key Features
Sentron[b]	Ag/AgCl	0–14	±0.01	259[d]	Handheld, stores dry
S1001	Fabric junction			695[e]	Man. 2 or 3 pt. cal
					Replaceable sensor tip, 0 °C–60 °C
					Min vol. = 50 µL (1 drop)
Meta/metal oxide electrodes					
Cypress Systems, Inc.[c]	Ag/AgCl	0–14	±0.01	1299[b]	Auto. temp. comp. 0 °C–100 °C
METOXY					Automatic two-point calibration
					Min vol. = 50 µL

[a] General-purpose meter.
[b] Integrated sensor and meter.
[c] Dedicated meter (specific to manufacturer).
[d] Replacement electrode.
[e] Electrode and converter for use with general-purpose meter.

59.2.1 Glass Membrane Indicator Electrode

The most widely used method for measuring pH is the glass membrane electrode. As illustrated schematically in Figure 59.1a, a pH meter measures the electrical potential difference (voltage) that develops between a glass membrane pH *indicator* electrode and a *reference* electrode immersed in the sample to be tested. The indicator and reference electrodes are commonly combined into a single, functionally equivalent, probe, referred to as a combination electrode. The glass membrane of the indicator electrode develops a pH-dependent potential, as a result of ion exchange between hydrogen ions in solution and univalent cations in the glass membrane. The sensitivity of the glass electrode

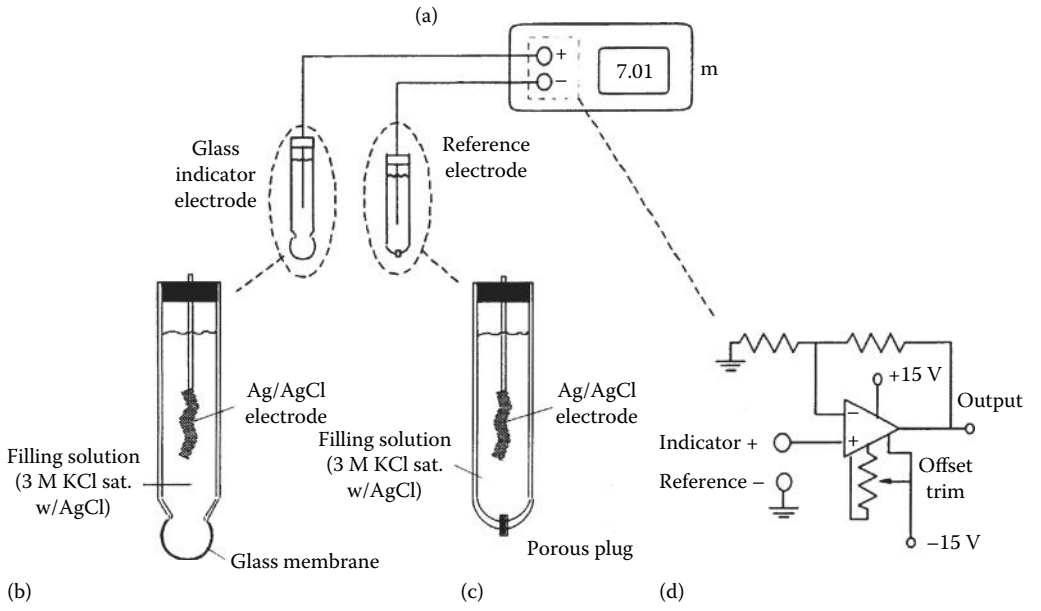


FIGURE 59.1 pH measurement using a glass membrane electrode: (a) measurement system comprising a pH meter, indicator, and reference electrodes; (b) indicator electrode construction; (c) reference electrode construction; and (d) amplifier circuit.

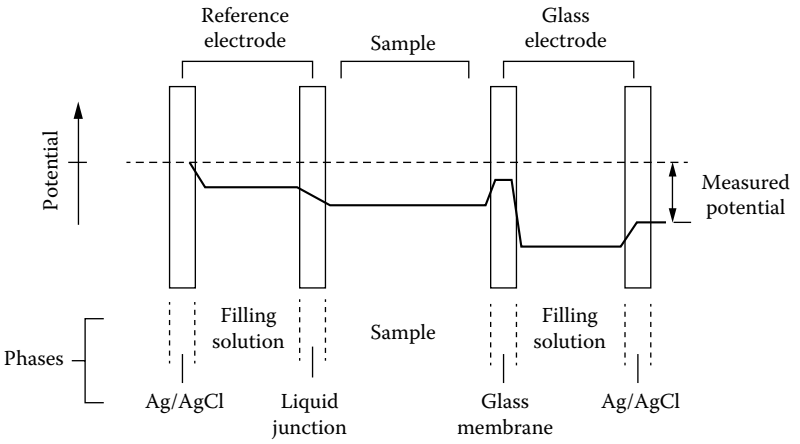


FIGURE 59.2 Schematic illustration of potential differences contributing to the measured potential between a glass membrane electrode and a reference electrode.

membrane potential to changes in pH is small, so a suitably designed reference electrode and a high input impedance meter are required in order for the potential to be precisely measured.

The construction of a typical pH indicator electrode is illustrated in Figure 59.1b. The glass membrane at the tip of the electrode, which acts as the transducer of pH, is approximately 0.1 mm thick. One side of the membrane contacts the sample, while the other contacts the electrode filling solution, an electrolyte of defined composition and pH. A reference element (e.g., a silver chloride–coated silver wire) immersed in the filling solution makes a stable electrical contact between the potential measuring circuitry of the pH meter and the electrolyte in contact with the inner side of the glass membrane.

The reference electrode illustrated in Figure 59.1c provides a means of making electrical contact between the meter and the sample that is in contact with the external side of the indicator electrode's pH-sensitive glass membrane. Like the indicator electrode, the body of the reference electrode is filled with an electrolyte into which a reference element is immersed. The electrode also has a means, such as a porous ceramic frit, by which the reference electrode filling solution can make physical and electrical contact to the sample. The importance of this "liquid junction" cannot be understated, as it is a potential source of sample contamination, measurement errors, and reliability problems.

The electrical potential difference between the pH indicator electrode and the reference electrode provides a measure of pH. Figure 59.2 provides an illustration of the potential differences, the origin of which will be discussed later, that contribute to the measured potential. The main function of the pH meter of Figure 59.1a is to measure this difference, with a precision of 0.1 mV or better. The amplifier circuit of Figure 59.1d illustrates how this may be done. Due to the high electrical resistance of the indicator electrode's glass membrane, the meter must have correspondingly high input impedance. Most pH meters currently sold contain built-in microprocessors that simplify pH measurement by performing and storing calibrations, doing diagnostics, and implementing temperature compensation.

59.2.1.1 Glass Membrane Indicator Electrode Construction

The hydrogen-ion-selective glasses used to construct pH indicator electrodes are formed by fusing silica, alkali metal oxides, and alkali earth oxides. The silica component, SiO_2, makes up approximately 70% of the glass. The alkali metal oxide contributes mobile ions such as sodium or lithium, which act as electrical charge carriers or are exchanged for protons in the hydrated glass layer. Components such as calcium oxide (CaO) are added to incorporate multivalent ions into the glass; these act to modify the network structure of the glass, imparting characteristics such as processability and chemical resistance.

59.2.1.2 Glass Membrane Indicator Electrode Response to pH

When the membrane glass is immersed in aqueous solution, a gel-like hydrated layer on the order of 100 nm thick forms at the surface of the glass. Within this hydrated layer, the exchange of sodium ions in the glass for hydrogen ions in the solution produces a potential difference between the hydrated layer and the solution that depends on the hydrogen ion activity. The net potential developed across the entire glass membrane, ϕ_{net}, is the sum of boundary potentials ϕ_b, at the inner (reference element side) and outer (sample side) extents of the membrane:

$$\phi_{net} = \phi_{b,inner} - \phi_{b,outer} \tag{59.4}$$

Under conditions where, at the outer surface, the exchange of sodium ions for hydrogen ions is complete (typically for pH less than 10), the hydrogen ion activity at the surface of the gel layer will be constant. The boundary potentials will then depend in a logarithmic manner on the hydrogen ion activity of the solution contacting the membrane, a function known as the Nernst equation:

$$\phi_b = \text{constant}' + \frac{RT}{F} \ln a_{H^+} \tag{59.5}$$

The net potential developed across the membrane is then proportional to the log of the ratio of the hydrogen ion activities of the sample and the electrode filling solution. The hydrogen ion activity of the filling solution is constant, so the membrane potential will depend linearly on pH:

$$\phi_{net} = \text{constant}'' + \frac{RT}{F} \ln \frac{a_{H^+,sample}}{a_{H^+,inner}} = \text{constant}''' - 2.3 \frac{RT}{F} \text{pH} \tag{59.6}$$

The dependence of membrane potential on pH is proportional to absolute temperature and at 25°C, the factor $2.3RT/F$ is equal to 59 mV per pH unit. Many pH meters that perform two-point calibrations calculate a slope and display it as a percentage of this theoretical value. For an electrode in good condition, this will be in the range of 90%–105%; values considerably different from this indicate a problem with the electrode.

While the membrane potential responds principally to changes in hydrogen ion activity, it also responds to sodium and other monovalent cations. These effects are most noticeable at high pH where hydrogen ion activity is small and is referred to as alkaline error. A more detailed analysis of the membrane response, which considers the effects of interfering ions, can be found in reference texts [3].

59.2.2 Reference Electrodes

A reference electrode, such as that pictured in Figure 59.1c, is used to make a stable, low-resistance electrical contact between the external measuring circuit and the sample, establishing a reference potential against which the indicator electrode can be referred. In the metallic conductors of the pH meter circuitry, current is carried by electrons, while in the sample electrolyte, current is carried by ions. The difference between reference electrodes is primarily the oxidation or reduction reactions that effect charge transfer across the reference element/electrolyte interface. The species participating in these reactions must be present in the electrolyte contained within the body of the reference electrode. As a result, the composition of the reference electrode filling solution is, in general, different from that of the sample, and the physical contact between the two dissimilar electrolytes forms what is known as a liquid junction. The two phase boundaries (reference element/filling solution and filling solution/sample) necessary to make electrical contact to the sample each introduce an additional potential, which adds to that of the indicator electrode, but proper design of the reference electrode ensures that these potentials remain constant and can be calibrated out.

59.2.2.1 Reference Elements

The most widely used reference electrode for pH measurement is the silver/silver chloride electrode. The electrode is constructed from a silver wire that has been coated with silver chloride and immersed in an electrolyte saturated with silver chloride. Current is readily passed across the electrode/electrolyte interface by the reduction of silver chloride to form silver metal and chloride ion:

$$AgCl + e^- \leftrightarrow Ag^0 + Cl^- \tag{59.7}$$

At equilibrium, the phase-boundary potential developed at the interface depends in a Nernstian manner on the chloride ion activity of the filling solution. Maintaining a constant chloride ion activity in the filling solution ensures that this potential remains constant. This type of electrode is easily and reproducibly constructed, and the phase-boundary potential has a smaller temperature coefficient than the calomel reference electrode. The main disadvantages of the Ag/AgCl electrode include the possibility that samples, biological samples in particular, may be contaminated with silver ion, and those precipitates (silver sulfide and nitrate) may form, which can clog the liquid junction.

Another commonly used reference electrode is the calomel electrode, based on the reduction of mercuric chloride to mercury:

$$HgCl_2 + 2e^- \leftrightarrow Hg^0 + 2Cl^- \tag{59.8}$$

Like the silver/silver chloride electrode, the potential of the calomel electrode depends on the chloride ion activity of the filling solution. It is more stable than the silver/silver chloride electrode, due to favorable reaction kinetics, and is therefore preferred for high-precision electrochemical determinations. However, the operating temperature of a calomel reference electrode is limited to 70 °C, compared to greater than 100 °C for the silver/silver chloride electrode.

A number of reference electrodes use a platinum reference element. Ross described a reference electrode based on the reduction of iodine to iodide [4]. The nature of the redox reaction differs from the Ag/AgCl and $Hg/HgCl_2$ electrodes, in that both the iodine and iodide are soluble in the filling solution, yielding an electrode potential that is relatively insensitive to temperature. The hydrogen reference electrode consists of a platinum electrode immersed in aqueous solution, over which hydrogen gas is bubbled. The platinum catalyzes the reduction of hydrogen ion to hydrogen gas. Historically, the standard hydrogen electrode has served as the primary reference electrode for precise electrochemical determinations. The role played by the hydrogen electrode in pH measurement, however, is primarily in its use as an indicator electrode for the determination of the hydrogen ion activity of primary pH standard solutions.

59.2.2.2 Liquid Junctions

The necessary contact between the filling electrolyte of the reference electrode and the sample forms a liquid junction and results in the development of a junction potential that depends on the compositions of the two electrolytes. The junction potential can be minimized by filling the reference electrode with an electrolyte having high concentrations of an anion and cation of comparable diffusivities. For Ag/AgCl and calomel reference electrodes, potassium chloride at concentrations ranging from 3 M to saturation is most commonly used as the filling electrolyte, as the diffusivities of the two ions are approximately equal. The stability of the liquid junction potential depends on the constancy of the interface between the filling electrolyte and the sample. A number of approaches exist to allow the filling electrolyte to controllably leak into the sample for this purpose. These include cracked bead, porous ceramic, and annular sleeve junctions. Leak rates range from less than 1 $\mu L\ h^{-1}$ (cracked bead) to greater than 100 (annular sleeve) $\mu L\ h^{-1}$.

59.2.2.3 Selection, Use, and Care of Reference Electrodes

Reference electrode selection is dictated by the sample to be studied and the measurement conditions, such as temperature. Silver/silver chloride electrodes are preferred over calomel for general-purpose use and at high temperatures (70 °C or greater). Calomel electrodes are preferred for high-precision determinations and where silver contamination of the sample presents a problem. An Ag/AgCl reference electrode incorporating a double junction is also another means of preventing silver contamination of the sample. The type of liquid junction and its flow rate is important; a high flow rate may contaminate the sample and deplete the electrode, while a low flow rate may lead to clogging. Viscous, semisolid samples or low ionic strength samples require a high flow rate junction. The reference electrolyte fill port should remain open during use of the electrode to ensure adequate flow through the junction. Leaving the fill port open at all times will prevent any contaminants from entering the body of the reference electrode and reduce the likelihood of junction clogging, at the expense of more frequent refilling. "Low maintenance" reference electrodes use a polymer gel saturated with electrolyte within the body of the electrode and will perform properly if stored in filling solution.

59.2.3 Instrumentation

The pH meter pictured in Figure 59.1a measures the potential developed between the pH indicator and reference electrodes, from which the pH of the sample is determined using a previously established calibration and possibly the sample temperature. The measurement of the potential, which may range in magnitude up to a few hundred millivolts, is complicated by the large electrical resistance presented by the glass membrane. This may range from 100 MΩ to greater than 1000 MΩ and necessitates the use of high input impedance amplifiers with FET input stages if the glass membrane potential is to be accurately measured. While the relatively low cost and high performance of commercially available meters obviates the need for amplifier construction, Figure 59.1d presents a simple amplifier suitable for use with a glass electrode. The amplifier circuitry of commercial pH meters incorporates additional functions to improve the accuracy and stability of the measurement, such as a driven shield to reduce noise pickup and chopper stabilization to reduce drift.

59.2.3.1 Calibration

The glass indicator electrode is calibrated using standard buffer solutions of known pH. A two-point calibration procedure is illustrated schematically in Figure 59.3. A pair of standards is chosen to bracket the pH range of interest. The response of the electrode is measured in each, and a calibration function is determined by linear interpolation. The slope can be compared to that expected from the Nernst equation to provide an indication of the performance of the electrode. The newest

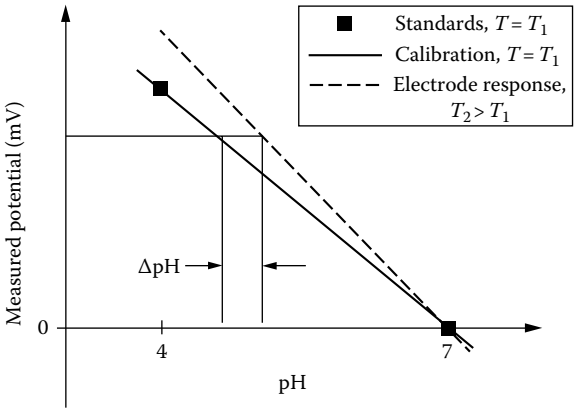


FIGURE 59.3 Calibration and temperature compensation of a glass pH electrode.

microprocessor-controlled pH meters simplify calibration by automatically recognizing the pH of standards and constructing calibration curves from as many as five buffers.

59.2.3.2 Temperature Compensation

The potential developed across the pH indicator glass electrode membrane is temperature dependent, with a temperature coefficient of approximately 0.3% per°C, as follows from Equation 59.6. The effect of temperature on the electrode calibration is illustrated in Figure 59.3. Most pH meters have provision for temperature compensation, which corrects the slope of the measured potential versus pH calibration. Manual compensation permits the user to dial in the temperature at which the measurement is going to be made. Meters equipped with automatic temperature compensation (ATC) use a platinum resistance thermometer to directly measure the temperature of the sample. The instrumentation within the meter then corrects the calibration function such that the millivolt reading is correctly interpreted as the pH of the sample at the measurement temperature.

59.2.3.3 pHFETs

A relatively recent development in pH measurement is the introduction of systems based on the use of ion-selective field-effect transistors (ISFETs) as the sensing element. ISFETs, of which the hydrogen ion-sensitive pHFET is one variant, are derived from the metal–oxide–semiconductor FET (MOSFET), the basic building block of integrated circuits. These silicon "chips" combine a pH-responsive membrane much like that of the glass electrode with the amplification of a field-effect transistor. The integral amplification and small size have led to the development of inexpensive, battery-powered, pocket-sized pH measurement systems. These devices have found unique and expanding niches, including the food industry where the measurement of pH using breakable glass electrodes presents an unacceptable safety hazard; the measurement of the pH of gels, pastes, and slurries; and for the measurement of strongly alkaline solutions where conventional glass bulbs respond to the sodium ions and give an erroneously low reading. The following sections cover the basic operation of these devices, instrumentation, and applications. Additional details can be found in reference texts [5].

59.2.3.4 Construction and Operation

A schematic cross section of a pH-sensitive ISFET is presented in Figure 59.4. The pHFET differs from a MOSFET in that the metal gate of the MOSFET is replaced by a pH-responsive membrane material

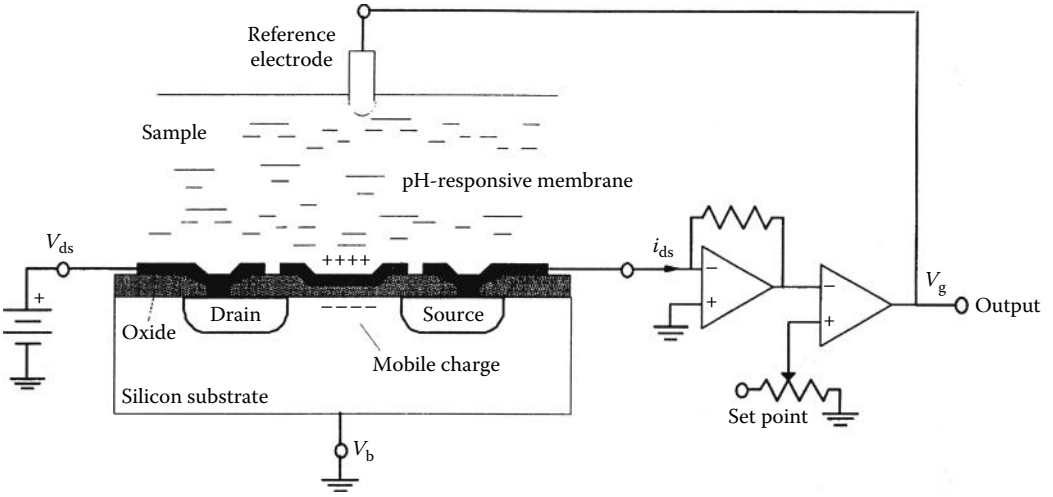


FIGURE 59.4 pH measurement using an ISFET, including an amplifier circuit for constant drain current operation.

such as silicon nitride, aluminum oxide, or tantalum oxide, which contacts the sample solution directly. As with the glass electrode, electrical contact is made to the sample through a reversible reference electrode. A suitable voltage applied to the reference electrode (relative to the silicon substrate) will charge the capacitor formed by the solution, insulating layers, and silicon substrate and create mobile charge in the channel region. A potential simultaneously applied between the drain and source electrodes will result in current flow. Using a "charge imaging" model [6], this drain current, i_{ds}, can be described by

$$i_{ds} = A\ V_{ds} Q_c = A\ V_{ds} C_2 \left(V_g - V_T - \frac{V_{ds}}{2} \right) \tag{59.9}$$

where

the constant A includes geometrical factors
V_{ds} is the voltage applied to the drain
Q_c is the mobile channel charge

The mobile charge is a function of V_g, the voltage applied to the reference electrode; V_T, the "threshold voltage" needed to produce mobile charge; and C_2, the capacitance of the gate region. The threshold voltage includes a number of terms, one of which is the phase-boundary potential at the interface between the sample and the pH-responsive insulating layer. Adsorption of protons at the surface of this layer leads to a Nernstian dependence of this potential on hydrogen ion activity with the result being that changes in pH modulate the drain current of the device:

$$i_{ds} = A V_{ds} C_2 \left(V_g - V_T + 2.3 \frac{RT}{F} \text{pH} - \frac{V_{ds}}{2} \right) \tag{59.10}$$

59.2.3.5 Instrumentation

Equation 59.10 illustrates the dependence of ISFET current on pH. The preferred method of operation of the pHFET is to operate at a constant drain current. A circuit for doing this is illustrated schematically in Figure 59.4. A transconductance amplifier converts the drain current to a voltage, which is referenced against a set point. The output of the comparator drives the reference electrode, V_g. Since the mobile channel charge Q_c is constant, as are V_{ds} and $V_{T'}$, changes in the comparator output directly reflect changes in the hydrogen ion activity. That is, a 59 mV change in output corresponds to a change in pH of one unit. Interface circuitry similar to that of Figure 59.4 allows the device to be connected directly to a glass electrode pH meter. As with the glass electrode, temperature compensation is required if the device is to be used at temperatures different from that of calibration.

59.2.4 Other Electrochemical pH Electrodes

There are a number of other approaches for electrochemical measurement of pH [1]. Two that deserve mention, as they are commercially available, either directly or incorporated within another product, are metal/metal oxide and liquid membrane electrodes.

59.2.4.1 Metal/Metal Oxide pH Sensors

A pH indicator electrode with application to measurement at high temperature and pressure can be constructed from metals coated with an oxide. These metal electrodes may take the form of a wire, polished disk, or sputtered thin film, on which an oxide has been formed through thermal oxidation, chemical vapor deposition, or electrochemical oxidation. Electrodes have been constructed from systems including W/W_2O_3, Sb/Sb_2O_3, Pt/PtO_2, Ru/RuO_2, Pb/PbO_2, and Ir/IrO_2. In aqueous solution, the metal oxide can be reduced to a lower oxidation state with the consumption of a proton in solution.

Using an iridium/iridium oxide (Ir/IrO_2) electrode as an example, this redox couple is believed to be represented by an Ir(III)/Ir(IV) half-cell reaction of the form

$$2\,IrO_2 + 2H^+ + 2e^- \leftrightarrow Ir_2O_3 + H_2O \tag{59.11}$$

The Ir/IrO_2 electrode shows a near Nernstian response of −59 mV per pH unit [7]. The chemical resistance, high-temperature and pressure performance, the nonglass construction, and the considerable potential for miniaturization suggest that these pH sensors will find application in areas not typically employed by glass membrane electrodes or pHFETs.

59.2.4.2 Liquid Membrane Electrodes

Another type of electrochemical pH electrode is based on polymeric "liquid membranes," most commonly used to construct ion-selective electrodes for ions such as potassium [8]. The membranes consist of a plasticized polymer film into which an ionophore has been incorporated. The ionophore is a molecule that selectively binds and transports a given ion across the membrane, making the membrane selectively permeable to the ion of interest. The ionophore *N*-tridodecylamine binds hydrogen ions and has been used in the construction of a pH sensor incorporated into a disposable cartridge device for measuring blood electrolytes [9].

59.3 Optical Methods of pH Measurement

For some applications, optical methods offer advantages over the use of the glass electrode or other electrochemical devices for the measurement of pH. Organic dye molecules with pH-dependent spectral properties have been routinely used for decades in acid–base titrations and in pH indicator papers. These dyes, many of natural origin, have more recently been put to use as indicators to measure localized pH within living cells and in the development of fiber-optic probes for measuring pH within the body. The following sections cover the basic principles and major applications of optical pH measurement; more detailed treatments may be found in Ref. [10].

59.3.1 Indicator Dyes

Optical measurement of pH is based on the use of organic dye molecules that are weak acids or bases. The loss or gain of a proton changes the electronic structure of the molecule, producing a measurable change in the manner in which the molecule interacts with light. This interaction can be the absorption of light at a particular wavelength or fluorescence by one form of the molecule. Equilibrium exists between the acid and base forms of the dye, whether free in solution or chemically attached to a supporting membrane, and can be described by the equilibrium

$$Dye^z + H^+ \xleftrightarrow{K_a} Hdye^{z+1} \quad K_a = \frac{a_{Dye^z} a_{H^+} a}{a_{Dye^{z+1}}} \tag{59.12}$$

where z represents the valence of the molecule. The interrogation of the solution or membrane with light will produce a response that is weighted by the relative proportions of the acid and base forms of the molecule, which in turn depends on pH. Inspection of Equation 59.12 shows that the ratio of the two forms, and hence color, varies continuously with hydrogen ion activity. The sensitivity of the measurement is greatest when the acid and base forms of the dye are present in approximately equal concentrations. This occurs when the pH of the solution is close to the pK_a ($-\log [K_a]$) of the dye. Limitations of the human eye restrict detectable changes in color to a tenfold excess of one species over the other. This corresponds to a change of ±1 pH unit. Thus, an indicator with a pK_a of 1×10^{-5} will display a color change if the solution in which it is dissolved changed from 4 to 6 pH units. The pH of interest therefore dictates selection of the particular dye.

An understanding of possible interactions between the indicator dye and the sample is important in the effective use of optical indicator dyes for pH measurement [11]. Factors such as temperature, electrolyte concentration, and the presence of organic solvents may cause a shift in p*K*, of the indicator dye; the shift could be as much as one or more pH units. As weak acids and bases, the addition of indicator dyes to a sample can change the pH, particularly in weakly buffered samples. This effect is important when using pH papers or optical sensors to measure small samples, where the potential exists for the buffering capacity of the sample to be exceeded by the amount of dye immobilized in the paper or on the sensor. As a general rule, potential errors can be minimized by calibrating in solutions similar in composition to the sample to be measured.

59.3.1.1 Absorption Indicator Dyes

The conjugate acid and base forms of absorption indicator dyes differ in the characteristic wavelengths at which they absorb light energy. A number of common absorption indicator dyes are listed in Table 59.3, and their pK_as span the range of pH from approximately 1 to 13. The absorption spectra of a solution of the dye phenol red at a number of different pHs is shown in the inset of Figure 59.5. The acid form of the dye has an absorption maxima at λ_{max} = 435 nm, while the base form absorbs maximally at λ_{max} = 565 nm. As the pH of a phenol red solution is increased from 6 toward the pK_a of 7.9, the equilibrium of Equation 59.12 shifts to the basic form of the dye. The relative absorption peaks change, reflecting that a shift in concentrations of the two is formed. For quantitative pH measurement, the absorption of light by the indicator is measured by a spectrometer or reflectometer at a specific wavelength or narrow set of wavelengths corresponding to the wavelength of maximum absorption λ_{max} for the acidic or basic form of the chromophoric dye molecule. Beer–Lambert's law describes the proportionality between absorbance, *A*, and the concentration of the indicator, c_{dye}, in moles per liter:

$$A = \log\left(\frac{I}{I_0}\right) = \varepsilon b c_{dye} \tag{59.13}$$

The quantity I/I_0 represents the ratio of the intensity of light transmitted through the sample to that incident on the sample, ε the extinction coefficient of the dye molecule, and *b* the path length. The magnitude of the absorption peak heights at 565 nm, plotted in Figure 59.5 as a function of pH, and reflects the changing concentration of the basic form of the dye as the equilibrium of Equation 59.12 shifts.

TABLE 59.3 Acid–Base Indicator Dyes

Indicator	pK_a	λ_{max}(nm) Acid Form	Base Form
Thymol blue	1.7	544	430
Methyl orange	3.4	522	464
Bromophenol blue	3.9	436	592
Bromocresol green	4.7	444	617
Methyl red	5.0	530	427
Chlorophenol red	6.0	—	573
Bromocresol purple	6.3	433	591
Bromothymol blue	7.1	433	617
Phenol red	7.9	433	558
Cresol red	8.2	434	572
Phenolphthalein	9.4	—	553
Thymolphthalein	10.0	—	598

Source: Dean J.A. and Lange N.A. (Eds.), *Lange's Handbook of Chemistry*, 14th edn., McGraw-Hill, New York, 1992.

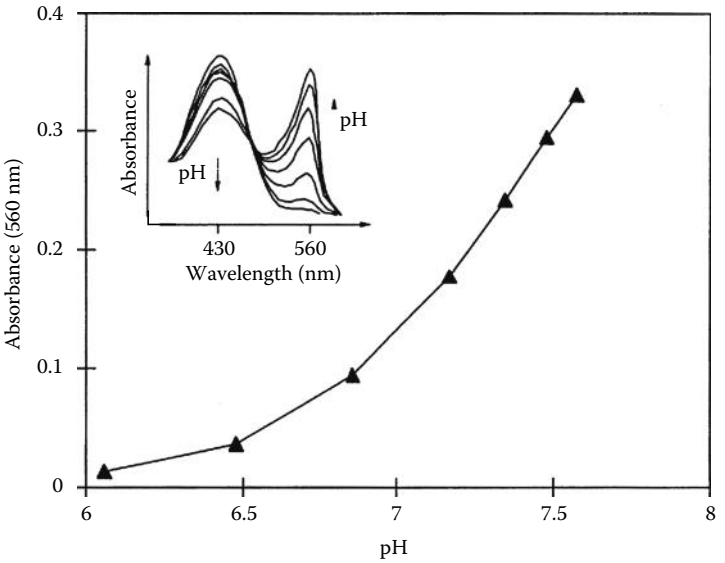


FIGURE 59.5 Magnitude of the absorption of phenol red at 565 nm as a function of pH. Inset: absorption spectra at different pHs.

59.3.1.2 Fluorescent Indicator Dyes

Fluorescent indicator dyes absorb light of a particular color (or wavelength) and reemit some of the absorbed energy as light of a different color. Absorption of light by the indicator promotes the molecule from the ground state energy, E_g, to a higher energy state, E_1. Subsequent processes such as molecular collisions lead to a transition to a lower energy excited state, E_2. The molecule can then emit a photon of energy E_2–E_g, resulting in the return of the molecule to the ground state from this intermediate state. The emitted light is at a longer wavelength than the exciting wavelength, and the difference is known as the Stokes shift. In principle, a fluorescence measurement is more sensitive than an absorption measurement because the only light measured by the detector originates from fluorescing molecules. Table 59.4 lists a number of common fluorescent pH indicators used for measurement of pH in the physiological range. Fluorescein is widely used because the absorption maximum at 490 nm of the fluorescent dianion is readily excited by the 488 nm emission of argon ion lasers.

59.3.1.3 Indicator Papers

Indicator papers are a simple, rapid, and inexpensive means of measuring pH when the precision of an instrumental measurement is not necessary. These are constructed from a strip of paper or plastic that has

TABLE 59.4 Fluorescent pH Indicator Dyes (for Physiological Applications)

Indicator	pK_a	Excitation Maximum (nm)	Emission Maximum (nm)
BCECF[a]	7.0	482	520
Fluorescein	6.4	490	515
HCC[b]	6.9–7.0	410	455
HPTS[c]	6.8–7.3	465	520

[a] 2′,7′-bis-(2-Carboxyethyl)-5-(and-6)-carboxyfluorescein.
[b] 7-Hydroxycoumarin-3-carboxylic acid.
[c] 8-Hydroxy-1,3,6-pyrenetrisulfonate.

been impregnated with one or more absorption indicator dyes chosen to span the pH range of interest. The dyes are generally covalently attached to the strip, to prevent contamination of the sample by leaching of the dyes. A strip for indicating pH in the range of 5.5–9.0 can be constructed using the dyes bromocresol purple, bromothymol blue, and phenol red [1]. The pH of a sample is determined to a precision of typically 0.5 units by comparing the color of the strip to a color calibration chart provided by the manufacturer.

59.3.2 Fiber-Optic pH Probes

Optical pH sensors, often referred to as optrodes, represent some of the most sophisticated pH sensors, finding use for remote sensing in the body or industrial plants due to their small size and lack of electrical connections. They are typically constructed by immobilizing an indicator dye at the tip of a light guide formed from one or more optical fibers, which are used to couple light between the indicator and the measurement instrumentation, as illustrated schematically in Figure 59.6a. The resulting probes can be made very small, and with high-quality optical fibers, pH can be measured over considerable distances in electrically noisy environments that would interfere with potentiometric-type electrodes.

59.3.2.1 Construction and Instrumentation

The pH indicator dye must be immobilized in close proximity to the fiber tip. The method of immobilization impacts the sensor's response, response time, long-term stability, and mechanical integrity. The dye can be immobilized to a solid support, such as a membrane or porous glass bead, which is then attached to the end of the fiber. Indicator dyes have been covalently linked to the tip of glass fibers or entrapped within polymer films formed at the tip using light energy from the fiber to initiate polymerization. The construction of probe tips and the instrumentation used for a fiber-optic pH probe depend on the type of indicator being used. Unlike the glass electrode, which has remained largely unchanged for more than

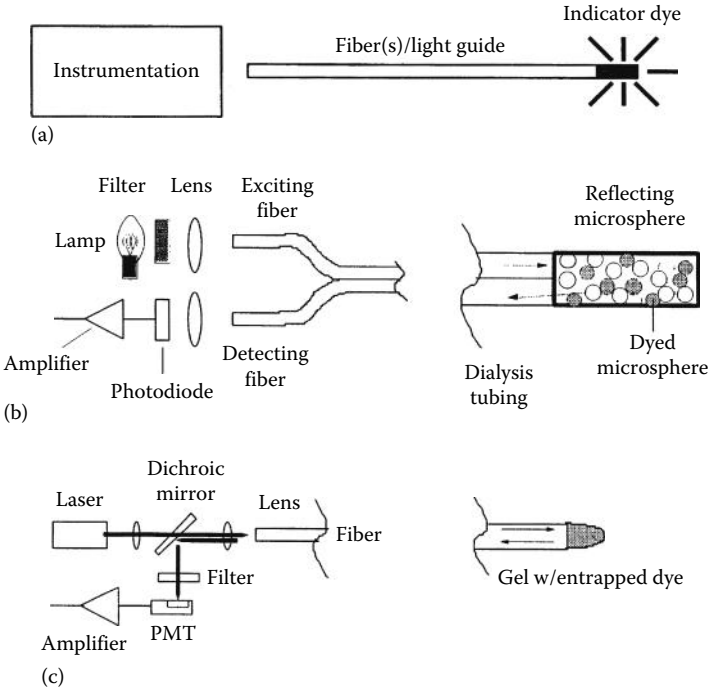


FIGURE 59.6 Optical fiber sensors: (a) Generic. (b) Based on absorption indicator dyes. (From Peterson, J.I. et al., *Anal. Chem.*, 52, 864, 1980.) (c) Based on fluorescent indicator dyes. (From Munkholm, C. et al., *Anal. Chem.*, 58, 1427, 1986.)

75 years, there is considerable variety in the design of fiber-optic pH sensors. While a complete review of the literature is beyond the scope of this work (interested readers may wish to consult Leinevand Wolfbeis [10] and a recent review of chemical sensors by Janata et al. [12]), the following examples illustrate the approaches to construct optical fiber pH sensors using absorption and fluorescent dyes.

Absorption optrodes measure the change in intensity of the light returned from the fiber tip/sensing region relative to the interrogating light of specific (or a narrow band) wavelength. These require a separate collection fiber or fibers and a means to reflect the light, such as a reflector or the use of scatterers such as polystyrene beads. In the example of Figure 59.6b [13], polyacrylamide beads containing the indicator dye phenol red were mixed with white polystyrene latex microspheres that served as scatterers. The beads were contained within dialysis tubing attached to the end of a pair of plastic optical fibers, one used for excitation, the other to collect scattered light. Light from a lamp was filtered to alternately select the λ_{max} of the base form of the dye (565 nm) and a wavelength (600 nm) where the absorbance is pH independent as a reference. (The isosbestic point at 480 nm could also have served as a reference.) The intensity of light scattered back into the second detector fiber was quantified by a photodiode. The ratio of the scattered intensities at the two wavelengths was then related to the pH of the sample.

Fluorescent indicator optrodes measure the Stokes-shifted fluorescence emission by the indicator and may use a single fiber to both interrogate and collect signal-carrying light. The amount of fluorescent pH indicator at the fiber tip must be maximized since fluorescence is emitted omnidirectionally and only a small fraction of the emitted fluorescence will be captured by the fiber. In the example of Figure 59.6c [14], the indicator dye was entrapped within a polymer gel formed at the tip of a glass fiber. The source of excitation light is typically a high-intensity lamp or a laser. If a lamp, its output is filtered to select a narrow band of wavelengths at the absorption maximum of the specific form of the indicator, which is then introduced into the fiber to excite the fluorophore at the tip. Emitted light collected by the fiber exits the fiber and is incident on a dichroic mirror, which reflects the long-wavelength light and passes shorter wavelengths. A filter then selects a narrow band of emission wavelengths, and finally, the light is detected. Due to the relatively small light intensities, the detector is typically a photomultiplier tube rather than a photodiode.

59.4 Frontiers of pH Measurements

Because of its wide-ranging importance, there is likely to be continued innovation in both approaches and opportunities for pH measurements. Some important developments, such as pHFETs and metal/metal oxide pH sensors, have already occurred and are now achieving noteworthy commercialization. Investigators continue to seek new materials for constructing transducers. Examples include electroconductive (electroactive and inherently conductive) polymers such as polyaniline [15] and polyelectrolyte hydrogels [16], which form pH-responsive membranes suitable for the construction of microsensors.

A noteworthy innovation is the application of pH measurement devices as transducers that are integrated into more complex analytical systems directed at other analytes. A pH indicator dye forms the basis of the pCO_2 measurement in an indwelling fiber-optic catheter [17]. Because protons are generated and/or consumed in many biological reactions involving enzymes, pH sensors have found use as transducers in biosensor devices.

An expanding area for pH measurement is likely to be the study of biological systems. The light-addressable potentiometric sensor is a device that measures the pH change resulting from the metabolic activity of cells, with application to examining cellular response to toxins and infectious agents [18]. The measurement of intracellular or subcellular pH using pH-sensitive fluorescent indicators [19] can provide insights into the physiology of the cell.

Acknowledgments

AGE thanks Allage Associates, Inc. and ABTECH Scientific, Inc. for financial support.

References

1. H. Galster, *pH Measurement: Fundamentals, Methods, Applications, Instrumentation*, New York: VCH Publishers, 1991.
2. D.A. Skoog, D.M. West, and F.J. Holler, *Fundamentals of Analytical Chemistry*, 7th edn., Philadelphia, PA: Saunders College Publishing, 1996.
3. A.J. Bard and L.R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, New York: John Wiley & Sons, 1980.
4. J.W. Ross, Temperature insensitive potentiometric electrode system. U.S. Patent No. 4,495,050, 1982.
5. J. Janata, *Principles of Chemical Sensors*, New York: Plenum Press, 1989.
6. S.D. Senturia, The role of the MOS structure in integrated sensors, *Sensors and Actuators*, 4, 507–526, 1983.
7. M.F. Yuen, I. Lauks, and W.C. Dautremont-Smith, pH dependent voltammetry of iridium oxide films, *Solid State Ionics*, 11(1), 19–29, 1983.
8. D. Ammann, *Ion-Selective Microelectrodes: Principles, Design and Applications*, New York: Springer, 1986.
9. S.N. Cozzette, G. Davis, I.R. Lauks, R.M. Mier, S. Piznik, N. Smit, P. Van Der Werf, and H.J. Wieck, Process for the manufacture of wholly microfabricated biosensors, U.S. Patent No. 5,466,575, 1995.
10. M.J.P. Leiner and O. Wolfbeis, Fiber optic pH sensors, in O. Wolfbeis (Ed.), *Fiber Optic Chemical Sensors and Biosensors*. Vol. 1, Boca Raton, FL: CRC Press, 1991.
11. J. Janata, Do optical sensors really measure pH? *Analytical Chemistry*, 59, 1351–1356, 1987.
12. J. Janata, M. Josowicz, and D.M. DeVaney, Chemical sensors, *Analytical Chemistry*., 66, 207R–228R, 1994.
13. J.I. Peterson, S.R. Goldstein, R.V. Fitzgerald, and D.K. Buchwald, Fiber optic pH probe for physiological use, *Analytical Chemistry*, 52, 864–869, 1980.
14. C. Munkholm, D.R. Walt, F.P. Milanovich, and S.M. Klainer, Polymer modification of fiber optic chemical sensors as a method of enhancing fluorescence signal for pH measurement, *Analytical Chemistry*, 58, 1427–1430, 1986.
15. A. Guiseppi-Elie, G.G. Wallace, and T. Matsue, Chemical and biological sensors based on electrically conducting polymers, in T. Skotheim, R. Elsenbaumer, and J.R. Reynolds (Eds.), *Handbook of Conductive Polymers*, 2nd edn., Chap. 34, p. 963, New York: Marcel Dekker, 1996.
16. N.F. Sheppard Jr., M.J. Lesho, P. McNally, and A.S. Francomacaro, Microfabricated conductimetric pH sensor. *Sensors and Actuators*. *B*, 28, 95–102, 1995.
17. J.B. Yim, G.E. Khalil, R.J. Pihl, B.D. Huss, and G.G. Vurek, Apparatus for continuously monitoring a plurality of chemical analytes through a single optical fiber and method of making, U.S. Patent No. 5,098,659, 1992.
18. J.W. Parce, J.C. Owicki, K.M. Kercso, G.B. Sigal, H.G. Wada, V.C. Muir, L.J. Bousse, K.L. Ross, B.I. Sikic, and H.M. McConnell, Detection of cell-affecting agents with a silicon biosensor, *Science*, 246, 243–247, 1989.
19. R. Haugland (Ed.), *Handbook of Fluorescent Probes and Research Chemicals,* 6th edn., Chap. 23, Eugene, OR: Molecular Probes, Inc., 1996.

VII

Medical, Biomedical, and Health

60

Wireless Sensor Networks in Health

Enrique Dorronzoro
University of Sevilla

Isabel Gómez
University of Sevilla

Ana Verónica Medina
University of Sevilla

Luis Fernández-Luque
University of Sevilla and Northem Research Institute

Jose Antonio Gómez
University of Sevilla

60.1 Introduction

Recent advances in telecommunications are changing our society. The main reasons for this evolution are the increase of connectivity and the monitoring advances, which provide mobility to humans and increase accessibility for the handicapped.

Wireless telecommunications allow transferring information between two or more devices that are not physically connected. Figure 60.1 shows wireless networks transmission technologies comparing bandwidth with power consumption. As shown in this figure, higher transmission speeds imply higher power consumption. In order to guarantee people mobility, there has to be a compromise between these two variables. It is required to use technologies that transmit with enough bandwidth but without high power consumption.

From all the standards defined for wireless communications, IEEE 802.11 and 802.15, standards have a higher impact. IEEE 802.11 describes a set of standards for implementing Wireless Local Area Networks (WLAN) and IEEE 802.15 is specialized in Wireless Personal Area Networks (WPAN).

WPAN is a network used to communicate devices in proximity to an individual's body using wireless technologies such as Bluetooth, ZigBee, and IrDA. This group of standards includes a wide range of devices.

A wireless sensor network (WSN) consists of a WPAN composed by autonomous sensors (sensing nodes) that cooperate to monitor physical or environmental conditions, such as temperature, sound, vibration, and pressure. WSN components are sensing nodes, transmission technology, and a standard to define the communication. WSN allows cooperation between systems where devices are close to the human body.

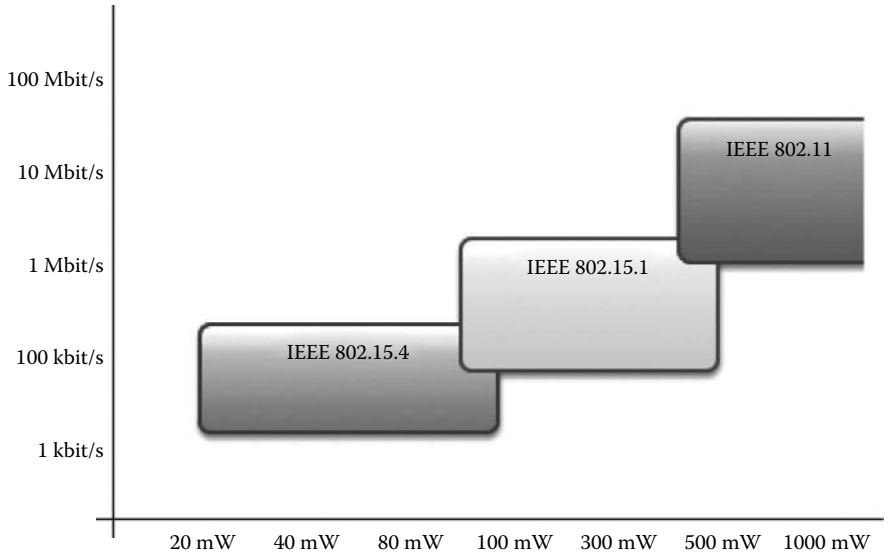


FIGURE 60.1 Body Area Network (BAN) power versus bandwidth. WLAN and IEEE 802.15 are specialized for WPANs.

60.1.1 Wireless Sensor Networks in Health

WSN are becoming increasingly important for monitoring patients both in the clinical setting and at home. Remote monitoring allows monitoring of sports activities, emergencies, catastrophe responses, etc.

Wireless communications also provide more comfort for the patients, with the absence of wires reducing costs and providing more flexibility. WSNs can integrate vital sign sensors and also environmental sensors such as air quality.

The following section describes the WSN transmission technologies including a state of the art of the publications that implement a WSN system in health. This review has been published in Ref. [45] and updated with recent progress.

60.2 Wireless Communication Standards

As previously presented, IEEE 802.11 and IEEE 802.15 family of standards are more relevant for wireless communications. But due to the high power consumption of IEEE 802.11, the IEEE 802.15 family selected for WSN. IEEE 802.15 is divided in different sections:

- IEEE 802.15.1, Bluetooth.
- IEEE 802.15.3 defines physical layer (PHY) and medium control access (MAC) layers for high-speed WPANs.
- IEEE 802.15.4 defines PHY and MAC layers. In some occasions, ZigBee is used to implement upper layers.
- IEEE 802.15.6, a task group is formed to be in charge of developing a communication standard optimized for low-power devices and operation on, in, or around the human body (but not limited to humans) to serve a variety of applications including medical, consumer electronics/personal entertainment, and others.

60.2.1 IEEE 802.15.1

Originally developed by Bluetooth SIG., the IEEE 802.15.1 standard is defined to operate with devices that provide a short range with medium transmission speed. In 1999, the committee came up the first Bluetooth specification, version 1.0. From this moment, it started spreading because of its low power consumption and low cost.

60.2.1.1 Frequency

Bluetooth operates on the 2.45 GHz frequency band (ISM band; Industrial, Scientific and Medical band). This band is shared with Wi-Fi, so to solve possible interference problems, it implements some key technologies. One of these solutions consists of using low power signals (1 mW). Even with the low power, Bluetooth devices do not need to be in the line of sight to communicate.

60.2.1.2 Range

The maximum distance that Bluetooth can reach is based on the power of the transmission. There are three classes with different ranges and signal power levels as shown in Table 60.1.

60.2.1.3 Bandwidth

Bluetooth devices can be also classified according to the bandwidth (Table 60.2).

60.2.1.4 Network Topologies

Bluetooth networks are called piconets. A piconet is composed of a master device that controls seven different devices. In this way, all the devices that belong to the same piconet share the same hop frequency and their clocks are synchronized with the master. One device can be part of different piconets at the same time irrespective of whether it is a master or a slave.

It is possible to connect different piconets in a scatternet in order to expand the physical size of the network beyond the Bluetooth's limited range. Having eight devices in less than 10 m causes interference between members. To prevent this problem, Bluetooth defines time slots to avoid two or more devices transmitting at the same time.

TABLE 60.1 RF Power Classification Levels

Class	Power (mW)	Maximum Allowed Range (m)
Class 1	100	100
Class 2	10	10
Class 3	1	1

TABLE 60.2 Bandwidth Classification

Version	Bandwidth (Mbit/s)
Version 1.2	1
Version 2.0 + EDR	3
Version 3.0 + HS	1
Version 4.0 + HS	1

However, this is not the only interference problem. The interference can also be caused by near piconet clusters in the vicinity. This problem is not resolved by time slots as this solution only works for the devices on the same piconet. This is the reason why Bluetooth implements a frequency-hopping spread spectrum (FHSS) mechanism to avoid interference between piconets. FHSS is a transmission process where the devices involved in the communication change their frequency in regularly hops according to a predetermined code.

Summarizing, time slots prevent interference between Bluetooth devices that belong to the same piconet and FHSS prevents interferences between Bluetooth devices of different piconets.

60.2.1.5 State of the Art

One of the advantages of Bluetooth is that it can be integrated in most mobile devices and laptops. It is common to use these devices as the master devices of the topology that is in control of the sensors that compose the network. Based on this topology, a laptop may be used as the master device on the piconet. Bluetooth publications are summarized in Table 60.3. Most of these reviewed publications are focused on monitoring parameters for a specific scenario. Some examples of the applications of Bluetooth in health are discussed later:

- A polysomnography sensor is used to remotely record sleep disorders [27].
- An accelerometer and optical sensor are used in sport activities [42]. An elastic belt encircles the user's chest and measures low frequency components of belt circumference change. Variations in this length are measured by an optical sensor and outputted as serial digital data. The accelerometer measures dynamic acceleration force produced by the user.
- Electrocardiogram (ECG) sensors, Electrooculography (EOG) sensors, and Electroencephalography (EEG) sensors are mounted on a helmet for general monitoring [28].

But there is also the possibility of using Bluetooth in a smartphone as the master device [26] to collect data from other sensors such as a glucose sensor to control diabetes.

Many medical systems are also using Bluetooth sensors [38]. One example uses a system with a wireless ocular telemetry sensor for glaucoma. Another example is found in [40] where the study of photoplethysmography using an accelerometer can provide valuable information about the cardiovascular system, such as heart rate, arterial blood oxygen saturation, blood pressure, cardiac output, and autonomic functions. A last example is a system for the detection of gait abnormalities or deteriorations in a patient's home environment by using accelerometers and gyroscopes [24].

Other classes of systems are focused on monitoring activities [18,43]. This kind of monitoring system is aimed for elderly people and rehabilitation therapies. Sensors used in these applications are presented in Table 60.4.

TABLE 60.3 IEEE 802.15.1

Reference	Health Focus	Health Focus	Battery Life
Greene et al. [24]	Gait monitoring	Accelerometer gyroscope	N/A
Istepanian et al. [26]	Diabetes monitoring	Glucose	N/A
Kayyali et al. [27]	Sleep disorders monitoring	Polysomnography	N/A
Kim et al. [28]	Monitoring	ECG, EOG, EEG	N/A
Mansouri and Shaaranyh[38]	Glaucoma monitoring	Intraocular pressure	N/A
Poh et al. [40]	Monitoring photoplethysmography	Accelerometer	N/A
Tawa et al. [42]	Breathing training	Optical accelerometer	N/A
Wagenaar et al. [43]	Monitoring of functional activities	Accelerometer, gyroscope	N/A
Au et al. [18]	Continuous activity monitoring	Accelerometer	N/A

TABLE 60.4 Sensors Used with IEEE 802.15.1

Sensors		
Greene et al. [24]	Gyroscope	SHIMMER add-on [15]
Istepanian et al. [26]	Glucose meter	OneTouch Ultra, LifeScan [11]
Kayyali et al. [27]	Polysomnography	Crystal Monitor PSG Series, CleveMed [8]
Kim et al. [28]	ECG, EOG, EEG	Electrical tape, 3M
Mansouri and Shaaranyh [38]	Intraocular pressure	SENSIMED Triggerfish [14]
Poh et al. [40]	Accelerometer	ADXL 330, Analog Devices [6]
Tawa et al. [42]	Accelerometer	KXM52–1050, Kionix
	Optical	BOMC2-USSP, Buffalo
Wagenaar et al. [43]	Accelerometer	MMA7260Q, Freescale [10]
	Gyroscope	Idg500 InvenSense [2]

60.2.3 IEEE 802.15.3

The IEEE 802.15.3 standard is designed to provide a high data rate and low power consumption solution for WPAN. It is designed to provide sufficient quality of services for the real-time distribution of content such as video and music. The original standard uses a traditional carrier-based 2.4 GHz radio as the physical transmission layer.

IEEE 802.15.3a is a follow-on standard still in the formative stages, which will define an alternative and improved physical layer. Current proposals based on ultra wide band (UWB) will provide more than 110 Mbit/s at a distance of about 10 m and 480 Mbit/s at 2 m. This will allow the streaming of high-definition video between media servers and high-definition monitors, as well as fast transfer of files among servers and portable devices.

Reference [44] compares the 802.11e hybrid coordination function (HCF) MAC and 802.15.3 time division multiple access (TDMA) MAC mainly in terms of throughput and power management. The conclusion after the comparison is that the throughput differences between them are quite small. On the other hand, the power management of 802.15.3 is easier than that of 802.11e. But unfortunately, its power consumption is similar to that in a Wi-Fi device, around 227 mA, when transmitting and receiving [31] so it is high for the telemedicine field when using wireless sensors.

60.2.4 IEEE 802.15.4

IEEE 802.15.4 is a standard that specifies the physical layer and MAC layer for low-rate wireless personal area networks (LR-WPANs). It offers short range and low bandwidth but with the benefit of low power consumption. It is mainly used for industrial control, embedded sensors, and it is also adequate for healthcare systems. Healthcare systems can benefit from this technology as a node can go to sleep and wake up when a new task is due thus saving battery power.

60.2.4.1 Characteristics

- Frequency: Uses ISM frequency band, mainly at the 2.4 GHz band, 868 MHz (Europe) and 902–928 MHz (North America).
- Range: 10–50 m, but it depends on the environment.
- Bandwidth: Data transmission range up to 250 kbps.

60.2.4.2 IEEE 802.15.4 Devices

Two different device types can participate in an LR-WPAN network; a full function device (FFD) and a reduced function device (RFD). The FFD can operate in three modes serving as a personal area network (PAN) coordinator, a coordinator, or a device. An FFD can talk to RFDs or other FFDs, while an RFD

can talk only to an FFD. An RFD is intended for applications that are extremely simple, such as a light switch or a passive infrared sensor; they do not have the need to send large amounts of data and may only associate with a single FFD at a time. Consequently, the RFD can be implemented using minimal resources and memory capacity.

60.2.4.3 Topologies

It is important to understand that the network layer is not defined for this standard. This means there is no possibility to route the data between the different nodes that compose the network.

There are two defined topologies. Both require a PAN coordinator that works as a network coordinator. This device must be an FFD. These topologies are illustrated in Figure 60.2.

- *Star topology*: Communication is established between a central device and other IEEE 802.15.4 devices. This central device is called a PAN coordinator and it is unique in the network.
- *Point to point topology*: IEEE 802.15.4 devices can form an ad hoc network. This standard does not define how to route data between nodes hence there must be a network layer implementator who is in charge of this task. Being able to route data allows the network to cover a wider area than the network topology but the power consumption is increased.

60.2.4.4 State of the Art

Recent publications in this area are in Table 60.5 and the sensors used are in Table 60.6. Two of the three publications focus on fall detection [22,41] and the other is a monitoring system [21].

It is important to mention that it is a common mistake to consider that IEEE 802.15.4 and ZigBee are the same, that is, Ref. [22]. As it will be presented next, the ZigBee is implemented over IEEE 802.15.4 in a different manner and with different characteristics.

The monitoring system described in Ref. [21] is composed by three different sensors, accelerometer, ECG, and saturation of peripheral oxygen (SpO_2). Sensors are located into a chest band (accelerometer and ECG) and a wristband (SpO_2). Although three different measures are acquired, the system is designed for general monitoring without a specific aim. References [22,41] present two systems for falling detection using a triaxial accelerometer. In Ref. [41], the information provided by the accelerometer is complemented with passive infrared (PIR) motion detectors. All three publications implement star topologies as they connect more than a sensor with a PAN coordinator.

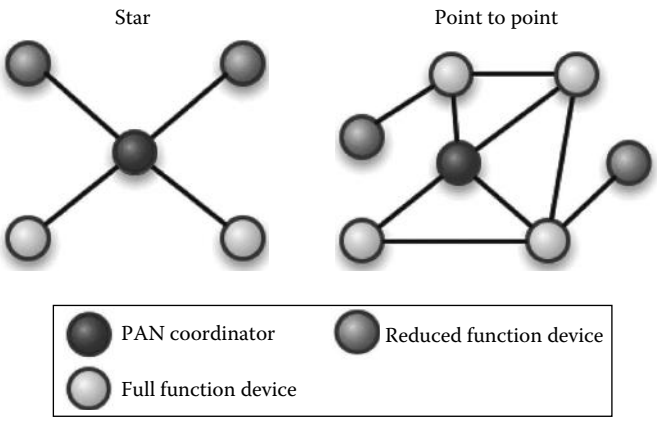


FIGURE 60.2 Topologies IEEE 802.15.4.

TABLE 60.5 Recent Publications on IEEE 802.15.4

Reference	Health Focus	Sensors	Battery Life
Chung et al. [21]	Monitoring	Accelerometer, ECG, SpO_2	N/A
Dinh et al. [22]	Fall detection	Accelerometer	N/A
Srinivasan et al. [41]	Fall detection	Accelerometer	N/A

TABLE 60.6 Sensors Used with IEEE 802.15.4

Sensors		
Chung et al. [21]	SpO_2, ECG, Accelerometer	Nonin OEM III module [12] Conductive fabric electrode MMA7260Q, Freescale [10]
Dinh et al. [22]	Accelerometer	ADIS16350/ADIS16355 iSensor [4]
Srinivasan et al. [41]	Accelerometer	N/A
	PIR motion detectors	N/A

60.2.5 IEEE 802.15.4/ZigBee

ZigBee is a specification set built for IEEE 802.15.4 protocol. ZigBee naming is derived from the waggle dance of honeybees after their return to the beehive.

ZigBee is designed to provide connectivity between high-efficiency devices with a reduced packet load. The aim of the standard is to offer low power consumption even if the bandwidth and range have to be sacrificed. It is mainly used in industrial control systems, embedded sensors, and healthcare. These applications benefit from the low node capacity and sleep modes. At this mode, power consumption is highly reduced. ZigBee is promoted by ZigBee Alliance, an international community supported by more than one hundred companies (Motorola, Mitsubishi, Phillips, Samsung, Honeywell, Siemens, etc.).

Because ZigBee is built using IEEE 802.15.4, it shares frequency, range, and bandwidth as described in the previous section.

60.2.5.1 ZigBee Devices

There are three different types of ZigBee devices with differing roles in the network:

- *ZigBee coordinator* (*ZC*): There must be one coordinator per network. It is in charge of controlling the network.
- *ZigBee router* (*ZR*): Connects separate devices in the network topology. It also provides a user application layer.
- *ZigBee end device* (*ZED*): End devices transmit information to ZR or ZC devices; they cannot route data (IEEE 802.15.4 RFD).

60.2.5.2 Topologies

There are three topologies defined for ZigBee networks, shown in Figure 60.3:

- *Star topology*: The communication is established between a central device and the other ZigBee devices, powered by batteries. This central device is called a PAN coordinator and it is unique in the network.
- *Mesh topology*: It is similar to the star topology. The difference between them is that any device can connect with other device, not just the PAN coordinator.

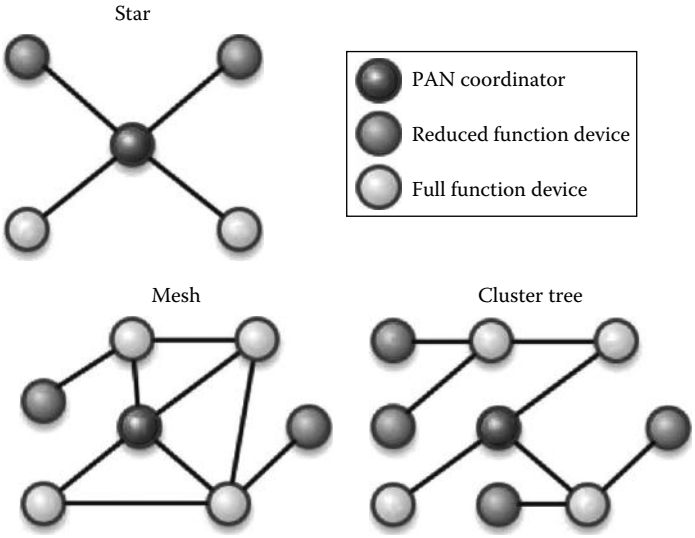


FIGURE 60.3 ZigBee topologies.

- *Cluster tree topology*: Cluster tree is a special scenario for a mesh topology where most of the devices are IEEE 802.15.4 FFD, and RFDs are connected to the network as leaves at the end of each branch. One of the IEEE 802.15.4 FFD is the PAN coordinator and the others can provide synchronization and coordination services.

60.2.5.3 State of the Art

ZigBee is one of the most popular transmission technologies in WSN in health. Its low power consumption increases the battery life of the systems. Table 60.7 summarizes the recent implementations that use this standard.

TABLE 60.7 Sensors Used with IEEE 802.15.4/ZigBee

Reference	Health Focus	Sensors	Battery Life	Device
Campo and Grangereaw [20]	Fall detection	Accelerometer Global positioning systems (GPS)	15 days	ADXL202 Analog Devices [5] SAM-LS U-BLOX
Kim et al. [28]	Respiratory rate monitoring	EMFi Ballistocardiogram	N/A	N/A N/A
Lai et al. [29]	Fall detection	Accelerometer Gyroscope	N/A	DEAMCC3D Design Engineering [9] IDG300 InvenSense [1]
Lee et al. [32]	Monitoring	ECG	24 h	N/A
Lee et al. [30]	Monitoring	PPG Accelerometer	N/A	BIOPAC systems [7] KXM52 series
Lee et al. [33]	Fall detection	ECG Accelerometer	10 h	N/A ADXL33 Analog Devices
Lou et al. [37]	Scoliosis monitoring	Force transducer	8 months	Honeywell FSO 1 [3]
Lou et al. [36]	Scoliosis monitoring	Force transducer	130 days	Honeywell FSO 1 [3]
Morris [et al. 39]	Body fluids analysis	pH Sodium	N/A	N/A N/A

Fall detection applications [20,29,33] are similar to the ones presented in the previous section. Falls are detected using an accelerometer but there are additional sensors that provide extra information in some cases. Global positioning system (GPS) [20] for location and ECG in Ref. [33] are collected for cardiac abnormality detection.

Monitoring is another typical focus on the application of WSN in health. References [36,37] are two interesting research papers about scoliosis monitoring. The trials of 24 h of continuous monitoring for 4–6 months are challenging to battery life of the sensors. At this point is where ZigBee shows its potential. Patients have to wear the sensors for long-time monitoring. As comfort is one of the main advantages of using WSN, patients shouldn't have to recharge sensors batteries quite often. Low power consumption of ZigBee becomes an important factor at these types of applications.

Other monitoring systems measure parameters such as respiratory rate [28], ECG [32], pulse wave signals [30] and bloody fluids analysis [39]. The battery life of these systems is important. Table 60.7 shows that there are ZigBee systems that can work continuously for more than a week [20] or even months [36,37]. Sensors used in these publications are also presented at Table 60.7.

60.2.5.4 Proprietary Solutions

Even if the most popular transmission technologies presented earlier are used extensively, there are other solutions where proprietary technologies are used as listed in Table 60.8. Some publications that use proprietary solutions are discussed later.

A radiotelemetry capsule is used to monitor pH, pressure, and temperature of the intestinal track [19]. The transceiver operates at the license-free 433.92 MHz industry, ISM band. It uses a frequency with ranges between 1 MHz and 1 GHz to avoid energy attenuation in human tissues. It covers a distance of about 2 m and has a battery life over 180 h.

A low-power wireless acquisition module is used within wearable health monitoring systems and ambient assisted living (AAL) applications [23]. The measures are acquired by three different sensors, ECG, accelerometer, and thermometer, but it does not aim to provide a specific solution. It focuses on general-purpose monitoring. The communications on this module is via a Toumaz Sensium system on chip (SoC) [16]. This solution operates at 868 MHz avoiding the ISM band because it is overly crowded. The battery life is high, over 90 h.

An example of fall detection system that implements a wireless proprietary solution is in Refs. [34,35]. The sensor used in this system is the usual one for fall detection, an accelerometer. The communication protocol in the link layer uses ShockBurst [13]. ShockBurst technology permits a low-cost microcontroller with a bit rate of 1 Mbit/s.

ShockBurst technology is also used in Ref. [25], a data acquisition system for diabetic and cardiac monitoring. Sensors used in these publications are presented at Table 60.8.

TABLE 60.8 Sensors Used with Proprietary Technologies

Reference	Health Focus	Sensor	Battery Life	Device
Biao et al. [19]	Monitoring intestinal motility	Pressure Thermometer pH	180 h	N/A N/A N/A
Figueiredo et al. [23]	Monitoring	ECG Accelerometer Thermometer	90 h	N/A LIS302DL, STMicroelectronics N/A
Lee et al. [34]	Fall detection	Accelerometer	N/A	ADXL202, Analog Device [5]
Lee and Lee [35]	Fall detection	Accelerometer	N/A	ADXL202, Analog Device [5]
Harvey et al. [25]	Monitoring	Glucose meter	N/A	OneTouch Ultra 2, LifeScan [11]

60.3 Conclusion

Several challenges need to be overcome to enrich the current implementations of WSN [45] in health applications. However, as presented in this chapter there are numerous implementations of WSN for health. The main applications focus on fall detection and general monitoring systems. It has been observed that IEEE 802.15.4 provides high battery life for WSN giving longer operational times and the ZigBee power consumption is much lower than Bluetooth.

In order to compare, the power consumption for IEEE 802.15.1 has been tested on a BlueCore2 and for IEEE 802.15.4 on a CC2430. Results are presented in Table 60.9, which shows that IEEE 802.15.4 power consumption is significantly smaller than Bluetooth both for transmission and reception. It is important to point out that the end devices in IEEE 802.15.4 networks can enter sleep modes, reducing their power consumption. In order to provide network transmission capabilities, they are usually attached to motes or a custom board.

Many different kinds of sensors have been used but some of them are not medical sensors, accelerometers, gyroscopes, etc. The sensors have not a smart status, which means they are not able to connect by themselves to a WSN. In order to provide them with network transmission capabilities they are usually attached to motes or a custom board.

In this chapter, it has been shown that there are several standards available in applications. But it is important to understand that the lack of any standard for the format of the data to be transmitted turns these solutions into isolated systems. The problem is the interoperability issues that reduce opportunities for integration in different platforms. A recent standard, IEEE 1451 [17], describes a set of open, common, network-independent communication interfaces for connecting transducers; it is not restricted to a single transmission technology, being able to operate with the most popular ones (ZigBee, Bluetooth, etc.). Other proposals use the X73, ISO/IEEE11073 standard, also defined by IEEE, which was originally designed to provide connectivity between medical devices.

Acronyms

AAL	ambient assisted living
BAN	body area network
ECG	electrocardiogram
EEG	electroencephalography
EOG	electrooculography
FFD	full function device
FHSS	frequency hopping spread spectrum
GPS	global positioning system
HCF	hybrid coordination function
ISM	industrial scientific medical
MAC	medium control access

TABLE 60.9 IEEE 802.15.4 and IEEE 802.15.1 Power Consumption

Standard	IEEE 802.15.4	IEEE 802.15.1
Chipset	CC2430	BlueCore2
VDD (V)	3.0	1.8
TX (mA)	24.7	57
RX (mA)	27	47
Bit rate (MB/s)	0.25	0.72

LR-WPANs	low-rate wireless personal area network
PAN	personal area network
PHY	physical layer
PPG	photoplethysmography
RFD	reduced function device
SpO_2	saturation of peripheral oxygen
TDMA	time division multiple access
UWB	ultra wide band
WLAN	wireless local area network
WPAN	wireless personal area network
WSN	wireless sensor network
ZC	ZigBee coordinator
ZED	ZigBee end device
ZR	ZigBee router

References

1. InvenSense Inc., http://invensense.com/mems/gyro/idg300.html
2. InvenSense Inc., http://invensense.com/mems/gyro/idg500.html
3. Honeywell S&C, http://sensing.honeywell.com/product20pageprid52508
4. Analog Devices, http://www.analog.com/en/mems-sensors/mems-inertial-sensors/adis16355/products/product.html
5. Analog Devices, http://www.analog.com/en/mems-sensors/mems-inertial-sensors/adxl202/products/product.html
6. Analog Devices, http://www.analog.com/en/mems-sensors/mems-inertial-sensors/adxl330/products/product.html
7. Biopac Systems, Inc. http://www.biopac.com/pulse-plethysmogram-amplifier
8. Cleveland Medical Devices Inc., http://www.clevemed.com/crystalmonitor/overview.shtml
9. Dimension Energy Services, http://www.dimensionengineering.com/datasheets/deaccm3d.pdf
10. Freescale TM, http://www.freescale.com/files/abstract/article/leadershipmma7260q.htmltidfs200505tsp6428senl
11. Lifescan, Johnson & Johnson Company, http://www.lifescan.es/
12. Nonin Medical, Inc., http://www.nonin.com/oemsolutions/oem-iii-module
13. Symmetry Electronics, http://www.semiconductorstore.com/pdf/newsite/nordic/wpnrf240xshockburst.pdf
14. Sensimed Company, http://www.sensimed.ch/en/products/sensimed-triggerfishr.html
15. Shimmer, http://www.shimmer-research.com
16. The Toumaz Group, http://www.toumaz.com/
17. IEEE Standard for a Smart Transducer Interface for Sensors and Actuators—Common Functions, Communication Protocols, and Transducer Electronic Data Sheet (TEDS) Formats. *IEEE Std* 1451.0–2007, pp. 1–335, 2007.
18. L. K. Au, A. A. T. Bui, M. A. Batalin, X. Xu, and W. J. Kaiser. CARER: Efficient dynamic sensing for continuous activity monitoring. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2011:2228–2232.
19. H. Biao, Y. Guozheng, and Z. Peng. Multi-sensor radiotelemetry system for intestinal motility measurement. *J. Med. Eng. Technol.*, 33(1):66–71, 2009.
20. E. Campo and E. Grangereau. Wireless fall sensor with GPS location for monitoring the elderly. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008:498–501.
21. W.-Y. Chung, Y.-D. Lee, and S.-J. Jung. A wireless sensor network compatible wearable u-healthcare monitoring system using integrated ECG, accelerometer and SpO2. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008:1529–1532, 2008.

22. A. Dinh, D. Teng, L. Chen, S. B. Ko, Y. Shi, J. Basran, and V. Del Bello- Hass. Data acquisition system using six degree-of-freedom inertia sensor and ZigBee wireless link for fall detection and prevention. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008:2353–2356, 2008.
23. C. P. Figueiredo, K. Becher, K. P. Hoffmann, and P. M. Mendes. Low power wireless acquisition module for wearable health monitoring systems. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 1:704–707, 2010.
24. B. R. Greene, D. McGrath, K. J. O'Donovan, R. O'Neill, A. Burns, and B. Caulfield. Adaptive estimation of temporal gait parameters using body-worn gyroscopes. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 1:1296–1299, 2010.
25. P. Harvey, B. Woodward, S. Datta, and D. Mulvaney. Data acquisition in a wireless diabetic and cardiac monitoring system. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2011:3154–3157, 2011.
26. R. S. H. Istepanian, K. Zitouni, D. Harry, N. Moutosammy, A. Sungoor, B. Tang, and K. A. Earle. Evaluation of a mobile phone telemonitoring system for glycaemic control in patients with diabetes. *J. Telemed. Telecare*, 15(3):125–128, 2009.
27. H. A. Kayyali, S. Weimer, C. Frederick, C. Martin, D. Basa, J. A. Juguilon, and F. Jugilioni. Remotely attended home monitoring of sleep disorders. *Telemed. J e-health: Theor. J. Am. Telemed. Assoc.*, 14(4):371–374, May 2008.
28. Y. S. Kim, H. J. Baek, J. S. Kim, H. B. Lee, J. M. Choi, and K. S. Park. Helmet-based physiological signal monitoring system. *Eur. J. Appl. Physiol.*, 105(3):365–372, February 2009.
29. D. T. H. Lai, E. Charry, R. Begg, and M. Palaniswami. A prototype wireless inertial-sensing device for measuring toe clearance. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008:4899–4902, 2008.
30. E. M. Lee, J. Y. Shin, J. H. Hong, E. J. Cha, and T. S. Lee. Glass-type wireless PPG measuring system. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 1:1433–1436, 2010.
31. J.-S. Lee, Y.-W. Su, and C.-C. Shen. A Comparative study of wireless protocols: Bluetooth, UWB, ZigBee, and Wi-Fi. *Industrial Electronics Soc. IECON 2007. 33rd Annu. Conf. IEEE*, pp. 46–51, Taipei, Taiwan, 2007.
32. S. Lee, J. Kim, J. Kim, and M. Lee. A design of the u-health monitoring system using a Nintendo DS game machine. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2009:1695–1698, 2009.
33. S. Lee, J. Kim, and M. Lee. The design of the m-health service application using a Nintendo DS game console. *Telemed. J. e-health Amer. Telemed. Assoc.*, 17(2):124–130, March 2011.
34. Y. Lee, J. Kim, M. Son, and M. Lee. Implementation of accelerometer sensor module and fall detection monitoring system based on wireless sensor network. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2007:2315–2318, 2007.
35. Y. Lee and M. Lee. Accelerometer sensor module and fall detection monitoring system based on wireless sensor network for e-health applications. *Telemed. J. e-health Amer. Telemed. Assoc.*, 14(6):587–592, August 2008.
36. E. Lou, D. Hill, D. Hedden, J. Mahood, M. Moreau, and J. Raso. An objective measurement of brace usage for the treatment of adolescent idiopathic scoliosis. *Med. Eng. Phys.*, 33(3):290–294, April 2011.
37. E. Lou, D. L. Hill, and J. V. Raso. A wireless sensor network system to determine biomechanics of spinal braces during daily living. *Med. Biol. Eng. Comput.*, 48(3):235–243, March 2010.
38. K. Mansouri and T. Shaarawy. Continuous intraocular pressure monitoring with a wireless ocular telemetry sensor: initial clinical experience in patients with open angle glaucoma. *Br. J. Ophthalmol.*, 95(5): 627–629, 2011.
39. D. Morris, B. Schazmann, Y. Wu, S. Coyle, S. Brady, C. Fay, J. Hayes, K. T. Lau, G. Wallace, and D. Diamond. Wearable technology for bio-chemical analysis of body fluids during exercise. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008:5741–5744, 2008.
40. M.-Z. Poh, N. C. Swenson, and R. W. Picard. Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography. *IEEE Trans. Inform. Technol. Biomed.*, 14(3):786–794, May 2010.

41. S. Srinivasan, J. Han, D. Lal, and A. Gacic. Towards automatic detection of falls using wireless sensors. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc*., 2007:1379–1382, 2007.
42. H. Tawa, Y. Yonezawa, H. Maki, H. Ogawa, I. Ninomiya, K. Sada, S. Hamada, and W. M. Caldwell. A wireless breathing-training support system for kinesitherapy. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc*., 2009:5179–5182, 2009.
43. R. C. Wagenaar, I. Sapir, Y. Zhang, S. Markovic, L. M. Vaina, and T. D. C. Little. Continuous monitoring of functional activities using wearable, wireless gyroscope and accelerometer technology. *Conf. Proc. IEEE Eng. Med. Biol. Soc*., 2011:4844–4847.
44. X. Wang, Y. Ren, J. Zhao, Z. Guo, and R. Yao. Comparison of IEEE 802.11e and IEEE 802.15.3 MAC. *Proc. IEEE 6th Circuits Syst. Symposium. Emerging Technologies: Frontiers Mobile Wireless Communication*, 2: 675–680, 2004.
45. E. D. Zubiete, L. F. Luque, A. V. M. Rodríguez, and I. G. González. Review of wireless sensors networks in health applications. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2011:1789–1793, 2011.

61

Sensors for Electrical Impedance Tomography

Tushar Kanti Bera
Indian Institute of Science, Bangalore

J. Nagaraju
Indian Institute of Science, Bangalore

61.1 Computed Tomography

In medical computed tomography (CT) [1] some kind of energy (such as x-rays, optical, electrical, sound, near infrared, etc.) is applied to the subject under test (SUT), and the modified or transformed form of the energy is collected or detected and compared with the incident one to reconstruct the spatial distribution of the tissue properties of the SUT called the parameters of interest (POI). The POI of the SUT gives anatomical, physiological (biophysiological or physiochemical), or pathological information. Hence, in medical tomographic imaging, the energy used to produce the image must penetrate the body tissues [2] to get information about the tissues. If the energy applied to the SUT does not experience any type of interaction (e.g., absorption, attenuation, scattering), the detected energy would not

contain any useful information about the internal anatomy of the SUT [2]. Therefore, it would not be possible to construct an image of the anatomy using that information. In the x-ray computed tomographic technique, x-rays are applied to the boundary of the object under test, and the attenuated x-rays at the boundary are collected for the different projection angles to obtain a complete scan around a 360° rotation. Thus, the x-ray CT involves gathering the output x-rays (X_{Out}) generated by applied x-ray signal (X_{In}) at the boundary from multiple directions and reconstructing the domain parameter relating the X_{Out} to X_{In} (POI) using a software called reconstruction algorithm in a personal computer (PC).

61.2 Electrical Impedance Tomography

Electrical impedance tomography (EIT) [3–7] is an image reconstruction technique in which the electrical conductivity or resistivity of a conducting domain (Ω) is reconstructed from the surface potentials developed by the current signals injected at the domain boundary ($\partial\Omega$). In EIT, the electrical signals (current or voltage) are applied to the boundary of the SUT through different pairs of the surface electrodes (yielding different current projections) attached to the SUT, and the signals generated at the boundary (voltage or current) are collected (Figure 61.1) across the pairs of electrodes for all current projections to obtain a complete scan (around the entire boundary or 360° angular orientation). Thus, the EIT involves gathering the output signal, Y (voltage [V] or current [I]), generated by applying an input signal X (current (I) or voltage (V)) at the boundary from multiple directions and reconstructing the domain parameter (POI) or the transfer function, T (impedance [Z] to admittance [Y]), using a software called reconstruction algorithm in a PC.

In EIT, if the input signal injected to the object boundary is voltage, the method is called electrical potential tomography or applied potential tomography (APT) [8]. On the other hand, if the current signal is applied to the object boundary as an input signal, the method is termed as electrical current tomography or applied current tomography (ACT). Generally, most of the literature used the abbreviation "EIT" to indicate the current tomography, although the term EIT does not indicate the type of excitation (current or voltage). Like most of the literature, in this chapter also, we will use the term EIT for indicating the electrical current tomography unless stated otherwise.

61.3 How It Works?

In EIT, a low-frequency constant sinusoidal current is injected to the domain under test (Ω) through the surface electrodes (shown in rectangles on the domain periphery in Figure 61.2) attached to the domain boundary ($\partial\Omega$), and the boundary potentials are measured on the other electrodes using a

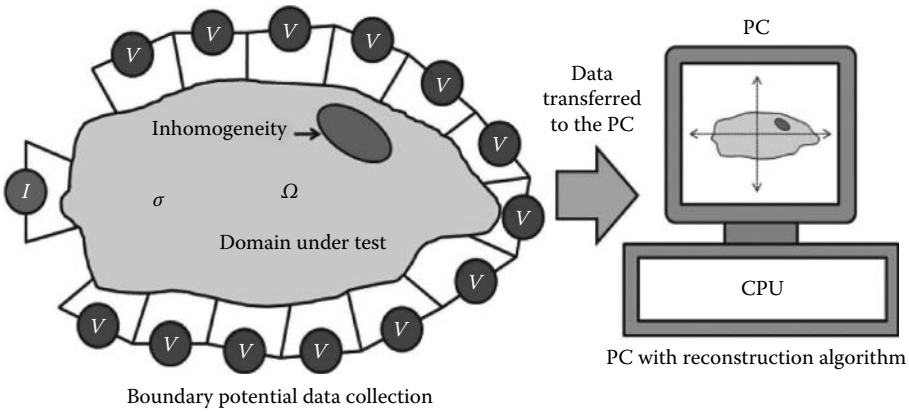


FIGURE 61.1 EIT schematic.

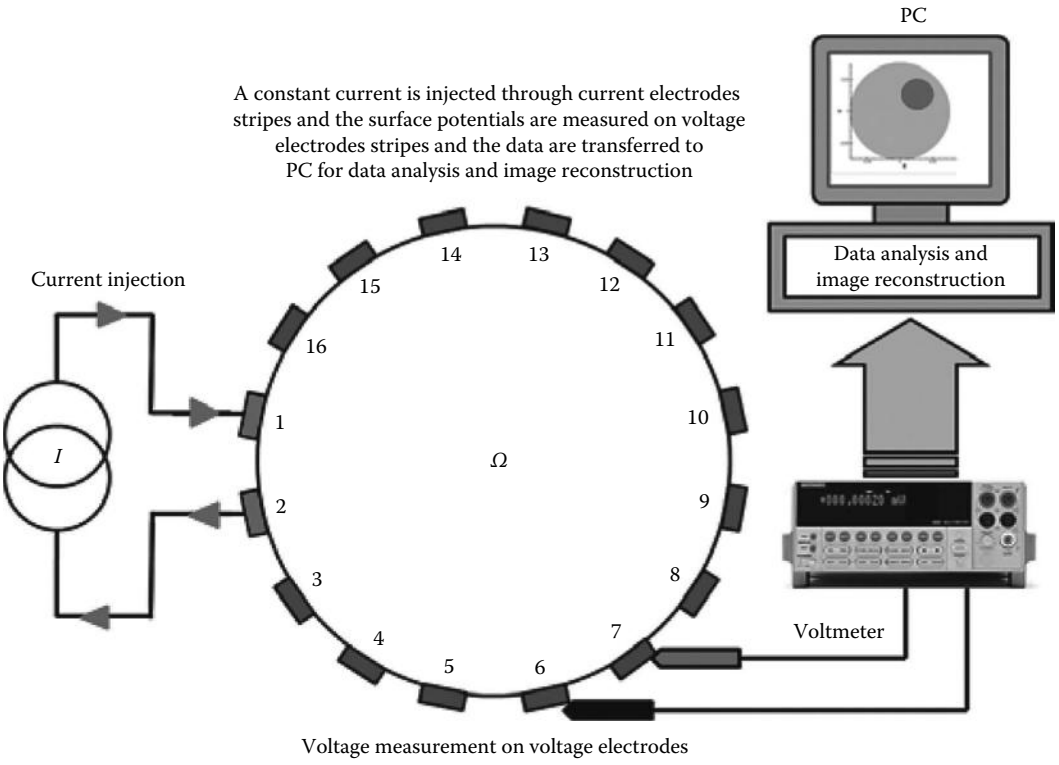


FIGURE 61.2 How EIT works.

voltmeter or any other voltage-measuring device. The electrodes, through which the current signal is injected, are called current electrodes (EC), shown as electrode 1 and 2 in Figure 61.2, and the electrodes on which the voltage data are measured are termed as voltage electrodes (EV), shown as electrode 3 through 16 in Figure 61.2. The voltage data are transferred to the PC for data analysis and image reconstruction (Figure 61.2). The spatial distribution of the electrical resistivity or the conductivity is reconstructed from the boundary data using a mathematical algorithm or computer program called the image reconstruction algorithm [1–4].

When a constant current is injected to the domain under test in EIT, the current fluxes are produced and the current signal interacts with the medium of the domain. The profile of the current fluxes produced within the domain depends on the electrical properties of the medium of the domain, that is, its impedance profile. If the medium is homogeneous (Figure 61.3a), the current flux lines will be symmetric; whereas the inhomogeneous domain (domain with inhomogeneity) will distort the current fluxes (Figure 61.3b).

As the current fluxes develop the voltages at the points interacted with the current fluxes, the potential distribution profile within the domain depends on the profile of the current flux. Similarly, the voltages developed at the domain boundary will also depend on the current flux profile. As the profile of the current flux depends on the profile of the impedance or conductivity distribution of the domain, the domain voltage profile will depend on its impedance profile. Therefore, the boundary potential profile will provide the information about the internal impedance profile of the domain under test. The boundary potential profile for a homogeneous medium will be different from a domain with inhomogeneity. If we have a boundary potential data set and an idea of the domain properties (initial guess of the impedance or the conductivity or resistivity), then we can reconstruct the impedance (or conductivity or resistivity) profile of the domain from the measured boundary data by comparing it from the calculated data.

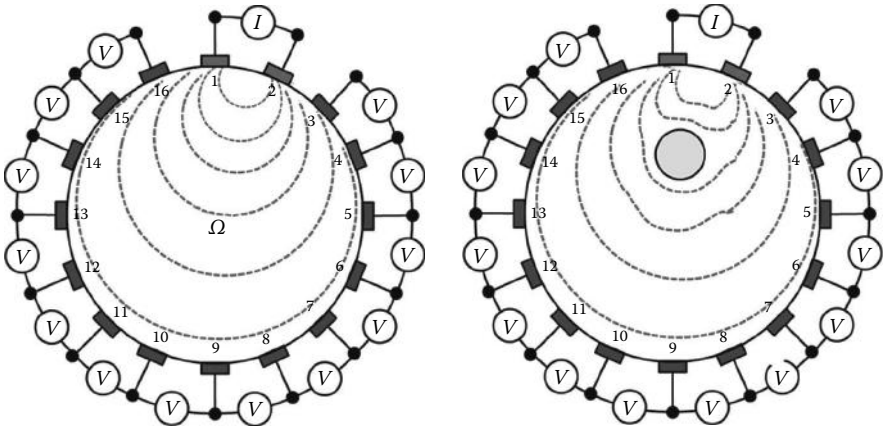


FIGURE 61.3 How EIT works: (a) current flux in homogeneous medium and (b) current flux in medium with inhomogeneity.

61.4 EIT Hardware

An EIT system is developed with mainly three parts: EIT sensors or surface electrodes [3,9], electronic instrumentation [10,11], and a PC (Figure 61.4) with reconstruction algorithm [12–16]. The object to be imaged in EIT is surrounded by the surface electrodes encircling the object boundary, and the constant current signal is injected through two particular electrodes (called current electrodes), and the surface

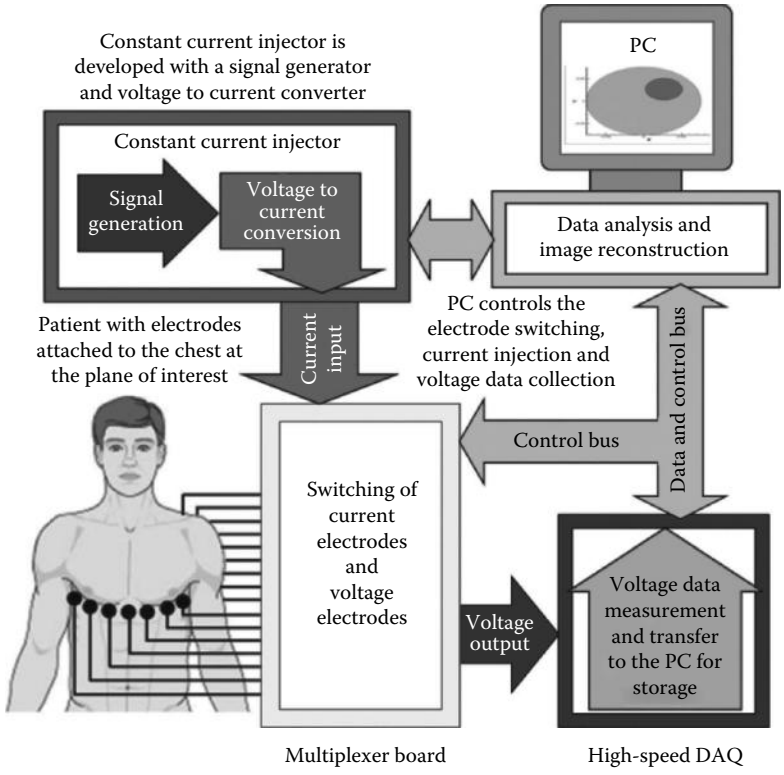


FIGURE 61.4 A medical EIT schematic.

potentials are measured on some particular electrodes (called voltage electrodes). The electrodes or sensors in EIT interface the object under test to the EIT instrumentation for current injection and voltage measurement.

61.5 Sensors and Electrodes

Sensors [17] are the signal sensing devices or signal detectors that measure a physical quantity and send by converting it (sometimes without converting although) into a readable and observable signal to an observer or analyzable and measurable signal to an instrument. For example, a medical thermometer [17] converts the temperature (input or measurand) into a volumetric change (observable output) of a mercury or other liquid, which can be observed and read on a calibrated transparent glass tube. Similarly, a thermocouple [18] converts temperature to an output voltage, which can be observed, measured, and read by an electronic or electrical voltmeter. Sensors are generally tested and calibrated against the known standards for reducing the measurement errors. Electrodes [19] are the sensors that are used to supply or sense the electrical signal (voltage or current) to or from any system. For example, the filaments of a general fluorescent lamp act as the electrodes.

A simple electrolytic cell contains two electrodes: one, called anode, supplies the electrical current (receives the electrons) and the other one, called cathode, receives the electrical current (supplies the electrons). In an x-ray tube, the heater filament and the tungsten target also work as the electrodes: filament supplies the electrons (called cathode) and tungsten target receives the electrons and generates the x-rays. The biomedical electrodes are the electrodes that are used to supply or sense the electrical signals (voltage or current) in a biomedical system. Biopotential electrodes are the transducers that are used to measure the surface potentials by converting the ionic current (caused by the movement of ions) in the body into the electronic current (caused by the movement of electrons). In some biomedical measurement or monitoring processes, current or voltage signals are injected to the body to observe, read, and measure the responses of the body tissue using biopotential or biocurrent electrodes.

61.6 Sensors and Electrodes for EIT

Surface electrodes [3,19] interfacing the EIT hardware and computer with the object under test (i.e., patient) are required to inject a constant current and measure the surface potentials. EIT electrodes generally play a very crucial role in the boundary data accuracy as a number of errors are contributed by the electrodes, electrode–patient interfaces/contacts, and electrode–wire connections. In EIT, generally, $16^{*}N$ ($N = 1, 2, 3, \ldots$) electrodes are used, and a constant current is injected to the object boundary through the different current electrode pairs (called current electrode pairs). The differential potentials are measured across the other pairs of electrodes (called voltage electrodes pairs) using the four-probe measurement technique.

61.6.1 Four-Probe Measurements in EIT

In EIT, a low-frequency constant sinusoidal current is injected to the domain under test through two particular electrodes, and the boundary potentials are measured between the other electrodes (Figure 61.5). The current is injected through different current electrode pairs, and for each current injection through the current electrode pairs, voltage data are collected from the voltage electrode pairs except the electrode pairs containing one or more current electrodes (Figure 61.6). Hence, in EIT, the current injection and voltage measurement are conducted using the repetitive four-electrode method (Figure 61.6) or tetrapolar method [8,20,21]. The four-electrode method is an impedance measurement technique in which a constant current is injected through two electrodes (outer electrodes) called current electrodes and the voltages are measured on other two electrodes (inner electrodes) named as voltage electrodes [20] as shown in Figure 61.6. On the contrary, the two-probe or two-electrode method uses only two electrodes for both

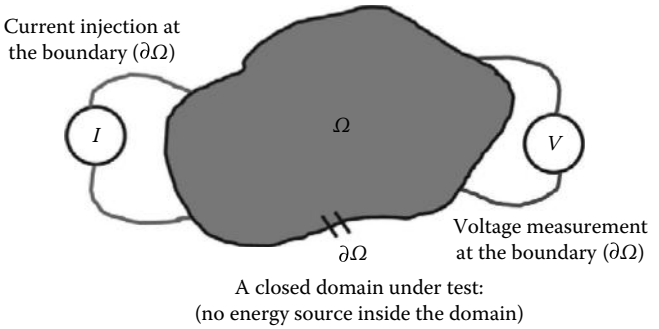


FIGURE 61.5 Four-electrode method in EIT.

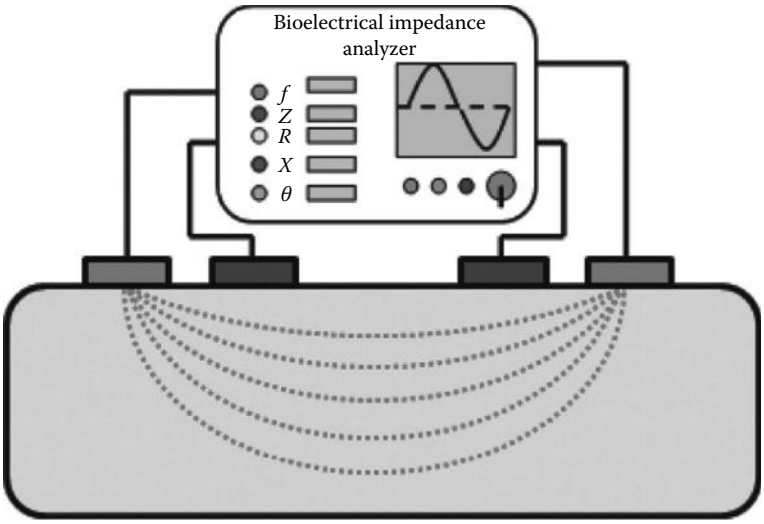


FIGURE 61.6 Four-electrode method for impedance measurement.

the current injection and the voltage measurement. As the same electrode pair is used for current injection and voltage measurement in the two electrode method, it has a contact impedance problem [8]. In two-electrode systems, as the voltages are measured on the current-carrying electrodes, the voltage drop (developed due to the high current magnitude) for the electrode contact impedance is also included with the measured data. Hence, in the two-probe method, the electrode contact impedance reduces the system accuracy. In the four-electrode method, the current injection and voltage measurement are done with two separate electrode pairs, and hence, the electrode contact impedance problem is neglected and the measurement accuracy is improved remarkably.

Boundary potentials in EIT are measured by injecting the current signal through the different current electrode pairs. The current injection through a particular current electrode pair and corresponding voltage measurement process is known as a current projection (P_n). Figure 61.7 shows the projection one (P_1). The voltage data are collected for different projections by injecting current through the different current electrode pairs to obtain a full data set from a complete scan of the object similar to the data collection process conducted at different projection angles in CT. Hence, in EIT also, the voltage data are collected for all the possible projections for obtaining the full set of boundary data.

For an EIT system with N number of electrodes, there will be N current projections (P_n: n = 1, 2, 3, …, N). Hence, a 16-electrode EIT system will yield 16 current projections. The current signal in a 16-electrode EIT system can be injected through 16 different current electrode pairs (E_1–E_2, E_2–E_3, E_3–E_4, E_4–E_5, E_5–E_6, E_6–E_7, E_7–E_8, E_8–E_9, E_9–E_{10}, E_{10}–E_{11}, E_{11}–E_{12}, E_{12}–E_{13}, E_{13}–E_{14}, E_{14}–E_{15}, E_{15}–E_{16}, and E_{16}–E_1),

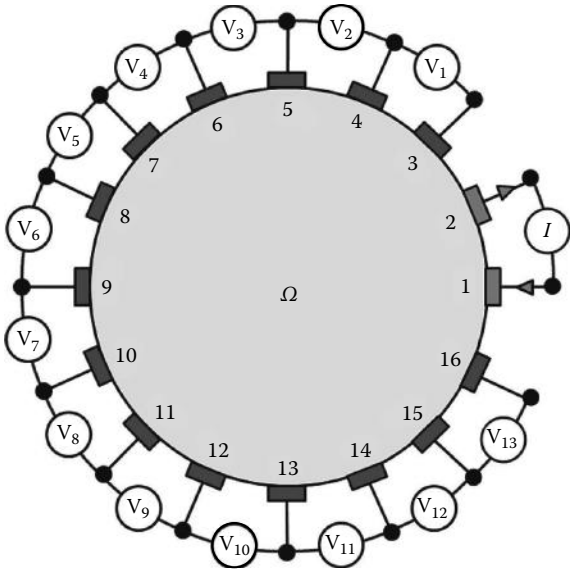


FIGURE 61.7 EIT data collection with repetitive four-electrode method.

which produce 16 current projections P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9, P_{10}, P_{11}, P_{12}, P_{13}, P_{14}, P_{15}, and P_{16}, respectively (P_1 has been shown in Figure 61.7). The voltage data are collected for all the projections (P_1 through P_{16}) for obtaining the full set of boundary data. In all the current projections, all the single voltage data measurement is conducted with a four-electrode method (Figure 61.6) as discussed earlier. For a particular current projection (say P_1: current injected through, say, E_1 and E_2), the voltage data are measured from all the possible voltage electrode pairs with four-electrode method keeping the current electrode unchanged. And after completing the P_1, current electrode pair is changed to another pair, say E_2 and E_3, to inject the current signal through them and start the data collection through the second projection P_2, and all the voltage data are collected with the four-electrode method. Hence, the EIT data collection is conducted using a repetitive four-electrode method (Figure 61.6).

61.6.2 Current Injection Methods in EIT

In EIT, a low-frequency constant sinusoidal current is injected to the domain under test through the electrodes called current electrodes. The boundary potentials are measured on the electrodes called voltage electrodes. The method or pattern in which the current is injected to the object under imaging is called the current pattern or current injection protocol [3,4,9]. The famous four different current injection protocols as reported by the researchers are summarized in the below. In general, the differential potentials are measured across the different electrode pairs containing one or both the current electrodes to avoid the contact impedance problem [3]. In some cases [9,22], however, in spite of the problem of skin impedance, to obtain the greatest sensitivity [22] to the changes in the resistivity of the body, voltages on current electrodes are also measured.

61.6.2.1 Neighboring Method

Brown and Segar [23] proposed a current injection method called the neighboring or adjacent current injection protocol (Figure 61.8). In this method the current is applied through two neighboring or adjacent electrodes, and the differential potentials are measured successively from all other adjacent electrode pairs excluding the pairs containing one or both the current electrodes. Figure 61.8 illustrates the neighboring method for a 16-electrode EIT system with a circular domain under imaging (DUI) surrounded by 16 surface electrodes denoted as electrode 1 through electrode 16 (E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8, E_9, E_{10}, E_{11}, E_{12}, E_{13}, E_{14}, E_{15}, and E_{16}). In the neighboring method, the current is injected through all the possible

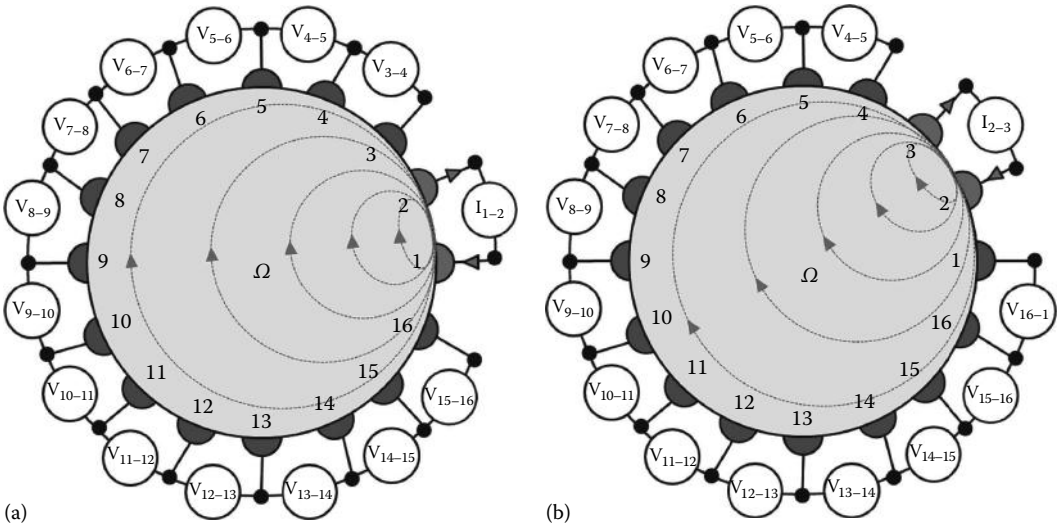


FIGURE 61.8 Neighboring or adjacent current method of boundary potential data collection illustrated for a cylindrical volume conductor and 16 equally spaced electrodes: (a) projection 1 and (b) projection 2.

adjacent electrode pairs and the differential voltage data are measured from the voltage electrode pairs. As discussed earlier, the current injection through a particular current electrode pair and corresponding voltage measurements on voltage electrode pairs is called a current projection. Hence, for a 16-electrode EIT system, there will be 16 current projections (P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9, P_{10}, P_{11}, P_{12}, P_{13}, P_{14}, P_{15}, and P_{16}) obtained by injecting the current signals through the 16 possible adjacent current electrode pairs: (E_1–E_2), (E_2–E_3), (E_3–E_4), (E_4–E_5), (E_5–E_6), (E_6–E_7), (E_7–E_8), (E_8–E_9), (E_9–E_{10}), (E_{10}–E_{11}), (E_{11}–E_{12}), (E_{12}–E_{13}), (E_{13}–E_{14}), (E_{14}–E_{15}), (E_{15}–E_{16}), and (E_{16}–E_1), respectively.

Hence, for a 16-electrode EIT system, the first current projection (P_1) of the neighboring data collection method (Figure 61.8a) injects the current signal through electrode 1 (E_1) and electrode 2 (E_2) and measures the 13 differential potentials (V_d: V_1, V_2, V_3, …, V_{13}) successively across the 13 electrode pairs E_3–E_4, E_4–E_5, …, and E_{15}–E_{16} (Figure 61.8a), respectively. Therefore, the first current projection yields 13 differential voltage data: V_1 or V_{3-4} (measured between E_3 and E_4), V_2 or V_{4-5} (measured between E_4 and E_5)…, and V_{13} or V_{15-16} (measured between E_{15} and E_{16}). The 13 differential voltage data obtained in this process are assumed to represent the impedance between the equipotential lines intersecting their corresponding measurement electrodes [3] and are all independent to each other [3]. In the neighboring method, the current density is, of course, highest between the current electrodes (E_1 and E_2 for the projection 1), decreasing rapidly as a function of distance [3]. Similarly, each of all the other current projections (P_2, P_3, …, P_{15}, and P_{16}) yields 13 voltage data and hence a 16-electrode system, and the adjacent current injection method yields $16 \times 13 = 208$ voltage measurements. Figure 61.8b shows the data collection process in P_2. But due to reciprocity [3], the measurements made on the boundary yield identical voltage data in which the current electrodes and voltage electrodes are interchanged [3]. Therefore, among the 208 differential voltage measurements, only 104 data are independent of each other. In the neighboring method, the measured voltage is at a maximum with adjacent electrode pairs, while in the opposite electrode pairs, the voltage is only about 2.5% of that [3].

61.6.2.2 Opposite Method

Hua et al. [24] proposed an alternative current injection method called the *opposite method* (Figure 61.9). In this method current is injected through the current electrodes placed diametrically opposite to each other, and hence, the method is known as the opposite method. Differential potentials are measured on the voltage electrodes with respect to the electrode (called as the voltage reference electrode) adjacent to the current-injecting electrode (Figure 61.9).

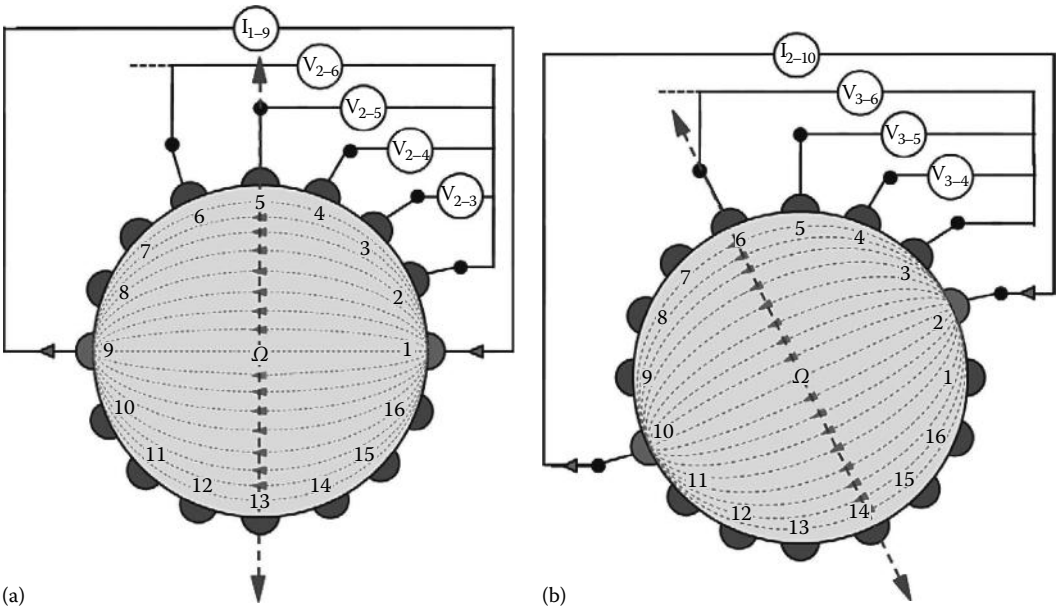


FIGURE 61.9 Opposite method of boundary potential data collection: (a) projection 1 and (b) projection 2.

Figure 61.9 illustrates the data collection strategies of opposite method for a 16-electrode EIT system with a circular DUI surrounded by 16 surface electrodes (E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8, E_9, E_{10}, E_{11}, E_{12}, E_{13}, E_{14}, E_{15}, and E_{16}). In opposite current injection method, the current is injected through all the possible opposite electrode pairs, and the differential voltage data are measured on all the voltage electrode pairs obtained with 13 voltage electrodes (E_V) and one voltage reference electrodes (E_{VR}). Hence, a 16-electrode EIT system will yield 16 current projections (P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9, P_{10}, P_{11}, P_{12}, P_{13}, P_{14}, P_{15}, and P_{16}) obtained by injecting the current signals through the 16 possible opposite current electrode pairs: (E_1–E_9), (E_2–E_{10}), (E_3–E_{10}), (E_4–E_{12}), (E_5–E_{13}), (E_6–E_{14}), (E_7–E_{15}), (E_8–E_{16}), (E_9–E_1), (E_{10}–E_2), (E_{11}–E_3), (E_{12}–E_4), (E_{13}–E_5), (E_{14}–E_6), (E_{15}–E_7), and (E_{16}–E_8), respectively.

Therefore, in the opposite current injection method for a 16-electrode EIT system, the first current projection (P_1) (Figure 61.9a) injects the current signal through electrode 1 (E_1) and electrode 9 (E_9). In P_1, the 13 differential potentials (V_d: V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, V_{10}, V_{11}, V_{12}, and V_{13},) are successively measured across the 13 voltage electrode pairs obtained by combining the 13 voltage electrodes (E_V: E_3, E_4, E_5, E_6, E_7, E_8, E_{10}, E_{11}, E_{12}, E_{13}, E_{14}, E_{15}, and E_{16}) with the voltage reference electrodes ($E_{VR} = E_2$). Therefore, the P_1 yields 13 differential voltage data $V_1 = V_{2-3}$, $V_2 = V_{2-4}$, $V_3 = V_{2-5}$, $V_4 = V_{2-6}$, $V_5 = V_{2-7}$, $V_6 = V_{2-8}$, $V_7 = V_{2-10}$, $V_8 = V_{2-11}$, $V_9 = V_{2-12}$, $V_{10} = V_{2-13}$, $V_{11} = V_{2-14}$, $V_{12} = V_{2-15}$, and $V_{13} = V_{2-16}$ across the electrode pairs E_2–E_3, E_2–E_4, E_2–E_5, E_2–E_6, E_2–E_7, E_2–E_8, E_2–E_{10}, E_2–E_{11}, E_2–E_{12}, E_2–E_{13}, E_2–E_{14}, E_2–E_{15}, and E_2–E_{16} (Figure 16.9a), respectively.

Similarly, in the opposite method, each of all the other current projections (P_2, P_3, …, P_{15}, and P_{16}) yields 13 voltage data and hence, for a 16-electrode system, a complete scan with opposite current injection method yields $16 \times 13 = 208$ voltage measurements. Figure 61.9b shows the data collection process in P_2. But, in the opposite method also, due to the reciprocity [3], the measurements made on the boundary yield identical voltage data in which the E_C and E_V are interchanged [3]. Therefore, among the 208 differential voltage measurements made in the opposite method, only 104 data are independent from each other [3]. In the opposite method, the current distribution is more uniform [3] and, therefore, has a good sensitivity [3].

61.6.2.3 Cross Method

Hua et al. [24] proposed a current injection method called the *cross method* (Figure 61.10) in which a more uniform current distribution is obtained when the current signal is injected between a pair of more distant electrodes [24]. As shown in Figure 61.10a, in the cross method, adjacent electrodes E_{16} and

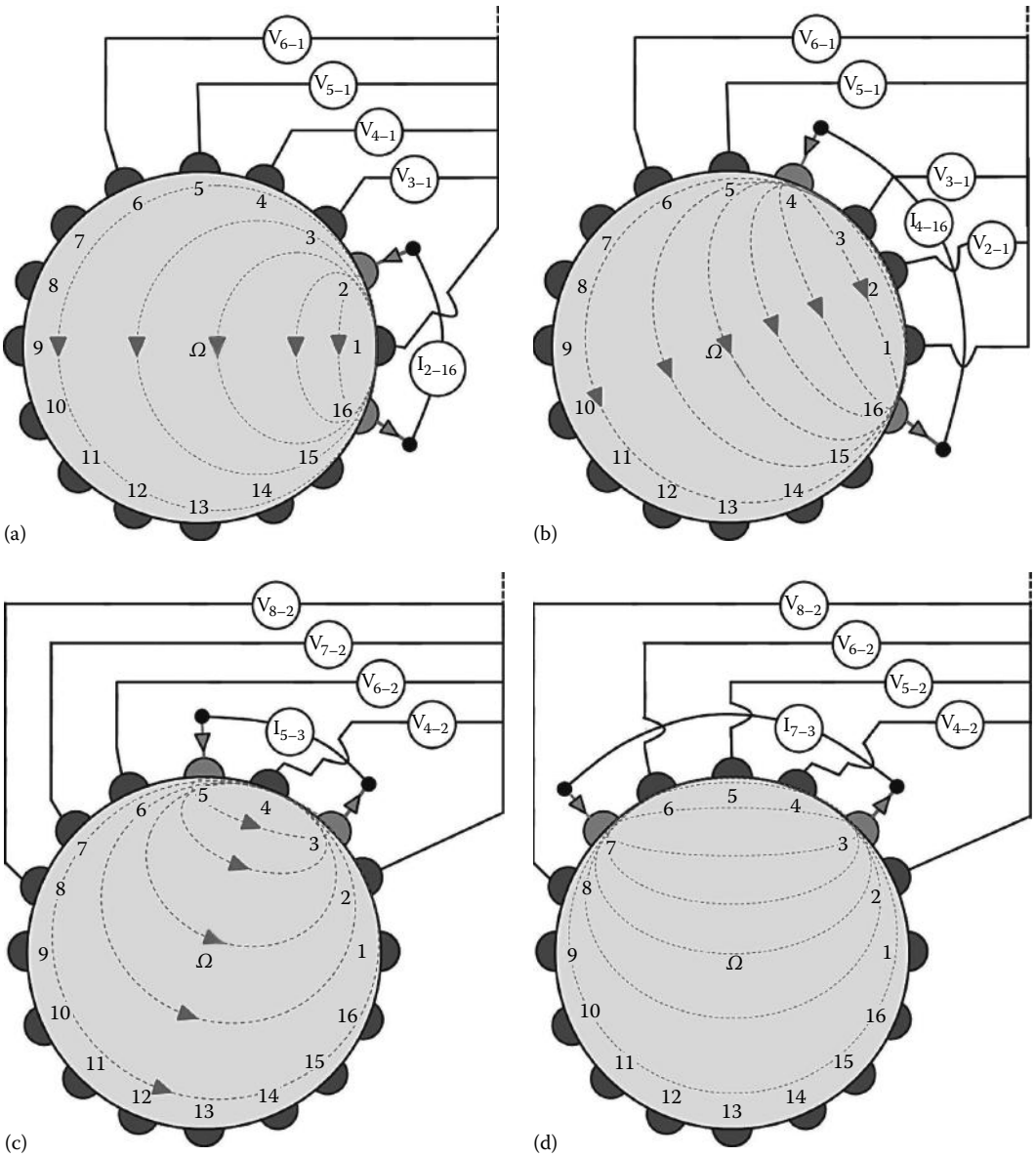


FIGURE 61.10 Cross method of boundary data collection: four different steps of this procedure are illustrated in (a) through (d).

E_1 are first selected as the current reference electrode (E_{CR}) and voltage reference electrode (E_{VR}), respectively. The positive terminal of the current source is connected to the electrode E_2, which acts as the positive current-injecting electrode. The 13 differential voltage data (V_d: V_1, V_2, V_3, … V_{13}) are measured successively across the 13 voltage electrode pairs obtained by combining 13 voltage electrodes (E_V: E_3, E_4, E_5, E_6, E_7, E_8, E_9, E_{10}, E_{11}, E_{12}, E_{13}, E_{14}, and E_{15}) with the voltage reference electrodes ($E_{VR} = E_1$) as shown in Figure 61.10a. Therefore, in the cross method, the P_1 yields 13 differential voltage data $V_1 = V_{1-3}$, $V_2 = V_{1-4}$, $V_3 = V_{1-5}$, $V_4 = V_{1-6}$, $V_5 = V_{1-7}$, $V_6 = V_{1-8}$, $V_7 = V_{1-9}$, $V_8 = V_{1-10}$, $V_9 = V_{1-11}$, $V_{10} = V_{1-12}$, $V_{11} = V_{1-13}$, $V_{12} = V_{1-14}$, and $V_{13} = V_{1-15}$ measured across the electrode pairs E_1–E_3, E_1–E_4, E_1–E_5, E_1–E_6, E_1–E_7, E_1–E_8, E_1–E_9, E_1–E_{10}, E_1–E_{11}, E_1–E_{12}, E_1–E_{13}, E_1–E_{14}, and E_1–E_{15} (Figure 61.10a), respectively.

In projection two (Figure 61.10b) another set of 13 differential voltage data are collected across the 13 voltage electrode pairs obtained by combining the 13 voltage electrodes (E_V: E_3, E_4, E_5, E_6, E_7, E_8, E_9, E_{10}, E_{11}, E_{12}, E_{13}, E_{14}, and E_{15}) with the voltage reference electrodes ($E_{VR} = E_1$). In this projection, the current injection is made by configuring the electrode 4 (E_4) as the positive current electrode and electrodes 16 (E_{16}) as the current reference (Figure 61.10b). With this current injection, the 13 differential potentials (V_d: V_1, V_2, V_3, … V_{13}) are measured successively across the 13 voltage electrode pairs obtained by combining the 13 voltage electrodes (E_V: E_2, E_3, E_5, E_6, E_7, E_8, E_9, E_{10}, E_{11}, E_{12}, E_{13}, E_{14} and E_{15}) with the voltage reference electrodes ($E_{VR} = E_1$) as shown in Figure 61.10b. Hence, in the cross method, the P_2 yields 13 differential voltage data $V_1 = V_{1-2}$, $V_2 = V_{1-3}$, $V_3 = V_{1-5}$, $V_4 = V_{1-6}$, $V_5 = V_{1-7}$, $V_6 = V_{1-8}$, $V_7 = V_{1-9}$, $V_8 = V_{1-10}$, $V_9 = V_{1-11}$, $V_{10} = V_{1-12}$, $V_{11} = V_{1-13}$, $V_{12} = V_{1-14}$, and $V_{13} = V_{1-15}$ measured across the electrode pairs E_1–E_2, E_1–E_3, E_1–E_5, E_1–E_6, E_1–E_7, E_1–E_8, E_1–E_9, E_1–E_{10}, E_1–E_{11}, E_1–E_{12}, E_1–E_{13}, E_1–E_{14}, and E_1–E_{15} (Figure 61.10b), respectively. Repeating this same procedure with a current injection by connecting the positive terminal of the current source to the electrodes E_6, E_8, …, E_{14}, other five projections (P_3, P_4, P_5, P_6, and P_7) are obtained. Hence, the cross method yields first seven projections (P_1 P_2 P_3 … P_6, and P_7) with the electrodes E_{16} and E_1 as current reference electrodes and voltage reference electrodes, respectively. Therefore, in this way the first seven current projections (P_1, P_2, P_3, …, P_6, and P_7) in cross method yield $7 \times 13 = 91$ measurements.

After collecting the 91 data from the projection P_1 through P_7, a similar measurement sequence is then repeated configuring electrodes E_3 and E_2 as current reference electrodes and voltage reference electrodes, respectively (Figure 61.10c). With positive terminal of the current source connected to the E_5, 13 voltage data (V_d: V_1, V_2, V_3, …, V_{13}) are measured successively across all the 13 voltage electrode pairs obtained by combining 13 voltage electrodes (E_V: E_4, E_6, E_7, E_8, E_9, E_{10}, E_{11}, E_{12}, E_{13}, E_{14}, E_{15}, E_{16}, and E_1) and the voltage reference electrodes ($E_{VR} = E_2$) as shown in Figure 61.10c. Therefore, in the cross method, this current projection (P_8) yields 13 differential voltage data $V_1 = V_{2-4}$, $V_2 = V_{2-6}$, $V_3 = V_{2-7}$, $V_4 = V_{2-8}$, $V_5 = V_{2-9}$, $V_6 = V_{2-10}$, $V_7 = V_{2-11}$, $V_8 = V_{2-12}$, $V_9 = V_{2-13}$, $V_{10} = V_{2-14}$, $V_{11} = V_{2-15}$, $V_{12} = V_{2-16}$, and $V_{13} = V_{2-1}$ measured across the electrode pairs E_2–E_4, E_2–E_6, E_2–E_7, E_2–E_8, E_2–E_9, E_2–E_{10}, E_2–E_{11}, E_2–E_{12}, E_2–E_{13}, E_2–E_{14}, E_2–E_{15} E_2–E_{16}, and E_2–E_1 (Figure 61.10b), respectively.

After collecting the set of 13 data from P_8, the procedure is just repeated by connecting the positive terminal of the current source to the electrodes E_7 (Figure 61.10d), and another set of 13 differential voltage data is collected (with the same current and voltage reference electrodes). Therefore, in P_9, the cross method yields 13 differential voltage data $V_1 = V_{2-4}$, $V_2 = V_{2-5}$, $V_3 = V_{2-6}$, $V_4 = V_{2-8}$, $V_5 = V_{2-9}$, $V_6 = V_{2-10}$, $V_7 = V_{2-11}$, $V_8 = V_{2-12}$, $V_9 = V_{2-13}$, $V_{10} = V_{2-14}$, $V_{11} = V_{2-15}$, $V_{12} = V_{2-16}$, and $V_{13} = V_{2-1}$ measured across the electrode pairs E_2–E_4, E_2–E_5, E_2–E_6, E_2–E_8, E_2–E_9, E_2–E_{10}, E_2–E_{11}, E_2–E_{12}, E_2–E_{13}, E_2–E_{14}, E_2–E_{15}, E_2–E_{16}, and E_2–E_1 (Figure 61.10d), respectively.

In the similar way, by connecting the positive terminal of the current source successively to the electrodes E_9, E_{11}, …, E_1, the cross method yields other five projections (P_{10}, P_{11}, P_{12}, P_{13}, and P_{14}), and similar sets of 13 voltage data are obtained with the aforementioned current and voltage reference electrodes. Hence, the cross method yields the other seven projections (P_8, P_9, P_{10}, …, P_{13}, and P_{14}) with the electrodes E_3 and E_2 as current reference electrodes and voltage reference electrodes, respectively. These seven projections give another set of $7 \times 13 = 91$ differential voltage measurements. Hence, the cross method has 14 projections yielding $91 + 91 = 182$ data. Among these 208 measurements, only 104 data are independent. Note that the cross method does not have as good a sensitivity in the periphery as does the neighboring method but has better sensitivity over the entire domain under test [3].

61.6.2.4 Adaptive or Trigonometric Method

In the aforementioned methods, current has been injected with a pair of current electrodes, and the differential voltage data have been measured between different pairs of voltage electrodes excluding the current electrodes. Gisser et al. [25] proposed a current injection method called the *adaptive method or trigonometric method* (Figure 61.11), in which the current is injected through all electrodes. Because current flows through all electrodes simultaneously, in the trigonometric method, as many independent

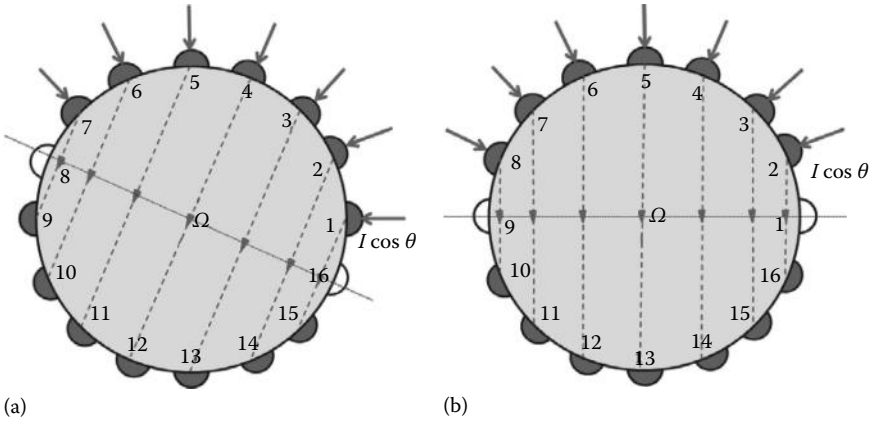


FIGURE 61.11 Adaptive or the trigonometric current injection method of voltage data collection: (a) projection 1 and (b) projection 16.

current injectors are needed as are the electrodes used in the system, and hence, a 16-electrode EIT system needs 16 current injectors. A current from $-I$ to $+I$ (I is any suitable RMS (root means square) amplitude of the current) can be fed to the electrodes, allowing different current distributions. Homogeneous current distribution may be obtained only in a homogeneous volume conductor. Hence, if the volume conductor is cylindrical with circular cross section, the injected current must be proportional to $\cos\theta$ to obtain a homogeneous current distribution (Figure 61.11). In the trigonometric method, the boundary potentials are measured with respect to a single grounded electrode and, hence, a 16-electrode EIT system, the trigonometric current injection method yields 15 voltage measurements in each current projections from projection 1 (Figure 61.11a) through projection 16 (Figure 61.11b). After collecting data from a particular projection (say P_1), the next projection (P_2) is obtained by rotating one electrode increment (22.5° for a 16-electrode system). Thus, for a 16-electrode EIT system, the trigonometric current injection method produces eight different current distributions yielding $8 \times 15 = 120$ independent voltage data.

Though the voltage measurements on current electrodes, sometimes, are avoided for contact impedance problem [26], in some present studies [9,22], measurements on current electrodes are also made to obtain the greatest sensitivity [22] to the resistivity changes in the domain under test. Hence, in this chapter EIT data collection strategies with the measurements on current electrodes in opposite and neighboring current injection are also discussed. Figure 61.12 presents the data collection procedures for both the current injection methods in which the boundary potentials are measured on all the electrodes including the current electrodes to obtain the greatest sensitivity to the resistivity changes in the domain.

Bera and Nagaraju [9] collected voltage data on all the surface electrodes for their practical phantom studies. In their data collection scheme with opposite pattern, current is injected through two diametrically opposed electrodes, and the potentials are measured on all the electrodes with respect to the ground point of the analog electronics (Figure 61.12a). Similarly, in the neighboring method, current is applied through two neighboring electrodes (Figure 61.12b) the potentials are measured on all the electrodes with respect to the ground point of the analog electronics (Figure 61.12b). Hence, for the 16-electrode system reported by Bera and Nagaraju, there will be 16 current projections (P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9 P_{10}, P_{11}, P_{12}, P_{13}, P_{14}, P_{15}, and P_{16}), each of which yields 16 voltage measurements (V_1, V_2, V_3, …, V_{16}) made on all the 16 electrodes (E_1, E_2, E_3, …, E_{16}) respectively. In opposite current method, RMS potentials on all the electrodes (including current electrodes) were measured for 16 current projections (Figure 61.12a), and the complete voltage data sets (containing 256 voltage data) were saved as a *.txt* file in a PC for computation. Bera and Nagaraju reported that the first 128 data collected from the first eight current projections (P_1 to P_8) were found independent from each other. Therefore, the boundary potentials were collected only for the first eight current projections (P_1 to P_8),

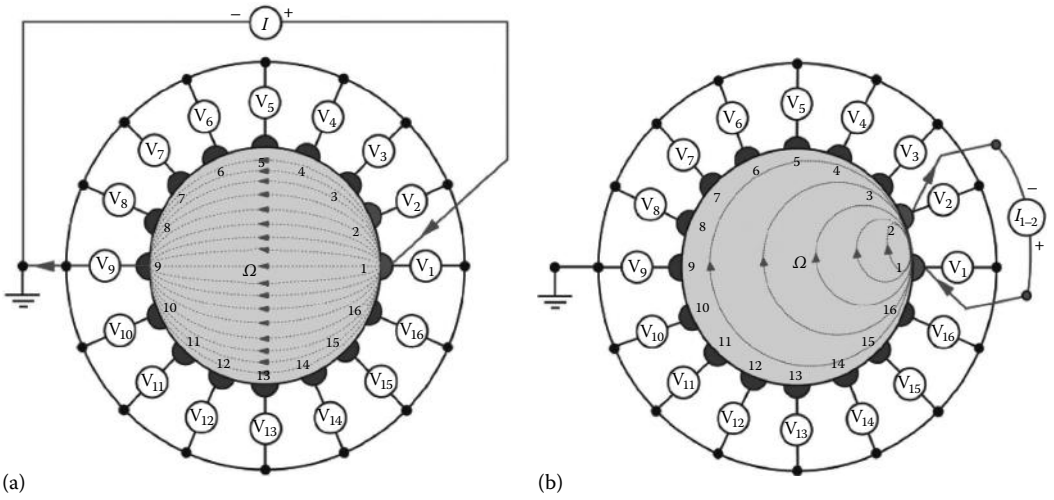


FIGURE 61.12 Current injection and voltage measurement at the phantom boundary: (a) opposite method and (b) neighboring method.

which yield $16 \times 8 = 128$ data. It was observed that the first 128 data collected in P_1 through P_8 are sufficient to reconstruct the resistivity images of the phantom domain [9]. Hence, in the opposite current injection method, instead of 256 voltage data, only 128 data (obtained from first current projections P_1, P_2, P_3, P_4, P_5, P_6, P_7, and P_8) are collected, and hence, the boundary data were not collected for the next eight current projections (P_9–P_{16}). Thus, the data collection time was reduced to 50% of the time required for a complete scan [9].

In the neighboring method with all electrode measurement, as reported by Bera and Nagaraju, the potentials were also measured on all the electrodes (Figure 61.12b) including the current electrodes to obtain the greatest sensitivity to the resistivity changes in the domain. Therefore, a 16-electrode system as reported by Bera and Nagaraju yields 16 current projections (P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9 P_{10}, P_{11}, P_{12}, P_{13}, P_{14}, P_{15}, and P_{16}), each of which gives 16 voltage measurements (V_1, V_2, V_3, …, V_{16}) made on all the 16 electrodes (E_1, E_2, E_3, …, E_{16}), respectively, as shown in (Figure 61.12b). Hence, in the neighboring current method, 16 current projections yield 256 voltage data that are used to reconstruct the resistivity images.

61.6.3 Electrode Materials

Electrodes in EIT are essentially required for injecting current and measuring potentials. Though different materials have been used for EIT electrodes, the stainless steel (SS) electrodes are found more common in phantoms and Ag/AgCl electrodes are found very popular for human body imaging in EIT. The sections that follow present a brief review of the electrode materials used for EIT sensors.

61.6.3.1 Stainless Steel Electrodes

Griffiths and Zhang [27] used SS electrodes in their phantom developed for a dual-frequency EIT system. Goble [28] used an SS electrode array for his 3D phantom developed for studying the 3D electrical impedance imaging. Boone and Holder [29] reported a saline-filled tank (88 mm diameter; 30 mm depth of 0.09% saline; SS electrodes of 8 mm diameter) developed for assessing their EIT system for imaging the neuronal depolarization in the brain using "direct current" electrical resistance tomography. Holder et al. [30] developed a number of practical biological phantoms with SS electrodes for calibrating multifrequency EIT. Goharian et al. [31] reported a phantom developed with a cylindrical phantom tank (made of acrylic plastic, 10 cm tall with a radius of 5 cm) housed with 48 circular SS electrodes (radius of each electrode was 0.3 cm) in three rings (each ring consisting of 16 electrodes with

interelectrode gap was 1.35 cm and located radially around the cylinder). The inhomogeneities were a metal rod and a plastic rod. Romsauerova et al. [32] developed a saline phantom with SS electrodes for studying the initial findings in brain tumors, arteriovenous malformations, and chronic stroke. Boyle et al. [33] developed a deformable phantom constructed with a sponge rubber plumbing gasket placed in a shallow pan and 16 SS wire electrodes. The gasket formed a thick rubber ring that was easily compressed when required, and it was rigid enough to return to its original shape easily.

61.6.3.2 Ag/AgCl Electrodes

Jossinet and Trillaud [34] investigated an EIT system using Ag/AgCl ECG electrodes. The boundary data measurements were obtained in vitro in a cylindrical tank filled with saline and in vivo using standard ECG electrodes. Holder et al. [35] reported a study on the superior surface of the cerebral cortex of paralyzed and artificially ventilated animals. The cerebral cortex was exposed by making a circular craniotomy, 20 mm in diameter; the dura mater was cut away to reveal the pia mater. A circular array of 16 Ag/AgCl spring-mounted ball electrodes, each about 1 mm in diameter, was placed on the surgically exposed dorsal surface of the cortex. Yerworth et al. [36] developed a spherical tank phantom (diameter 20 cm) filled with 0.2% saline and imaged a cylinder of banana. The phantom was built with 31 Ag/AgCl disk electrodes and imaged with the new UCLH Mk2 system. Tidswell et al. [37] tested the performance of the three types of commercially available headnet electrode arrays, designed for use in EEG, and conventional EEG Ag/AgCl cup electrodes on human subjects and a saline-filled head-shaped tank. The phantom tank was prepared with vegetable skin to simulate human skin in order to determine the optimum electrode system for EIT of the human head. Gilad et al. [38] reported the studies on the standard Ag/AgCl EEG cup electrodes (11 mm diameter) applied to abraded skin under the electrodes area using Ten20 conductive EEG paste (D.O. Weaver and Co., Aurora, CO). Author employed a conical recessed electrode design, which was made from polyoxymethylene (POM), filled with Ten20 conductive EEG paste, and had an Ag/AgCl disk, 14 mm diameter, at the vertex and circular opening to abraded skin of diameter 21 mm. Rahal et al. [39] investigated the impedance of various clinical electrodes as a function of frequency to identify the optimum electrode type. They studied six different types of self-adhesive electrodes commonly used in general and neonatal cardiology. Robitaille et al. [40] developed a practical phantom with Ag/AgCl electrodes for studying their EIT system developed for human brain function imaging. They constructed a realistically head-shaped tank using a silicone rubber cast of a head-shaped model made from a human skull covered with clay. Thirty-one Ag/AgCl electrodes (1 cm diameter each) were embedded into the inner walls of the tank filled with 0.2% saline to represent the brain. Xu et al. [41] studied the performance of the five common types of Ag/AgCl bioelectrodes by measuring the forearm and the brain of 10 healthy adult volunteers and evaluating those data in frequency or time domain in terms of contact impedance, uniformity, signal-to-noise ratio (SNR), and stability. They reported that the Ag/AgCl powder electrode has an overall best performance with as low contact impedance as commercial ECG electrodes ($p > 0.05$), high SNR (60.3 ± 4.5 dB), better uniformity (coefficient of correlation 0.95 ± 0.03), and greater stability (slope 0.68 ± 0.03).

61.6.3.3 Carbon Electrodes

Huang et al. [42] developed a rotating electrode phantom for EIT using carbon electrodes to increase the number of independent measurements. The rotating carbon electrode system increased the number of independent measurements, which enhanced the resolution of the impedance image and improves the quality arising from the ill-posed condition.

61.6.3.4 Brass Electrodes

Thomas et al. [43] reported a phantom with brass electrodes to study the correction of the nonuniform spatial sensitivity of EIT images. In their studies they developed a cylindrical Perspex tank (40 cm high with an inside diameter of 17 cm). The tank was drilled to accept 16 equi-spaced brass electrodes around the circumference of the tank, 15 cm from its threaded end.

61.6.3.5 Platinum–Iridium Electrodes

Gissert et al. [44] designed and developed a platinum–iridium electrode array for EIT for application to biological systems like the thorax. They developed the system containing 32 electrodes fashioned from a titanium plate having a platinum–iridium surface. The electrodes were attached to the interior surface of a circular plastic container 30 cm in diameter, 6.5 cm deep. The electrodes were 28.4 mm wide (with a 1 mm interelectrode gap) and extend throughout the depth of the container.

61.6.3.6 Textile Electrodes

Rahal et al. [39] investigated impedance imaging with textile electrodes and other types of Ag/AgCl electrodes for identifying the optimum electrode type in EIT.

61.6.4 Thin-Film-Based Flexible Electrodes

EIT electrodes are commonly made up of metals or metal composites. But as we have seen in the previous sections, SS electrodes and Ag/AgCl electrodes are most commonly used as surface electrodes in practical EIT systems. SS electrodes have been frequently used in practical phantoms due to their easy availability, profound rigidity, chemical and electrochemical stability, noncorrosiveness, and biocompatibility. Ag/AgCl electrodes are found very popular in medical EIT systems for their improved electrical performance characteristics and low half-cell potentials. But a special electrochemical process is to be carried out to coat the silver electrodes by AgCl layers for making an Ag/AgCl electrode. Moreover, AgCl is relatively a poor conductor [45]. Hence, increasing its layer thickness causes an increase in contact impedance, which adversely affects the boundary data accuracy in high-frequency EIT [45]. Bera and Nagaraju [46,47] proposed a flexible thin-film gold electrode array for impedance imaging in biomedical EIT for improved impedance imaging. They reported their research work on gold electrode phantoms and compared the gold electrode array performance with an identical SS electrode array.

61.6.4.1 Thin-Film-Based EIT Sensor Development

Bera and Nagaraju developed a flexible gold electrode array by depositing a gold thin film on a flexible FR4 sheet with a critical bending angle (θ_c) of 45° using an electrodeposition process. The copper patterns (Figure 61.13) of 16 identical electrodes (equally spaced with an interelectrode gap of 6.75 mm)

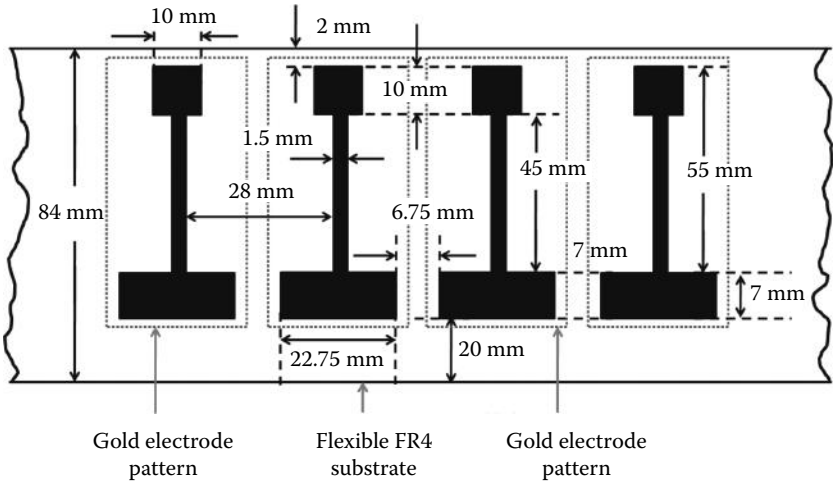


FIGURE 61.13 Electrode layout design: gold layer patterns.

with rectangular active electrode area (width 22.75 mm, height 7 mm) are made on a copper-cladded (copper layer thickness 35 μm) flexible FR4 sheet (200 μm thickness) using UV light lithography process with a proper electrode mask (Figure 61.14). Copper patterns made on the FR4 sheet are electroplated by a Ni layer of 2 μm thickness (Figure 61.15). Finally, the patterns are electroplated with a 2 μm thick gold layer and the gold electrode array is obtained (Figure 61.16a).

Thus, the electrode patterns in the gold electrode array (Figure 61.16a) are composed of a 2 μm thick gold layer, 2 μm thick Ni layer, and 35 μm thick copper layer (Figure 61.15). For comparison studies an SS electrode array (Figure 61.16b) was also developed with 16 rectangular electrodes identical to the electrodes in the gold array. Sixteen rectangular strips (width = 22.75 and height = 95 mm) were cut from an SS sheet (type-304, thickness = 50 μm). The SS strips are sandwiched between two flexible plastic sheets keeping the electrode areas (width = 22.75 mm, height = 7 mm) and the connecting pad open (Figure 61.16b).

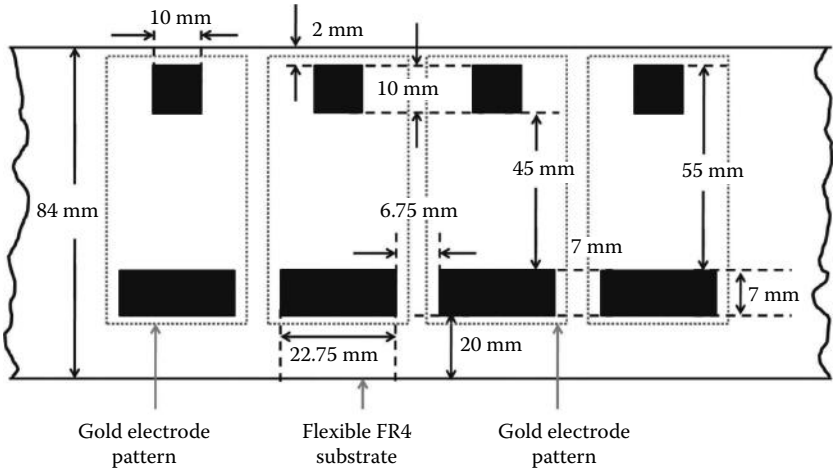


FIGURE 61.14 Schematic of the electrode patterns: mask design for lithography.

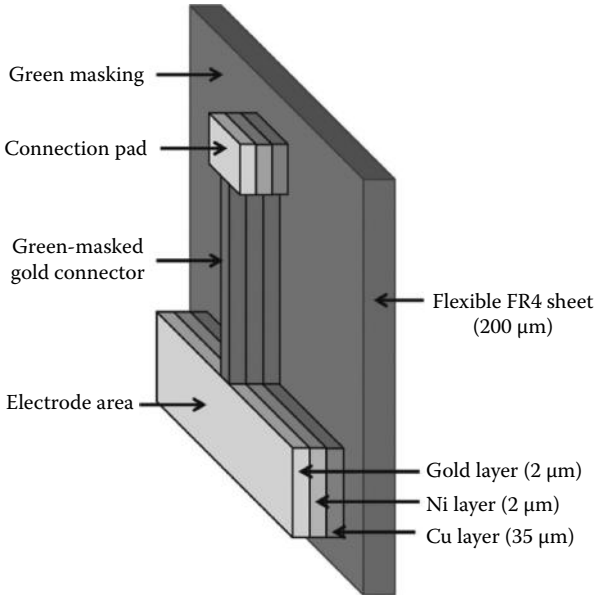


FIGURE 61.15 Schematic of the electrode patterns.

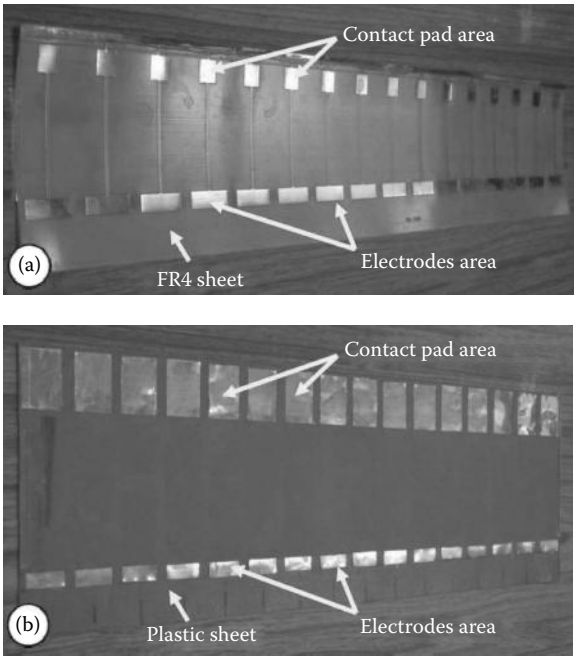


FIGURE 61.16 Developed flexible electrode arrays: (a) electroplated gold electrode array and (b) SS electrode array.

61.6.4.2 Gold Electrode Phantoms

Bera and Nagaraju reported the resistivity imaging studies with a gold electrode phantom (Figure 61.17a), and they compared the results with identical SS electrode phantom (Figure 61.17b). They reported that the gold electrodes were found suitable to connect directly to the copper wire by soldering, and hence, the contact impedance problem was reduced. The gold electrode array was put inside nylon tanks (diameter = 150 mm), and a thin layer of silicone glue was applied at the junction of the electrode array. The tank was filled by 0.9% (w/v) NaCl solution as the bathing solution of the EIT phantom. A common-mode electrode (CME) [46,47] was placed at the phantom center (Figures 61.17a and b) and connected to the ground point of the EIT hardwires to reduce the common-mode error [48] of the electronic circuits.

Nylon cylinders (diameter = 35 mm) were placed inside the phantom, and the boundary potentials were measured by injecting 1 mA, 50 kHz sinusoidal current signal using the opposite current injection protocol. Boundary potential data and electrode contact impedance parameters were analyzed [46,47].

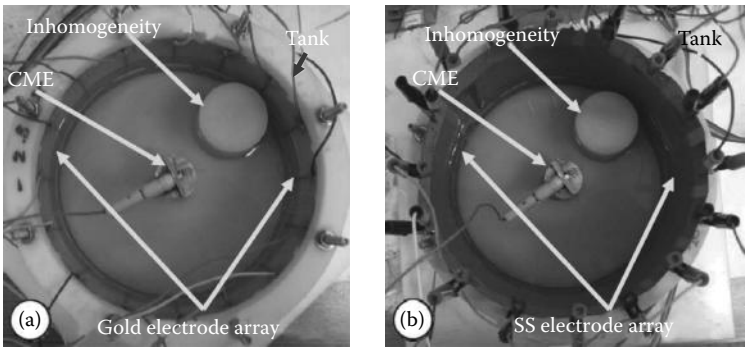


FIGURE 61.17 Phantoms with flexible electrodes: (a) SS electrode array inside the tank and connected with steel alligator clips and (b) gold electrode array inside the tank and soldered with the connecting wires.

Resistivity images were reconstructed using the Electrical Impedance and Diffuse Optical Reconstruction Software (EIDORS) [49] with a FEM (finite element method) mesh [50] containing 1968 elements and 1049 nodes. Boundary potential profiles and the resistivity images obtained with gold electrode array were compared with the SS electrode phantom with identical configurations.

Results reported by Bera and Nagaraju [46,47] demonstrated that the gold electrode array was found highly conductive, flexible, chemically stable, biocompatible, and noncorrosive in all natural circumstances compared to SS electrodes. Geometrically, gold electrode array was found more accurate to the required design exhibiting uniform and low contact impedances compared to the SS one. Electrodes in the SS electrode array were connected to the wires with steel clips as the soldering was found difficult for SS connecting pads. Hence, the electrode–wire contact impedances were found large and dissimilar for the individual electrodes in the SS electrode array. Electrode-to-electrode impedance ($Z_{\text{E-E}}$) [46] and electrode-to-CME impedance ($Z_{\text{E-CME}}$) [46] are measured, and it is observed that both the impedances are more uniform and small for gold electrode array [46] compared to the SS electrodes.

Result reported by Bera and Nagaraju [46] also showed that the standard deviation (STDV) in the $Z_{\text{E-CME}}$ [46] and $Z_{\text{E-E}}$ [46] was less for gold electrode array. Phase angles of $Z_{\text{E-CME}}$ ($\theta_{\text{E-CME}}$) and $Z_{\text{E-E}}$ ($\theta_{\text{E-E}}$) and their electrode-to-electrode variations were found also less for gold electrode array compared to the SS array [46]. They reported that the gold electrode pairs offered more uniform and low impedance path compared to that of the SS array [46] by assuming the NaCl solution has uniform resistivity throughout the whole medium.

Boundary data were collected with homogeneous medium for both the SS electrode array and the gold electrode array. It was observed that the boundary data collected from the SS electrode array were noisier than that of the gold electrode array [46]. Results showed that the boundary potential data as well as the mean boundary potential data obtained with gold electrode array were more symmetric [46] for all the 16 current projections than the data obtained from the SS electrode system. It was also noticed that the SNR of the gold electrode system was greater than that of the SS electrode system [46]. The resistivity images with gold electrode array were found with better image contrast and better spatial resolution [46]. The background noise in the reconstructed images was also found less for the gold electrode array [46]. The diametric resistivity plots (DRP) [46] of the resistivity images obtained with gold electrode system showed that the reconstructed resistivity profiles were found better and closer to the original domain resistivity profiles [46] compared to the SS electrode array.

Bera and Nagaraju also studied the gold electrode array performance for real tissue imaging and compared the results with an identical real tissue phantom with SS electrodes [47]. Bera and Nagaraju calculated the image parameters from the elemental resistivity profiles and assessed the resistivity images. IR_{Mean} [9,51,52], BR_{Mean} [9,52], CNR [9,52,53], PCR [9,52,53], COC [9,52,53], and DRP [46,52] were studied, and the reconstructed images were evaluated for both the phantoms. The reported results demonstrated that the resistivity images with gold electrode array were found with better image contrast and better spatial resolution with low background noise [47]. The DRP of the resistivity images obtained with gold electrode real tissue phantom showed that the reconstructed resistivity profiles were found closer to the original DRPs [47].

61.7 Electrode Size

Reconstructed image quality depends on the surface electrode performance and the electrode performance is dependent on electrode material as we discussed in the earlier section. One more factor that affects the electrode performance is electrode geometry, that is, electrode shape, electrode size or electrode area, and electrode position in the phantom or object boundary. Hence, the boundary potential profile is found dependent on the electrode size in EIT studies. A number of researchers have tried to find the optimum electrode size for better performance. The following section will present a literature survey conducted on the research work carried out on the electrode size and other electrode geometric parameters.

Luo et al. [54] investigated the best type, size, and location of electrodes for EIT for ventilation measurements. They compared the electrode performance and reported that adhesive-gel electrodes are easy to apply, make good body contact, and do not slip during the experimentation. They also reported that higher SNRs were obtained when electrode area was increased by connecting several small electrodes together rather than by using a single electrode with a larger area [54]. Yan and Hong [55] presented a method for EIT electrode structure and parameter optimizing design. They took an example for a compound electrode that was based on the coercive equipotential node model of line electrode. They also developed a simulation research software platform for electrode structure and parameter optimizing design. They anticipated that by the method proposed, the different influences of electrode structure and parameters on the reconstructed image can be obtained along with an estimation of the system sensitivity. They also expected that their method would provide the theory basis for the optimum design of the surface electrode structure in real EIT systems. Yan et al. [56] presented a method to optimize the EIT electrode structure and parameters based on coercive equipotential node models. The coercive equipotential node model of the compound electrode was established based on that of the line electrode. Using a simulation software platform, the line electrode and the compound electrode of EIT were studied. The authors also studied the influences of different electrode structures and parameters on measurement sensitivity and the image reconstruction quality. From simulation studies they reported that a narrower electrode was found helpful in improving the imaging quality and using too wide an electrode causes the measurement sensitivity to decrease, although it is known that a wider electrode is beneficial in decreasing the contact impedance. They also reported that the electrode width leading to the best measurement sensitivity is different for different measurement depths. The author expressed the design aspects of the compound electrodes with four parameters, which have mutual restrictions and complex influences on each other. The four parameters mentioned in their report were the excitation electrode width, the measurement electrode width, the space between the excitation electrode and the measurement electrode, and the distance between two adjacent compound electrodes. Authors concluded that it is not advised to optimize the design of a compound electrode by only using the overlay rate of electrodes.

61.8 Compound Electrode

Hua et al. [57] described the use of compound electrodes to improve the conditioning and sensitivity of our EIT imaging system. They discussed effects of electrode–skin contact impedance by showing the reconstructed results of data from phantom measurements. Hua et al. [58] developed a compound electrode composed of two electrodes: a large outer electrode to inject current and a small inner electrode to sense voltage. They measured the surface potentials from a physical phantom using these compound electrodes and showed that the measured voltages from the compound electrodes were smaller in amplitude than those from conventional electrodes, which demonstrated that the compound electrode can minimize contact impedance voltage drop from the measured data.

61.9 Active Electrode

Rigaud et al. [59] proposed active current electrodes for in vivo EIT. To increase the measurement accuracy, the authors studied the active approach for limiting the buffer input stray capacitance by positive feedback. Limitations of this technique were also discussed. Jossinet et al. [60] developed active current electrodes for in vivo EIT by designing and fabricating accurate and reproducible current generators for high-frequency EIT. The authors reported that the frequency range used was limited to 2048 kHz. They reported that sorting the components resulted in low, reproducible output conductances (0.85 μS), which reduced the influence of electrode mismatch. They also reported that the low variability (±2%) of the output admittance provided a facility of accurate current injection and reduced the common-mode signal applied to the differential amplifiers. Lead capacitance problems were removed due to the encapsulation of

the current source within the electrode shell, and therefore, it enabled the use of a high-frequency signal up to 2048 kHz. Li et al. [61] developed a wide-band high-speed EIT data acquisition system with active electrodes and its application was studied in cardiac imaging. They developed fast data acquisition with 32 active electrodes, half of them as receive electrodes and the other half as drive electrodes. The active electrode system was developed by providing a buffer mounted on the back of each receive electrode and a current source on each drive electrode. Gaggero et al. [62] presented the design and initial tests of an active electrode-based system to address the difficulties associated with the electrode contact impedance problem and the electrode placements on human subject. They developed an active electrode belt for EIT, which was designed and developed with 32 active electrodes containing the electronic amplifiers, switches, and associated logic. Experimental results showed stable device performance with a convenient ease of use and good imaging ability in volunteer tests. Gaggero et al. [63] developed and presented an active electrode-based EIT electrode belt system. The active electrode EIT system incorporated an active electrode belt, a central voltage-driven current source, central analog-to-digital converters and digital-to-analog converters, and a central FPGA-based demodulator and controller. The electrode belt was designed with 32 active electrodes containing the electronic amplifiers, switches, and associated logic. A stable device performance with a convenient ease of use and good imaging ability was observed in practical experimentation.

61.10 Electrode Number

An EIT system collects voltage data developed for a constant current injection through a set of electrodes or EIT sensors attached to the domain under test. Spatial resolution depends on the number of electrodes, but the optimum number of electrodes is decided by the computation facilities available and the computation time permissible. The electrode number is also restricted by the circumferential area of the domain to be imaged. The minimum number of electrodes used for the EIT as reported by the researchers is eight [64]. The 16-electrode EIT system is very popular in EIT imaging. Though most of the other EIT systems have an electrode number that is $2^{(3+N)}$ ($N = 0, 1, 2, 3...$), other EIT systems are also reported with electrodes other than $16*N$. The following section will present a brief literature survey conducted on the EIT system with different electrode numbers proposed by the different research groups. Though the survey shows that the electrode array of a practical EIT system generally contains $2^{(3+N)}$ where N is an positive integer ($N = 1, 2, 3, \ldots$), a few systems have electrodes other than this number series.

61.10.1 8-Electrode System

Yu et al. [64] proposed an 8-electrode rotative EIT system with eight carbon rods working as the electrodes dipped inside the saline phantom. Huang [65] developed an 8-electrode EIT system with eight carbon rods working as the electrodes dipped inside the phantom filled with saline. The electrode system was attached to a rotating system developed with an electrical motor, and the boundary data measurements were conducted each after rotation of the electrode system.

61.10.2 12-Electrode System

Ito et al. [66] proposed a 12-electrode EIT system for measuring of body fat distribution by using EIT. Alme et al. [67] developed and studied a 12-electrode EIT system for analyzing 3D and conductivity effects in electrical tomography systems using COMSOL multiphysics EM module.

61.10.3 13-Electrode System

Linderholm et al. [68] proposed a 13-electrode EIT system for probing of multilayered tissue phantoms. They developed the electrode array system of the proposed EIT system using microelectrodes.

61.10.4 14-Electrode System

Kimoto and Shida [69] developed and studied a 14-electrode EIT system for imaging of temperature-change distribution in the brain phantom by means of capacitance measurement.

61.10.5 16-Electrode System

The 16-electrode EIT system has been developed and studied by several groups such as Huang et al. [42], Shuai et al. [70], Hartinger et al. [71,77], Robitaille et al. [40], Griffithst et al. [22,72], Boyle et al. [33], Hahn et al. [29,73,78], Holder et al. [30], Thomas et al. [43], Yu et al. [64], Alme et al. [67], McArdlet et al. [74], Sadleiry and Foxz [75], Bera and Nagaraju [7,9,11,46–47,50–53], and Vauhkonen et al. [76].

61.10.6 31-Electrode System

Bagshaw et al. [79] developed and studied a 31-electrode EIT system for the imaging of human brain function. Yerworth et al. [36] developed a spherical tank phantom (diameter 20 cm) filled with 0.2% saline and imaged a cylinder of banana. The phantom system was developed with 31 Ag/AgCl disk electrodes and imaged with the new UCLH Mk2 system.

61.10.7 32-Electrode System

Gissert et al. [44] developed the system containing 32 electrodes fashioned from a titanium plate having a platinum–iridium surface, and the electrodes were attached to the interior surface of a circular plastic container 30.0 cm in diameter, 6.5 cm deep. The electrodes were 28.4 mm wide (with a 1.0 mm interelectrode gap) and extended throughout the depth of the container. Cook et al. [80] developed a 32-electrode EIT system, ACT3, for high-speed, high-precision EIT.

61.10.8 60-Electrode System

Boverman et al. [81] developed a 60-electrode test phantom for 3D mammography geometry and studied EIT imaging and formulated, implemented, and experimentally demonstrated a procedure using the complete electrode model (CEM) to compensate for poor electrode contact.

61.10.9 64-Electrode System

Yerworth et al. [82] reported a 64-electrode 3D EIT system called the UCLH Mark 1b. UCLH Mark 1b is a portable EIT system that can address up to 64 electrodes, which has been designed for imaging brain function with scalp electrodes.

61.10.10 128-Electrode System

Xu et al. [83] developed a 128-electrode 3D EIT system for impedance change detection and 3D imaging of the human thorax. The proposed system was made up of several modules, including multifrequency current source, driving, measuring, data acquisition, and controlling and signal processing modules. A high-speed digital signal processor (DSP) was used as the controller. The 128-electrode system was developed with 64 driving electrodes and 64 measuring electrodes, and all the electrodes were positioned uniformly in four planes around the surface of a cylindrical phantom filled with a saline solution and objects of varying conductivities.

61.11 Rotating Electrodes in EIT

Huang et al. [84] developed a high-performance rotational EIT (REIT) system capable of producing better-quality EIT images by increasing the number of independent measurements. The REIT system was constructed with the electrodes attached to a rotational phantom tank driven by a stepper motor. The REIT phantom was developed with a plastic tank (110 mm in height and 180 mm in diameter) and 16 compound electrode lines placed on the inner surface of the tank. The compound electrode contained a voltage electrode and a current electrode. Since the accuracy of measurements depends on the electrode geometry, in EIT it is essential to optimize the design parameters, especially their size and shape. The current electrodes were designed with a large surface area required in obtaining the uniform current density, and the voltage electrodes were designed with a small surface area that is optimal so as to avoid the influence from other equipotentials. Huang et al. constructed their current electrode from a rectangular copper slice (20 mm in width, 90 mm in height, and 0.5 mm in thickness) and voltage electrode from a copper bar (3 mm in width, 50 mm in height, and 0.5 mm in thickness) isolated by insulating tape. Using the stepper motor, the system was capable of obtaining a precise measurement, and the number of independent measurements was increased for enhancing the resolution of the impedance image. The experimental results demonstrated that the quality of the REIT images was improved by expanding the independent measurements.

61.12 Pseudo Electrodes in EIT

The quality of impedance image in EIT is limited by the number of electrodes attached to the object boundary. Yu et al. [85] proposed pseudo electrodes for EIT for enhancing the image quality. They obtained the expanded pseudo electrodes by changing the measuring sites of the physical electrodes. The number of independent measurement was increased. The result demonstrated that the pseudo electrode system improves the image quality in EIT. However, the authors reported that the large number of measurement data obtained from pseudo electrode driven pattern would cause serious problem in image reconstruction [85].

61.13 Other EIT Electrodes

61.13.1 Esophageal Electrode

Tehrani et al. [86] developed esophageal electrodes to investigate the improvements of the measurement on EIT and initial conductivity estimation of the internal electrode by modeling an internal electrode inside the esophagus.

61.13.2 Monopolar Needle Electrodes

Martinsen et al. [87] presented a study on invasive EIT for blood vessel detection using monopolar needle electrodes. They presented a novel method for localization of large blood vessels using a bioimpedance-based needle positioning system on an array of ten monopolar needle electrodes.

61.13.3 Silver Ink Electrodes

McAdams et al. [45] reported a study on the optimal electrolytic chloriding of silver ink electrodes for use in EIT. They reported that there is an optimal AgCl layer thickness which minimizes the interelectrode impedance at a given frequency. Authors reported that as this thickness is less for high frequencies, the EIT electrode systems require only relatively thin layers of AgCl.

61.13.4 Nonstationary Electrodes

Murphy et al. [88] developed an EIT system with nonstationary electrodes to improve the levels of recoverable image content by attaching electrodes to a rotating central impeller. They incorporated the rotational motion of the electrode array into an optimized measurement strategy, which provided more independent measurements without compromising the stability of a consistently regularized inverse solution.

61.14 Electrode Models

Different electrode models have been presented by researchers to explore the optimum electrode model providing better performance. They are continuum model [89], gap model [89], shunt model [89], and complete electrode model [89].

61.15 Electrodes of 3D EIT

In 2D electrical impedance tomography (2D EIT), the resistivity distribution is reconstructed by assuming that the electrical current is confined within a 2D electrode plane (plane of interest). But in the real case the electrical current is conducted in a 3D space within the volume conductor. Hence, the boundary data measured for 2D EIT are extensively influenced by the 3D conduction of electrical current. As a result 2D EIT imposes one more assumption, which reduces the accuracy of the resistivity assessment by solving the inverse ill-posed problem of EIT. Therefore, 3D EIT (Figure 61.18) was introduced for better assessment of the resistivity distribution. Moreover, a 3D EIT [90] provides 3D impedance distributions with a 3D visualization of the object interiors, which helps the doctors to get a more clear and more accurate insight of the body of the patient during diagnosis and treatment.

In 3D EIT (Figure 61.18) the boundary data from the 3D SUT using multiple electrode arrays placed at the different planes of interest within the SUT. Generally, three separate planes are chosen for data collection in 3D EIT though more planes give more boundary data, which facilitate the inverse solution with more computation time and extra cost. The data collection procedure from the surface electrodes in 3D EIT is similar to the 2D EIT only, and the number of arrays is more in 3D EIT. In the medical imaging procedure, the surface electrodes are stacked on the patient body in three different planes within the volume of interest [90]. Similarly in 3D EIT applied in multiphase flow imaging in process tomography, the fluid pipe is provided with three or more different electrode arrays placed at three or more different planes.

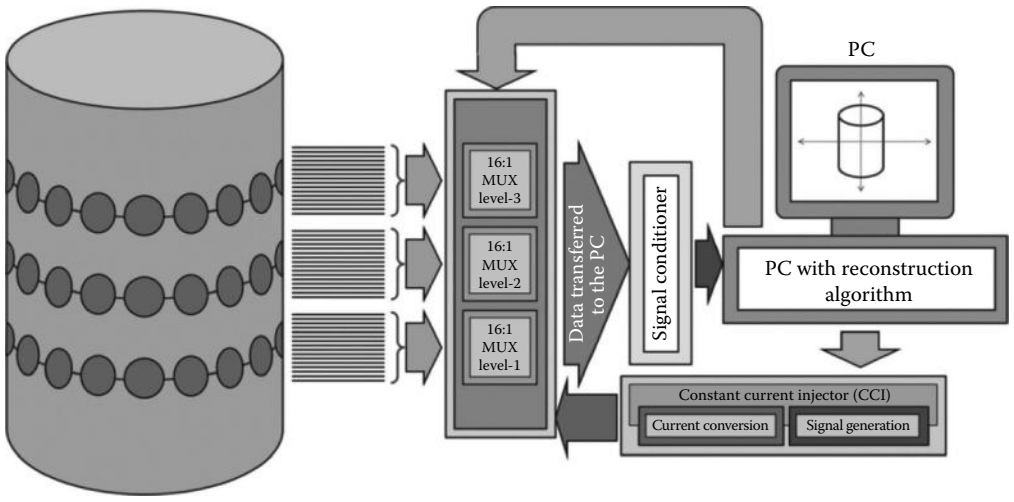


FIGURE 61.18 3D EIT schematic: phantom with multiple layers of electrodes and the data acquisition system.

TABLE 61.1 Various Errors Produced by the Electrodes and EIT Hardware

Authors	Most Significant Error Factor
Brown and Seagar [91]	Input stray capacitance
Sakamoto et al. [92]	Amplifier noise
Jossinet and Trillaud [93], Trillaud and Jossinet [94]	Mismatches between amplifiers, inadequate CMMR
Record et al. [95]	Multiplexers capacitance
Rigaud et al. [96]	Common-mode current injection
McLeod et al. [97]	Multiplexer capacitance
Riu et al. [98]	Common-mode current injection
Sansen et al. [99]	Interchannel stray capacitance
Record and Hargreaves [100]	Unequal electrode impedance

61.16 Errors of EIT Electrodes and Hardware

The EIT systems always suffer from several errors associated to the electrodes, phantoms and the electronic hardware. Boone and Holder [10] presented a summary of research work reporting on the various errors produced by the different parts of the EIT hardware.

The most significant source of errors produced by the electrodes and the associated EIT hardware as suggested by researchers are given in Table 61.1 [10].

61.17 Conclusions

EIT electrodes or EIT sensors are required to interface the EIT hardware with the object under test (i.e., patient) for current injection and boundary data collection. Therefore, in EIT, the surface electrodes play a very crucial role in the boundary data accuracy as they contribute a number of errors produced by electrode geometry, electrode number, electrode materials, improper electrode modeling, electrode–patient interfaces/contact impedance, and electrode–wire connections. The electrode size and electrode number are considered as the major influencing factors for the image quality and image resolution, respectively. It is reported that the electrodes with large contact area are preferred for current injection to minimize contact impedance, and the voltage-measuring electrodes are preferred with small contact area to obtain the maximum SNR. The wider electrodes can decrease the contact impedance between the electrode and skin, but the maximum possible number of electrodes for a particular domain will be reduced by increasing the electrode width. Hence, the wider electrode geometry will decrease the number of electrodes and the number of independent boundary data, which in turn reduces the image resolution. On the other hand, though a large number of electrodes provide more independent boundary data and better image resolution, the system cost and computation time and cost increase. Moreover, the wider electrodes reduce the interelectrode distance, and hence, an electrode array with too small interelectrode gap will reduce the measurement sensitivity, especially at the central area of the domain. It is reported that the influence of contact impedance between electrode and skin is reduced maximally when the total width of electrode array is about 80%–90% of the imaged area's perimeter. It is also reported that the optimum area underneath the electrodes as a percentage of the overall boundary area depends on the current pattern used (80% for the adjacent patterns and 60% for the opposite ones). Though varieties of electrode materials have been used in different EIT systems with their own advantages, the Ag/AgCl electrodes are found more popular in human body imaging, and the SS electrodes are widely preferred in saline phantoms for their easy availability, profound stability, chemical inertness, and low cost. Recently, thin-film-based gold electrode arrays have been found suitable for both saline phantoms and real tissue phantom imaging. The materials of EIT electrode can be

varied from application to application and with the atmosphere surrounding the system. A few studies on EIT electrode arrays have been reported, but a lot of experimental studies are still required to be conducted to obtain a proper design of electrode geometry with suitable materials.

References

1. Kak A. C. and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, New York, 1988.
2. Bushberg J. T., Seibert J. A., Leidholdt Jr. E. M., and Boone J. M., *The Essential Physics of Medical Imaging*, 2nd edn., Chapter 13, Lippincott Williams & Wilkins, Philadelphia, PA, ISBN-10: 0683301187, 2002.
3. Webster J. G., *Electrical Impedance Tomography*. Adam Hilger Series of Biomedical Engineering, Adam Hilger, New York, 1990.
4. Holder D. S., *Electrical Impedance Tomography: Methods, History and Applications* (Series in Medical Physics and Biomedical Engineering), 1st edn., Institute of Physics Publishing Ltd., Bristol, U.K., 2005.
5. Bayford R. H. Bioimpedance tomography (electrical impedance tomography), *Annual Review of Biomedical Engineering*, 8: 63–91, 2006.
6. Cheney M., D. Isaacson, and J. C. Newell, Electrical impedance tomography, *SIAM Review*, 41(1): 85–101, 1999.
7. Bera T. K. and J. Nagaraju, Resistivity imaging of a reconfigurable phantom with circular inhomogeneities in 2D-electrical impedance tomography, *Measurement*, 44(3): 518–526, March 2011.
8. Barber D. C. and B. H. Brown, Applied potential tomography, *Journal of Physics E: Scientific Instruments*, 17: 723–733, 1984.
9. Bera T. K. and J. Nagaraju, Studying the resistivity imaging of chicken tissue phantoms with different current patterns in electrical impedance tomography (EIT), *Measurement*, 45: 663–682, 2012, doi:10.1016/j.measurement.2012.01.002
10. Boone K. G. and D. S. Holder, Current approaches to analogue instrumentation design in electrical impedance tomography, *Physiological Measurement*, 17: 229–247, 1996.
11. Bera T. K and J. Nagaraju, A multifrequency constant current source for medical electrical impedance tomography, *Proceedings of IEEE International Conference in Systems Medicine and Biology*, (IEEE ICSMB 2010), December 16–18 2010, Kharagpur, India, pp 278–283, 2010. doi: 10.1109/ICSMB.2010.5735387
12. Yorkey T. J., Comparing reconstruction methods for electrical impedance tomography, PhD thesis, University of Wisconsin at Madison, Madison, WI, 1986.
13. Lionheart W. R. B., EIT reconstruction algorithms: Pitfalls, review article, challenges, and recent developments, *Physiological Measurements* 25: 125–142, 2004.
14. Breckon W. R., Image reconstruction in electrical impedance tomography, PhD thesis, Oxford Polytechnic, Oxford, U.K., 1990
15. Bera T. K., S. K. Biswas, K. Rajan, and J. Nagaraju, Improving image quality in electrical impedance tomography (EIT) using projection error propagation-based regularization (PEPR) technique: A simulation study, *Journal of Electrical Bioimpedance*, 2: 2–12, 2011. doi:10.5617/jeb.158
16. Bera T. K., S. K. Biswas, K. Rajan, and J. Nagaraju, Improving conductivity image quality using block matrix-based multiple regularization (BMMR) technique in EIT: A simulation study, *Journal Electrical Bioimpedance*, 2, 33–47, 2011. doi:10.5617/jeb.170
17. Webster J. G., *Measurement, Instrumentation, and Sensors Handbook*, CRC Press, Boca Raton, FL, 1999.
18. Morris A. S., *Measurement and Instrumentation Principles*, 3rd edn., Butterworth-Heinemann, Linacre House, Jordan Hill, Oxford, U.K., 2001, OX2 8DP, MA.
19. Webster J. G., *Medical Instrumentation: Application and Design*, 4th edn., John Wiley & Sons, Hoboken, NJ, 2010.

20. Bera T. K. and J. Nagaraju, Electrical impedance spectroscopic study of broiler chicken tissues suitable for the development of practical phantoms in multifrequency EIT, *Journal Electrical Bioimpedance*, 2: 48–63, 2011. doi:10.5617/jeb.174
21. Guimera A., G. Gabriel, M. Plata-Cordero, L. Montero, M. J. Maldonado, and R. Villa, A non-invasive method for an in vivo assessment of corneal epithelium permeability through tetrapolar impedance measurements, *Biosensors Bioelectronics*, 31: 55–61, 2012.
22. Cheng K. S., S. J. Simske, D. Isaacson, J. C. Newell, and D. G. Gisser, Errors due to measuring voltage on current-carrying electrodes in electric current computed tomography. *IEEE Transactions on Biomedical Engineering*, 37(60): 60–65, 1990.
23. Brown B. H. and A. D. Segar, The Sheffield data collection system. *Clinical Physics and Physiological Measurement*, 8(Suppl. A): 91–97, 1987.
24. Hua P., J. G. Webster, and W. J. Tompkins, Effect of the measurement method on noise handling and image quality of EIT imaging. *Proceedings of Ninth International Conference IEEE Engineering in Medicine and Biology Society*, Vol. 2, pp. 1429–1430, IEEE, New York, 1987.
25. Gisser D. G, D. Isaacson, and J. C. Newell, Current topics in impedance imaging. *Clinical Physics and Physiological Measurement*, 8(Suppl. A): 39–46, 1987.
26. Malmivuo J. and R. Plonsey, *Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields*, Chapter-26, Section-26.2.1, Oxford University Press, New York, 1995.
27. Griffiths H. and Z. Zhang, A dual-frequency electrical impedance tomography system, *Physics in Medicine and Biology*, 34(10): 1465–1476, 1989.
28. Goble J. C., The three-dimensional inverse problem in electric current tomography, PhD thesis, Rensselaer Polytechnic Institute, New York, 1990.
29. Boone K. G. and D. S. Holder, Design considerations and performance of a prototype system for imaging neuronal depolarization in the brain using "direct current" electrical resistance tomography, *Physiological Measurement*, 16: A87–A98, 1995.
30. Holder, D. S., Y. Hanquan, and A. Rao, Some practical biological phantoms for calibrating multifrequency electrical impedance tomography, *Physiological Measurement* 17, A167–A177, 1996.
31. Goharian M., M. Soleimani, A. Jegatheesan, K. Chin, and G. R. Moran, A DSP based multi-frequency 3D electrical impedance tomography system, *Annals Biomedical Engineering*, 36(9): 1594–1603, September 2008. doi: 10.1007/s10439-008-9537-5
32. Romsauerova A., A. McEwan, L. Horesh, R. Yerworth, R. H. Bayford, and D. S. Holder, Multi-frequency electrical impedance tomography (EIT) of the adult human head: Initial findings in brain tumours, arteriovenous malformations and chronic stroke, development of an analysis method and calibration, *Physiological Measurement*, 27: S147–S161, 2006. doi:10.1088/0967-3334/27/5/S13
33. Boyle A., W. R. B. Lionheart, and A. Adler. Evaluating deformation corrections in electrical impedance tomography, *9th Conference on Electrical Impedance Tomography*, Dartmouth College, Hannover, NH, pp. 175–178, 2008.
34. Jossinet J. and C. Trillaud, Imaging the complex impedance in electrical impedance tomography, *Physiological Measurement*, 13(Suppl. A): 47–50, 1992.
35. Holder D. S., A. Rao, and Y. Hanquan, Imaging of physiologically evoked responses by electrical impedance tomography with cortical electrodes in the anaesthetized rabbit, *Physiological Measurement*, 17: A179–A186, 1996.
36. Yerworth R. J., R. H. Bayford, B. Brown, P. Milnes, M. Conway, and D. S. Holder, Electrical impedance tomography spectroscopy (EITS) for human head imaging, *Physiological Measurement*, 24: 477–489, 2003. PII: S0967–3334(03)54284–8
37. Tidswell A. T., A. P. Bagshaw, D. S. Holder, R. J. Yerworth, L. Eadie, S. Murray, L. Morgan, and R. H. Bayford, A comparison of headnet electrode arrays for electrical impedance tomography of the human head, *Physiological Measurement*, 24: 527–544, 2003. PII: S0967–3334(03)54286–1

38. Gilad O., L. Horesh, and D. S. Holder, Design of electrodes and current limits for low frequency electrical impedance tomography of the brain, *Medical and Biological Engineering Computing*, 45: 621–633, 2007. doi 10.1007/s11517-007-0209-7.
39. Rahal M., J. M. Khor, A. Demosthenous, A. Tizzard, and R. Bayford, A comparison study of electrodes for neonate electrical impedance tomography, *Physiological Measurement*, 30: S73–S84, 2009. doi:10.1088/0967-3334/30/6/S05.
40. Robitaille N., R. Guardo, I. Maurice, A. E. Hartinger, and H. Gagnon, A multi-frequency EIT system design based on telecommunication signal processors, *Physiological Measurement*. 30: S57–S71, 2009. doi:10.1088/0967-3334/30/6/S04.
41. Xu, S., M. Dai, C. Xu, C. Chen, M. Tang, X. Shi, and X. Dong, Performance evaluation of five types of Ag/AgCl bio-electrodes for cerebral electrical impedance tomography, *Annals of Biomedical Engineering*, 39(7): 2059–2067, July 2011. doi: 10.1007/s10439-011-0302-9.
42. Huang, C. N., F. M. Yu, and H. Y. Chung. The scanning data collection strategy for enhancing the quality of electrical impedance tomography. *IEEE Transactions Instrumentation and Measurement* 57(6), 1193–1198, 2008.
43. Thomas D. C., Siddall-Allum J. N., Sutherland I. A., and Beard R. W., Correction of the non-uniform spatial sensitivity of electrical impedance tomography images, *Physiological Measurement*, 15: A147–A152, 1994.
44. Gissert D. G., D. Isaacson, and J. C. Newells, Theory and performance of an adaptive current tomography system, *Clinical Physics and Physiological Measurement*, 9(Suppl. A), 35–41, 1988.
45. McAdams E. T., P. Henry, J. M. Anderson, and J. Jossinet, Optimal electrolytic chloriding of silver ink electrodes for use in electrical impedance tomography, *Clinical Physics and Physiological Measurement*, 13(Suppl. A), 19–23, 1992.
46. Bera, T. K. and J. Nagaraju, Gold electrode sensors for electrical impedance tomography (EIT) studies, *IEEE Sensors Application Symposium 2011* (*IEEE SAS 2011*), Lafayette, LA, pp 24–28, February 22–24, 2011. doi: 10.1109/SAS.2011.5739810.
47. Bera T. K. and J. Nagaraju, A gold sensors array for imaging the real tissue phantom in electrical impedance tomography, *International Conference on Materials Science and Technology 2012* (*ICMST 2012*), Department of Physics, St.Thomas College Pala, Kottayam, Kerala, India.
48. Rosell J. and P. Riu, Common-mode feedback in electrical impedance tomography, *Clinical Physics and Physiological Measurement*, 13(Suppl. 4): 11–14, 1992.
49. Vauhkonen M., W. R. B. Lionheart, L. M. Heikkinen, P. J. Vauhkonen, and J. P. Kaipio, A MATLAB package for the EIDORS project to reconstruct two dimensional EIT images, *Physiological Measurement*, 22: 107–111, 2001.
50. Bera T. K. and J. Nagaraju, A study of practical biological phantoms with simple instrumentation for electrical impedance tomography (EIT), *Proceedings of the IEEE International Instrumentation and Measurement Technology Conference* (*I2MTC2009*), Singapore, May 5–7, 2009, pp. 511–516.
51. Bera T. K. and J. Nagaraju, Surface electrode switching of a 16-electrode wireless EIT system using RF-based digital data transmission scheme with 8 channel encoder/decoder ICs, *Measurement*, 45: 541–555, 2012, doi:10.1016/j.measurement.2011.10.012
52. Bera T. K. and J. Nagaraju, Studying the elemental resistivity profile of electrical impedance tomography (EIT) images to assess the reconstructed image quality, *Communications in Computer and Information Science*, 157(CCIS-157): 621–630, 2011. ISBN 978-3-642-22785-8, doi: 10.1007/978-3-642-22786-8_78.
53. Bera T. K. and J. Nagaraju, A multifrequency electrical impedance tomography (EIT) system for biomedical imaging, *International Conference on Signal Processing Communications* (*SPCOM 2012*), IISc-Bangalore, Karnataka, India.

54. Luo S., V. X. Afonso, J. G. Webster, and W. J. Tompkins, The electrode system in impedance-based ventilation measurement, *IEEE Transactions on Biomedical Engineering*, 39(11), 1130–1141, November 1992, doi: 10.1109/10.168692.
55. Yan. W., S. Hong, and R. Chaoshi, An optimizing design method for electrode structure parameter of electrical impedance tomography, *Proceedings of 27th Annual Conference on IEEE Engineering and Medical Biology*, Shanghai, China, September 1–4, 2005.
56. Yan W., S. Hong, and R. Chaoshi, Optimum design of electrode structure and parameters in electrical impedance tomography, *Physiological Measurement* 27: 291–306, 2006 doi: 10.1088/0967-3334/27/3/007.
57. Hua P., E. J. Woo, J. G. Webster, and W. J. Tompkins, An electrical impedance tomography using compound electrodes, *11th Annual International Conference of IEEE Engineering Medical and Biological Society*. Ce2770-6/09/0000-0472 $01.00 C 1989 IEEE.
58. Hua P., E. J. Woo, J. G. Webster, and W. J. Tompkins, Using compound electrodes in electrical impedance tomography, *IEEE Transactions Biomedical Engineering*, Seattle, WA, 40(1): 29–34, January 1993.
59. Rigaud B., P. M. Record, J. Anah, and J.-P. Morucci, Active electrodes for electrical impedance tomography the limitation of active stray capacitance compensation, *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 13(4): 1587–1588, 1991. CH3068-4/91/oooO-158870 1.00 Q 1991 IEEE.
60. Jossinet J., C. Tourtel, and R. Jarry, Active current electrodes for in vivo electrical impedance tomography, *Physiological Measurement*, 15: A83–A90, 1994.
61. Li J. H., C. Joppek, and U. Faust, Fast EIT data acquisition system with active electrodes and its application to cardiac imaging, *Physiological Measurement*, 17: A25–A32, 1996.
62. Gaggero P. O., A. Adler, J. X. Brunner, S. Böhm, and P. Seitz, Active electrode based electrical impedance tomography system, EIT 2011, http://www.sce.carleton.ca/faculty/adler/publications/2011/gaggero-EIT2011-active-electrodes.pdf
63. Gaggero P. O., A. Adler, J. Brunner, and P. Seitz, Electrical impedance tomography system based on active electrodes, *Physiological Measurement*, 33: 831–847, 2012. http://www.sce.carleton.ca/faculty/adler/talks/2011/gaggero-EIT2011-active-electrode-pres-EIT.pdf.
64. Yu F.-M., C.-N. Huang, F.-W. Chang, and H.-Y. Chung, A rotative electrical impedance tomography reconstruction system, *Journal Physics: Conference Series*, 48: 542–549, 2006.
65. Huang C.-N., Study of rotational electrical impedance tomography, PhD Thesis, Department of Electrical Engineering, National Central University, Taiwan, Republic of China, 2008.
66. Ito T., S. Yamada, and H. Kazama, Measurement of body fat distribution by using electrical impedance tomography, *Proceedings of SICE-ICASE International Joint Conference*, Busan, Korea, pp. 2551–2554, 2006.
67. Alme K. J. and S. Mylvaganam, Analyzing 3D and conductivity effects in electrical tomography systems using COMSOL multiphysics EM module, *Proceedings of the Nordic Comsol Conference 2006*, Copenhagen, Denmark, November 1–2, 2006.
68. Linderholm P., A. Bertsch, and P. Renaud, Resistivity probing of multi-layered tissue phantoms using microelectrodes, *Physiological Measurement*, 25: 645–658, 2004.
69. Kimoto A. and K. Shida, Imaging of temperature-change distribution in the brain phantom by means of capacitance measurement, *IEEE Transactions on Instrumentation and Measurement*, 49(3): 591–595, 2000
70. Zhang S, G. Xu, H. Wu, D. Geng, and W. Yan, Multi-frequency EIT hardware system based on DSP, *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Suppl: 6677–6680, 2006.
71. Hartinger A. E., H. Gagnon, and R. Guardo, A method for modelling and optimizing an electrical impedance tomography system, *Physiological Measurement* 27: S51–S64, 2006. doi:10.1088/0967-3334/27/5/S0
72. Griffithst H., Z. Zhangf, and M. Watts, A constant-perturbation saline phantom for electrical impedance tomography, *Physics in Medicine and Biology*, 34(8): 1063–1071, 1989.

73. Hahn G., A. Just, J. Dittmar, and G. Hellige, Systematic errors of EIT systems determined by easily-scalable resistive phantoms, *Physiological Measurement*, 29: S163–S172, 2008.
74. McArdlet F. J., B. H. Brownt, R. G. Pearse, and D. C. Barbert, The effect of the skull of low-birthweight neonates on applied potential tomography imaging of centralised resistivity changes, *Clinical Physics and Physiological Measurement*, 9(Suppl. A): 55–60, 1988.
75. Sadleiry R. and R. Foxz, Quantification of blood volume by electrical impedance tomography using a tissue-equivalent phantom, *Physiological Measurement*, 19: 501–516, 1998.
76. Vauhkonen P. J., M. Vauhkonen, T. Makinen, P. A. Karjalainen, J. P. Kaipio, Dynamic electrical impedance tomography—Phantom studies, *Inverse Problems in Engineering*, 8: 495–510, 2000.
77. Hartinger A. E., H. Gagnon, and R. Guardo, Accounting for hardware imperfections in EIT image reconstruction algorithms, *Physiological Measurement*, 28: S13–S27, 2007.
78. Hahn G., A. Just, and G. Hellige, Determination of the dynamic measurement error of EIT systems, *ICEBI 2007, IFMBE Proceedings*, 17: 320–323, 2007.
79. Bagshaw A. P., A. D. Liston, R. H. Bayford, A. Tizzard, A. P. Gibson, A. T. Tidswell, M. K. Sparkes, H. Dehghani, C. D. Binnie, and D. S. Holdera, Electrical impedance tomography of human brain function using reconstruction algorithms based on the finite element method, *NeuroImage*, 20: 752–764, 2003.
80. Cook R. D., G. J. Saulnier, D. G. Gisser, J. C. Goble, J. C. Newell, and D. Isaacson, ACT3: A high-speed, high-precision electrical impedance tomography, *IEEE Transactions on Biomedical Engineering*, 41(8), 713–722, August 1994.
81. Boverman, G, D. Isaacson, G. J. Saulnier, and J. C. Newell, Methods for compensating for variable electrode contact in EIT, *IEEE Transactions on Biomedical Engineering*, 56(12): 2762–2772, December 2009.
82. Yerworth R. J., R. H. Bayford, G. Cusick, M. Conway, and D. S. Holder, Design and performance of the UCLH Mark 1b 64 channel electrical impedance tomography (EIT) system, optimized for imaging brain function, *Physiological Measurement*, 23: 149–158, 2002. PII: S0967–3334(02)30765-2.
83. Xu G., R. Wang, S. Zhang, S. Yang, G. A. Justin, M. Sun and W. Yan, A 128-Electrode three dimensional electrical impedance tomography system, *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biological Society*, Cité Internationale, Lyon, France, August 23–26, 2007.
84. Huang C.-N., F.-M. Yu, and H.-Y. Chung, Rotational electrical impedance tomography, *Measurement Science and Technology*, 18: 2958–2966, 2007. doi:10.1088/0957-0233/18/9/028.
85. Yu F.-M., C.-N. Huang, F.-M. Hsu, and H.-Y. Chung, Pseudo electrodes driven patterns for electrical impedance tomography, *SICE Annual Conference 2007*, Sept. 17–20, 2007, Kagawa University, Kagawa, Japan.
86. Tehrani J. N., C. Jin, and A. L. McEwan, Modelling of an oesophageal electrode for cardiac function tomography, *Computational and Mathematical Methods in Medicine*, Volume 2012, Article ID 585786, 10p., http://dx.doi.org/10.1155/2012/585786.
87. Martinsen O. G., H. Kalvøy, S. Grimnes, B. Nordbotten, P. K. Hol, E. Fosse, H. Myklebust and L. B. Becker, Invasive electrical impedance tomography for blood vessel detection, *Open Biomedical Engineering Journal*, 4: 135–137, 2010.
88. Murphy S. C. and T. A. York, Electrical impedance tomography with non-stationary electrodes, *Measurement Science and Technology*, 17: 3042–3052, 2006. doi:10.1088/0957-0233/17/11/025
89. Vauhkonen, M. Electrical impedance tomography and prior information, PhD thesis, Kuopio University Publications C. Natural and Environmental Sciences 62, 110p., 1997. ISBN 951-781-700-2, ISSN 1235-0486.
90. Metherall P., D. C. Barber, R. H. Smallwood, and B. H. Brown, Three-dimensional electrical impedance tomography, *Nature*, 380: 509–512, April 11, 1996. doi:10.1038/380509a0.
91. Brown D. C. and A. D. Seagar, Limitations in hardware design *Clinical Physics and Physiological Measurement*, 8(Suppl. A) 91–98, 1987.

92. Sakamoto K., T. J. Yorkey and J. G. Webster. Some physical results from an impedance camera, *Clinical Physics and Physiological Measurement*, 8(Suppl. A) 71–76, 1987.
93. Jossinet J and C. Trillaud, A high contrast dual frequency multi-electrode system for electrical impedance tomography, *Proceedings of Copenhagen Meeting Electrical Impedance Tomography*, ed. T J Hames (Sheffield: Sheffield University), South Yorkshire, U.K., pp 144–149, 1990.
94. Trillaud C. and J. Jossinet, An improved design of voltmeter for semi-parallel data acquisition, *Clinical Physics and Physiological Measurement*, 13(Suppl. A): 5–10, 1992.
95. Record P. M., R. Gadd, and P. Rolfe, A signal-conditioning electrode for electrical impedance tomography, *Proceedings of Copenhagen Meeting Electrical Impedance Tomography*, ed. T J Hames (Sheffield: Sheffield University), South Yorkshire, U.K., pp. 168–174, 1990.
96. Rigaud B., J. Anah, P. Givelin, P. Grazoitin, and J. P. Morucci, A multi-function electrode module for electrical impedance tomography, *Proceedings of Copenhagen Meeting Electrical Impedance Tomography*, ed. T. J. Hames (Sheffield: University of Sheffield), South Yorkshire, U.K., pp. 175–182, 1990.
97. McLeod C. N., W. R. Breckon, and D. Murphy, OXPACT: The development of an adaptive current tomograph, *Proceedings of Copenhagen Meeting Electrical Impedance Tomography* ed T J Hames (Sheffield: Sheffield University), South Yorkshire, U.K., pp. 191–197, 1990.
98. Riu P., J. Rossel, A. Lozano, and R. Pallas-Areny, A broadband system for multifrequency static imaging in electrical impedance tomography, *Clinical Physics and Physiological Measurement*, 13(Suppl. A) 61–66, 1992.
99. Sansen W., B. Geeraerts, W. Van Petegem, and M. Stetaert, Electrical impedance tomography systems based on voltage drive, *Clinical Physics and Physiological Measurement*, 13(Suppl. A): 35–42, 1992.
100. Record P. M. and R. W. Hargreaves, Frequency-independent common-mode voltage compensation for EIT applications, *Innovation and Technology in Biology and Medicine*, 15(Suppl. 1): 9–15, 1994.

62

Patient-Care Sensing and Monitoring Systems

Akihiro Kajiwara
University of Kitakyushu

Ryohei Nakamura
University of Kitakyushu

62.1 Introduction

Wireless sensors that have more advantages relative to wired sensors are playing an important role in improving the quality of life in daily indoor space. The applications include home security systems to improve the safety of our home and vital signs monitoring systems to improve our health. The sensor systems can remotely detect the echo reflected from a person and then estimate the state and/or vital signs such as respiration rate and heart rate, for example. Several types of wireless sensor device have thus far been developed such as video camera, passive infrared (IR), and microwave. The video camera is unacceptable for some applications from a privacy protection point of view, although it can monitor a wide area of the indoor space. Passive IR sensors are generally developed to be dedicated to a relatively small area such as the home entrance and toilet area. Microwave sensors such as Doppler and frequency modulation continuous wave (FM-CW), which are capable of penetrating a variety of nonmetallic materials such as inner walls, provide a wider coverage area when compared with video cameras and IR, but it is susceptible to interference from other wireless radio systems, thereby causing false alarms or false detection. Ultrawideband impulse-radio (UWB-IR) has lately attracted considerable attention in short-range remote sensor applications since it offers high-ranging accuracy, multipath reduction, and environmental friendliness due to the very low energy emission [1–4]. However, it requires very-high-speed analog-to-digital converter (ADC) and high-level processors in order to synchronize and detect the received nanosecond pulse. It may also cause interference to existing or future wireless systems using the same or nearby bands because it occupies a bandwidth wider than 500 MHz. Therefore, the UWB device operated in the lower band of 3.4–4.8 GHz is required to implement the detect-and-avoid (DAA) technique in many countries such as Europe, Korea, and Japan [5].

To solve these problems, the use of the stepped-FM scheme has been suggested for the UWB sensor. The scheme has the following main advantages relative to UWB-IR [6–10]:

- *Lower-speed ADC and lower-level processor*: It transmits a series of bursts of narrowband pulses where each burst is a sequence consisting of many pulses shifted in frequency from pulse to pulse with a fixed frequency step. Each received narrowband pulse is phase-detected and then combined into the large effective bandwidth (sequentially over many pulses). Therefore, the hardware requirement is less stringent relative to UWB-IR. The detector bandwidth is smaller, resulting in lower noise bandwidth and higher SN ratio when compared with UWB-IR.
- *Inherent DAA function*: It is capable of coexisting with other narrowband wireless systems operating in the same frequency range. It detects interference radio potential by searching and is then designed to have some spectrum hole (nonactivated within a portion of the wide radio spectrum) to prevent any conflict and to coexist with the other narrowband wireless systems [11]. For example, the frequency band of interference can be detected by the phase detector, while the spectrum hole is adaptively assigned according to the interference band.

We have fabricated the sensor setup where the algorithm detecting and avoiding the interference previously mentioned has also been developed. Note that the ADC speed is 10 ks/s for the 2 GHz bandwidth. Measurements were conducted for some scenarios and the results are presented. From the results, the scheme is shown to be useful as a wireless sensor and can also coexist with other wireless systems operating in an overlaid frequency band.

62.2 Stepped-FM UWB Sensor

62.2.1 UWB-IR Sensor

UWB-IR uses nanosecond pulses spreading out all over a continuous wide bandwidth as shown in Figure 62.1a. It offers many advantages such as low-cost implementation, low transmission power, ranging, multipath immunity, and low interference. Due to the ultrashort-duration pulses, subcentimeter ranging is possible, thereby resulting in the correct identification of the complex-shaped target. It does not require up and down conversion; thus, it may reduce the implementation cost and low power consumption. In spite of the advantages, however, there are several engineering issues that need to be considered. Due to the very short pulses, accurate synchronization and detection may be difficult. Interference to existing or future radio systems using the same band is one of the difficult issues to be solved. Designing wideband RF components is also a big challenge, which includes high-speed ADCs.

Now, the received 1D signal, which is referred to as range profile, is generally presented by multiple impulses with gains $\{\beta_k\}$ and propagation delays $\{\tau_k\}$, where k is the impulse index. Suppose for a nanosecond pulse of $s(t)$, the range profile, $y(\tau,t)$, is the time convolution of $s(t)$ and the impulse echo response $\sum \beta_k \delta(t-\tau_k)$ is as follows:

$$y(\tau,t) = \sum_k \beta_k s(t-\tau_k) \tag{62.1}$$

Figure 62.1b shows an example of received power range profile for a bandwidth of 1 GHz in a room (corresponding to 1 ns pulse).

62.2.2 Stepped-FM UWB Sensor

A different approach of UWB, which solves the previous problems, is the use of stepped-FM pulse sequences instead of transmitting a nanosecond short pulse. We have fabricated an experimental setup of a stepped-FM UWB sensor. The block diagram is illustrated in Figure 62.2a. Waveforms at each stage and the external photo are given in Figure 62.2b and c, respectively. The sensor transmits a series of bursts of narrowband pulses, where each burst is a sequence consisting of N narrowband pulses shifted

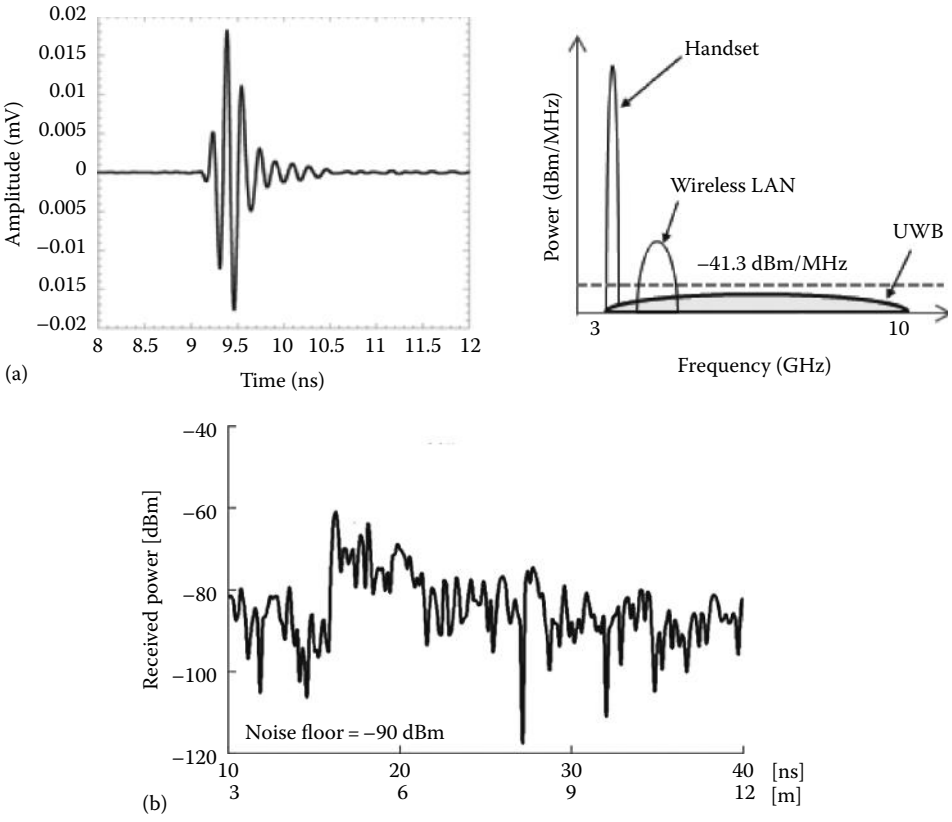


FIGURE 62.1 UWB-IR: (a) Waveform and the spectrum and (b) example of power range profile (1 GHz bandwidth).

in frequency from pulse to pulse with a fixed frequency step Δf [4]. As shown in Figure 62.2a through c, the received echo from an object is phase-detected with the transmitted stepped signal (homodyne detection) and is then In-phase and quadrature-phase (I-Q) sampled by a relaxed speed of ADC.

Then, the nth complex sample R_n is given by

$$R_n = A_n \exp(-j\theta_n) \tag{62.2}$$

$$\theta_n = 2\pi(f_c + (n-1)\Delta f) \cdot \frac{2d}{c} \tag{62.3}$$

where

A_n is the amplitude of nth pulse (A_n can be approximated by A for a stationary object)
f_c is the fundamental frequency
c is the velocity of light

Next, each complex sample is applied to the inverse discrete Fourier transformation (IDFT) device in order to obtain an N-element synthetic range profile, which is called range spectrum. The N-element range spectrum is given by

$$R(\phi) = \left| \sum_{n=1}^{N} R_n \cdot \exp\left(j\frac{2\pi}{N}(n-1)\cdot\phi \right) \right|$$

$$= N \cdot A \cdot \left| \frac{\sin c[\pi(\phi - N\Delta f(2d/c))]}{\sin c[(\pi/N)(\phi - N\Delta f(2d/c))]} \right| \tag{62.4}$$

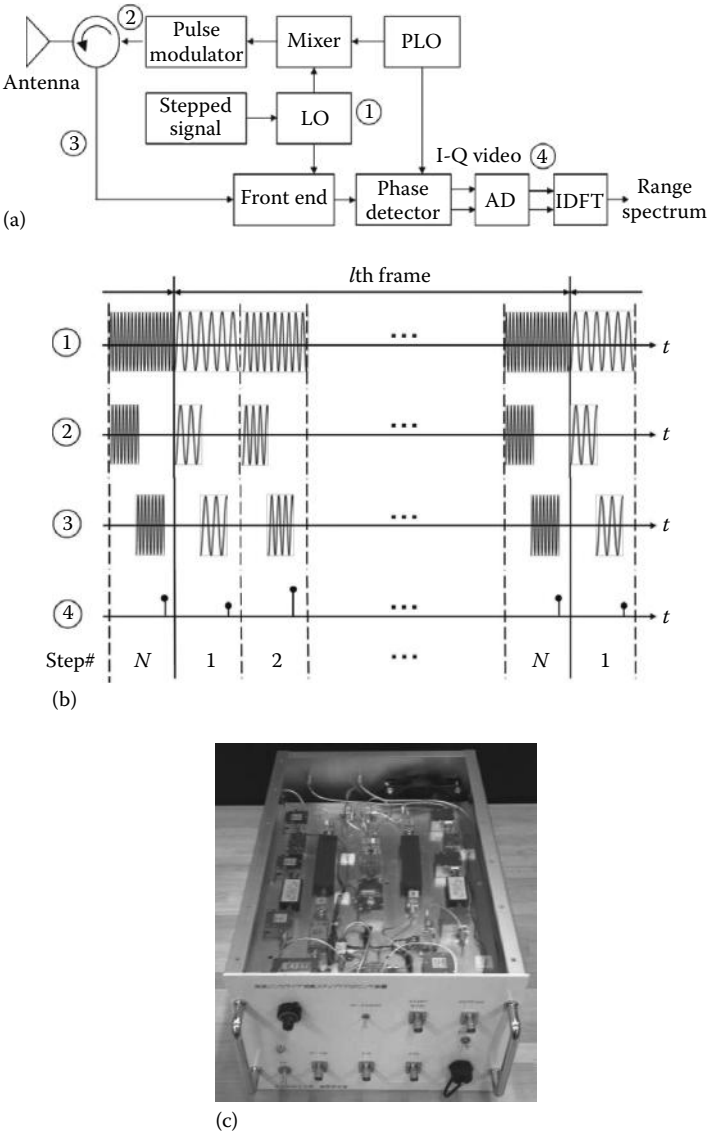


FIGURE 62.2 Stepped-FM UWB sensors: (a) Block diagram, (b) signal waveforms, and (c) external photo of setup.

$$\phi = \frac{2dN\Delta f}{c} \tag{62.5}$$

It is clear that the range resolution ΔR is approximately $1/N\Delta f$. For example, suppose $\Delta f = 34.5$ MHz and $N = 30$, the resolution is approximately 30 cm, which is equivalent to a UWB-IR with 1 GHz. Hence, it does not require high-speed ADC devices and a high-level processor at the receiver. Note that the unambiguous range (maximum detectable range) $R_{\max}$ is given by $c/2\Delta f$.

An example of the range spectrum is shown in Figure 62.3 where $\Delta f = 34.5$ MHz and $N = 30$ and zero-padding was also used for the IDFT operation. The phase detector should be generally AD sampled by 10 times of stepped pulse repetition rate or more. For example, when the pulse repletion rate of the setup is 1 kHz, the used ADC is 10 kS/s. Therefore, the hardware requirements become less stringent when compared with UWB-IR.

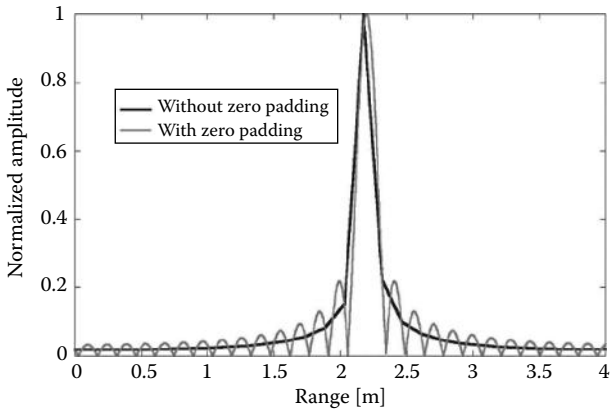


FIGURE 62.3 Example of range spectrum with zero padding ($\Delta f = 34.5$ MHz and $N = 30$).

62.2.3 DAA and Spectrum Hole

The stepped-FM UWB sensor can detect radio interference by monitoring the phase detector output in passive mode (transmitter power off/receiver on) and is then designed to have some spectrum hole (nonactivated within a portion of the wide radio spectrum) at the transmitter in order to prevent any conflict and to coexist with the other narrowband wireless systems. Note that the transmit radio consists of independent narrowband pulses with a different frequency. As such, it provides the flexibility to support regulatory measures in different areas of the world and ease concerns about interference. This section describes an interference avoidance technology called spectrum hole.

Figure 62.4a shows the power spectrum for $\Delta f = 34.5$ MHz and $N = 30$ where some stepped pulses from 3.655 to 3.724 GHz are not transmitted corresponding to the spectrum hole of 6.6% (simulation result). The band is seen to be suppressed less than –13 dB. Consider the federal communications commission (FCC)

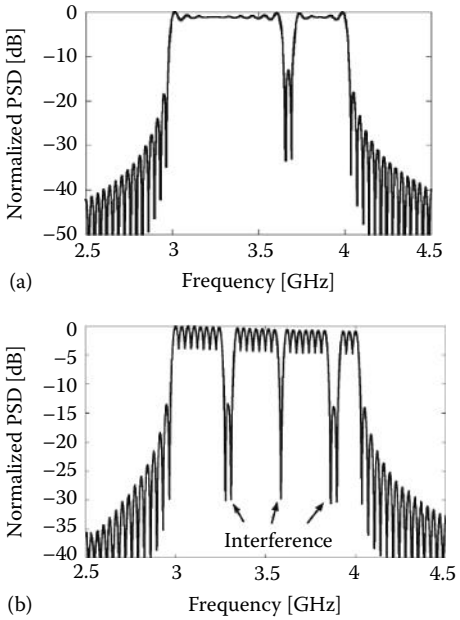


FIGURE 62.4 Power spectrums of transmit signal with spectrum hole: (a) Spectrum hole of 6.6% (b) spectrum hole of 16.6%.

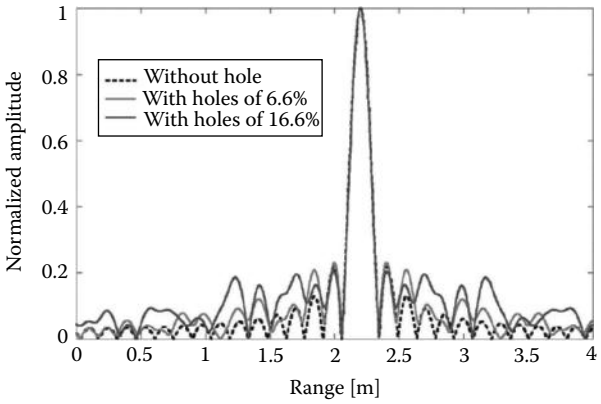


FIGURE 62.5 Range spectrums with and without spectrum hole.

regulation of −41.3 dBm/MHz, the band is suppressed to less than −55 dBm. Therefore, it does not interfere with existing radios. Figure 62.4b also shows the power spectrum against three narrowband interferences where we assumed three spectrum holes of 3.310–3.379 GHz, 3.621–3.655 GHz, and 3.897–3.966 GHz (corresponding to 16.6%) as shown in Figure 62.4b. Figure 62.5 shows the range spectrums for 6.6% and 16.6% where a zero-padding was used for the IDFT processing (1024 points).

62.3 Detection and Avoid Technology

The UWB device operating in the 3.1–4.2 GHz band is required to implement a DAA technology (narrowband signal detection and avoidance function) that allows it to detect an active wireless system operating in the same frequency range. The stepped-FM UWB sensor has inherently a DAA function unlike UWB-IR.

The flow chart of the suggested DAA algorithm is shown in Figure 62.6. Prior to transmitting the UWB radio, the receiver is activated (passive mode) and detects some narrowband radios such as

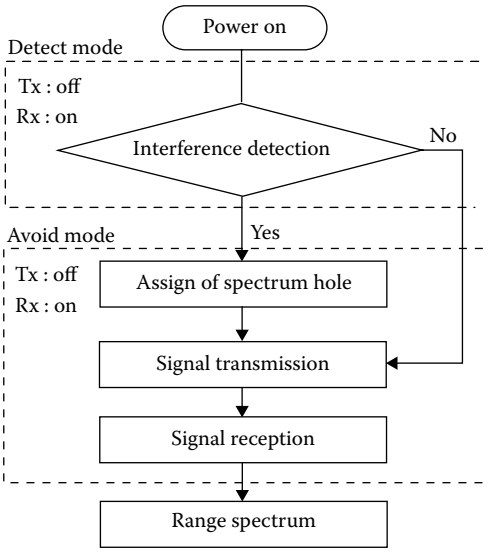


FIGURE 62.6 Flow chart of DAA algorithm.

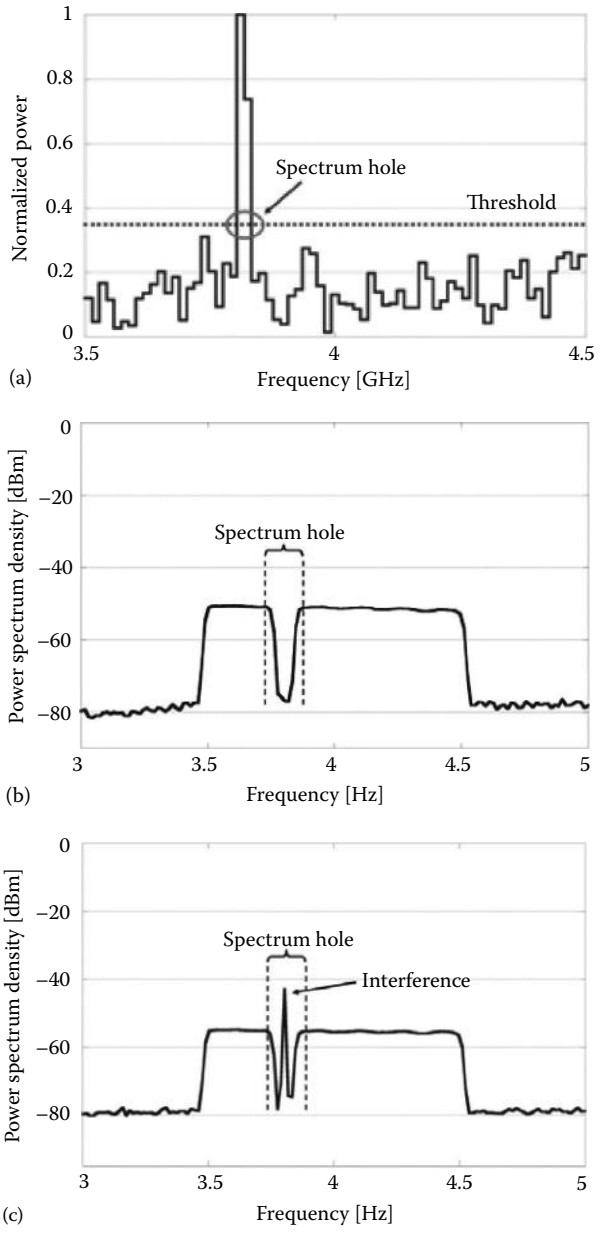


FIGURE 62.7 DAA technology: (a) Phase detector output, (b) power spectrum of transmit signal, and (c) power spectrum of received signal.

communications and broadcast signals. Figure 62.7a shows the output of the phase detector where a narrowband radio of 3.84–3.85 GHz is assumed. In the measurement, the radio was generated by an Agilent E8254A signal generator. It is seen that the radio is detected corresponding to the step number of 23 and 24 where $\Delta f = 14.5$ MHz and $N = 70$. Based on the result, the transmitter is then activated. Figure 62.7b shows the power spectrum where 7 stepped pulses from 3.76 to 3.85 GHz are not transmitted. Figure 62.7c shows the transmit signal together with the interference in frequency domain. It is seen that the UWB radio can coexist with the radio.

62.4 Patient-Care Sensing and Monitoring System

A patient-care sensing and monitoring system has been developed using the stepped-FM UWB sensor. The measurements were conducted and the results are presented.

62.4.1 Sensing and Monitoring Algorithm

The requirement for monitoring the state of the elderly person in care facilities and hospitals is increasing year by year since the increase in accidents involving the elderly persons is of great concern. When the elderly attempt to leave the bed alone, for example, it has been reported that the fall accident occurs frequently. Thereby, it is important to monitor the state of the person. The flow chart is shown in Figure 62.8 where the sensor detects various states or state in a room. The received range spectrum consists of echoes from various obstructions such as the bed, person, and walls where the state should be estimated by employing a "ranging filter" and a "motion filter."

The ranging filter $\Delta R_{pl}(\phi)$ is to detect the range from the sensor to a person, which is given by

$$\Delta R_{pl}(\phi) = \left| R_l(\phi) - R_0(\phi) \right| \tag{62.6}$$

where

$R_0(\phi)$ is a reference range spectrum that represents the range spectrum in a static room without a person
$R_l(\phi)$ is the range spectrum of lth frame

For example, when a person moves, the range spectrum would also fluctuate. Also, some motion in a room is expected to be detected by the motion filter $\Delta R_{vl}(\phi)$, which is given by

$$\Delta R_{vl}(\phi) = \left| R_l(\phi) - R_{l-1}(\phi) \right| \tag{62.7}$$

Note that $R_l(\phi)$ and $R_{l-1}(\phi)$ represent the range spectrum of lth and $(l-1)$th frame, respectively.

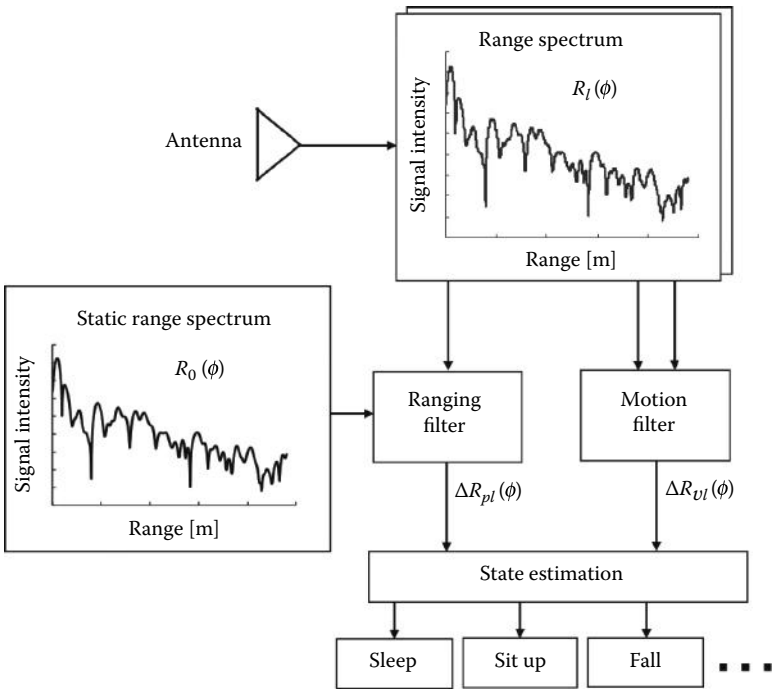


FIGURE 62.8 Flow chart of monitoring algorithm.

The motion filter is to detect some motion in a room, while the ranging filter is to estimate the range of a person. The trajectory of a moving person can also be estimated; thereby the state should be estimated such as "walk in room" and "fall." It is also possible to observe his state in bed such as "tossing about in bed" and "sitting up in bed" using the motion filter without invasion of privacy. Therefore, the state of a person can be detected that includes "out of room," "static," "walk in room," "sleep in bed," "tossing about in bed," "sitting up in bed," and "fall" in this algorithm.

62.4.2 Measurement Results

The measurements were conducted in a care room shown in Figure 62.9 and the results are presented. The specification is shown in Table 62.1. The algorithm has been developed for six states of "out of room ('out' for short)," "walk in room (walk)," "sleep in bed (sleep)," "tossing about in bed (toss about)," "sitting up in bed (sit up)," and "fall." Figure 62.10 represents an example of the measured results where the solid line is the sensor's estimated state, while the dashed line is the actual state identified by a video camera as shown in Figure 62.9. The measurements were conducted for different subjects and the detection rate was investigated. Next, it is important to investigate the detection performance for spectrum hole. The result of "fall," "sit up," "sleep," and "out" is shown in Figure 62.11. Table 62.2 summarizes the results. A detection rate of more than 88% is seen to be attained for a spectrum hole of 10%. Note that the rate is approximately 100% without a spectrum hole.

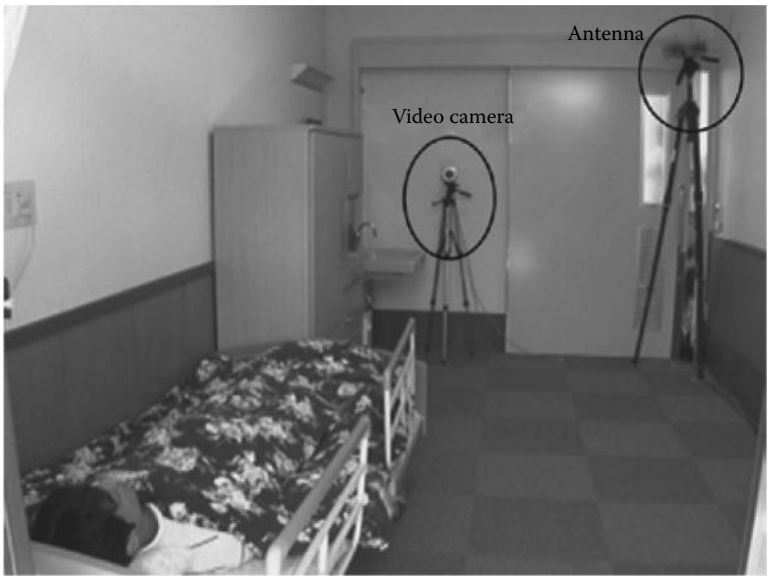


FIGURE 62.9 Scene for measurement environment.

TABLE 62.1 Measurement Specification

Transmit power	−12 dBm
Frequency	3.5–4.5 GHz
Stepped bandwidth Δf	14.5 MHz
Number of steps N	70
Frame period	0.1 s
ADC device	70 kS/s
Antenna	Horn (9.88 dBi)

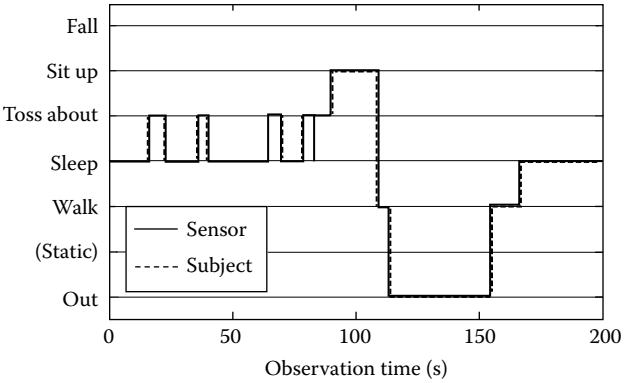


FIGURE 62.10 An example of measured results.

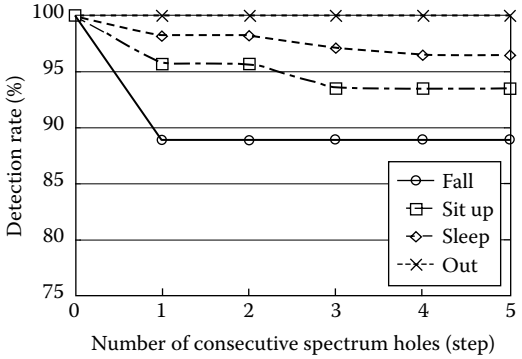


FIGURE 62.11 Detection rate as a function of spectrum hole number.

TABLE 62.2 Detection Rate

State	Detection Rate (%)
Fall	100
Sit up	100
Sleep	100
Out	100
Toss about	98.80
Walk	100

62.5 Conclusions

The increase in accidents involving the elderly patient becomes a great concern, and the requirement for monitoring their activity and state is especially increasing in care facilities. This chapter presents a bed state monitoring sensor of the elderly patient using a stepped-FM UWB scheme, and the performance has been investigated for each spectrum by measurements in care facilities. The sensor provides the following advantages:

- *Lower-speed ADC and lower-level processor*: It transmits a series of bursts of narrowband pulses where each burst is a sequence consisting of many pulses shifted in frequency from pulse to pulse with a fixed frequency step. Each received narrowband pulse is phase-detected and then combined with the large effective bandwidth (sequentially over many pulses). Therefore, the hardware requirement is less stringent relative to UWB-IR.

- *Inherent DAA function*: It is capable of coexisting with other narrowband wireless systems operating in an overlaid frequency band. It can be designed to have any spectrum hole that does not cause interference with other wireless systems and medical equipment. For example, the location of the spectrum hole can be adaptively assigned according to the interference band detected by the DAA.

From the results, the scheme has been found to be useful and can also coexist with other wireless systems operating in the overlaid frequency band. Also, it has been shown that various states of the patient can be detected including "sleeping in bed," "sitting up in bed," "fall," "walk in room," "static," and "going out and in at the door."

References

1. I. Immoreev, S. Samkov, and T. Tao, Short-distance ultra wideband radars, *IEEE Aerospace Electron. Syst. Mag*., 20, 6, 9–14, June 2005.
2. K. Ota, Y. Ota, M. Otsu, and A. Kajiwara, Elderly-care motion sensor using UWB-IR, *Proc. IEEE Sensors Appl. Symp. 2011*, San Antonio, TX, pp. 159–162, February 2011.
3. J. Taylor, *Introduction to Ultra-Wideband Radar Systems*, CRC Press, Boca Raton, FL, 1995.
4. J. Lansford, UWB radio issues and opportunities, *USC UWB Symposium*, April, 2006.
5. Electronic Communications Committee (ECC), ECC REPORT 120, Technical requirements for UWB DAA devices to ensure the protection of radiolocation services in the bands 3.1–3.4 GHz and 8.5–9 GHz and BWA terminals in the band 3.4–4.2 GHz, Kristiansand, Norway, June 2008.
6. D. Wehner, *High Resolution Radar*, 2nd edn., pp. 197–237, Artech House, Boston, MA, 1995.
7. M. Otsu, R. Nakamura, and A. Kajiwara, Remote respiration monitoring sensor using stepped-FM, *Proc. IEEE Sensors Appl. Symp. 2011*, San Antonio, TX, pp. 159–162, February 2011.
8. A. Kajiwara and H. Yamaguchi, Clutter suppression characteristics of stepped-FM radar with MUSIC algorithm, *Trans. IEICE (in Japanese)*, J84-B, 10, 1848–1856, October 2001.
9. M. Otsu, R. Nakamura, and A. Kajiwara, Elderly-care monitoring sensor using stepped-FM UWB scheme, *Proc. IEEE Sensor Appl. Symp. (SAS2012)*, Brescia, Italy, February 2012.
10. Y. Ohta, R. Nakamura, and A. Kajiwara, UWB stepped-FM sensor for home security, *Proc. IEEE Sensor Appl. Symp. (SAS2012)*, Brescia, Italy, February 2012.
11. R. Nakamura and A. Kajiwara, Empirical study of stepped-FM UWB microwave sensor, *Proc. IEEE Radio and Wireless Symp. (RWS2011)*, Phoenix, AZ, pp. 363–366, January 2011.

63

Sensor-Based Activity Recognition Techniques

Donghai Guan
Harbin Engineering University

Weiwei Yuan
Harbin Engineering University

Sungyoung Lee
Kyung Hee University

63.1 Introduction

Understanding human activities and behavior has long been a research goal. In terms of daily lives, awareness of human activities can support many diverse applications. For example, in a home environment, users can be reminded to perform activities (e.g., taking medicine). In a hospital environment, a patient can be reminded to complete an exercise rehabilitation program. Employing a person to monitor other persons' activity 24 h a day is unrealistic; therefore, the automated, automatic recognition of human activities is important and necessary.

Although it is natural for humans to recognize their activities, it is not an easy task for a computer. Computers need to analyze the information gathered from sensors and infer the ongoing activity. Sensors' noises and variances make activity recognition (AR) even more complex for computers. For the recognition of a cooking activity, people might utilize various cues, such as recognizing a meal time, smelling the food being cooked, or seeing that the stove is on. Humans easily use evidence and their past experiences to infer activities. However, all the functions involved in sensing the environment, learning from past experience, and applying knowledge for inference are still a challenge for computers. The goal of AR research is to enable computers to have capabilities comparable to people for recognizing human activity. Only if computers can reliably recognize people's various activities can intelligent systems (such as healthcare and security systems) actually work. Moreover, research on AR has a huge impact on behavioral, social, and cognitive sciences.

To build an automatic AR system, the first thing to consider is the system's sensing ability, how the system recognizes the state of the physical world. Sensing ability comes from low-level sensors (Figure 63.1).

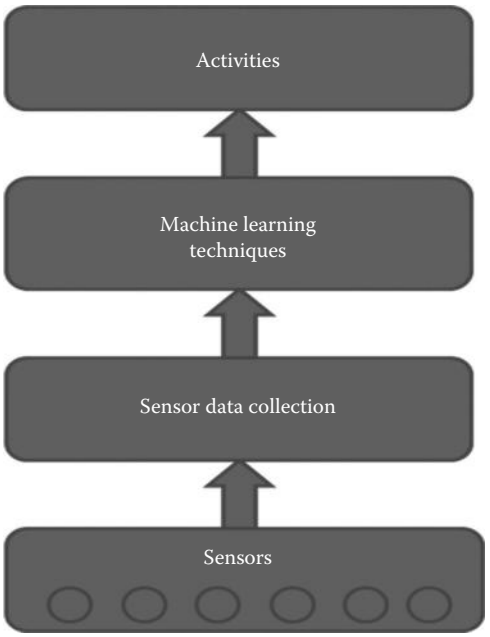


FIGURE 63.1 Workflow of AR system.

Sensed information is processed by machine learning techniques (e.g., noise filtering, feature extraction, and classification techniques) to infer the human activity (e.g., sitting, raising hand, or raising foot). Due to the variations in sensing techniques, a considerable number of AR systems have been developed.

According to the sensing techniques, the existing AR can be roughly divided into three categories:

1. *Video sensor–based activity recognition* (*VSAR*). This is the classical AR approach, which attempts to capture a human's activity information using appropriately placed cameras. Video analysis is the most important and challenging aspect of VSAR.
2. *Wearable sensor–based activity recognition* (*WSAR*). In this AR, physical sensors are attached to the body. This is a relatively new approach, which emerged with the development of wearable computing. Motion sensors (accelerometers and gyroscopes) are the most common wearable sensors. They are attached to relevant parts of the body, in particular hands and arms, to provide information about the limbs' motions. Other sensors, which are usually used as complements to the motion sensors, include microphones, Global Positioning System (GPS) devices, and light sensors. They provide additional information about the user's environment to help improve the accuracy of AR.
3. *Object usage–based activity recognition* (*OUAR*). In this AR, sensors are attached to objects. With this design, a system can recognize the objects accessed by the users and what the users do with them. Although OUAR does not explicitly monitor human activity, it is applicable, because many activities can be implicitly inferred by knowing the objects utilized. For example, if we know that the stove is on, we might infer that someone is cooking. Binary sensors and radio frequency identification (RFID)-based sensors are commonly used sensors in OUAR. They are cheap and can be ubiquitously deployed.

Of the different types of sensor-based AR systems, no one is the best. Each of them has its own characteristics and suitable applications. In this chapter, we outline the strong/weak points, main techniques, and suitable applications of each type of system, ultimately comparing them.

63.2 Video Sensor–Based Activity Recognition

VSAR is the classical AR method and is based on computer vision techniques (Figure 63.2). It is most common among the three AR approaches. Currently, it is widely used in various domains including industry, academia, security, and consumer agencies.

63.2.1 Applications

In this section, we present some typical applications of VSAR systems:

1. *Surveillance systems*. Surveillance systems can support many applications, such as healthcare and unusual event alarms. Traditional surveillance systems require a human operator to constantly monitor the scene, which is tedious and inefficient. Video-based activity recognition systems can replace or help human operators monitor anomalies and interesting activity.
2. *Human–machine interaction*. Human–machine interaction needs to understand a human's interactions. People's activity information is particularly important in understanding intentions and purposes. For example, researchers have proposed vision-based activity recognition to develop alternatives to the traditional mouse- and keyboard-controlled Graphic User Interface (GUI).
3. *Sports analysis*. In many kinds of sports (e.g., basketball, soccer), during a match or training, a player's performance and behavior is recorded. This has many potential usages. For example, the coach might adjust a soccer player's shooting action after analyzing the videos. Or during the match, if a penalty dispute happens, a referee can analyze the video. Due to the explosive increase of video use, it has been necessary to develop efficient indexing schemes, and sports players' activity information is one of the most popular schemes for video indexing.

63.2.2 Feature Extraction in VSAR

The major steps involved in VSAR include (1) video (or image) gathering, (2) feature extraction from images, and (3) AR based on the extracted features.

Among these three steps, feature extraction is important and challenging. Extracting good features is the mark of good recognition performance. The good features should generalize small variations in the images, such as personal appearance, background, and viewpoint and action execution, while being sufficient enough to support a robust recognition of activities.

The features in VSAR can be divided into two categories: global features and local features. Global features need the procedure of localizing a person and encoding a region of interest (ROI) and have powerful expressing abilities when rich information is encoded. However, they need accurate localization,

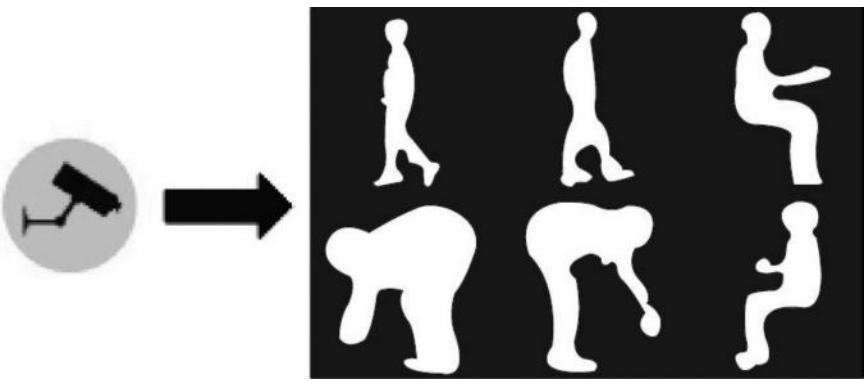


FIGURE 63.2 Video sensor–based activity recognition.

background subtraction, and tracking. Conversely, local features describe observations as a collection of independent patches, detecting patio-temporal interest points, and calculating the local patches around these points. Finally, the patches are integrated into a final representation. Compared with global features, local features are less sensitive to noises and do not require high-quality background subtraction or tracking; however, they require more preprocessing efforts to extract a sufficient amount of relevant interest points.

63.2.2.1 Global Features of VSAR

Global features represent the ROI of a person as a whole. The ROI is usually defined using background subtraction or tracking. Common global features are derived from silhouettes, edges, optical flow, or trajectories.

Background subtraction segments a scene into a background and a foreground to isolate the moving parts. The segmentation evaluates the difference in pixel features for the current scene versus the reference background image. Background subtraction is sensitive to environmental changes, like illumination variations.

In general, human silhouette information is commonly employed to represent human activities [1]. In fact, human activities can be considered as temporal variations of human silhouettes. In addition to silhouette shape, motion information is also used. Optical flow is defined as the apparent motion of individual pixels on the image plane. It often serves as a good approximation of true physical motion projected onto an image plane.

Trajectories of moving objects have also been used as features to infer the activities of objects [2]. The image-plane trajectory itself is not very useful, as it is sensitive to translations, rotations, and scale changes. Alternative representations such as trajectory velocities, trajectory speeds, and relative motion have been proposed.

Although the features extracted using previous methods are informative, they are sensitive to noise, partial occlusions, and variations in viewpoint. To overcome these limitations, grid-based features and space–time volumes have been proposed. Global grid-based features divide the ROI into a fixed spatial or temporal grid. Each cell in the grid describes the image observation locally, thus partial occlusions and changes in viewpoints can be partially solved [3]. Multiple images over time can be stacked to form a 3D space–time volume. A space–time volume was originally proposed in [4] and has since been used by other research.

63.2.2.2 Local Features of VSAR

Local features describe the observations as a collection of local patches, which can be samples, either densely spaced or at space–time interest points. Superior to global features, local features do not require accurate localization and background subtraction. The first type of local features is space–time interest point detectors, which are the locations in space and time where sudden changes of movement occur in the video. The locations are viewed as the most informative source for human AR. The research that uses space–time interest points includes [5,6]. The main limitation with using these features is that the number of stable interest points is normally small. Fortunately, this limitation has been partially addressed in [7,8].

Local descriptors are the second type of local features. They summarize an image or video patch in a representation that is ideally invariant to background clutter, appearance and occlusions, and possibly to rotation and scale. The research that has used this type of features includes [9,10]. Comparing sets of local descriptors is not straightforward due to the possibly of different numbers and the high dimensionality of the descriptors. Therefore, a codebook is usually generated by clustering patches and selecting the cluster centers of the closest patches as codewords.

Similar to the global grid-based features, local grid-based features also exist. These features are the third type of local features and have theoretical similarities to global grid-based features. Related works include [11,12]. The last type of local features is a correlation between local descriptors. Grid-based

representations model spatial and temporal relationships between local descriptors to some extent, but they are often redundant and contain unimportant or uninformative features. Therefore, some researchers exploit correlations between local descriptors as features.

63.2.3 Recognition Techniques in VSAR

Recognition techniques are used to infer activity information based on the aforementioned extracted features. An activity label can be predicted for each frame or sequence of frames. Four types of recognition techniques have been widely used in VSAR: nonparametric, volumetric, temporal-independent, and temporal-based techniques.

63.2.3.1 Nonparametric Techniques

Nonparametric approaches typically extract a set of features from each video frame. The features are then matched to a stored template. The template can be either 2D or 3D. When using nonparametric techniques, the typical procedures consist of motion detection and human tracking in the scene, which enables the construction of a sequence. Then, a periodicity index is computed and the periodicity sequence is segmented into individual cycles for recognition.

63.2.3.2 Volumetric Techniques

Volumetric techniques do not extract features on a frame-by-frame basis. Instead, they use a video as a 3D volume of pixel intensities and extend standard image features to a 3D case. Volumetric techniques mainly include four types of methods: spatiotemporal filtering, which filters video volume using a large filter bank; part-based approach, which considers video volume as a collection of local parts; subvolume matching, which matches videos by matching subvolumes between videos and templates; and a tensor-based approach, which is a generalized collection of matrices for multiple dimensions.

63.2.3.3 Temporal-Independent Techniques

Temporal-independent techniques neglect the information in the temporal domain. They summarize all of the frames for an observed sequence into a single representation, subsequently performing AR for each frame. The main techniques in this category include *k*-nearest neighbors (*k*-NN) and support vector machines (SVMs).

When using *k*-NN, the distance measurement can affect recognition performance. Euclidean distance with global features and histograms of codewords has been used [13]. Other measures, like Mahalanobis distance, have also been utilized.

Another popular temporal-independent recognition method in VSAR is SVM. It determines a hyperplane in feature space, described by a weighted combination of support vectors. SVM has been used jointly with local features of fixed lengths, such as histograms of codewords, for activity learning [14].

63.2.3.4 Temporal-Based Techniques

Temporal-based techniques fully utilize the temporal relationships of activities. They can be either generative or discriminative. Dynamic time warping (DTW) is a generative learning method. It models highly nonlinear warping functions based on the distance measure between two sequences. Hidden Markov models (HMMs) are the other common generative learning method in VSAR. They use hidden states that correspond to different phases in the performance of an action. They model state transition probabilities and observation probabilities. Although HMM is widely used, it is limited to observations that are time independent. Discriminative models overcome this issue by modeling the conditional distribution of activity labels given the observations. Conditional random fields (CRFs) are the most common used discriminative models in VSAR. They use multiple overlapping features. A detailed comparison between CRF and HMM is given in [15].

63.3 Wearable Sensor–Based Activity Recognition

WSAR occurred with the development of wearable computing. The goal is to create personal applications that can adapt and react to the user's current context. The scope of context is diverse and, in WSAR, refers to the current activity information.

Unlike VSAR, WSAR requires more from the sensors; therefore, it is a newer technology, only dating back to the 1990s, when hardware (including sensing, display, and computing equipment) became lightweight enough that an integrated mobile system could be "worn" by a single person for an extended period of time. With the development of hardware techniques, there now exist a large number of WSAR systems that use various sensors, and recognition techniques, and are used for diverse applications.

63.3.1 Applications of WSAR

An important application of WSAR is in healthcare and assisted living. There also exist a number of other applications, such as industrial applications and applications for entertainment and gaming:

1. *Healthcare and assisted living*: WSAR can help elderly people live more independent lives and thus reduces the burden of caregivers (Figure 63.3). By detecting the activities of elderly people, we can implicitly detect potentially dangerous situations in daily life. The typical functions of these systems are to detect when the people fall and to monitor vital signs.
2. *Prevention of diseases*: Some diseases can be detected from activity information prior to being diagnosed (Figure 63.4). For example, Alzheimer's can be detected through early symptoms related to human activity. These applications accumulate and summarize statistics about daily activities and continuous recordings of physiological parameters. These applications can help physicians and caregivers estimate the physical well-being of a person.
3. *Promoting exercise*: This type of system aims to aid elderly or disabled people performing everyday activities and thus generate a healthier lifestyle (Figure 63.5). With the increasing intelligence and processing power of mobile phones, many such applications use mobile phones to remind people to perform certain activities.

63.3.2 Sensors in WSAR

VSAR uses only video sensors; however, WSAR involves many different kinds of sensors. Each type of sensor has its own strengths and suitable applications.

FIGURE 63.3 WSAR for healthcare.

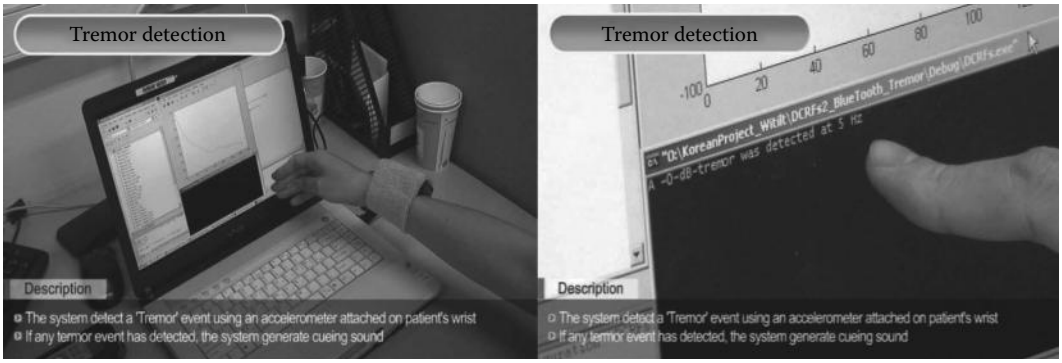


FIGURE 63.4 Tremor detection based on WSAR.

FIGURE 63.5 Gait detection based on WSAR.

The types of sensors used in WSAR range from relatively simple sensors with discrete output, such as ball switches, to sensors with continuous output such as accelerometers, to more complex sensing methods such as audio processing. Video sensors are wearable, with accelerometers being the most commonly used type of sensor, as they usually give good results in terms of physical AR. The most commonly used accelerometers and their vendor information are summarized in Table 63.1.

In addition to the aforementioned sensors, there are some sensors that are less commonly used. Fiber-optical sensors are used to measure posture. Foam pressure sensors are used to measure respiration rates. Force-sensitive resistors are used to measure muscle contractions. Moreover, various kinds of physiological sensors, such as oximetry sensors, skin conductivity sensors, electrocardiographs, and body temperature sensors, have also been used for various applications. As different types of sensors can provide complementary information, using many various sensors simultaneously usually gives more accurate results. For example, the Ubiquitous Computing Lab of Kyung Hee University has developed the Mobile Activity Sensor Logger (MASoL). MASoL consists of

TABLE 63.1 Accelerometers Used in WSAR

Sensor	Vendor
CDXL04M3 [16]	Crossbow technologies
ADXL210E [17]	Analog devices
LIS302DL [18]	ST Microelectronics
MMA7260Q [19]	Freescale semiconductor
ADXL330 [20]	Analog devices

FIGURE 63.6 Mobile Activity Sensor Logger.

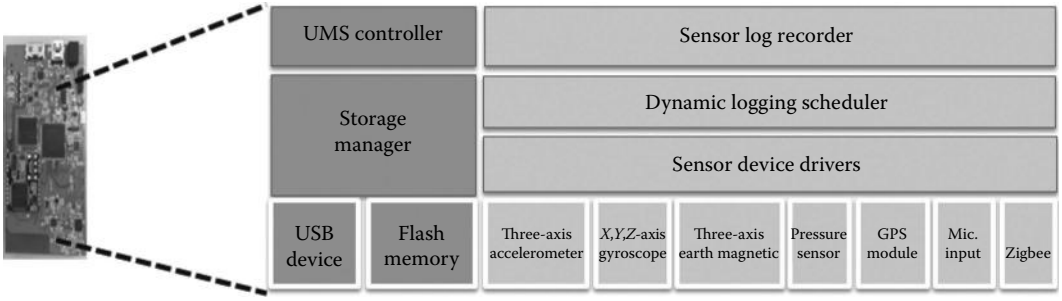


FIGURE 63.7 Key components in MASoL.

various sensors that record a personal activity log. All of the sensors in the MASoL platform are embedded in one chip, so as to decrease user discomfort (Figures 63.6 and 63.7).

Depending on the type of activity, recognition performance can be improved by using the same type of sensor at multiple body locations, employing networks of heterogeneous sensors, or integrating a variety of sensors on a single device. Combining two or more complementary types of sensor data can also help in recognizing activities, by combining motion and audio data, motion and proximity data, or motion and location data.

63.3.3 Recognition Techniques for WSAR

Quite a range of different recognition techniques have been used for WSAR. There is no one best recognition technique for WSAR, as considerations include the kind of data to be processed and the types of activities to be recognized. The recognition techniques for WSAR can be classified into two types: supervised and unsupervised. Supervised learning requires labeled data. Conversely, unsupervised learning tries to directly construct models from unlabeled data, either by estimating the properties of its underlying probability density or by discovering groups of similar examples. When using labeled data, supervised learning usually outperforms unsupervised learning; therefore, supervised learning is the predominant approach for WSAR.

63.3.3.1 Supervised Recognition Techniques

There exists a wide range of algorithms and models for supervised learning. Commonly used methods in the WSAR include Naive Bayes (NB), C4.5 decision trees, and nearest neighbors. HMMs are well-suited for capturing temporal patterns in the data but can be difficult to train due to an abundance of parameters [21,22]. Other methods that have been applied include SVMs [23,24] and more recently string-matching methods. Boosting uses multiple classifiers to boost the classification performance and has also been used in WSARs.

63.3.3.2 Unsupervised Recognition Techniques

In Ref. [25], hierarchies of HMM are used to learn the locations and scenes, such as walking through a supermarket, from audio and video data in an unsupervised fashion. Graphical model–based

unsupervised learning schemes were used in [38,39], wherein the focus was on inferring the transportation modes (such as bus, car, walking) and destinations of the users. One work [26] combines discrete string-matching techniques with continuous HMM classifiers to discover short recurring motifs in acceleration data. They aimed to discover and model short-term motion primitives, such as those occurring during physical exercise. Other research [27] used the concept of multiple eigenspaces for the unsupervised learning of activities such as walking or juggling.

63.4 Object Usage–Based Activity Recognition

In daily life, people usually perform an activity by interacting with a series of objects. For example, when bathing people may interact with a door, light, exhaust fan, shower faucet, etc. The strategy of OUAR is to attach sensors to these objects so that it is possible to determine the state of that object when a person interacts with it.

OUAR [28–30] can be particularly useful in domains such as cooking, which involves a relatively small number of repeated actions (chopping, pouring, spreading, etc). Object use information can help discriminate between activities, such as making toast and making a sandwich, which may be similar from the standpoint of the activity alone. Such distinctions can be important for application domains, such as health monitoring or memory aids.

63.4.1 Sensors in OUAR

There are two main types of sensors used in OUAR systems: RFID-based sensors and simple binary sensors.

63.4.1.1 RFID-Based Sensors

To use an RFID-based sensor, one needs RFID tags and an RFID reader. RFID tags are small and are attached to objects. Subjects need to wear the short-range RFID reader (or equipment with the RFID reader integrated). An RFID-based sensor will generate object use events when a tagged object is manipulated during an activity. Whenever the user's hand is in close proximity to a tagged object, the reader indicates this. The tags are inexpensive and maintenance free given that they do not contain batteries. For example, using an RFID-based sensor, we can detect that a user is reading (Figure 63.8).

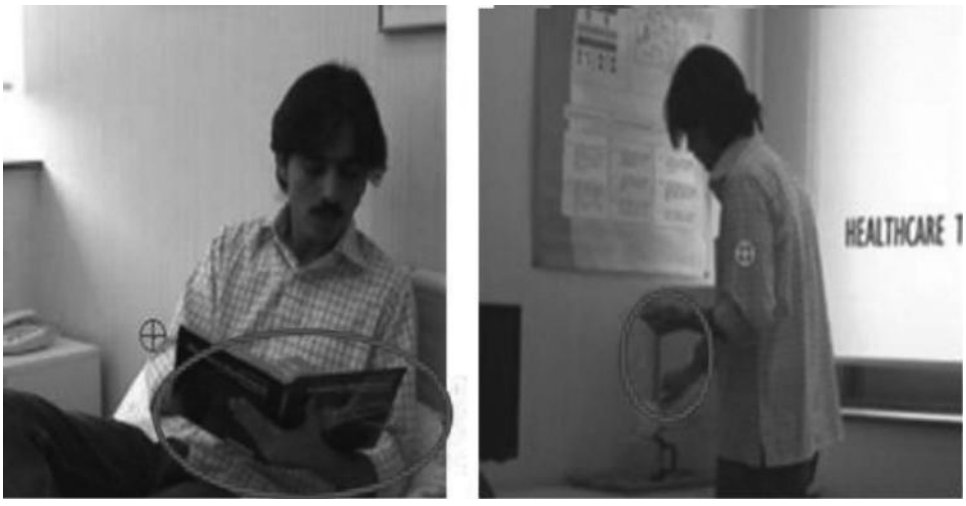


FIGURE 63.8 Using RFID for AR.

63.4.1.2 Binary Sensors

The other type of sensor that has received widespread acceptance in OUAR is the binary sensor. Binary sensors are usually simple and anonymous. They are unable to directly identify occupants; however, at any given time, a binary value may be obtained from them. Whenever the state of a certain context (object, movement) associated with a binary sensor changes, the value of the sensor changes to "1" from "0" (indicating its static state). There are many, diverse types of binary sensors, including movement detectors, contact switches, and pressure mats. They are applicable to different scenarios. For example, contact switches may be installed on the doors of cupboards, the fridge, the microwave, etc. Pressure mats may be discreetly installed in objects such as chairs, sofas, and beds and, in some instances, may be used to locate specific movements within rooms. They have the advantage that they do not require the person being monitored to wear or carry any new technology.

63.4.2 Recognition Algorithms

Temporal probabilistic models are most common in OUAR. Sets of probabilistic models have been proposed in OUAR: the NB in [31], the HMMs in [32], and the CRF in [33]. HMMs are generative probabilistic models consisting of a hidden variable and an observable variable at each time step. For OUAR, the hidden variable is the activity performed and the observable variable is the vector of the sensor readings.

A CRF is a discriminative probabilistic model. Kasteran et al. [32] compared CRF with HMMs; findings found that CRF outperformed HMMs in all cases with respect to time slice accuracy, but that HMMs achieved the overall highest accuracy. This is due to the way both models maximize their parameters. HMMs make use of a Bayesian framework in which a separate model is learned for each class. A CRF uses a single model for all classes. A comparison of HMMs and CRF was also discussed in [34], where the authors found that CRF was able to easily incorporate a wide variety of computed features, which allowed domain knowledge to be added to the models. They also showed that due to the independence assumptions inherent in HMMs, such computed features are not nearly as effective in improving classification accuracy. Thus, CRF's classification accuracy has shown to be consistently higher than HMMs.

63.5 Comparisons of VSAR, WSAR, and OUAR

In this section, we discuss the strengths and limitations of VSAR, WSAR, and OUAR. Based on our comparisons, fusing these three approaches seems to be the most promising solution for complex AR applications.

63.5.1 VSAR

VSAR is the traditional AR method. Among these three approaches, a video system with appropriately placed cameras provides the richest information and in principle could facilitate the most detailed analysis. However, the signal processing involved in extracting this information can be computationally intensive and, in general cases, remains an open problem. In general, VSAR often works well in a laboratory or well-controlled environment. However, it fails to achieve the same level of accuracy for real home settings due to the clutter, variable lighting, and highly varied activities that take place in natural environments. Video sensor–based activity recognition is complex to implement because it requires processing highly multidimensional data. Additionally, video information may violate the user's privacy.

63.5.2 WSAR

The WSAR approach tries to recognize a user's activities by employing sensors such as body-worn accelerometers and microphones to capture characteristic repetitive motions, postures, and sounds of the activities. Using these types of wearable sensors, sensing studies have successfully recognized such

activities as walking, bicycling, brushing teeth, speaking and laughing, and workshop activities such as sawing and drilling that have characteristic motions and/or sounds. An advantage of this approach is that it does not require environment-embedded sensors. That is, this approach incurs no cost in terms of money or time for embedding sensors in indoor objects and furniture. Also, users can easily turn off their wearable devices when they want to preserve their privacy. Wearable sensors are capable of measuring mobility directly. They are well suited to collecting data on daily activity patterns over an extended period of time, as they can be integrated into clothing or worn as wearable devices. Since they are attached to the subjects they are monitoring, wearable sensors can therefore measure physiological parameters, which may not be measureable using ambient sensors.

Although body-attached sensors are promising in identifying primitive sequences of movements, such as walking and running, it is difficult to identify goal-oriented activities (e.g., making tea and taking medicine). In addition, designing of wearable systems is complicated by size, weight, and power consumption requirements.

63.5.3 OUAR

OUAR is the third type of AR approach. It infers a person's activity by analyzing the person's interactions with various objects. This approach is newer than VSAR and WSAR and was proposed to overcome the limitations of VSAR and WSAR. Many people are uncomfortable living with cameras. Moreover, people are often unwilling or forget to wear sensors. However, OUAR does have some drawbacks. For extensive and detailed recognition, OUAR requires a large number of objects to be attached to sensors. This is often either infeasible or too expensive. The cost of the sensors and sensor acceptance are pivotal issues, especially in homes.

Each of the three types of AR approaches has its own advantages and disadvantages; however, given the specific application and associated requirements of this study, there is a most suitable approach (Table 63.2).

As shown in Table 63.2, the three types of AR are complementary. For example, VSAR violates user privacy; however WSAR and OUAR do not violate privacy. OUAR is usually limited to the home environment, whereas WSAR and VSAR can work outside home. In our u-healthcare project, we tried to combine these three approaches for our u-healthcare applications [35].

Our system needs to constantly monitor human activity information to improve health conditions and determine potential health problems as early as possible; therefore, all three types of AR approaches were used together (Figure 63.9).

TABLE 63.2 Analysis of VSAR, WSAR, and OUAR

Categories	Strengths	Limitations	Applications
VSAR	1. Video includes richest information; therefore, its applied range is widest 2. Setting up the environmental is easy	1. Violates privacy 2. Sensitive to environment (e.g., light condition and viewpoint variation)	1. Healthcare 2. Security 3. Interactive applications 4. Content-based video analysis
PSAR			
WSAR	1. Setting up the environmental is not required 2. Good for applications that require explicit motion analysis	1. Wearing sensors is a burden. 2. Unable to separate similar actions (e.g., making tea and coffee)	1. Healthcare 2. Interactive applications
OUAR	1. Does not violate privacy 2. Good at recognizing goal-oriented activities	1. Recognizable activities need to be related to objects 2. Setting up the environmental takes more effort than for VSAR and WSAR	1. Healthcare 2. Security

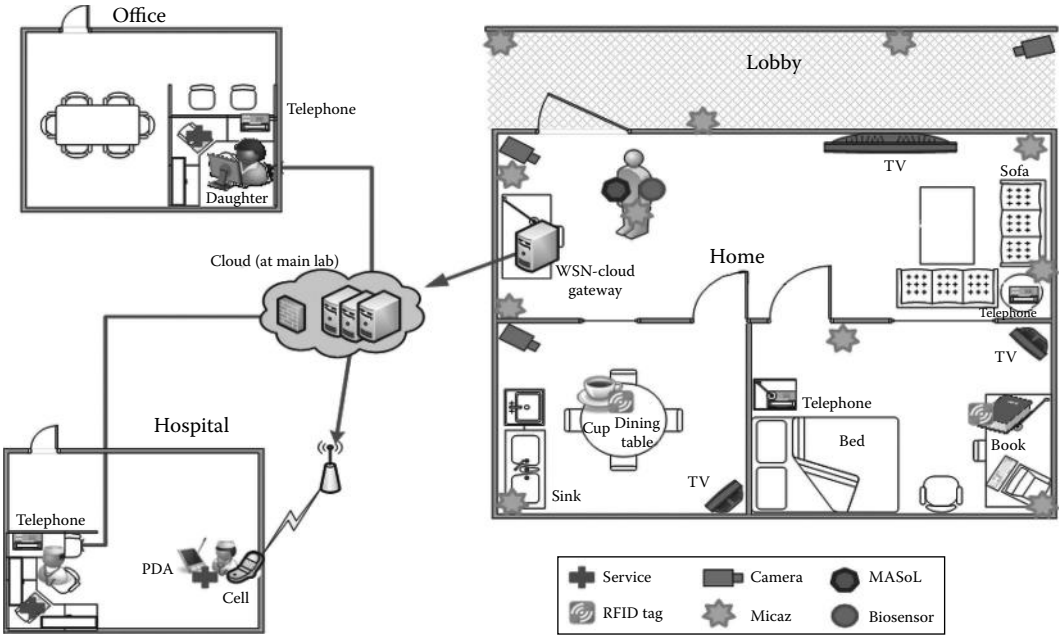


FIGURE 63.9 AR for u-healthcare application. (From http://www.tr.ietejournals.org/article.asp?issn=02564602;year=2011;volume=28;issue=5;spage=418;epage=433;aulast=Guan)

The cameras used for VSAR were in the lobby, living room, and kitchen. Due to privacy issues, they were not used in the bedroom. For WSAR, we developed the MASoL, which is a low-cost, low energy consumption sensor device. MASoL contains 13 axis sensors, gathering many kinds of behaviors and storing them. In addition, we also used the OUAR approach with an RFID-based sensor, with RFID tags attached to the book and cup.

Combining the three AR approaches enabled our system to accurately recognize many kinds of activities. Although it is logical that the combination of three AR approaches would outperform an individual AR approach, their fusion is not trivial, as we developed a sophisticated probabilistic fusion model for this task.

The experience of building our system suggests that each AR approach has some limitations and that combining them results in a promising approach for complex recognition.

63.6 Challenges in Sensor-Based Activity Recognition

The significant potential of AR systems has been recognized by many researchers who have attempted different approaches, including video sensor–based, wearable sensor–based, and object usage–based approaches. In the last twenty years, many AR systems have been developed; however, AR systems still have limited functions. The main challenges include

1. *Real-world conditions*: Most existing systems are designed and tested in constrained conditions. Many factors in the real world can severely limit the applicability of a system. For example, in VSAR, these factors can include noise, occlusions, and shadows. Another challenge is the robustness of system. In real-world conditions, there might be a wide variability within the same activity class. Therefore, we need to find methods that can explain and withstand the wide variability of features that are observed within the same activity class. Overall, more research needs to be done to address these practical issues in real-world conditions.

2. *Evaluation of systems*: A common problem in the AR community is the lack of annotated reference data and standardized test-beds that could help researchers compare the performance of their approaches. Although there exist several benchmark datasets, most of these datasets consist of simple activities, such as walking, running, sitting, and sleeping. Very few common datasets exist for evaluating higher-level complex activities and reasoning algorithms. A lack of evaluation standards is a challenge, which makes it difficult to compare the performances of different systems.
3. *Hardware challenge*: With the development of the sensor industry, sensors have gotten stronger processing power and smaller sizes. However, the energy supply is still a problem, particularly in long-term AR. Moreover, making sensors easier to use and less obtrusive also needs to be addressed.
4. *Privacy and security:* In order for AR to become widely used, the users' privacy must be considered. Therefore, comprehensive privacy and security mechanisms must be developed for activity recognition systems.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 61100007, 61100081). It was also supported by the collaborative research project under the NSFC-NRF cooperative program (Grant No. 613111015).

References

1. A. F. Bobick, Movement, activity and action: The role of knowledge in the perception of motion, *Philos. Trans. R. Soc. B Biol. Sci*. 352(1358), 1257–1265, 1997.
2. C. Cedras and M. Shah, Motion-based recognition: A survey, *Image Vis. Comput*., 13(2), 129–155, 1995.
3. V. Kellokumpu, G. Zhao, and M. Pietikäinen, Human activity recognition using a dynamic texture based method, *Proceedings of British Machine Vision Conference* (*BMVC'08*), Leeds, U.K., pp. 885–894, 2008.
4. M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri, Actions as space-time shapes, *Proceedings of the International Conference on Computer Vision*, Beijing, China, pp. 1395–1402, 2005.
5. I. Laptev and T. Lindeberg, Space-time interest points, *Proceedings of the International Conference on Computer Vision*, Nice, France, pp. 432–439, 2003.
6. C. Harris and M. Stephens, A combined corner and edge detector, *Proceedings Alvey Vision Conference*, Manchester, U.K., pp. 147–151, 1988.
7. P. Dollár, V. Rabaud, G. Cottrell, and S. Belongie, Behavior recognition via sparse spatio-temporal features, *Proceedings of the International Workshop in Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, Beijing, China, pp. 65–72, 2005.
8. O. Chomat, J. Martin, and J. L. Crowley, A probabilistic sensor for the perception and recognition of activities, *Proceedings of the European Conference on Computer Vision*, Dublin, Ireland, pp. 487–503, 2000.
9. C. Schüldt, I. Laptev, and B. Caputo, Recognizing human actions: A local SVM approach, *Proceedings of the International Conference Pattern Recognition*, Cambridge, U.K., pp. 32–36, 2004.
10. J. C. Niebles, H. Wang, and L. Fei-Fei, Unsupervised learning of human action categories using spatial-temporal words, *Int. J. Comput. Vision*, 79(3), 299–318, 2008.
11. N. Ikizler and P. Duygulu, Histogram of oriented rectangles: A new pose descriptor for human action recognition, *Image Vision Comput*., 27(10), 1515–1526, 2009.
12. Z. Zhao and A. Elgammal, Human activity recognition from frame's spatiotemporal representation. *Proceedings of International Conference and Pattern Recognition*, Anchorage, AK, pp. 1–4, 2008.
13. D. Batra, T. Chen, and R. Sukthankar, Space-time shapelets for action recognition, *Proceedings of the Workshop on Motion and Video Computing*, pp. 1–6, 2008.

14. M. D. Rodriguez, J. Ahmed, and M. Shah, Action MACH: A spatio-temporal maximum average correlation height filter for action recognition. *Proceedings of the Conference Computer Vision and Pattern Recognition*, Anchorage, Alaska, pp. 1–8, 2008.
15. J. D. Lafferty, A. McCallum, and F. C. Pereira, Conditional random fields: Probabilistic models for segmenting and labeling sequence data, *Proceedings of the International Conference Machine Learning*, Jiuzhaigou, China, pp. 282–289, 2001.
16. N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman. Activity recognition from accelerometer data. *AAAI*, Menlo Park, CA, pp. 1541–1546, 2005.
17. L. Bao. Physical activity recognition from acceleration data under semi naturalistic conditions. M Eng thesis, Massachusetts Institute of Technology, Cambridge, MA, 2003.
18. J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Activity recognition using cell phone accelerometers. *Proceedings of Fourth International Workshop on Knowledge Discovery Sensor Data*, Washington, DC, pp. 10–18, 2010.
19. J. B. Chong, Activity recognition processing in a self-contained wearable system, Master's thesis, Virginia Tech, Blacksburg, VA, 2008.
20. P. Zappi, T. Stiefmeier, E. Farella, D. Roggen, L. Benini, and G. Troster. Activity recognition from on-body sensors by classifier fusion: Sensor scalability and robustness. *Proceedings of ISSNIP*, Melbourne, Victoria, Australia, pp. 281–286, 2007.
21. J. A. Ward, P. Lukowicz, G. Tröster, and T. E. Starner. Activity recognition of assembly tasks using body-worn microphones and accelerometers. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(10), 1553–1567, 2006.
22. J. Lester, T. Choudhury, N. Kern, G. Borriello, and B. Hannaford. A hybrid discriminative/generative approach for modeling human activities. *Proceedings of IJCAI*, Edinburgh, Scotland, pp. 766–772, 2005.
23. T. Hu'ynh, U. Blanke, and B. Schiele. Scalable recognition of daily activities with wearable sensors. *Third International Symposium on Location and Context-Awareness*, Berlin, Germany, pp. 50–67, 2007.
24. G. Loosli, S. Canu, and A. Rakotomamonjy. Détection des activités quotidiennes à l'aide des séparateurs à Vaste Marge. *RJCIA*, Upper Normandy, France, pp. 139–152, 2003.
25. B. Clarkson and A. Pentland. Unsupervised clustering of ambulatory audio and video. *ICASSP*, Cambridge, MA, pp. 3037–3040, 1999.
26. D. Minnen, T. Starner, I. Essa, and C. Isbell. Discovering characteristic actions from on-body sensor data. *Proceedings of ISWC*, Atlanta, GA, pp. 11–18, 2006.
27. T. Hu'ynh and B. Schiele. Unsupervised discovery of structure in activity data using multiple eigenspaces. *Second International Workshop on Location and Context-Awareness*, pp. 151–167, 2006.
28. J. R. Smith, K. P. Fishkin, B. Jiang, A. Mamishev, M. Philipose, A. D. Rea, S. Roy, and K. Sundara-Rajan. RFID-based techniques for human activity detection. *Commun. ACM*, 48(9), 39–44, 2005.
29. M. Philipose, K. P. Fishkin, M. Perkowitz, D. J. Patterson, D. Fox, H. Kautz, and D. Hahnel. Inferring activities from interactions with objects, *IEEE Pers. Comput.*, 3(4), 50–57, 2004.
30. D. H. Wilson and C. Atkeson, Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors, *Pervasive Computing*, Pittsburgh, PA, pp. 329–334, 2005.
31. E. M. Tapia, S. S. Intille, and K. Larson, Activity recognition in the home using simple and ubiquitous sensors, *Proceedings of Pervasive*, Vol. 3001, pp. 158–175, 2004.
32. T. van Kasteren, A. Noulas, G. Englebienne, and B. Krose, Accurate activity recognition in a home setting, *Proceedings of Ubiquitous Computing*, New York, pp. 1–9, 2008.
33. D. H. Hu, S. J. Pan, V. W. Zheng, N. N. Liu, and Q. Yang, Real world activity recognition with multiple goals, *Proceedings of Ubiquitous Computing*, New York, pp. 30–39, 2008.
34. D. H. Hu, S. J. Pan, V. W. Zheng, N. N. Liu, and Q. Yang, Real world activity recognition with multiple goals. *Proceedings 10th International Conference on Ubiquitous Computing*, Seoul, Korea, pp. 30–39, 2008.
35. A. M. Khattak, P. T. H. Truc, L. X. Hung, L. T. Vinh, V.-H. Dang, D. Guan, Z. Pervez, M. Han, S. Lee, and Y.-K. Lee, Towards smart homes using low level sensory data, *J. Sensors*, 1424–8220, 2011.

64

Biopotentials and Electrophysiology Measurement

Nitish V. Thakor
Johns Hopkins University

64.1 Introduction

This chapter reviews the origins, principles, and designs of instrumentation used in biopotential measurements, in particular for the electrocardiogram (ECG), the electroencephalogram (EEG), the electromyogram (EMG), and the electrooculogram (EOG). These biopotentials represent the activity of the respective organs: the heart, brain, muscle, and eyes. The biopotentials are acquired with the help of specialized electrodes that interface to the organ or the body and transduce low-noise, artifact-free signals. The basic design of a biopotential amplifier consists of an instrumentation amplifier. The amplifier should possess several characteristics, including high amplification, input impedance, and the ability to reject electrical interference, all of which are needed for the measurement of these biopotentials. Ancillary useful circuits are filters for attenuating electric interference, electrical isolation, and defibrillation shock protection. Practical considerations in biopotential measurement involve electrode placement and skin preparation, shielding from interference, and other good measurement practices.

64.2 Origins of Biopotentials

Many organs in the human body, such as the heart, brain, muscles, and eyes, manifest their function through electric activity [1]. The heart, for example, produces a signal called the ECG (Figure 64.1a). The brain produces a signal called an EEG (Figure 64.1b). The activity of muscles, such as contraction and relaxation, produces an EMG (Figure 64.1c). Eye movement results in a signal called an EOG (Figure 64.1d), and the retina within the eyes produces the electroretinogram (ERG). Measurements of these and other electric signals from the body can provide vital clues as to normal or pathological functions of the organs. For example, abnormal heart beats or arrhythmias can be readily diagnosed from an ECG. Neurologists interpret EEG signals to identify epileptic seizure events. EMG signals can be helpful in assessing muscle function as well as neuromuscular disorders. EOG signals are used in the diagnosis of disorders of eye movement and balance disorders.

The origins of these biopotentials can be traced to the electric activity at the cellular level [2]. The electric potential across a cell membrane is the result of different ionic concentrations that exist inside and outside the cell. The electrochemical concentration gradient across a semipermeable membrane results in the Nernst potential. The cell membrane separates high concentrations of potassium ion and low concentrations of sodium ions (along with other ions such as calcium in less significant proportions) inside a cell and just the opposite outside a cell. This difference in ionic concentration across the cell membrane produces the resting potential [3]. Some of the cells in the body are excitable and produce what is called an action potential, which results from a rapid flux of ions across the cell membrane in response to an electric stimulation or transient change in the electric gradient of the cell [4]. The electric excitation of cells generates currents in the surrounding volume conductor manifesting itself as potentials on the body.

Figure 64.2 illustrates the continuum of electrophysiological signals from the (a) heart cells, (b) myocardium (the heart muscle), and (c) the body surface. Each cell in the heart produces a characteristic action

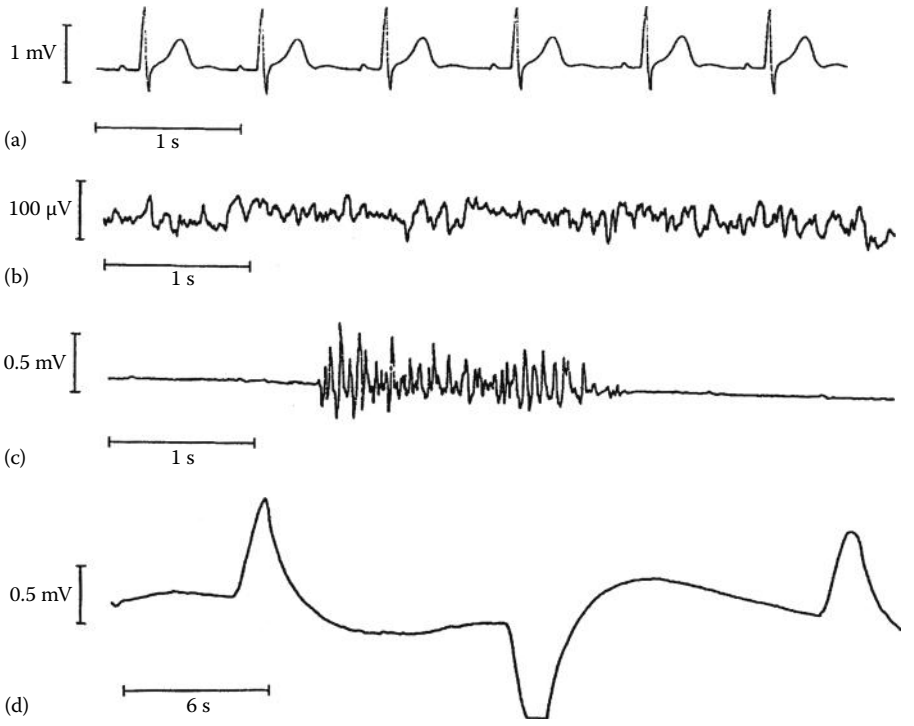


FIGURE 64.1 Sample waveforms: (a) ECG, normal sinus rhythm; (b) EEG, normal patient with open eyes; (c) EMG, flexion of biceps muscles; (d) EOG, movement of eyes from left to right.

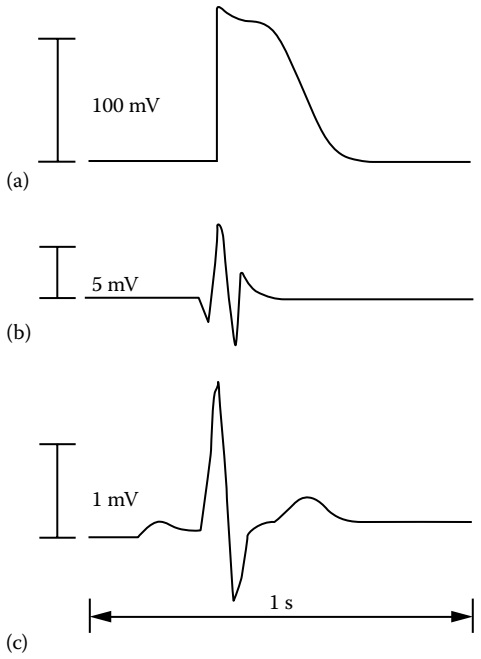


FIGURE 64.2 Schematic showing origins of biopotentials: (a) An action potential from a heart cell (recorded using a microelectrode) and (b) the electrogram from the heart surface (recorded using an endocardial catheter), and (c) the ECG signal at the chest (recorded using surface electrodes).

potential [4]. The activity of cells in the sinoatrial node of the heart produces an excitation that propagates from the atria to the ventricles through well-defined pathways and eventually throughout the heart; this electric excitation produces a synchronous contraction of the heart muscle [5]. The associated biopotential is the ECG. Electric excitation of a neuron produces an action potential that travels down its dendrites and axon [4]; activity of a massive number of neurons and their interactions within the cortical mantle results in the EEG signal [6]. Excitation of neurons transmitted via a nerve to a neuromuscular junction produces stimulation of muscle fibers. Constitutive elements of muscle fibers are the single motor units, and their electric activity is called a single motor unit potential [7]. The electric activity of large numbers of single motor unit potentials from groups of muscle fibers manifests on the body surface as the EMG. Contraction and relaxation of muscles is accompanied by proportionate EMG signals. The retina of the eye is a multilayered and rather regularly structured organ containing cells called rods and cones, cells that sense light and color. Motion of the eyeballs inside the conductive contents of the skull alters the electric potentials. Placing the electrode in the vicinity of the eyes (on either side of the eyes on the temples or above and below the eyes) picks up the potentials associated with eye movements called EOGs. Thus, it is clear that biopotentials at the cellular level play an integral role in the function of various vital organs.

64.3 Biopotentials

Biopotentials from organs are diverse. Table 64.1 lists some of these biopotentials, their representative clinical applications, and their key measurement indices and associated sensors. Note that all acquisitions are made with the aid of specialized electrodes in which actual design may be customized for specific needs. The most noteworthy features of biopotentials are [1,8]

- Small amplitudes (10 μV–10 mV)
- Low-frequency range of signals (dc to several hundred hertz)

TABLE 64.1 Biopotentials, Specifications, and Applications

Source	Amplitude (mV)	Bandwidth (Hz)	Sensor (Electrodes)	Measurement Error Source	Selected Applications
ECG	1–5	0.05–100	Ag–AgCl disposable	Motion artifact, 50/60 Hz power line interference	Diagnosis of ischemia, arrhythmia, conduction defects
EEG	0.001–0.001	0.5–40	Gold-plated or Ag–AgCl reusable	Thermal (Johnson) RF noise, 50/60 Hz	Sleep studies, seizure detection, cortical mapping
EMG	1–10	20–2000	Ag or carbon, stainless steel, needle	50/60 Hz, RF	Muscle function, neuromuscular disease, prosthesis
EOG	0.01–0.1	dc–10	Ag–AgCl	Skin potential motion	Eye position, sleep state, vestibuloocular reflex

The most noteworthy problems of such acquisitions are

- Presence of biological interference (from skin, electrodes, motion, etc.)
- Noise from environmental sources (power line, radio frequency [RF], electromagnetic, etc.)

These signal acquisition challenges and problems for each of the biopotentials are considered in greater detail in the following texts.

64.3.1 ECG

ECG signals are acquired by placing electrodes directly on the torso, arms, and legs (Figure 64.3a). The activity on the body surface is known to reflect the activity of the heart muscle underneath and in its proximity. A clinically accepted lead system has been devised and is called the 12-lead system [9,10]. It comprises a combination of electrodes taking measurements from different regions designated limb leads, the precordial leads, and the chest leads. Limb leads derive signals from electrodes on the limbs and are designated as leads I, II, and III. Precordial leads are designated aVR, aVL, and aVF and are derived by combining signals from the limb leads. The remaining six leads, V1, V2, …, V6, are chest leads. Together, ECGs from these various leads help define the nature of the activity on a specific part of the heart muscle: for example, ischemia (impaired oxygen supply to the muscle) or infarction (damage to the muscle) on the left side of the chest may be noticeable in lead III.

The ECG signals at the surface of the body are small in amplitude, which make the measurements susceptible to artifacts [11], generated by the relative motion of the electrode and the skin as well as by the activity of the nearby muscles. An important consideration in good ECG signal acquisition is the use of high-quality electrodes [12]. Electrodes made out of silver coated with silver chloride or of sintered Ag–AgCl material are recommended. An electrolytic gel is used to enhance conduction between the skin and the electrode metal. Artifacts at the electrode–skin contact as well as electromagnetic interference from all sources must be minimized [13]. Since ECG instruments are often used in critical-care environments, they must be electrically isolated for safety [14] and protected from the high voltages generated by defibrillators [15].

ECG biopotential amplifiers find use in many monitoring instruments, pacemakers, and defibrillators [16]. ECG signal acquisition is also useful in many clinical applications including diagnosis of arrhythmias, ischemia, or heart failure.

64.3.2 EEG

EEG signals are characterized by their extremely small amplitudes (in the microvolt range). Gold-plated electrodes are placed very securely on the scalp to make a very low-resistance contact. A clinically

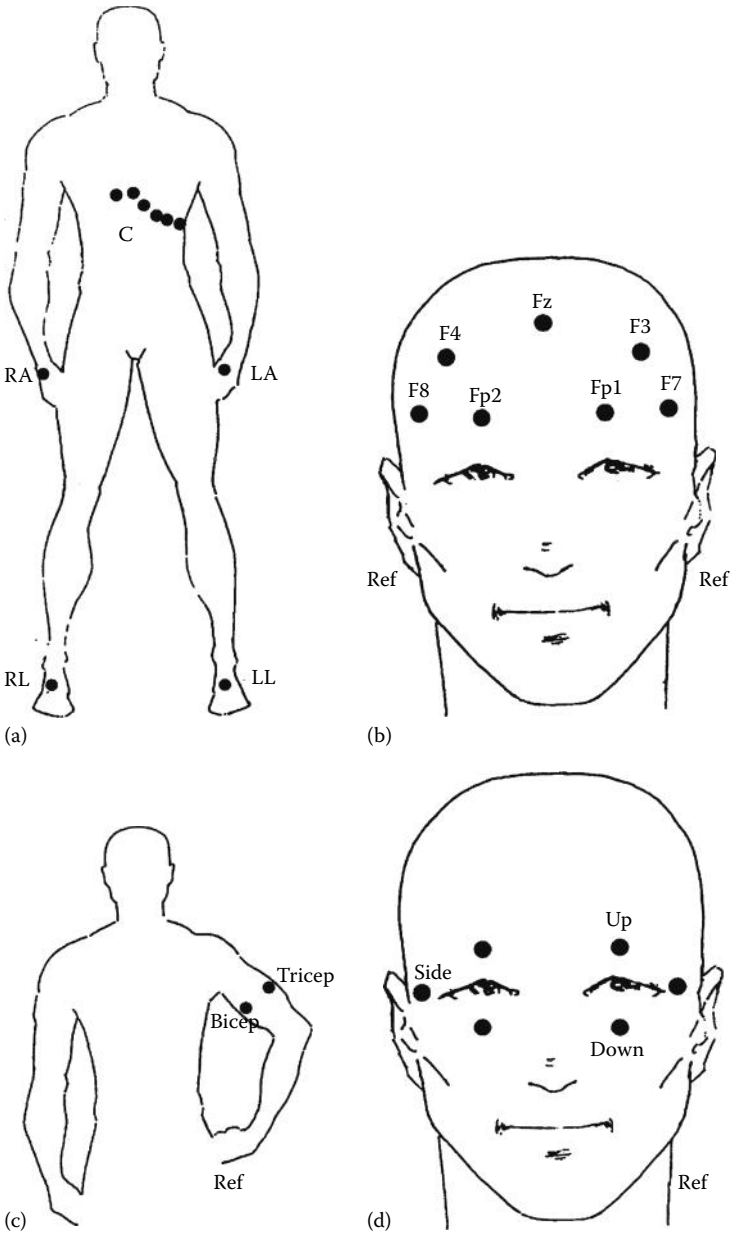


FIGURE 64.3 Schematics showing how biopotential signals are recorded from the human body: (a) ECG: 12-lead ECG is recorded using right arm (RA), left arm (LA), left leg (LL), right leg reference (RL), and six chest (C) electrodes, (b) EEG: selected electrode locations from the standard 10–20 EEG lead system with ears used as reference, (c) EMG: recording electrodes on the biceps and triceps with an independent reference, and (d) EOG: electrodes above or below (up–down) and the sides of the eyes along with an independent reference.

accepted lead system [17], which includes several electrodes placed uniformly around the head, is called the 10- to 20-lead system (Figure 64.3b). This comprehensive lead system allows localization of diagnostic features, such as seizure spikes, in the vicinity of the electrode [18].

EEG signals are difficult to interpret since they represent the comprehensive activity of billions of neurons transmitted via the brain tissues, fluids, and scalp [18]. Nevertheless, certain features can be interpreted.

In the waveform itself, it is possible to see interictal seizure spikes or a full seizure (such as petit mal and grand mal epilepsy) [18]. Analysis of the frequency spectrum of the EEG can reveal changes in the signal power at different frequencies being produced during various stages of sleep, as a result of anesthetic effects and sometimes as a result of brain injury [17].

Practical problems and challenges associated with EEG signal recordings arise from physiological, environmental, and electronic noise sources. Physiological sources of interference are motion artifact, muscle noise, eye motion or blink artifact, and sometimes even heartbeat signals. Electrical interference arises from the usual sources: 60 Hz power lines, RFs, and electrically or magnetically induced interference. Moreover, the electronic components in the amplifier also contribute noise. Good design and measuring techniques can mitigate the effects of such noise and interference.

64.3.3 EMG

Muscle fibers generate electric activity whenever muscles are active [19]. EMG signals are recorded by placing electrodes close to the muscle group (Figure 64.3c). For example, a pair of electrodes placed on the biceps and another pair placed on the triceps can capture the EMG signals generated when these muscles contract. EMG signals recorded in this manner have been shown to give a rough indication of the force generated by the muscle group [8]. Electrodes used for such applications should be small, securely attached, and should provide recordings free of artifacts. Either silver–silver chloride or gold-plated electrodes perform quite well, although inexpensive stainless steel electrodes may also suffice.

Since the frequency range of EMG signals is higher than that of ECG and EEG signals, and since the signals are of comparable or larger amplitudes, the problem of motion artifact and other interference is relatively less severe. Filtering can reduce the artifact and interference: for example, setting the bandwidth to above 20 Hz can greatly reduce the skin potentials and motion artifacts.

Recording activity directly from the muscle fibers themselves can be clinically valuable in identifying neuromuscular disorders [19]. Therefore, invasive electrodes are needed to access the muscle fibers or the neuromuscular junction. Fine-needle electrodes or thin stainless steel wires are inserted or implanted to obtain local recording from the fibers or neuromuscular junctions [7].

64.3.4 EOG

Electric potentials are generated as a result of movement of the eyeballs within the conductive environment of the skull. The generation of EOG signals can be understood by envisaging dipoles (indicating separated positive and negative potential sources) located in the eyeballs. Electrodes placed on either side of the eyes or above and below them pick up the potentials generated by the motion of the eyeball (Figure 64.3d). This potential varies approximately in proportion to the movement of the eyeballs, and hence EOG is sometimes used to study eye positions or disorders of eye movement and balance (a reflex called vestibuloocular reflex affects the nystagmus of the eye). Similarly, saccades inherent in eye motion as well as blinking of the eyelids can produce changes in the EOG signal.

This signal is small (10–100 μV) and has low frequencies (dc to 10 Hz) [8]. Hence, an amplifier with a high gain and good low-frequency response and dc stability is desirable. Additionally, the electrode–gel combination should be such that it produces low levels of junction potential, motion artifacts, and drift in the dc signal [20]. Practical problems associated with dc drift, motion artifacts, and securing electrodes in the vicinity of the eyes make their long-term use problematic. Nevertheless, EOG signals can be useful clinically in acute studies of human disorders and therefore careful acquisition of the signal followed by appropriate analysis is used to interpret the EOG potentials.

Other biopotential recording techniques follow similar principles of measurements. The electrode design should be specifically adapted to the source of the signal. A thorough effort is required to minimize the noise and interference by improving electrode design and placement and optimizing

the amplifier circuit. Good electrode attachment along with selective filtering at the amplifier can help obtain relatively noise-free recording. The design principles and practical considerations are described later.

64.4 Principles of Biopotential Measurements

The unifying principles of biopotential recordings involve

- Electrode design and its attachment suited to the application
- Amplifier circuit design for suitable amplification of the signal and rejection of noise and interference
- Good measurement practices to mitigate artifacts, noise, and interference

64.5 Electrodes for Biopotential Recordings

Electrodes for biopotential recordings are designed to obtain the signal of interest selectively while reducing the potential to pick up artifact. The design should be pragmatic to reduce cost and allow for good manufacturing and reliable long-term use. These practical considerations determine whether high-quality but reusable electrodes made of silver or gold or cheaper disposable electrodes are used [20].

64.5.1 Silver–Silver Chloride Electrodes

The classic, high-quality electrode design consists of highly conductive metal, silver, interfaced to its salt, silver chloride, and connected via an electrolytic gel to the human body [21]. Silver–silver chloride-based electrode design is known to produce the lowest and most stable junction potentials [1,20]. Junction potentials are the result of the dissimilar electrolytic interfaces and are a serious source of electrode-based motion artifacts. Therefore, additionally, an electrolytic gel typically based on sodium or potassium chloride is applied to the electrode. A gel concentration in the order of 0.1 M (molar concentration) results in a good conductivity and low junction potential without causing skin irritation.

Reusable silver–silver chloride electrodes (Figure 64.4a) are made of silver disks coated electrolytically by silver chloride [1], or, alternatively, particles of silver and silver chloride are sintered together

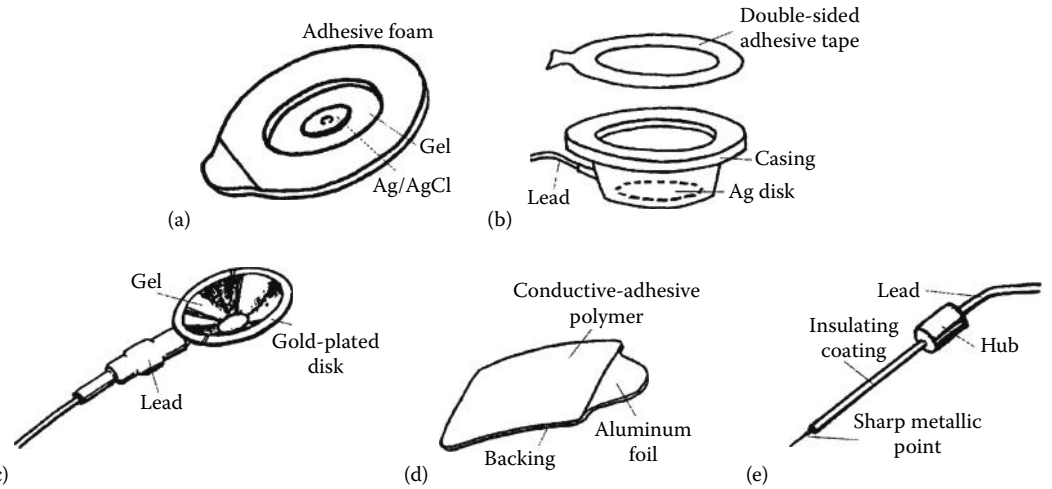


FIGURE 64.4 Examples of electrodes used in biopotential recordings: (a) Disposable Ag–AgCl electrode, (b) reusable Ag–AgCl disk electrode, (c) gold disk electrode, (d) disposable conductive polymer electrode, and (e) needle electrode.

to form the metallic structure of the electrode. The gel is typically soaked into a foam pad or is applied directly in a pocket produced by the electrode housing. The electrode is secured to the skin by means of nonallergenic adhesive tape. The electrode is connected to the external instrumentation typically via a snap-on connector. Such electrodes are well suited for acute studies or basic research investigations.

Disposable electrodes are made similarly, although the use of silver may be minimized (e.g., the snap-on button itself may be silver coated and chlorided). To allow for a secure attachment, a large foam pad attaches the electrode body with adhesive coating on one side (Figure 64.4b). Such electrodes are particularly suited for ambulatory or long-term use.

64.5.2 Gold Electrodes

Gold-plated electrodes (Figure 64.4c), which have the advantages of high conductivity and inertness desirable in reusable electrodes, are commonly used in EEG recordings [1]. Small reusable electrodes are designed so that they can be securely attached to the scalp. The electrode body is also shaped to make a recessed space for electrolytic gel, which can be applied through a hole in the electrode body [18]. The electrodes are attached in hair-free areas by use of a strong adhesive such as collodion or securely attached with elastic bandages or wire mesh. Similar electrodes may also be used for recording EMG, especially when a great deal of motion is expected. Disadvantages of using gold electrodes over silver–silver chloride electrodes include greater expense, higher junction potentials, and greater susceptibility to motion artifacts [20]. On the other hand, gold electrodes maintain low impedance, are inert and reusable, and are good for short-term recordings as long as a highly conductive gel is applied and they are attached securely.

64.5.3 Conductive Polymer Electrodes

It is often convenient to construct an electrode out of a material that is simultaneously conductive and adhesive [20]. Certain polymeric materials have adhesive properties and by attaching monovalent metal ions can be made conductive. The polymer is attached to a metallic backing made of silver or aluminum foil, which allows electric contact to external instrumentation (Figure 64.4d). This electrode does not need additional adhesive or electrolytic gel and hence can be immediately and conveniently used. The conductive polymeric electrode performs adequately as long as its relatively higher resistivity (over metallic electrodes) and greater likelihood of generating artifacts are acceptable. The higher resistivity of the polymer makes these electrodes unsuitable for low-noise measurement. The polymer does not attach as effectively to the skin as does the conventional adhesive on disposable ECG electrodes built with a foam base, and furthermore, the potentials generated at the electrode–skin interface are more readily disturbed by motion. Nevertheless, when the signal level is high and when restricting the subject movement minimizes artifact, the polymeric electrode offers a relatively inexpensive solution to biopotential recording.

64.5.4 Metal or Carbon Electrodes

Although other metals such as stainless steel or brass electrodes [21] are used rather infrequently now because high-quality noble metal electrodes or low-cost carbon or polymeric electrodes are so readily available, historically these metallic electrodes were used in laboratory or clinical settings because of their sturdy construction and reusability. Electrode gel is applied to the metal electrode that is fastened to the body by means of a rubber band. These electrodes have the potential for producing very high levels of artifact and are bulky and awkward to use but do offer the advantage of being reusable and tend to be inexpensive. Carbon or carbon-impregnated polymer electrodes are also used occasionally (although they are mainly used as electrical stimulation electrodes) [20]. These electrodes have a much higher resistivity and are noisier and more susceptible to artifacts, but they are inexpensive,

flexible, and reusable and are thus chosen for applications such as electric stimulation or impedance plethysmography. For these applications, gel is usually not applied and the electrodes are used in "dry" form for easy attachment and removal.

64.5.5 Needle Electrodes

Needle electrodes (Figure 64.4e) comprise a small class of invasive electrodes, used when it is absolutely essential to record from the organ itself. The most common application is in recording from muscles or muscle fibers [8]. A metallic, typically steel, wire is delivered via a needle inserted at the site of the muscle fiber. The wire is hooked and hence fastens to the muscle fiber, even as the needle is removed. Small signals such as motor unit potentials can be recorded in this manner [7]. For research applications, similar needle or wire electrodes are sometimes connected directly to the heart muscle. Since such electrodes are invasive, their use is limited to only highly specialized and supervised clinical or research applications.

64.6 Biopotential Amplifier

Biopotentials exhibit small amplitudes and low frequencies [22]. Moreover, biopotential measurements are corrupted by environmental and biological sources of interference. Therefore, the essential, although not exhaustive, design considerations include proper amplification and bandwidth, high input impedance, low noise, and stability against temperature and voltage fluctuations. The key design component of all biopotential amplifiers is the instrumentation amplifier [21]. However, each biopotential acquisition instrument has a somewhat differing set of characteristics, necessitating some specialization in the design of the instrumentation amplifier. Table 64.2 summarizes the circuit specialization needed in various biopotential amplifiers, with the ECG amplifier used as the basic design.

64.6.1 Instrumentation Amplifier

The instrumentation amplifier is a circuit configuration that potentially combines the best features desirable for biopotential measurements [8], namely, high differential gain, low common-mode gain, high common-mode rejection ratio (CMRR), and high input resistance [23]. Figure 64.5 shows the design of the basic instrumentation amplifier. The basic circuit design principles have been described elsewhere [23,24]. The instrumentation amplifier is constructed from operational amplifiers, or op amps, which have many of the desirable features listed earlier [25]. The front end of the amplifier has two op amps, which consists of two noninverting amplifiers that have been coupled together by a common resistor R_1. The gain of the first stage is $(1 + 2R_2/R_1)$. The second stage is a conventional differential amplifier with gain of $-(R_4/R_3)$. This design results in the desired differential gain distributed over two stages of the amplifier.

TABLE 64.2 Distinguishing Features and Design Consideration for Biopotentials

Biopotential	Distinguishing Feature	Exclusive Amplifier Design Consideration	Additional Features Desired
ECG[a]	1 mV signal, 0.05–100 Hz BW[b]	Moderate gain, BW, noise, CMRR, input *R*	Electrical safety, isolation, defibrillation protection
EEG	Very small signal (microvolts)	High gain, very low noise, filtering	Safety, isolation, low electrode–skin resistance
EMG	Higher BW	Gain and BW of op amps	Post-acquisition data processing
EOG	Lower frequencies, small signal	dc and low drift	Electrode–skin junction potential, artifact reduction

[a] The ECG signal acquisition is considered as the standard against which the other acquisitions are compared.
[b] BW, bandwidth.

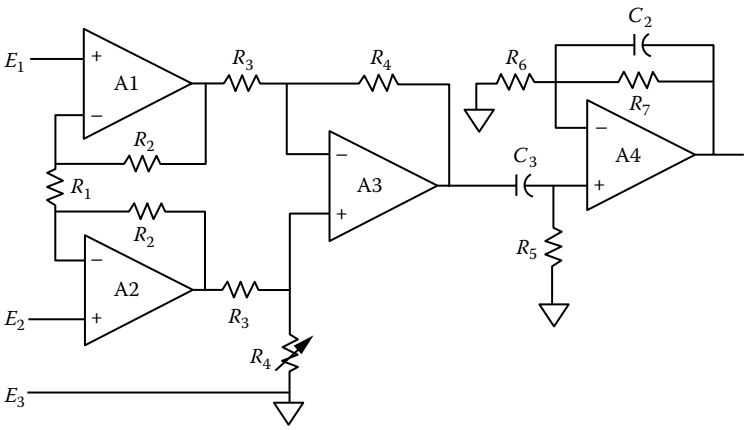


FIGURE 64.5 The instrumentation amplifier. This amplifier has a very high input impedance, high CMRR, and a differential gain set by the resistors in the two amplifier stages. The gain of the first stage (amplifiers Al and A2) is $1 + 2R_2/R_1$, the second stage (amplifier A3) is R_4/R_3, and the third stage (amplifier A4) is $1 + R_7/R_6$. The lower corner frequency is $1/(2\pi R_5 C_1)$ and the upper corner frequency is $1/(2\pi R_7 C_2)$. The variable resistor R is adjusted to maximize the CMRR. Electrodes E_1 and E_2 are the recording electrodes, while E_3 is the reference or the ground electrode.

It also achieves a very high input resistance as a result of the noninverting amplifier front end. It exhibits a very high CMRR as a result of the differential first stage followed by a second-stage differential amplifier. The CMRR is enhanced by adjusting one of the matching resistors and by selecting high CMRR op amps. This instrumentation amplifier is a key design component universal to many biosensor interfaces and almost all biopotential instruments [22]. The LT1167 can replace the front end, while the AD8232 provides a complete ECG amplifier.

64.6.2 ECG Amplifier

The ECG amplifier can readily be designed using the instrumentation amplifier as the principal building block. Active filters with a lower corner frequency of 0.05 Hz and an upper corner frequency of 100 Hz are also typically added [8].

ECG amplifiers are needed in many applications, such as monitoring in cardiac intensive-care units, where safety and protection are of paramount importance. Because the possibility of a direct or low-resistance access to the heart via catheters or intravenous lines exists in such settings, very small electric leakage currents can be fatal. Consequently, leakage from the amplifier is required to be below the safety standard limit of 10 μA [14]. Additionally, safety of the patient is achieved by providing electrical isolation from the power line and the earth ground, which prevents passage of leakage current from the instrument to the patient under normal conditions or under reasonable failure conditions. Electrical isolation is achieved by using transformer or optical coupling components [9], although it is important to remember that any such design should preserve the bandwidth and linearity of the amplifier. ECG amplifiers are also likely to be operated in circumstances where defibrillators might be used; thus, the amplifier circuit must be protected against the high defibrillation voltages and must be augmented by circuit components such as current-limiting resistors, voltage-limiting diodes, and spark gaps [15].

64.6.3 EEG Amplifiers

The distinguishing feature of an EEG amplifier is that it must amplify very small signals [8]. The amplifier gain must be suitably enhanced to deal with microvolt or lower levels of signals. Furthermore, all

components of the amplifier must have a very low thermal noise and in particular low electronic (voltage and current) noise at the front end of the amplifier. EEG amplifiers used in clinical applications again must be electrically isolated and protected against high defibrillation voltages, similar to the ECG amplifier.

64.6.4 EMG Amplifiers

EMG amplifiers are often used in the investigation of muscle performance, neuromuscular diseases, and in building certain powered or smart prostheses. In such applications, slightly enhanced amplifier bandwidth suffices. In addition, postprocessing circuits are almost always needed. For example, a rectified and integrated EMG signal has been shown to give a rough indication of the muscle activity, approximately related to the force being generated at the location of the EMG electrode [8].

64.6.5 EOG Amplifiers

The EOG signal is small in amplitude and consists of very low frequencies. Therefore, an EOG amplifier must not only have a high gain but also a very good low frequency, or even dc, response. This frequency response also makes the amplifier potentially susceptible to shifts in the junction potential at the skin–electrode interface and to drift in the electronic circuit characteristics. In addition to using good electrodes (Ag–AgCl) and gel (high conductivity), some type of active dc or drift cancellation or correction circuit design may be necessary.

64.7 Circuit Enhancements

The basic biopotential amplifier described earlier, along with the specific design considerations for each biopotential, can yield a signal acquisition of acceptable quality in most laboratory settings. In practice, however, further enhancements are always necessary to achieve acceptable clinical performance in novel applications. These enhancements include circuits for reducing electric interference, filtering noise, reduction of artifacts, electrical isolation of the amplifier, and electrical protection of the circuit against defibrillation shocks [9].

64.7.1 Electrical Interference Reduction

Environmental electric interference is always present, especially in urban hospital environments. It is desirable to eliminate interference before it enters the amplifier, for example, by proper shielding of the subject, leads, and the instrument and by grounding the subject and the instrument. Sources of interference include induced signals from power lines and electric wiring; RF from transmitters, electric motors, and other appliances; and magnetically induced currents in lead wires [13]. Interference induced on the body common to the biopotential sensing electrodes is called the common-mode interference (as distinguished from the biopotential that is differential to the sensing electrodes). If the induced current is i_d and the resistance to ground is R_0, then the common-mode interference potential is $V_c = i_d R_0$. The common-mode interference is principally rejected by a differential or instrumentation amplifier with a high CMRR. Further improvement is possible by use of the "driven right leg circuit." The right leg lead, by standard convention, is used as the ground or the circuit reference. The driven right leg circuit employs the clever idea of negative feedback of the common-mode signal into this lead. The common-mode signal is sensed from the first stage of the instrumentation amplifier, amplified and inverted, and fed back into the right leg lead (Figure 64.6a). At this stage the common-mode signal is reduced to $(i_d R_0)/(1 + 2R_2/R_1)$. Thus, the common-mode interference is greatly reduced at its source. The driven right leg circuit along with a high CMRR of the amplifier and filtering permit very high-quality biopotential measurements.

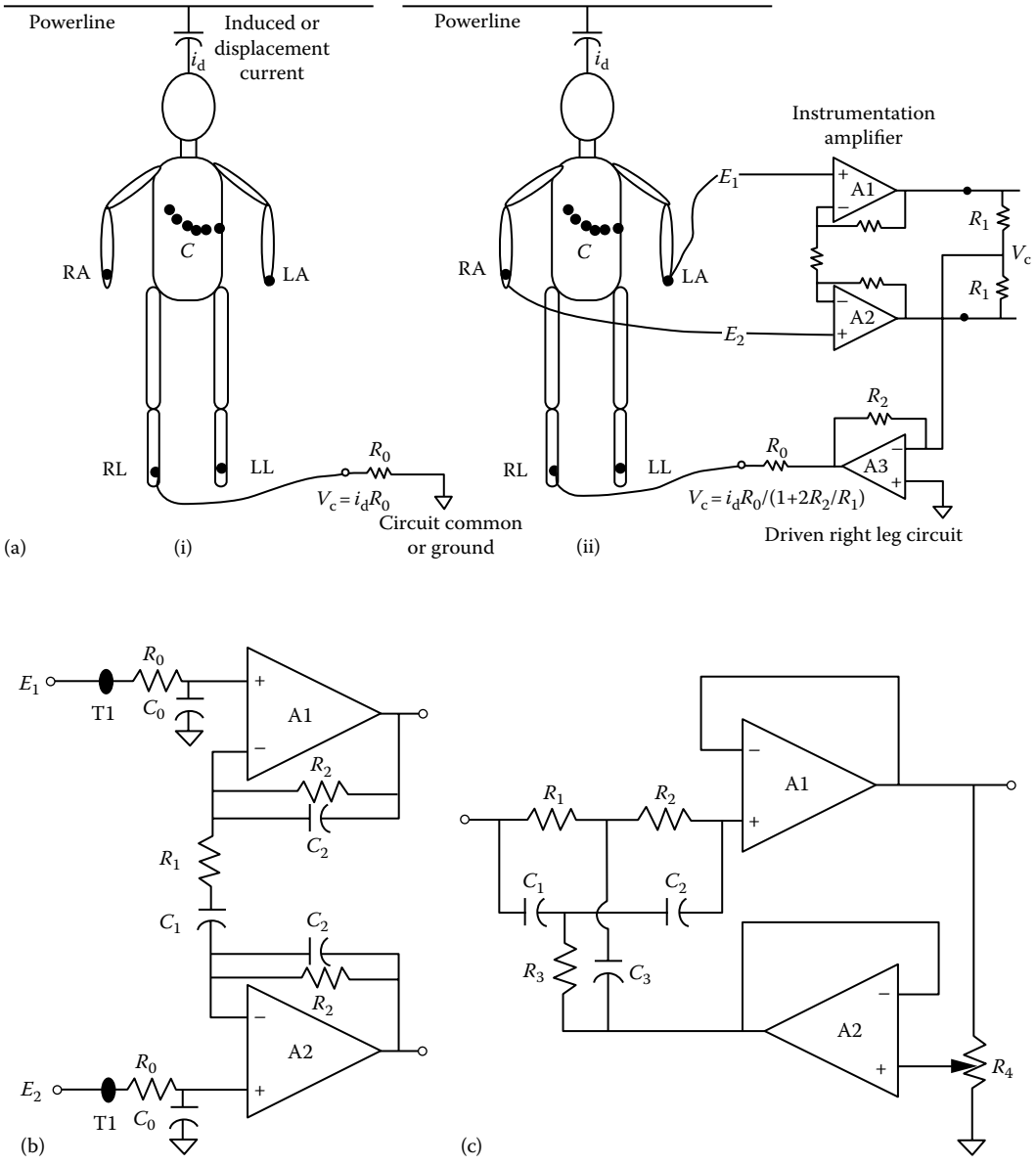


FIGURE 64.6 Circuit enhancements for biopotential measurements. (a) The schematic on the left shows electric interference induced by the displacement current i_d from the power line. This current flows into the ground electrode lead generating common-mode voltage V_c. The driven right leg circuit on the right uses negative feedback into the right leg electrode to reduce the effective common-mode voltage. (b) Amplifier front end filters—T1: RF choke; R_0 and C_0: RF filter; R_1 and C_1: high-pass filter; R_2 and C_2: low-pass filter. (c) Notch filter for power line interference (50 or 60 Hz): twin T notch filter in which notch frequency is governed by R_1, R_2, R_3, C_1, C_2, and C_3 and notch tuning by R_4.

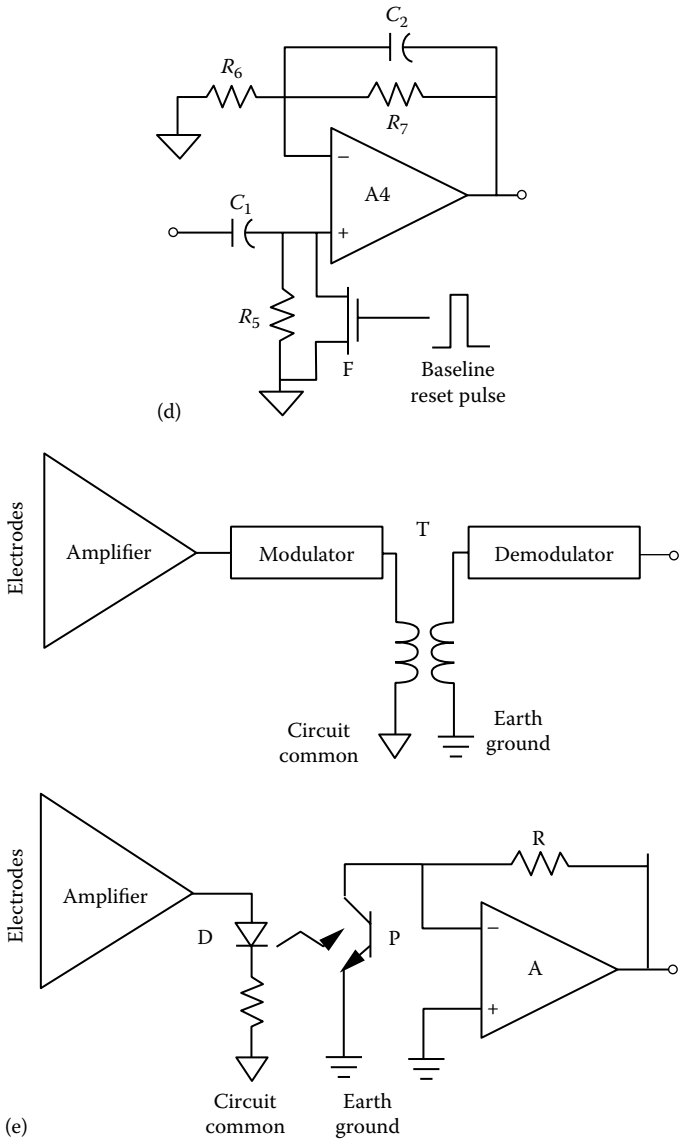


FIGURE 64.6 (continued) Circuit enhancements for biopotential measurements. (d) Baseline restoration circuit: the high-pass filter capacitor C1 is discharged by field effect transistor F when activated manually or automatically by a baseline restoration pulse. (e) Electrical isolation: transformer coupled using the transformer T (top) or optical using the diode D and the photodetector P (bottom). Note that the isolator separates circuit common on the amplifier side from the earth ground on the output side.

(*continued*)

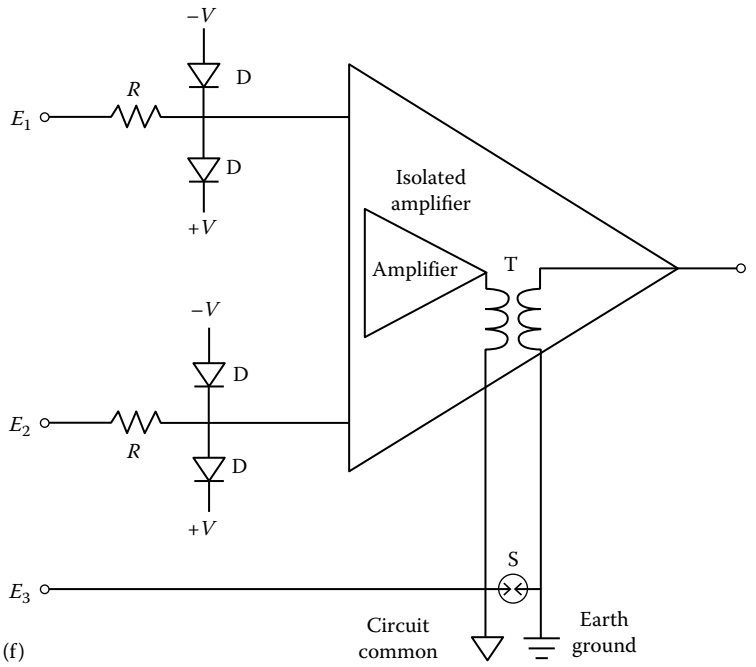


FIGURE 64.6 (continued) Circuit enhancements for biopotential measurements. (f) Electrical protection circuit: resistance R limits the current, reverse-biased diodes D limit the input voltage, and the spark gap S protects against defibrillation pulse-related breakdown of the isolation transformer T.

64.7.2 Filtering

After following the precautions described earlier, filtering at the front end of the amplifier and limiting the bandwidth of the biopotential amplifier can further help to reduce the interference (Figure 64.6b). Small inductors or ferrite beads in the lead wires help to block very high-frequency electromagnetic interference. Small capacitors between each electrode lead and ground filter the RF interference. Bandwidth limitation can be imposed at each stage of the amplifier. Because dc potentials arising at the electrode–skin interface must be blocked well before the biopotential is amplified greatly (otherwise, the amplifier could saturate), use of high-pass filtering in the early stages of amplification is recommended. Low-pass filtering at several stages of amplification is recommended to attenuate residual RF interference as well as muscle signal interference. Power line interference at 50 or 60 Hz and their harmonics clearly pose the biggest problem in biopotential measurement [11,13]. Sometimes it may be desirable to provide a 50 or 60 Hz notch filter to remove the power line interference (Figure 64.6c), an option that is often available with low-level signal (EEG, EOG) measuring instruments. The risk of a distorted biopotential signal arises when a notch filter is used and this may affect diagnosis. Filtering should, therefore, be used selectively.

64.7.3 Artifact Reduction

One principal source of artifact is the potential arising at the electrode–skin interface [11]. Slow changes in the baseline can arise due to changes in the junction potential at this interface and, in some instances, can cause a temporary saturation of the amplifier [9]. This event is detected manually or automatically (by quickly discharging the high-pass capacitor in the amplifier to restore the baseline; Figure 64.6d). Movement of the subject or disturbance of the electrode can produce motion artifacts [11], which can be reduced by filtering the signal, but as suggested earlier, such filtering, typically high pass, can severely distort the biopotential being measured. Alternatively, computerized processing may be necessary to

identify an artifact and delete it from display and processing. Of note, a biopotential source could be the desired one in one case but an unwanted artifact in another case. For example, EOG signal resulting from blinking of eyes can produce a rather significant artifact in EEG recordings. Similarly, EMG signals become unwanted artifacts in all other non-EMG biopotential measurements. ECG monitoring must especially account for EMG artifact for high-fidelity recording. Another example is the pacemaker pulse. Since a pacemaker pulse can be detected and amplified as a short (about 2 ms) pulse preceding a QRS complex, it can be mistakenly interpreted as a heartbeat by some circuits for automatically determining heart rate. Special circuits must be designed to identify and delete this artifact [9].

64.7.4 Electrical Isolation

Electrical isolation limits the possibility of the passage of any leakage current from the instrument in use to the patient [22]. Conversely, patient safety must be ensured by electrical isolation to reduce the prospect of leakage of current from any other sensor or instrument attached to the patient to the earth ground of the instrument being tested [8]. Passage of leakage current through the patient could be harmful or even fatal if this current were to leak to the heart via a catheter or intravenous line. Electrical isolation can be done electrically by inserting a transformer in the signal path or optically by introducing an optical coupler (Figure 64.6e). Since the primary and the secondary of the transformer remain electrically isolated, no direct path to ground can exist. One problem with this approach is that the transformer is inherently an ac high-frequency device. Therefore, a suitable solution is to modulate the biopotential signal using a high-frequency carrier preferred by the transformer. An alternative solution is to use optical isolation. The electric signal from the amplifier is first converted to light by a light-emitting diode (LED). This optical signal is modulated in proportion to the electric signal and transmitted to the detector. A photodetector (photodiode or a phototransistor) then picks up the light and converts it into an electric signal, which is then demodulated to recover the original signal. The optical signal is typically pulse code modulated to circumvent the inherent nonlinearity of the LED–phototransistor combination.

64.7.5 Defibrillation Protection

Biopotential measuring instruments can encounter very high voltages, such as those from electric defibrillators that can damage the instrument [9]. For example, electric shocks in the order of 1500–5000 V may be produced by an external defibrillator [1]. Other high-voltage sources are electrocautery (used in surgery) and power lines (inadvertent short circuits in the instrument). Therefore, the front end of the biopotential instrument must be designed to withstand these high voltages (Figure 64.6f). Use of resistors in the input leads can limit the current in the lead and the instrument. Protection against high voltages is achieved by the use of diodes or Zener diodes. These components conduct at 0.7 V (diode conduction voltage) or 10–15 V (depending on the Zener diode breakdown voltage), thus protecting the sensitive amplifier components. Since it is more likely that protection against higher voltages will be needed, low-pressure gas discharge tubes such as neon lamps are also used. They break down at voltages on the order of 100 V, providing an alternative path to ground for the high voltages. As a final line of protection, the isolation components (optical isolator or transformer) must be protected by a spark gap that activates at several thousand volts. The spark gap ensures that the defibrillation pulse does not breach the isolation.

64.8 Measurement Practices

Biopotential measurements are made feasible, first of all, by good amplifier designs. High-quality biopotential measurements require use of good electrodes and their proper application on the patient, along with good laboratory or clinical practices. These practices are summarized later.

64.8.1 Electrode Use

Various electrodes best suited for each biopotential measurement were described earlier. First, different electrodes by virtue of their design offer distinguishing features: more secure (use of strong but less-irritant adhesives), more conductive (use of noble metals such as silver and gold), and less prone to artifact (use of low-junction-potential materials such as Ag–AgCl). Electrode gel can be of considerable importance in maintaining a high-quality interface between the electrode metal and the skin. High-conductivity gels, in general, help reduce the junction potentials along with the resistance (they tend, however, to be allergenic or irritating and hence a practical compromise in terms of electrolyte concentration must be found) [20]. Movement of the electrode with respect to the electrode gel and the skin is a potential source of artifact (Figure 64.7a). Such movements can change the electrode junction to skin potentials, producing motion artifacts [21]. Placement above bony structures where there is less muscle mass can reduce unwanted motion artifact and EMG interference (Figure 64.7b). Electrodes must be securely attached, for example, with stress loops secured away from the electrode site, so that motion artifact can be reduced. In certain instances, the electrodes may be essentially glued to the skin, as in the case of EEG measurements.

64.8.2 Skin Preparation

The potentials existing at the skin surface, attributable to potentials at the membranes of cells in the epidermal layers of the skin, can result in a large dc potential (which can be a significant problem in EOG measurements). Any disturbance of the skin by motion, touching, or deformation can cause this potential to change and result in motion artifacts (Figure 64.7a). Sweat glands in the epidermis can also contribute varying extents of skin resistance and skin potential. Such potentials and artifacts can be reduced by abrading the epidermal skin. A mild abrasion by sandpaper or its equivalent can significantly reduce skin resistance and skin potential and thereby reduce artifact [26]. A less traumatic but somewhat less effective approach is to use an alcohol swab or similar skin-cleansing solution to wet and clean the skin surface to remove debris, oils, and damaged or dead epidermal cells. Sometimes, as with EEG measurements where very low signals are recorded and very low noise is permitted, skin resistance must be significantly lowered, perhaps to below 2 kΩ [18]. Obviously, reduced motion or muscle activity while measurement is carried out also helps.

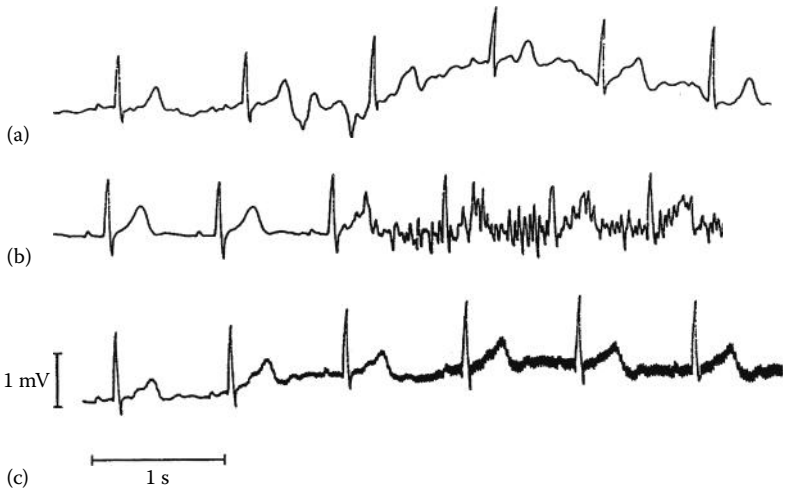


FIGURE 64.7 Examples of electric interference in biopotential recordings: (a) ECG signal with baseline changes and motion artifacts, (b) muscle signal interference, and (c) electromagnetic interference (60 Hz power line and RF).

64.8.3 Reduction of Environmental Interference

Electromagnetic interference radiated from the power lines, RF interference from machines, induced magnetic field in the leads, and electric currents induced on to the body are all potential sources of environmental interference (Figure 64.7c). Shielding of the amplifier along with the electrode and the lead, and in certain extreme conditions shielding of the subject (e.g., when taking magnetic field measurements from the body), can greatly help reduce the signals picked up by or induced into the amplifier. The electrode leads can be shielded or at the very least twisted together to reduce induced electromagnetic interference.

The amplifier circuit should also have extensive filtering of unwanted electromagnetic interference. To eliminate RF interference, filter capacitors should be used in the front end of the amplifier as well as at various stages of the amplifier. Very high frequencies can be blocked by the use of a choke or an inductor at the input leads. The effect of electrostatic interference can be minimized or eliminated by grounding the instrument.

Electric interference in the environment induces current into the body, which is then picked up by the biopotential amplifier as a common-mode voltage [27]. The CMRR property of the amplifier is essential for reduction of the common-mode voltage [24]. Finally, the driven right leg design [27], described earlier, can be optionally used to reduce further the common-mode voltage and the effective interference.

64.9 Conclusions

Biopotential acquisition is a well-developed science, and acceptable engineering design solutions do exist. It is apparent that each biopotential source presents its own distinct challenge in terms of electrode interface, amplifier design, pre- or postprocessing, and practical implementation and usage. ECG signals can be best acquired using Ag–AgCl electrodes, although good experimental/clinical practice is needed to reduce biological and environmental interference. Further circuit protection and isolation are necessary in clinical usage. EEG signals are distinguishable by their very low amplitude, and hence EEG electrodes must be securely attached via a very small electrode–skin resistance and the amplifier must exhibit exceptionally low noise. For EMG acquisition, electrodes are needed that can be attached for long periods of time to the muscle groups under study. The EMG signal inevitably needs postprocessing, such as integration, to derive a measure of muscle activity. EOG signals have small amplitudes and are characterized by dc or low frequencies. Skin–electrode potentials and dc drift of the amplifier are, therefore, important considerations.

These biopotential measurement principles are applicable to a variety of conventional as well as emerging applications. For example, although ECG acquisition is used mainly in cardiac monitors, it is also of interest and importance in implantable pacemakers and defibrillators. EEG acquisition is useful in the detection of seizure spikes and study of sleep patterns and it may also be used to identify cortical dysfunction after trauma or stroke. EMG acquisition is used in diagnosing neuromuscular diseases. Interesting attempts have been made to use EMG for controlling prostheses. EOG has been helpful in diagnosing vestibuloocular disorders and also has been studied as a way of operating communication devices (pointing) used by quadriplegics. The measurement and instrumentation principles described in this chapter would be applicable, with some modifications, to these emergent applications.

References

1. L. A. Geddes and L. E. Baker, *Principles of Applied Biomedical Instrumentation*, 3rd edn., New York: Wiley, 1989.
2. R. Plonsey, *Bioelectric Phenomena*, New York: McGraw-Hill, 1969.
3. R. Plonsey and R. C. Barr, *Bioelectricity*, New York: Plenum, 1988.
4. R. C. Barr, Basic electrophysiology, in *The Biomedical Engineering Handbook*, Bronzino J. (Ed.), Boca Raton, FL: CRC Press, pp. 101–118, 1995.
5. D. Durrer et al., Total excitation of the isolated human heart, *Circulation*, 41, 899–912, 1970.

6. P. L. Nunez, *Electric Fields of the Brain*, New York: Oxford University Press, p. 484, 1981.
7. K.-A. Henneberg, Principles of electromyography, in *The Biomedical Engineering Handbook*, Bronzino J. D. (Ed.), Boca Raton, FL: CRC Press, pp. 191–200, 1995.
8. J. G. Webster, (Ed.), *Medical Instrumentation: Application and Design*, 4th edn., Hoboken, NJ: Wiley, 2010.
9. N. V. Thakor, Electrocardiographic monitors, in *Encyclopedia of Medical Devices and Instrumentation*, Webster J. G. (Ed.), New York: Wiley, pp. 1002–1017, 1988.
10. H. V. Pipberger et al., Recommendations for standardization of leads and specifications for instruments in electrocardiography and vector cardiography, *Circulation*, 52, 11–31, 1975.
11. J. G. Webster, Reducing motion artifacts and interference in biopotential recording, *IEEE Trans. Biomed. Eng.*, 31, 823–826, 1984.
12. N. V. Thakor and J. G. Webster, Electrode studies for the long-term ambulatory ECG, *Med. Biol. Eng. Comput.*, 23, 116–121, 1985.
13. J. C. Huhta and J. G. Webster, 60-Hz interference in electrocardiography, *IEEE Trans. Biomed. Eng.*, 20, 91–101, 1973.
14. Anonymous, American national standard safe current limits for electromedical apparatus, *ANSI/AAMI*, Washington, DC. Vol. SCL 12/78, 1978.
15. Anonymous, American national standard for diagnostic electrocardiographic devices, *ANSI/AAMI*, Vol. EC11–1982, 1984.
16. N. V. Thakor, From Holter monitors to automatic defibrillators: developments in ambulatory arrhythmia monitoring, *IEEE Trans. Biomed. Eng.*, 31, 770–778, 1984.
17. A. S. Gevins and M. J. Aminoff, Electroencephalography: Brain electrical activity, in *Encyclopedia of Medical Devices and Instrumentation*, Webster, J. G. (Ed.), New York: Wiley, pp. 1084–1107, 1988.
18. E. Niedermeyer and F. Lopes da Silva, *Electroencephalography*, Baltimore, MD: Urban, Schwarzenberg, 1987.
19. C. J. De Luca, Electromyography, in *Encyclopedia of Medical Devices and Instrumentation*, Webster J. G. (Ed.), New York: Wiley, pp. 1111–1120, 1988.
20. H. Carim, Bioelectrodes, in *Encyclopedia of Medical Devices and Instrumentation*, Webster J. G. (Ed.), New York: Wiley, pp. 195–226, 1988.
21. M. R. Neuman, Biopotential electrodes, in *Medical Instrumentation: Application and Design*, Webster J. G. (Ed.), 4th edn., Hoboken, NJ: Wiley, 2010.
22. J. H. Nagle, Biopotential amplifiers, in *The Biomedical Engineering Handbook*, Bronzino J. D. (Ed.), Boca Raton, FL: CRC Press, pp. 1185–1195, 1995.
23. S. Franco, *Design with Operational Amplifiers*, New York: McGraw-Hill, 1988.
24. W. J. Jung, *IC Op Amp Cookbook*, 3rd edn., Indianapolis, IN: Howard W. Sams, 1986.
25. P. Horowitz and W. Hill, *The Art of Electronics*, 2nd edn., Cambridge, U.K.: Cambridge University Press, 1989.
26. H. W. Tam and J. G. Webster, Minimizing electrode motion artifact by skin abrasion, *IEEE Trans. Biomed. Eng.*, 24, 134–139, 1977.
27. M. R. Neuman, Biopotential amplifiers, in *Medical Instrumentation: Application and Design*, Webster J. G. (Ed.), 4th edn., Hoboken, NJ: Wiley, 2010.

65

Blood Pressure Measurement

Shyam Rithalia
University of Salford

Mark Sun
NeoPath. Inc.

Roger Jones
Primary Children's Medical Center

65.1 Introduction

Blood pressure measurements have been part of the basic clinical examination since the earliest days of modern medicine. The origin of blood pressure is the pumping action of the heart, and its value depends on the relationship between cardiac output and peripheral resistance. Therefore, blood pressure is considered as one of the most important physiological variables with which to assess cardiovascular hemodynamics. Venous blood pressure is determined by vascular tone, blood volume, cardiac output, and the force of contraction of the chambers of the right side of the heart. Since venous blood pressure must be obtained invasively, the term *blood pressure* most commonly refers to arterial blood pressure, which is the pressure exerted on the arterial walls when blood flows through the arteries. The highest value of pressure, which occurs when the heart contracts and ejects blood to the arteries, is called the systolic pressure (SP). The diastolic pressure (DP) represents the lowest value occurring between the ejections of blood from the heart. Pulse pressure (PP) is the difference between SP and DP, that is, PP = SP–DP. The period from the end of one heart contraction to the end of the next is called the cardiac cycle. Mean pressure (MP) is the average pressure during a cardiac cycle.

Mathematically, MP can be decided by integrating the blood pressure over time. When only SP and DP are available, MP is often estimated by an empirical formula:

$$\mathrm{MP} \approx \mathrm{DP} + \frac{\mathrm{PP}}{3} \tag{65.1}$$

Note that this formula can be very inaccurate in some extreme situations. Although SP and DP are most often measured in the clinical setting, MP has particular importance in some situations because it is the driving force of peripheral perfusion. SP and DP can vary significantly throughout the arterial system, whereas MP is almost uniform in normal situations.

The values of blood pressure vary significantly during the course of 24 h according to an individual's activity [1]. Basically, three factors, namely, the diameter of the arteries, the cardiac output, and the state or quantity of blood, are mainly responsible for the blood pressure level. When the tone increases in the muscular arterial walls so that they narrow or become less compliant, the pressure becomes higher than normal. Unfortunately, increased blood pressure does not ensure proper tissue perfusion, and in some instances, such as certain types of shock, blood pressure may seem appropriate when peripheral tissue perfusion has all but stopped. Nevertheless, observation or monitoring of blood pressures affords dynamic tracking of pathology and physiology affecting the cardiovascular system. This system in turn has profound effects on the other organs of the body.

65.2 Measurement Techniques

The basis of any physiological measurement is the biological signal, which is first sensed and transduced or converted from one form of energy to another. The signal is then conditioned, processed, and amplified. Subsequently, it is displayed, recorded, or transmitted (in some ambulatory monitoring situations). Blood pressure sensors often detect mechanical signals, such as blood pressure waves, to convert them into electric signals for further processing or transmission. They work on a variety of principles, for example, resistance, inductance, and capacitance. For accurate and reliable measurements, a sensor should have good sensitivity, linearity, and stability [2].

65.3 Indirect Blood Pressure Measurement

Indirect measurement is often called noninvasive measurement because the body is not entered in the process. The upper arm, containing the brachial artery, is the most common site for indirect measurement because of its closeness to the heart and convenience of measurement, although many other sites may have been used, such as forearm or radial artery and finger. Distal sites such as the wrist, although convenient to use, may give much higher SP than brachial or central sites as a result of the phenomena of impedance mismatch and reflective waves [3]. An occlusive cuff is normally placed over the upper arm and is inflated to a pressure greater than the systolic blood pressure. The cuff is then gradually deflated, while a detector system simultaneously employed determines the point at which the blood flow is restored to the limb. The detector system does not need to be a sophisticated electronic device. It may be as simple as manual palpation of the radial pulse. The most commonly used indirect methods are auscultation and oscillometry, each is described later.

65.3.1 Auscultatory Method

The auscultatory method most commonly employs a mercury column, an occlusive cuff, and a stethoscope. The stethoscope is placed over the blood vessel for auscultation of the Korotkoff sounds, which defines both SP and DP. The Korotkoff sounds are mainly generated by the pulse wave propagating through the brachial artery [4]. The Korotkoff sounds consist of five distinct phases. The onset of phase I Korotkoff sounds (first appearance of clear, repetitive, tapping sounds) signifies SP, and the onset of phase V Korotkoff sounds (sounds disappear completely) often defines DP [5].

Observers may differ greatly in their interpretation of the Korotkoff sounds. Simple mechanical error can occur in the form of air leaks or obstruction in the cuff, coupling tubing, or Bourdon gage. Mercury can leak from a column gage system. In spite of the errors inherent in such simple systems, more mechanically complex systems have come into use. The impetus for the development of more

elaborate detectors has come from the advantage of reproducibility from observer to observer and the convenience of automated operation. Examples of this improved instrumentation include sensors using plethysmographic principles, pulse-wave velocity sensors, and audible as well as ultrasonic microphones [6].

The readings by auscultation do not always correspond to those of intra-arterial pressure [5] The differences are more pronounced in certain special occasions such as obesity, pregnancy, arteriosclerosis, and shock. Experience with the auscultation method has also shown that determination of DP is often more difficult and less reliable than SP. However, the situation is different for the oscillometric method where oscillations caused by the pressure pulse amplitude are interpreted for SP and DP according to empirical rules [7].

65.3.2 Oscillometric Method

In recent years, electronic pressure and pulse monitors based on oscillometry have become popular for their simplicity of use and reliability. The principle of blood pressure measurement using the oscillometric technique is dependent on the transmission of intra-arterial pulsation to the occluding cuff surrounding the limb. An approach using this technique could start with a cuff placed around the upper arm and rapidly inflated to about 30 mmHg above the systolic blood pressure, occluding blood flow in the brachial artery. The pressure in the cuff is measured by a sensor. The pressure is then gradually decreased, often in steps, such as 5–8 mmHg. The oscillometric signal is detected and processed at each step of pressure. The cuff pressure can also be deflated linearly in a similar fashion as the conventional auscultatory method.

Figure 65.1 illustrates the principle of oscillometric measurement along with auscultatory measurement. Arterial pressure oscillations are superimposed on the cuff pressure when the blood vessel is no longer fully occluded. Separation of the superimposed oscillations from the cuff pressure is accomplished by filters that extract the corresponding signals. Signal sampling is carried out at a rate determined by the pulse or heart rate [7]. The oscillation amplitudes are most often used with an empirical algorithm to estimate SP and DP. Unlike the Korotkoff sounds, the pressure oscillations are detectable throughout the whole measurement, even at cuff pressures higher than SP or lower than DP. Since many oscillometric devices use empirically fixed algorithms, variance of measurement can be large across a wide range of blood pressures [8]. Significantly, however, MP is determined by the lowest cuff pressure of maximum oscillations [9] and has been strongly supported by many clinical validations [10,11].

65.3.3 Self-Measurement

From the growing number of publications on the topic in recent years, it is evident that the interest in self-measurement of blood pressure has increased dramatically. There is also evidence that the management of patients with high blood pressure can be improved if clinic measurements are supplemented by home or ambulatory monitoring. Research has shown that blood pressure readings taken in the clinic can be elevated, by as much as 75 mmHg in SP and 40 mmHg in DP, when taken by a physician. The tendency for blood pressure to increase in certain individuals in the presence of a physician due to stress response is generally known as "white-coat" hypertension [12]. When reasonably priced and easy to use, oscillometric devices became commonly available in the early 1970s, public interest in the self-measurement of blood pressure increased, and this has made it possible for greater patient involvement in the detection and management of hypertension [13]. Health-care costs may also be reduced by home monitoring. Indeed, a recent study found that costs were almost 30% lower for patients who measured their own blood pressure than those who did not [14]. Measurements taken at patient's home are more highly correlated to 24 h blood pressure levels than clinic readings are. It has also been shown that most patients are able to monitor their blood pressure and may be more relaxed as well as assured by doing so, particularly when experiencing symptoms [15].

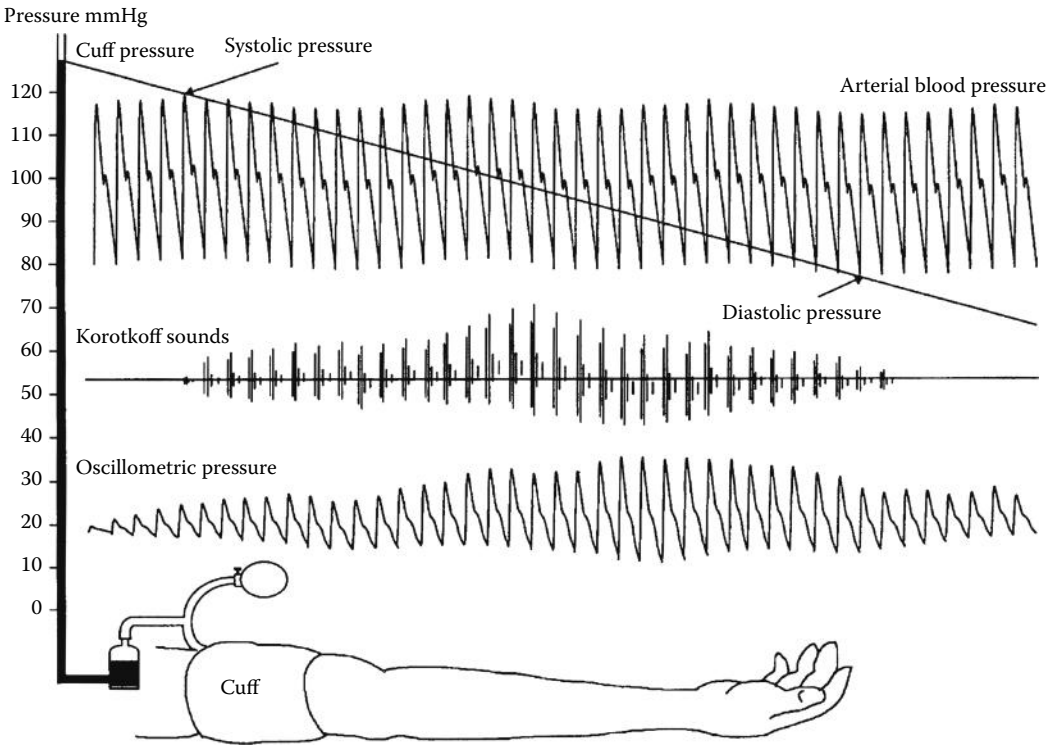


FIGURE 65.1 Indirect blood pressure measurements: oscillometric measurement and auscultatory measurement. (Adapted from Current technologies and advancement in blood pressure measurements—review of accuracy and reliability, *Biomed. Instrum. Technol.*, AAMI, Arlington, VA. With permission.)

65.3.4 Ambulatory Monitoring

There is great significance for ambulatory monitoring of blood pressure. Over a period of 24 h, blood pressure is subject to numerous situational and periodic fluctuations [1]. The pressure readings have a pronounced diurnal rhythm in an individual, with a decrease from 10 to 20 mmHg during sleep and a prompt increase on getting up and walking in the morning. Readings tend to be higher during working hours and lower at home, and they depend on the pattern of activity. After a bout of vigorous exercise or strenuous work, blood pressure may be reduced for several hours. The readings may be raised if the patient is talking during the measurement period. Smoking a cigarette and drinking coffee, especially if they are combined, may both raise the pressure [16]. When assessing the efficacy of antihypertensive drugs, ambulatory blood pressure monitoring can provide considerable information and validation of the drug treatment [17].

Although the technique of noninvasive ambulatory blood pressure monitoring was first described more than three decades ago, it has only recently become accepted as a clinically useful procedure for evaluation of patients with abnormal regulation of blood pressure. It gives the best evaluation for patients who have white-coat hypertension. Technical advances in microelectronics and computer technology have led to the introduction of ambulatory monitors with improved accuracy and reliability, small size, quiet operation, and reasonable low price. They can take and store several hundred readings over a period of 24 h while patients may not be compromised with their normal activities, thus becoming usable for purposes of clinical diagnosis [18]. Theoretically, ambulatory monitoring can provide information about the level and variability covering the full range of blood pressure experienced during day-to-day activities. It is now recognized to be a very useful procedure in clinical

practice since blood pressure varies significantly during the course of 24 h, especially useful in detecting white-coat hypertension. However, many studies have found that accuracy of monitoring using current ambulatory monitors is acceptable only when patients are at rest but not during physical activity [19] or under truly ambulatory conditions. Report of error codes during operation in the latter situations is much higher [20].

65.3.5 Cuff Size

Both the length and width of an occluding cuff are important for accurate and reliable measurement of blood pressure by indirect methods. A too-short or too-narrow cuff results in false high blood pressure readings. Several studies have shown that a cuff of inappropriate size in relation to the patient's arm circumference can cause considerable error in blood pressure measurement [21]. The cuff should also fit around the arm firmly and comfortably. Some manufacturers have designed cuffs with a fastener spaced so that a cuff of appropriate width only fits an arm of appropriate diameter. With this design, the cuff will not stay on the arm during inflation unless it fits accordingly.

According to the American Heart Association (AHA) recommendations [5], the width of the cuff should be 40% of the midcircumference of the limb and the length should be twice the recommended width. Table 65.1 presents the AHA cuff sizes covering from neonates to adults.

65.3.6 Recommendations, Standards, and Validation Requirements

The AHA has published six editions of the AHA recommendations for indirect measurement of arterial blood pressure. The most recent edition [5] included the recommendations of the joint national committee on the diagnosis, evaluation, and treatment of hypertension for classifying and defining blood pressure levels for adults (age 18 years and older) [22], as shown in Table 65.2. The "Report of the Second Task Force on Blood Pressure Control in Children" [23] offered classification of hypertension in young age groups from newborns to adolescents, as shown in Table 65.3.

The AHA recommendations provide a systemic step-by-step procedure for measuring blood pressure, including equipment, observer, subject, and technique. It extends considerations of blood pressure recording in special populations such as infants and children, elderly, and pregnant and obese subjects.

TABLE 65.1 AHA Acceptable Bladder Dimensions for Arm of Different Sizes[a]

Cuff	Bladder Width (cm)	Bladder Length (cm)	Arm Circumference Range at Midpoint (cm)
Newborn	3	6	≤6
Infant	5	15	6–15[b]
Child	8	21	16–21[b]
Small adult	10	24	22–26
Adult	13	30	27–34
Large adult	16	38	35–44
Adult thigh	20	42	45–52

Source: Adapted from the *Recommendations for Human Blood Pressure Determination by Sphygmomanometers*, American Heart Association, Dallas, TX, 1993. With permission.

[a] There is some overlapping of the recommended range for arm circumferences in order to limit the number of cuffs; it is recommended that the larger cuff be used when available.

[b] To approximate the bladder width/arm circumference ratio of 0.40 more closely in infants and children, additional cuffs are available.

TABLE 65.2 Recommendations of the Joint National Committee on the Diagnosis, Evaluation, and Treatment of Hypertension for Classifying and Defining Blood Pressure Levels for Adults (Age 18 Years and Older)[a]

Category	SP (mmHg)	DP (mmHg)
Normal[b]	<130	<85
High normal	130–139	85–89
Hypertension[c]		
Stage 1 (mild)	140–159	90–99
Stage 2 (moderate)	160–179	100–109
Stage 3 (severe)	180–209	110–119
Stage 4 (very severe)	≥210	120

Source: Adapted from The fifth report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure (JNCW), *Arch. Intern. Med.*, 153, 154, 1993.

[a] Not taking antihypertensive drugs and not acutely ill. When SPs and DPs fall into different categories, the higher category should be selected to classify the individual's blood pressure status. For instance, 160/92 mmHg should be classified as stage 2, and 180/120 mmHg should be classified as stage 4. Isolated systolic hypertension is defined as a systolic blood pressure of 140 mmHg or more and a diastolic blood pressure of less than 90 mmHg and staged appropriately (e.g., 170/85 mmHg is defined as stage 2 isolated systolic hypertension). In addition to classifying stages of hypertension on the basis of average blood pressure levels, the clinician should specify presence or absence of target-organ disease and additional risk factors. For example, a patient with diabetes and a blood pressure of 142/94 mmHg plus left ventricular hypertrophy should be classified as having "stage 1 hypertension with target-organ disease (left ventricular hypertrophy) and with another major risk factor (diabetes)." This specificity is important for risk classification and management.

[b] Optimal blood pressure with respect to cardiovascular risk is less than 120 mmHg systolic and less than 80 mmHg diastolic. However, unusually low readings should be evaluated for clinical significance.

[c] Based on the average of two or more readings taken at each of two or more visits after an initial screening.

It also provides recommendations of self-measurement or home measurement of blood pressure, as well as ambulatory blood pressure measurement.

The Association for the Advancement of Medical Instrumentation (AAMI) and American National Standard Institute (ANSI) published and revised a national standard [24,25] for evaluating electronic or automated sphygmomanometers. This standard established labeling requirements, safety and performance requirements, and referee test methods for electronic or automated sphygmomanometers used in indirect measurement of blood pressure. Specific requirements for ambulatory blood pressure monitors were also included. Recently, AAMI/ANSI amended this SP10 standard to include neonatal devices as well [26]. Some of the specific requirements, procedures, and limits were modified to fit neonatal applications, such as the maximum cuff pressure, ranges of age and weight, reference standards for validation, and minimum sample size of data. The overall system efficacy for both neonatal and adult devices requires that for SPs and DPs treated separately, the mean difference between the paired measurements of the test system and the reference standard shall be ±5 mmHg or less, with a standard deviation of 8 mmHg or less.

For manual or nonautomated indirect blood pressure measuring devices, ANSI/AAMI SP9 standard [27] applies.

The British Hypertension Society (BHS) also published and revised a protocol for assessing accuracy and reliability of blood pressure measurement using automatic and semiautomatic devices [28,29]. Many automatic and semiautomatic devices, including ambulatory devices, have been evaluated according to the BHS protocol. Such evaluation provided a quality-control mechanism for manufacturers and an objective comparison for customers. However, there are many more devices available on the market,

TABLE 65.3 Classification of Hypertension in the Young by Age Group

Age Group	High Normal (90th–94th Percentile) (mmHg)	Significant Hypertension (95th–99th Percentile) (mmHg)	Severe Hypertension (>99th Percentile) (mmHg)
Newborns (SBP)		96–105	≥106
7 days	—	104–109	≥110
8–30 days	—		
Infants (≥2 year)			
SBP	104–111	112–117	≥118
DBP	70–73	74–81	82
Children			
3–5 year			
SBP	108–115	116–123	≥124
DBP	70–75	76–83	≥84
6–9 year			
SBP	114–121	122–129	≥130
DBP	74–77	78–85	≥86
10–12 year			
SBP	122–125	126–133	≥134
DBP	78–81	82–89	≥90
13–15 year			
SBP	130–135	136–143	≥144
DBP	80–85	86–91	≥92
Adolescents			
16–18 year			
SBP	136–141	142–149	≥150
DBP	84–91	92–97	≥98

Source: Adapted from the Report of the Second Task Force on Blood Pressure Control in Children—1987, *Pediatrics*, 79, 1, 1987. With permission.

SBP, systolic blood pressure; DBP, diastolic blood pressure.

which have not been accordingly evaluated. Different from the AAMI SP10 standard in which either indirect or direct blood pressure may be used as a reference standard, the BHS protocol relies exclusively on references of sphygmomanometric blood pressure measurement and does not recommend comparison with intra-arterial blood pressure values [30]. This could make accurate validation of ambulatory devices difficult because sphygmomanometric measurements during exercise and under ambulatory conditions are not accurate [31].

Significantly, the BHS protocol emphasized the need on special-group validation, such as children, pregnancy, and the elderly for the intended use. It also emphasized the need for validation under special circumstances, such as exercise and posture. The accuracy criteria use a grading system based on the percentages of test instrument measurements differing from the sphygmomanometric measurements by ≤5, ≤10, and ≤15 mmHg for systolic and diastolic blood pressure, respectively, as shown in Table 65.4.

65.3.7 Manufacturer, Product, Price, Efficacy, and Technology

The annual publication of the *Medical Device Register* is a comprehensive reference work that provides a wealth of detailed information on U.S. and international medical devices, medical device companies, OEM suppliers, and the key personnel in the industry. Blood pressure devices are listed in the sphygmomanometer directory. Price information of specific models for some providers is also published. For example, A&D Engineering, Inc. listed price from $51.95 (model UA701) to $179.95 (model UA-751) for a

TABLE 65.4 Grading Criteria of the 1993 BHS Protocol[a,b]

Grade	Absolute Difference between Standard and Test Device (mmHg)		
	≤5	≤10	≤15
	Cumulative Percentage of Readings		
A	60	85	95
B	50	75	90
C	40	65	85
D	Worse than C		

Source: Adapted from The British Hypertension Society protocol for the evaluation of blood pressure measuring devices, *J. Hypertension*, 11 (Suppl. 2), S43, 1993. With permission.

[a] Grades are derived from percentages of readings within 5, 10, and 15 mmHg. To achieve a grade, all three percentages must be equal to or greater than the tabulated values.

[b] Grading percentages changed from the 1990 BHS protocols due to changes in sequential assessment of blood pressure references. See original publications for details (1990, 1993).

whole line of sphygmomanometers in the 1997 *Medical Device Register* [32]. Since technology and market can change rapidly, models, features, specifications, and prices may change accordingly. More specific and updated information may be available by contacting the manufacturers or distributors directly.

Table 65.5 lists only a limited number of indirect blood pressure devices from a literature review. Many of the listed blood pressure devices have multiple evaluation studies, and only a few study results are presented here. In view of reference standards for comparison, although direct and indirect methods yield similar measurements, they are rarely identical because the direct method measures pressure and the indirect method are more indicative of flow [5]. Egmond et al. [33] evaluated the accuracy and reproducibility of 30 home blood pressure devices in comparison with a direct brachial arterial standard. They found average offsets of all tested devices amounted to –11.7 mmHg for systolic and 1.6 mmHg for diastolic blood pressure, which were close to those of the mercury sphygmomanometer (–14.2 mmHg for SP and –0.1 mmHg for DP), indicating a significant difference between the two assessment standards. When selecting a blood pressure device for a specific application, the evaluation using the reference that is of a common practice in the intended population or environment may be practically more informative, since that reference has been the common basis for decision making in blood pressure diagnosis and treatment.

Different evaluation results for the same brand product can also be due to different versions of a model used for validation, where a later version may have performance improvement over the earlier one [34]. Another source of discrepancies can come from utilizing different study protocols or only partially following the same protocols. It is recommended that the original clinical evaluation report be carefully examined in determining the desired efficacy that may meet the users' requirements. If the devices were FAD approved for marketing in the United States, one may request a copy of their clinical validation report directly from the manufacturer.

In addition to the fundamental categories such as intended use, efficacy, and acquisition technology, listed in Table 65.5, many other categories are also very important in evaluating, selecting, purchasing, using, and maintaining blood pressure devices. These include but are not limited to the following items: measurement range of each pressure (systolic, diastolic, and mean) for each mode of intended use (i.e., neonates, children, adults); maximum pressure that can be applied by the monitor and cuff for each

TABLE 65.5 Survey of Indirect Blood Pressure Device Manufacturer, Product, Intended Use, and Efficacy

				Efficacy								
				Reference Standard	BHS Protocol Grading		AAMI SP10 Comparison (Device—Reference) (mmHg)		Other Validations (Device—Reference) (mmHg)			
Manufacturer	Model	Technology	Intended Use	MC/AC[a]	SBP	DBP	SBP	DSP	SBP	DBP	MBP	Ref.
A&D, Tokyo, Japan	TM-2420/ TM-2020	Korotkoff	Health care: ambulatory	MC	D	D	−4 ± 11	−2 ± 11				35
	TM-2420 version 7	Korotkoff	Health care: ambulatory	MC	B	B	−1.8 ± 5.0	−3.5 ± 6.8				36
Colin Medical Instruments, Plainfield, NJ	ABPM 630	Korotkoff (primary mode)	Health care: ambulatory	AC					1.4 ± 7.1	−0.1 ± 5.6		37
				MC					−0.4 ± 4.6	−6.0 ± 5.9		
		Oscillometry (backup mode)	Health care: ambulatory	AC					4.5 ± 6.6	−1.2 + 6.3		
				MC					1.9 ± 4.0	−6.9 + 5.1		
Del Mar Avionics, Irvine, CA	Pressurometer IV	Korotkoff (ECG R-wave gating)	Health care: ambulatory	MC	C	D	−2 ± 11	−3 ± 11				38
				MC					1.2 ± 7.3	−2.2 ± 5.7		39
Disetronic Medical Systems AG Burgdorf, Switzerland	CH-Druck/ Pressure Scan ERKA	Korotkoff	Health care: ambulatory	MC	A	A	−3 ± 4	−2 ± 4				40
Novacor, France	DIASYS 200	Korotkoff	Health care: ambulatory	MC	C	C	−1 ± 8	0 ± 8				41
Oxford Medical, Abingdon, Oxford, United Kingdom	Medilog ABP	Korotkoff	Health care: ambulatory	AC			−8 ± 8	6 ± 6				42
				MC			−4 ± 6	−2 ± 8				
Disetronic Medical Systems AG, Burgdorf, Switzerland	Profilomat	Korotkoff	Health care: ambulatory	MC	B	A	−3 ± 5	−1 ± 5				43
SpaceLabs Medical, Redmond, WA	90207	Oscillometry	Health care: ambulatory	MC	B	B	−1 ± 7	−3 ± 6				44
Suntech Medical Instruments, Raleigh, NC	Accutracker II (v30/23)	Korotkoff (ECG R-wave gating)	Health care: ambulatory	MC	A	C	−1.3 ± 6.5	−4.5 ± 7.3				45[b]

(*continued*)

TABLE 65.5 (continued) Survey of Indirect Blood Pressure Device Manufacturer, Product, Intended Use, and Efficacy

				Efficacy								
				Reference Standard	BHS Protocol Grading		AAMI SP10 Comparison (Device—Reference) (mmHg)		Other Validations (Device—Reference) (mmHg)			
Manufacturer	Model	Technology	Intended Use	MC/AC[a]	SBP	DBP	SBP	DSP	SBP	DBP	MBP	Ref.
Tycos-Welch-Allyn, Arden, NC	QuietTrak	Korotkoff	Health care: ambulatory				0.3 ± 5.0	−1.5 ± 7.5				46[c]
Colin Medical Instruments, San Antonio, TX	BP8800MS	Oscillometry	Health care: children	MC			3.2 ± 6.0	−0.8 + 5.2				47
		Oscillometry	Health care: adults	MC			2.8 ± 5.4	0.0 ± 4.9				
Critikon, Tampa, FL	Dinamap 1846SX	Oscillometry	Health care: neonates, children, adults	AC					−8.8 ± 11.2	1.6 ± 8.9	−1.8 ± 9.7	48
	Dinamap portable monitor	Oscillometry	Health care: neonates, children, adults	MC	B	D	−1 ± 7	−6 ± 7				49[d]
SpaceLabs Medical, Redmond, WA	Oscillometric blood pressure monitor	Oscillometry	Health care: neonates,	AC	B	B	0.1 ± 4.3	2.7 ± 4.8				50[e]
			Health care: children, adults	MC	B	B	−0.6 ± 5.9	0.9 ± 6.4				
Ohmeda, Denver, CO	Finapres 3700	Volume clamping	Health care: continuous monitoring	AC					−8.4 ± 8.6	−1.1 ± 7.0	−6.8 ± 6.7	48
Terumo, Tokyo, Japan	ES-H51[f]	Korotkoff (primary mode)	Health care: routine	MC	A	A	0.7 ± 2.9	0.3 ± 2.6				51[g]
		Oscillometry (backup mode)	Clinical	MC	B	A	−0.3 ± 5.7	−0.3 ± 4.3				

Matsushita, Osaka, Japan	Denko EW 160	Oscillometry	Self-care: home measurement	MC	1.8 ± 5.2	−1.7 ± 5.5	52
Nissei, Tokyo, Japan	DS 91[f]	Korotkoff	Self-care: home measurement	MC	−2.5 ± 7.4	2.8 ± 10.8	52
Omron, Tokyo, Japan	HEM 439[f]	Korotkoff	Self-care: home measurement	MC	−0.2 ± 5.3	6.2 ± 9.9	52
	HEM 719K	Korotkoff	Self-care: home measurement	MC	−2.3 ± 5.6	2.4 ± 4.7	52
	401C[f]	Oscillometry	Self-care: home measurement	MC	−1.6 ± 7.7	2.4 ± 6.1	52
Sharp, Osaka, Japan	MB 305H[f]	Korotkoff	Self-care: home measurement	MC	0.5 ± 4.5	9.6 ± 14.3	52
	MB 500A	Oscillometry	Self-care: home measurement	MC	−1.8 ± 6.7	0.7 ± 6.3	52
A&D, Tokyo, Japan	Takeda UA 751	Oscillometry	Self-care: home measurement	MC	−4.1 ± 5.6	0.4 ± 7.8	52

Source: Adapted from Current technologies and advancement in blood pressure measurements-review of accuracy and reliability, *Biomed. Instrum. Technol.*, AAMI, Arlington, VA (publication pending). With permission.

[a] MC, mercury column; AC, arterial catheter.

[b] Data quoted for the standing position; grading was the same as for pooled data of three positions (supine, seated, and standing).

[c] Data quoted for the three positions of supine, seated, and standing.

[d] Efficacy quoted was determined in adult population.

[e] Efficacy quoted was determined in neonate and adult populations, respectively.

[f] Semiautomatic; all other listed are automatic.

[g] Only partially followed AAMI and BHS protocol and only validated one size (median) of three cuffs (small, median, large).

mode of intended use; cuff size range for the target population of the intended use; cost; measurement and record failure rate; noise and artifact rejection capability; mode of operation (manual, automatic, semiautomatic); data display; recording, charting, reporting, and interfacing; physical size and weight; power consumption; operation manual; service manual; and labels and warnings.

65.3.8 Advancement of Indirect Blood Pressure Measurement

Since the introduction of Dinamap™, an automated blood pressure monitor based on the oscillometric principle [9], many variants of oscillometric algorithms were developed. However, the fixed or variable fractions of the maximum oscillations are still the fundamental algorithms of the oscillometry [10,53,54]. Typically, mean blood pressure was determined by the lowest cuff pressure with greatest average oscillation [11]. Systolic and diastolic blood pressure was determined by the cuff pressure with the amplitude of oscillation being a fixed fraction of the maximum. Performance of the algorithms may be improved by introducing a greater level of complexity or variables into considerations. The Dinamap 1846SX (Critikon, Tampa, FL) oscillometric device offered two measurement modes. The normal mode uses two matching pulse amplitudes at each cuff pressure step to establish an oscillometric envelope or curve. Therefore, measurement time is heart rate dependent. The second mode, which the manufacturer refers to as "stat mode," is capable of faster determination by disabling the dual pulse-matching algorithm that was designed for artifact rejection. The stat mode does not appear to compromise accuracy in anesthetized patients [55], in which rapid measurement of blood pressure is often more desirable, particularly during induction and management of anesthesia.

Another variant of the oscillometric algorithm was developed by Protocol Systems [56]. In addition to using pulse amplitude for primary artifact rejection, it further calculated impulse value, a principal area of pulse waveform, in constructing an oscillometric curve. This curve is smoothed by employing a Kalman filter that also provides an expected mean and acceptable upper and lower bounds of prediction for the principal area of subsequent pulse waveform. Smoothing of the oscillometric curve is accomplished by using the difference between the predicted and calculated area data of pulse waveform for each cuff pressure step. Blood pressures are derived from the final smoothed oscillometric curve.

In more recent study, oscillometric algorithms using an artificial neural network have been reported to produce better estimates of reference blood pressures than the standard oscillometric algorithm [57]. By using neural network training and processing, subtle features and nonlinear relationships of the oscillometric envelope have been modeled. Empiricism of the oscillometric fixed fraction criteria is overcome, and variances of measurements are greatly reduced.

Because of its low risk and cost, noninvasive continuous blood pressure monitoring represents another need in critical-care monitoring to supplement invasive arterial catheterization. A significant development in this field is the arterial counterpulsation principle, proposed by Penaz [58] and further developed by two major groups of people [59,60]. Finapres™, a continuous finger arterial blood pressure monitor, was engineered and developed by Ohmeda, Denver, CO. Many clinical evaluation reports of these devices have been published since then.

Recently, a number of other continuous blood pressure monitors have been made commercially available. Examples of these are Cortronic APM770 [61], which monitors pulsation of the brachial artery with a slightly pressurized arm cuff and calibrates it to a continuous pressure waveform; Sentinel ARTRAC 7000 [62], which monitors pulse transit time and correlates that to pressure change; and Colin CBM-3000 and Jentow (Colin Electronics, Komaki, Japan) [63,64] and Nellcor NCAT N-500 (Nellcor, Hayward, CA) [65], which are tonometric devices monitoring the radial artery pulse waveform by a matrix pressure sensor. All of these monitors require a frequent calibration reference. Except for a few favorable reports with the tonometric method and devices, many reports so far are unfavorable. Nevertheless, noninvasive continuous monitoring represents an important and growing field of biomedical sensor and instrumentation research and development. Continuous monitors, which maintain cuff pressure, must periodically relieve pressure to prevent the risk of venous congestion, edema, swelling, and tissue damage.

65.4 Direct Blood Pressure Measurement

Direct measurement is also called invasive measurement because bodily entry is made. For direct arterial blood pressure measurement, an artery is cannulated. The equipment and procedure require proper setup, calibration, operation, and maintenance [66]. Such a system yields blood pressures dependent upon the location of the catheter tip in the vascular system. It is particularly useful for continuous determination of pressure changes at any instant in dynamic circumstances. When heavy blood loss is anticipated, powerful cardiovascular medications are suddenly administered, or when a patient is induced to general anesthesia, continuous monitoring of blood pressure becomes vital.

Most commonly used sites to make continuous observations are the brachial and radial arteries. The femoral or other sites may be used as points of entry to sample pressures at different locations inside the arterial tree, or even the left ventricle of the heart. Entry through the venous side of the circulation allows checks of pressures in the central veins close to the heart, the right atrium, the right ventricle, and the pulmonary artery. A catheter with a balloon tip carried by blood flow into smaller branches of the pulmonary artery can occlude flow in the artery from the right ventricle so that the tip of the catheter reads the pressure of the left atrium, just downstream. These procedures are very complex and there is always concern of risk of hazard as opposed to benefit [67].

Invasive access to a systemic artery involves considerable handling of a patient. The longer a catheter stays in a vessel, the more likely an associated thrombus will form. The Allen's test can be performed by pressing on one of the two main arteries at the wrist when the fist is clenched and then opening the hand to see if blanching indicates inadequate perfusion by the other artery. However, it has proved an equivocal predictor of possible ischemia [68]. In the newborn, when the arterial catheter is inserted through an umbilical artery, there is a particular hazard of infection and thrombosis, since thrombosis from the catheter tip in the aorta can occlude the arterial supply to vital abdominal organs. Some of the recognized contraindications and complications include poor collateral flow, severe hemorrhage diathesis, occlusive arterial disease, arterial spasm, and hematoma formation [69].

In spite of well-studied potential problems, direct blood pressure measurement is generally accepted as the gold standard of arterial pressure recording and presents the only satisfactory alternative when conventional cuff techniques are not successful. This also confers the benefit of continuous access to the artery for monitoring gas tension and blood sampling for biochemical tests. It also has the advantage of assessing cyclic variations and beat-to-beat changes of pressure continuously and permits assessment of short-term variations [70,71].

65.4.1 Catheter–Tubing–Sensor System

A large variety of vascular catheters exist. Catheter materials have undergone testing to ensure that they have a minimal tendency to form blood clots on their surface. The catheter chosen may be inserted percutaneously over a hollow stylet into the blood vessel. Guide wires can be useful to facilitate longer or larger-diameter catheters into vessels, after the guide wires have been placed through a smaller catheter or needle. Less often, entry to a vessel requires a "cutdown," a direct exposure of the vessel after a skin incision. Ultrasonic devices may assist locating the vessels not readily apparent at the skin surface.

Although pressure sensors can be located at the catheter tip, this presents a problem for calibration if left in place and a clot forms near the tip of the catheter, damping the pressure signal. Instead, most catheters connect to an external pressure sensor via fluid-filled low-compliance tubing. The signal from the sensor then undergoes transformation for display or recording. The sensor may take one of several forms, from a variable resistance diaphragm to a silicon microchip. A basic system can consist of an intravascular catheter connected to a rigid fluid-filled catheter and tubing that communicates the pressure to an elastic diaphragm, the deflection of which is detected electrically. There is a direct relationship between the deflection of the diaphragm and the voltage.

The higher the voltage is the greater the pressure. Continuous low-rate infusion of heparinized saline is carried out to keep the catheter patent or free from coagulation. The advent of disposable sensor kits has greatly simplified the clinical use of intravascular monitoring [72]. With the cost continually being lowered with the development of semiconductor industry, disposable sensors become more and more cost-effective.

Although direct recording is considered the most accurate method, its accuracy may be limited by variations in the kinetic energy of the fluid in the catheter or dynamic frequency response of the measurement system. The hydraulic link between the patient and the sensor is the major source of potential errors and hazard for the monitoring. Damping and degrading the system's natural frequency, caused by trapped air bubbles, small catheters, various narrow connections, compliant and too long tubing, and too many components connected, are the two characteristic problems with a pressure sensor system. Extreme care should be exercised to eliminate all air bubbles from the fluid to provide adequate dynamic response. The sensor should be zeroed at the level of the heart to eliminate hydrostatic error [73]. A fast flush testing is easy to use for inspection of the dynamic response of the whole system of catheter–tubing–sensor. It can also help direct adjustments for the system to minimize dynamic artifacts [74,75].

65.5 Reproducibility, Accuracy, and Reliability Issues and Recommendations for Corrective Measures

For each blood pressure assay technique, there is an issue of reproducibility of measurements given approximately similar conditions. Reproducibility quantifies the internal uncertainty of an individual method and instrument, whereas accuracy quantifies the external uncertainty when compared with a reference. Table 65.6 presents estimated uncertainties of reproducibility for three blood pressure measuring techniques: auscultation, oscillometry, and umbilical arterial catheter [50]. When dealing with blood pressure measurement, it is important to bear in mind that even for standard methods, there is a certain amount of nonrepeatable random error. Consequently, taking the average of repeated measurements or multiple readings is always advised before any serious recommendation or management is made.

Table 65.7 presents a review of common problems associated with accuracy and reliability in both indirect and direct blood pressure measurements. Consequences of these problems are analyzed and recommendation of preventive action or alternative solutions is provided. Hazard or safety analyses and review are also very important.

TABLE 65.6 Estimated Uncertainties of Reproducibility for Blood Pressure Measuring Techniques of Auscultation, Oscillometry, and Umbilical Arterial Catheter

	Auscultation	Oscillometry[a]	Umbilical Arterial Catheter
Neonate (mmHg)			
SP	N/A	3.3	2.2
DP	N/A	3.4	1.8
Adult			
SP	2.8	3.2	N/A
DP	2.2	3.5	N/A

Source: From Sun, M. et al., *Biomed. Instrum. Technol.*, 30, 439, 1996.

[a] Evaluated from SpaceLabs Medical Oscillometric monitor [50].

TABLE 65.7 Common Issues of Accuracy and Reliability in Blood Pressure Measurement and Recommendations of Preventive Action or Alternative Solution in Both Indirect and Direct Measurements

	Problem	Result	Recommendation
Source	Indirect Measurement		
Subject	Obesity, peripheral edema, peripheral vascular disease	Weak Korotkoff sounds and diminished sound transmission may reduce the accuracy and reliability of auscultatory measurement; oscillometric measurement may also be affected	Verify with a second indirect method such as oscillometry; direct blood pressure measurement may be elected to use in severe conditions that indirect measurement does not warrant sufficient accurate and reliable measurement
	Shock, severe peripheral vasoconstriction, diminished peripheral circulation resulting from shunting of blood to central organs; the Korotkoff sounds and pulses may be absent even in the presence of normal pressure [76]	Any of the indirect methods, including auscultatory, oscillometric, and Doppler techniques, may not provide accurate and reliable reading; indirect measurement may be impossible or may give misleading results	Direct measurement should be considered
	Arrhythmias, respiratory effect	Pronounced variation in beat-to-beat blood pressure and waveform	Take multiple measurements and average
	Subject shivering, pain, anxiety, discomfort, motion artifact	Shivering and motion artifact may cause either false high or false low reading, whereas pain, anxiety, and discomfort may cause false high reading	Minimize pain, anxiety, and discomfort; reduce shivering and movement
	Physical activity within 5 min of measurement; talking, moving, arm unsupported, back unsupported, legs dangling, and any other isometric activities	False high reading that does not reflect subject's resting blood pressure	Subject should rest at least 5 min in the same position that blood pressure is going to be taken; subject should not talk and involve any isometric activities during measurement; arm should be supported at heart level
	Arm supported at above heart level	Hydrostatic pressure causes false low reading by 0.78 mmHg for each centimeter of offset [77]	Support the arm with midpoint of upper arm at heart level
	Arm supported at below heart level	Hydrostatic pressure causes false high reading by 0.78 mmHg for each centimeter of offset	Support the arm with midpoint of upper arm at heart level
	White-coat hypertension during clinical measurement	Psychological or stress response causes blood pressure temporarily elevated and unrepresentative of subject's true condition	Take multiple self-measurements at home or ambulatory monitoring as desired, and provide record to care providers
	"Pseudohypertension" with calcified or stiffened arteries	Reduced arterial compliance, often occurring in the elderly, causes cuff blood pressure falsely too high or unable to be measured accurately	Use Osler maneuver for screening; direct method is recommended for those who test positive [78,79]

(*continued*)

TABLE 65.7 (continued) Common Issues of Accuracy and Reliability in Blood Pressure Measurement and Recommendations of Preventive Action or Alternative Solution in Both Indirect and Direct Measurements

	Problem	Result	Recommendation
Source	Indirect Measurement		
Operator	Hose kinked	Will cause reading error or operation failure	Rearrange hose to avoid kink
	Cuff used too narrow for arm	Will cause false high reading	Select appropriate cuff size that its width encircles 40% of arm circumference
	Cuff used too wide for arm	May cause false low reading; may not fit on arm	Select appropriate cuff size that its width encircles 40% of arm circumference
	Cuff wrapped too loosely	Will cause false high reading; may introduce artifact of inter cuff-arm abrasion if placed for long-term monitoring	Cuff should be snugly applied; one should not be able to insert two fingers between the cuff and arm for adult
	Cuff wrapped too tightly	May cause false low reading; will restrict and impair limb circulation if placed for long-term monitoring	Cuff should be snugly but not restrictively applied; one should be able to insert one finger between the cuff and arm for adults
	Cuff pressure inflated too high	Patient discomfort; may induce increase in systolic blood pressure during inflation period, so-called cuff-inflation hypertension [80]	Inflate cuff pressure to 30 mmHg above palpatory blood pressure
	Cuff pressure inflated too low	Will either miss or have false low SP reading	Inflate cuff pressure to 30 mmHg above palpatory blood pressure
	Cuff pressure deflated too fast	May degrade the accuracy of the reading	Deflate cuff pressure at 2–4 mmHg per heart beat or 3 mmHg/s
	Cuff pressure deflated too slow	May cause discomfort or forearm congestion	Deflate cuff pressure at 2–4 mmHg per heart beat or 3 mmHg/s
	Repeated cuff pressure measurement too frequently	May cause discomfort and forearm congestion	A sufficient time should lapse (at least 60 s) before the next reading to allow the return of normal circulation
	Miss identifying auscultatory gap between SP and DP	Will cause false low SP or false high DP	Listen to the Korotkoff sounds carefully for a wide pressure deflation range or use oscillometric method
	Stethoscope head or sensor not over the brachial artery	Will not hear clear sounds or detect sufficient signal for blood pressure determination	Place the stethoscope head or sensor over the brachial artery at least 1.5 cm above the antecubital fossa
	Noise and artifact created by accidentally touching or bumping the cuff, hose, stethoscope, or sensor	May cause inaccurate reading or failure of reading	Avoid incidence of extraneous noise and artifact
Equipment	Leaky hose, bladder/cuff, or pneumatic components	Will cause inaccurate reading or failure in operation	Require service or replace equipment
	Faulty valves	Will cause inaccurate reading or failure in operation	Require service or replace equipment
	Limited selection for different size of cuffs	Will cause false low or false high reading if cuff is too large or too small, respectively	Manufacturer should provide appropriate label/labeling for the intended use and arm size; blood pressure measurement beyond the intended use of the device should be warned against and prohibited
	Device zero-shifted, out of calibration	Will create systematic bias or uncertainty in blood pressure reading	Require routine calibration and maintenance

TABLE 65.7 (continued) Common Issues of Accuracy and Reliability in Blood Pressure Measurement and Recommendations of Preventive Action or Alternative Solution in Both Indirect and Direct Measurements

	Problem	Result	Recommendation
Source	Direct Measurement		
Subject	Subject position change (e.g., body position change and bed lowered or elevated) in relation to pressure sensor	Subject heart level change in relation to pressure sensor will introduce bias of hydrostatic pressure in blood pressure recording	Move the sensor zero port to the heart level, and zero the sensor/monitor
	Catheter whip in pulmonary artery, catheter impact in aorta or ventricle	Catheter whip can result in superimposed waves of ±10 mmHg; catheter impact can cause high-frequency transients to occur in waveform [81]	Catheter whip and catheter impact are difficult to prevent; evaluation of pressure waveform and reading should consider the effect of these events
	Subject severe shivering, pain, anxiety, discomfort, moving	Severe shivering and moving may cause artifact on blood pressure waveform, whereas pain, anxiety, and discomfort may elevate blood pressure	Minimize pain, anxiety, and discomfort, and reduce shivering and moving
Operator	Tubing kinked	Will change dynamic response of tubing system and distort pressure waveform	Use short and low compliant tubing, and place tubing appropriately to avoid kink
	Sensor zero port higher than heart level when zeroing	Hydrostatic pressure causes false low pressure measurement by 0.78 mmHg for each centimeter of offset	Move the sensor zero port to heart level, and zero the sensor/monitor
	Sensor zero port lower than heart level when zeroing	Hydrostatic pressure causes false high pressure measurement by 0.78 mmHg for each centimeter of offset	Move the sensor zero port to heart level and zero the sensor/monitor
	Air bubbles entrapped in the tubing system	Air bubbles will decrease natural frequency and increase damping coefficient; therefore they damp and distort the waveform, causing high-frequency components to loss in pressure waveform	Eliminate air in both tubing system and flush solution bag; light tapping while fluid is filling the tubing system is an effective method for removing air
	Tubing too long, too thin, and with too many components	All of these will degrade the system dynamic response and result in distorted waveform and erroneous reading	Use tubing of large inner diameter and short length, and reduce the number of components as much as possible
	Connectors not tightly connected	Will decrease natural frequency of tubing system and cause pressure waveform to be distorted	Check loose luer-lock connection and cracked connection; replace cracked components and secure tight connection of all components
	Failure to flush the arterial line adequately after blood draw	May cause the catheter tip partially clotted by the blood and pressure waveform over damped and distorted	Flush the arterial line adequately; may need to replace with a new catheter if dynamic response cannot be improved to meet the minimum requirement
	Failure to zero the sensor/monitor after subject position change in relation to pressure sensor	Subject heart level change in relation to pressure sensor will introduce bias of hydrostatic pressure in blood pressure recording	Move the sensor zero port to the heart level and zero the sensor/monitor

(*continued*)

TABLE 65.7 (continued) Common Issues of Accuracy and Reliability in Blood Pressure Measurement and Recommendations of Preventive Action or Alternative Solution in Both Indirect and Direct Measurements

Source	Problem	Result	Recommendation
	Direct Measurement		
	Failure to provide constant infusion of anticoagulation/saline solution	May cause catheter tip partially clotted by the blood and pressure waveform overdamped and distorted	Check the constant infusion device to have sufficient flow rate; flush the arterial line adequately; may need to replace with a new catheter if dynamic response cannot be improved to meet the minimum requirement
	Failure to test dynamic response at least once a shift and anytime after blood draw or component change	This leaves the system dynamic performance unknown, which may affect the accuracy of SP the most and DP the second; MP is hardly affected	Routinely perform the fast flush test to evaluate the dynamic response visually according to Gardner's chart of natural frequency vs. damping coefficient [73]
Equipment	Not equipped with an appropriate flush device	May not be able to generate quality test waveform to evaluate the adequacy of dynamic response of the catheter–tubing–sensor system	Select appropriate flush device that permits fast flush test for the system dynamic response
	Tubing or component not transparent	Unable to see entrapped air bubbles	Use transparent tubing and components
	Tubing, sensor, or constant flush device too compliant	Will decrease natural frequency of the system and cause pressure waveform to be distorted	Use only high-quality and low-compliance tubing, sensor and constant flush device
	Stopcocks not tightly sealed	Will decrease natural frequency of the system and cause pressure waveform to be distorted	Replace with tightly sealed, high-quality stopcocks
	Monitor failure to zero the sensor electronically, sensor zero drift, or pressure amplifier zero drift	Will introduce unknown offset or bias in pressure measurement	Require service or replacement of the equipment
	Natural frequency and damping coefficient of the catheter–tubing–sensor system failure to meet minimum dynamic response requirement	Fidelity of pressure waveform recording suffers and accuracy of SP and DP measurement degrades	Need to optimize the catheter–tubing–sensor system by replacing part or all of the components; use low-compliance pressure sensor, tubing, and all other components; use short and large tubing and reduce the number of components as much as possible
	Blood pressure monitor failure to identify special events such as sensor zeroing, fast flush testing, blood drawing, as well as artifacts	Blood pressure monitor displays false digital reading of blood pressure without warning sign or error message	Health-care provider needs to exercise care in viewing the digital results with waveform display; quality-control or screening process is needed in dealing with monitoring database

Source: Adapted from Current technologies and advancement in blood pressure measurements-review of accuracy and reliability, *Biomed. Instrum. Technol.*, AAMI, Arlington, VA. With permission.

65.6 Blood Pressure Impact, Challenge, and Future

Hypertension is one of the most common and important risk factors of health in industrialized countries [82]. It is the leading cause of death in the United States. It is treatable by a variety of effective medications. It can cause serious damage to the heart and arteries leading to cardiac infarct, stroke, or renal failure. Significant sudden changes in blood pressure may also precede a major physiological catastrophe such as cardiac arrest. There is now almost universal acceptance that basic physiological parameters such as blood pressure should always be monitored in the clinical setting.

There has been increasing interest in automatic blood pressure monitoring devices in recent years, and some clinicians are now advising patients to record their blood pressure at home over a period of up to 3 months before starting antihypertensive medication [83]. Self-monitoring of blood pressure has become very common with the development of microchip technology and oscillometric monitors. The patients no longer have to learn how to listen for the Korotkoff sounds. This has also removed bias and observer errors, allowing more accurate measurement than by conventional techniques using a stethoscope and a mercury sphygmomanometer [84].

Special populations have unique blood pressure assessment requirements. Newborns require miniaturized equipment. The act of taking a blood pressure in a newborn may stimulate a series of movements causing motion artifact. The very obese may be hard to fit properly with a cuff at the upper arm, if the upper arm is too conical rather than cylindrical. In pregnancy, auscultatory and oscillometric methods, although useful to follow trends, may correspond poorly with central pressures [85], and even the proper Korotkoff sound (IV or V) to designate as DP is uncertain [86].

Observing blood pressures has limitations. It may suggest what is happening with blood volume, but sometimes does not reveal that blood volume has become inadequate until circulatory collapse has occurred. Venous and left atrial pressures are often used in an attempt to clarify blood volume problems but with uncertain results [87]. Similarly, a satisfactory blood pressure does not always indicate adequate tissue perfusion. Some medications that increase blood pressure can do so at the expense of general perfusion. Since blood pressure is measured at specific sites in the arterial tree, if circulation has become nonhomogeneous (such as what can happen in arteriosclerosis) the region distal to the arteriosclerosis can be compromised without warning from blood pressure readings sampled at another site. Even mean blood pressure, so useful otherwise, can fail in these circumstances.

In spite of inherent problems, observation of blood pressure through both old and new technologies retains more than enough usefulness to have remained an essential aspect of patient care. The promise of improved technology to solve problems such as those of motion artifacts, noninvasive continuous monitoring, long-distance telemetry, rapid analysis of accumulated or concurrent data, and assessment of new inaccessible regions of blood flow represents continued challenges for future biomedical research and development.

Recently, exciting research has revealed that comparing pressures taken at the arm and the ankle results in a simple but extremely useful index for assessment of lower extremity vascular disease, with implications for general cardiovascular risk factors [88]. The possibility of obtaining noninvasive blood pressures from arteries in the forehead by stick-on oscillometric patches has also been proved. At least in anesthetized patients, the forehead noninvasive blood pressure corresponded reasonably well with central arterial pressures [89]. Finger blood pressure monitors have found some applications in continuous ambulatory and sleep blood pressure assessments [90]. A technology that is capable of continuously monitoring brachial or even central blood pressure continues to be a clinical demand and future challenge.

References

1. T. G. Pickering, G. A. Harshfield, H. D. Klienert, S. Blank, and J. H. Laragh, Blood pressure during normal daily activities, sleep, and exercise, *J. Am. Med. Assoc.*, 247, 992–996, 1982.
2. L. A. Geddes, *The Direct and Indirect Measurement of Blood Pressure*, Chicago, IL: Year Book Medical Publishers, 1970.
3. Y. Saul, F. Aristidou, D. Klaus, A. Wiemeyer, and B. Losse, Comparison of invasive blood pressure measurement in the aorta with indirect oscillometric blood pressure measurement at the wrist and forearm, *Z. Kardiol.*, 84(9), 675–685, 1995.
4. W. Dock, Occasional notes—Korotkoff sounds, *N. Engl. J. Med.*, 302, 1264–1267, 1980.
5. Anonymous, *Recommendations for Human Blood Pressure Determination by Sphygmomanometers*, Dallas, TX: American Heart Association, 1993.
6. S. J. Meldrum, Indirect blood pressure measurement, *Br. J. Clin. Equip.*, 1, 257–265, 1976.
7. K. Yamakoshi, Non-invasive techniques for ambulatory blood pressure monitoring and simultaneous cardiovascular measurement, *J. Ambulat. Monit.*, 4, 123–143, 1991.
8. P. G. Loubser, Comparison of intra-arterial and automated oscillometric blood pressure measurement methods in postoperative hypertensive patients, *Med. Instrum.*, 20, 255–259, 1986.
9. M. Ramsey, Non-invasive automatic determination of mean arterial pressure, *Med. Biol. Eng. Comput.*, 17, 11–18, 1979.
10. L. A. Geddes, M. Voelz, C. Combs, D. Reiner, and C. F. Babbs, Characterization of the oscillometric method for measuring indirect blood pressure, *Ann. Biomed. Eng.*, 10, 271–280, 1982.
11. M. Ramsey, Blood pressure monitoring: Automated oscillometric devices, *J. Clin. Monit.*, 7, 56–67. 1991.
12. S. D. Pierdomenico, A. Mezzetti, D. Lapenna, M. D. Guglielmi, L. Mancini, L. Salvatore, T. Antidormi, F. Costantini, and F. Cuccurullo, "White-coat" hypertension in patients with newly diagnosed hypertension: Evaluation of prevalence by ambulatory monitoring and impact on cost of health care, *Eur. Heart J.*, 16, 692–697, 1995.
13. P. R. Wilkinson and E. B. Raftery, Patients' attitudes to measuring their own blood pressure, *Br. Med. J.*, 1, 824, 1978.
14. T. G. Pickering, Utility of 24 h ambulatory blood pressure monitoring in clinical practice, *Can. J. Cardiol.*, 11 (Suppl H), 43H–48H, 1995.
15. P. E. Nielsen and J. Badskjaer, Assessment of blood pressure in hypertensive subjects using home readings, *Dan. Med. Bull.*, 28, 197–200, 1981.
16. S. Mann, R. I. Jones, M. W. Millar-Craig, C. Wood, B. A. Gould, and E. B. Raftery, The safety of ambulatory intra-arterial pressure monitoring: A clinical audit of 1000 studies, *Int. J. Cardiol.*, 5, 585–597, 1984.
17. J. M. Grin, E. J. McCabe, and W. B. White, Management of hypertension after ambulatory blood pressure monitoring, *Ann. Intern. Med.*, 118, 833–837, 1993.
18. M. Bass, Ambulatory blood pressure monitoring and the primary care physician, *Clin. Invest. Med.*, 14, 256–259, 1991.
19. W. B. White, P. Lund-Johansen, and P. Omvik, Assessment of four ambulatory blood pressure monitors and measurements by clinicians versus intraarterial blood pressure at rest and during exercise, *Am. J. Cardiol.*, 65, 60–66, 1989.
20. J. A. Staessen, R. Fagard, L. Thijs, and A. Amery, and participants in the fourth international consensus conference on 24-hour ambulatory blood pressure monitoring, *Hypertension*, 26(1), 912–918, 1995.
21. H. Alexander, M. L. Cohen, and L. Steinfeld, Criteria in the choice of an occluding cuff for the indirect measurement of blood pressure, *Med. Biol. Eng. Comput.*, 15, 2–10, 1977.
22. The fifth report of the Joint National Committee on Detection, evaluation, and treatment of high blood pressure (JNCW), *Arch. Intern. Med.*, 153, 154–183, 1993.

23. Task Force on Blood Pressure Control in Children, Report of the second task force on blood pressure control in children—1987, *Pediatrics*, 79, 1–25, 1987.
24. Anonymous, *American National Standard for Electronic or Automated Sphygmomanometers ANSI/AAMI SP10—1987*, Arlington, VA: Association for the Advancement of Medical Instrumentation, 1987.
25. Anonymous, *American National Standard for Electronic or Automated Sphygmomanometers ANSI/AAMI SP10—1992*, Arlington, VA: Association for the Advancement of Medical Instrumentation, 1992.
26. Anonymous, *Amendment to ANSI/AAMI SP10—1992: American National Standard for Electronic or Automated Sphygmomanometers, ANSI/AAMI SP10A—1996*, Arlington, VA: Association for the Advancement of Medical Instrumentation, 1996.
27. Anonymous, *American National Standard for Non-Automated Sphygmomanometers ANSI/AAMI SP9—1986*, Arlington, VA: Association for the Advancement of Medical Instrumentation, 1986.
28. E. O'Brien, J. Petrie, W. Littler, M. Sweit, P. L. Padfield, K. O'Malley, M. Jamieson, D. Altman, M. Bland, and N. Atkins, The British Hypertension Society protocol for the evaluation of automated and semi-automated blood pressure measuring devices with special reference to ambulatory systems, *J. Hypertens.*, 8, 607–619, 1990.
29. E. O'Brien, J. Petrie, W. Littler et al., The British Hypertension Society protocol for the evaluation of blood pressure measuring devices, *J. Hypertens.*, 11 (Suppl. 2), S43–S62, 1993.
30. G. Mancia and G. Parati, Commentary on the revised British Hypertension Society protocol for the evaluation of blood pressure measuring devices: A critique of aspects related to 24-hour ambulatory blood pressure measurement [commentary], *J. Hypertens.*, 11, 595–597, 1993.
31. J. Conway, Home blood pressure recording, *Clin. Exp. Hypertens.*, 8, 1247–1274, 1986.
32. Anonymous, *Medical Device Register*, Montvale, NJ: Medical Economics Company, 1, III-997, 1995.
33. J. Egmond, J. Lenders, E. Weernink, and T. Thien. Accuracy and reproducibility of 30 devices for self-measurement of arterial blood pressure, *Am. J. Hypertens.*, 6, 873–879, 1993.
34. Y. Imai, S. Sasaki, N. Minami et al., The accuracy and performance of the A&D TM 2421, a new ambulatory blood pressure monitoring device based on the cuff-oscillometric method and the Korotkoff sound technique, *Am. J. Hypertens.*, 5, 719–726, 1992.
35. E. O'Brien, F. Mee, N. Atkins, and K. O'Malley, Accuracy of the Takeda TM-2420/TM-2020 determined by the British Hypertension Society Protocol, *J. Hypertens.*, 9, 571–572, 1991.
36. P. Palatini, M. Penzo, C. Canali, and C. Pessina, Validation of the A&D TM-2420 Model 7 for ambulatory blood pressure monitoring and effect of microphone replacement on its performance, *J. Ambulatory Monit.*, 4, 281–288, 1991.
37. W. White, P. Lund-Johansen, and J. McCabe, Clinical evaluation of the Colin ABPM 630 at rest and during exercise: An ambulatory blood pressure monitor with gas-powered cuff inflation, *J. Hypertens.*, 7, 477–483, 1989.
38. E. O'Brien, F. Mee, N. Atkins, and K. O'Malley, Accuracy of the Del Mar Avionics pressurometer IV determined by the British Hypertension Society protocol, *J. Hypertens.*, 9, 567–568, 1991.
39. S. Santucci, E. Cates, G. James, Y. Schussel, D. Steiner, and T. Pickering, A comparison of two ambulatory blood pressure monitors, the Del Mar Avionics pressurometer IV and the SpaceLabs 90202, *Am. J. Hypertens.*, 2, 797–799, 1989.
40. E. O'Brien, F. Mee, N. Atkins, and K. O'Malley, Short report: Accuracy of the CH-Druck/Pressure Scan ERKA ambulatory blood pressure measuring system determined by the British Hypertension Society protocol, *J. Hypertens.*, 10, 1283–1284, 1992.
41. E. O'Brien, F. Mee, N. Atkins, and K. O'Malley, Accuracy of the Novacor DIASYS 200 determined by the British Hypertension Society protocol, *J. Hypertens.*, 9, 569–570, 1991.
42. G. Manning, S. Vijan, and M. Millar-Craig, Technical and clinical evaluation of the Medilog ABP non-invasive blood pressure monitor, *J. Ambulatory. Monit.*, 7, 255–264, 1994.
43. E. O'Brien, F. Mee, N. Atkins, and K. O'Malley, Accuracy of the Profilomat determined by the British Hypertension Society protocol [short report], *J. Hypertens.*, 10, 1285–1286, 1992.

44. E. O'Brien, F. Mee, N. Atkins, and K. O'Malley, Accuracy of the SpaceLabs 90207 determined by the British Hypertension Society protocol [short report], *J. Hypertens*., 9, 573–574, 1991.
45. R. Taylor, K. Chidley, J. Goodwin et al., Accutracker II (version 30/23) ambulatory blood pressure monitor: Clinical validation using the British Hypertension Society and Association for the Advancement of Medical Instrumentation standards, *J. Hypertens*., 11, 1275–1282, 1993.
46. W. White, W. Susser, G. James et al. Multicenter assessment of the QuietTrak ambulatory blood pressure recorder according to the 1992 AAMI guidelines, *Am. J. Hypertens*., 7, 509–514, 1994.
47. J. Ling, Y. Ohara, Y. Orime et al., Clinical evaluation of the oscillometric blood pressure monitor in adults and children based on the 1992 AAMI SP-10 standards, *J. Clin. Monit*., 11, 123–130,1995.
48. M. Gorback, T. Quill, and M. Lavine, The relative accuracies of two automated noninvasive arterial pressure measurement devices, *J. Clin. Monit*., 7, 13–22, 1991.
49. E. O'Brien, F. Mee, N. Atkins, and K. O'Malley, Short report: Accuracy of the Dinamap portable monitor, model 8100 determined by the British Hypertension Society Protocol, *J. Hypertens*., 11, 761–763, 1993.
50. M. Sun, J. Tien, R. Jones, and R. Ward, A new approach to reproducibility assessment: Clinical evaluation of SpaceLabs Medical oscillometric blood pressure monitor, *Biomed. Instrum. Technol*., 30, 439–448, 1996.
51. Y. Imai, J. Hashimoto, N. Minami et al., Accuracy and performance of the Terumo ES-H51, a new portable blood pressure monitor, *Am. J. Hypertens*., 7, 255–260, 1994.
52. Y. Imai, K. Abe, S. Sasaki et al., Clinical evaluation of semiautomatic and automatic devices for home blood pressure measurement: Comparison between cuff-oscillometric and microphone methods, *J. Hypertens*., 7, 983–990, 1989.
53. P. H. Fabre, Determination de la pression arterielle maxima par la methode oscillometrique, *C. R. Soc. Biol*. (Paris), 951–952, 1922.
54. H. Benson and J. A. Herd, Oscillometric measurement of arterial blood pressure, *Circulation*, Suppl. 3, 39–40, 1969.
55. M. Gorback, T. Quill, and D. Graubert, The accuracy of rapid oscillometric blood pressure determination, *Biomed. Instrum. Technol*., 24, 371–374, 1990.
56. C. Nelson, T. Dorsett, and C. Davis, *Method for Noninvasive Blood Pressure Measurement by Evaluation of Waveform-Specific Area Data*, U.S. Patent 4,889,133, December 26, 1989.
57. S. Narus, T. Egbert, T. K. Lee, J. Lu, and D. Westenskow, Noninvasive blood pressure monitoring from the supraorbital artery using an arterial neural network oscillometric algorithm, *J. Clin. Monit*., 11, 289–297, 1995.
58. J. Penaz, Photoelectric measurement of blood pressure, volume and flow in the finger, *Digest of 10th International Conference on Medical and Biological Engineering*, p. 104, Dresden, Germany, 1973.
59. K. H. Wesseling, B. de Wit, J. J. Settels, and W. H. Klawer, On the indirect registration of finger blood pressure after Penaz, *Funkt. Biol. Med*., 1, 245–250, 1982.
60. K. Yamakoshi and A. Kamiya, Noninvasive automatic monitoring of instantaneous arterial blood pressure using the vascular unloading technique, *Med. Biol. Eng. Comput*., 21, 557–565, 1983.
61. J. R. de Jong, R. Tepaske, G. J. Scheffer, H. H. Ros, P. P. Sipkema, and J. J. de Lange, Noninvasive continuous blood pressure measurement: A clinical evaluation of the Cortronic APM 770, *J. Clin. Monit*., 1, 18–24, 1993.
62. C. Young, J. Mark, W White et al., Clinical evaluation of continuous noninvasive blood pressure monitoring: Accuracy and tracking capabilities, *J. Clin. Monit*., 11, 245–252, 1995.
63. O. Kemmotsu, M. Ueda, H. Otsuka, T. Yamamura, D. C. Winter, and J. S. Eckerle, Arterial tonometry for noninvasive, continuous blood pressure monitoring during anesthesia, *Anesthesiology*, 75, 333–340, 1991.

64. O. Kemmotsu, M. Ueda, H. Otsuka, T. Yamamura, A. Okamura, T. Ishikawa, D. C. Winter, and J. S. Eckerle, Blood pressure measurement by arterial tonometry in controlled hypertension, *Anesth. Analg*., 73, 54–58, 1991.
65. N. R. Searle, J. Perrault, H. Ste-Marie, and C. Dupont, Assessment of the arterial tonometer (N-CAT) for the continuous blood pressure measurement in rapid arterial fibrillation, *Can. J. Anaesth*., 40, 388–393, 1993.
66. S. V. S. Rithalia, Measurement of arterial blood pressure and pulse, *J. Tissue Viability*, 4, 44–47, 1994.
67. J. E. Dalen and R. C. Bone, Is it time to pull the pulmonary artery catheter? *J. Am. Med. Assoc*., 276, 916–918, 1996.
68. E. V. Allen, Methods of diagnosis of chronic occlusive arterial lesions distal to the wrist with illustrative cases, *Am. J. Med. Sci*., 178, 237–244, 1929.
69. F. M. Ducharme, M. Gauthier, J. Lacroix, and L. Lafleur, Incidence of infection related to arterial catheterization in children: A prospective study, *Crit. Care Med*., 16, 272–276, 1988.
70. S. S. Moorthy, R. K. Stoelting, and R. D. King, Delayed cyclic variations (oscillations) in pressure in a critically ill patient, *Crit. Care Med*., 11, 476–477, 1983.
71. S. V. S. Rithalia, Non-invasive measurement of blood gases in critically III adults, PhD thesis, University of London, London, U.K., 1982.
72. S. Cunningham and N. McIntosh, Blood pressure monitoring in intensive care neonates, *Br. J. Intens. Care*, 2, 381–388, 1992.
73. R. M. Gardner, Equivalence of fast flush and square wave testing of blood pressure monitoring systems, Direct blood pressure measurement-dynamic response requirements, *Anesthesiology*, 54, 227–236, 1981.
74. B. Kleinman, S. Powell, P. Kumar, and R. M. Gardner, The fast flush does measure the dynamic response of the entire blood pressure monitoring system, *Anesthesiology*, 77, 1215–1220, 1992.
75. R. M. Gardner and K. W. Hollingsworth, Optimizing ECG and pressure monitoring, *Crit. Care Med*., 14, 651–658, 1986.
76. J. N. Cohn, Blood pressure measurement in shock: Mechanism of inaccuracy in auscultatory and palpatory methods, *J Am. Med. Assoc*., 199, 118–122, 1967.
77. M. Sun and R. Jones, A hydrostatic method assessing accuracy and reliability while revealing asymmetry in blood pressure measurements, *Biomed. Instrum. Technol*., 29, 331–342, 1995.
78. F. H. Messerli, H. O. Ventura, and C. Amodeo, Osier's maneuver and pseudohypertension, *N. Engl. J. Med*., 312, 1548–1551, 1985.
79. F. H. Messerli, The age factor in hypertension, *Hosp. Prac*., 15, 103–112, 1986.
80. J. Kugler, N. Schmitz, H. Seelbach, J. Rollnik, and G. Kruskemper, Rise in systolic pressure during sphygmomanometry depends on the maximum inflation pressure of the arm cuff, *J. Hypertens*., 12, 825–829, 1994.
81. A. R. Nara, M. P. Burns, and W. G. Downs, *Biophysical Measurement Series: Blood Pressure*, Redmond, WA: SpaceLabs Medical Inc., 1993.
82. J. A. Blumenthal, W. C. Siegel, and M. Appelbaum, Failure of exercise to reduce blood pressure in patients with mild hypertension. *J Am. Med. Assoc*., 266, 2098–2104, 1991.
83. N. M. Kaplan, Misdiagnosis of systemic hypertension and recommendations for improvement, *Am. J. Cardiol*., 60, 1383–1385, 1987.
84. D.W. McKay, N. R. C. Campbell, A. Chockalingam, L. Ku, C. Small, and F. Washi, Self-measurement of blood pressure: Assessment of equipment. *Can. J. Cardiol*., 11, 29H–34H, 1995.
85. M. A. Brown, M. L. Buddle, M. Bennett, B. Smith, R. Morris, and J. A. Whitworth, Ambulatory blood pressure in pregnancy: Comparison of the Spacelabs 90207 and Accutracker II monitors with intra-arterial recordings, *Am. J. Obstet. Gynecol*., 173(1), 218–223, 1995.
86. M. A. Brown, L. Reiter, B. Smith, M. L. Buddle, R. Morris, and J. A. Whitworth, Measuring blood pressure in pregnant women; a comparison of direct and indirect methods, *Am. J. Obstet. Gynecol*., 171(3), 661–667, 1994.

87. A. Hoeft, B. Schorn, A. Weyland et al., Beside assessment of intravascular volume status in patient undergoing coronary bypass surgery, *Anesthesiology*, 81, 76–86, 1994.
88. A. B. Newman, K. Sutton-Tyrrell, and L. H. Kuller, Lower extremity arterial disease in older hypertensive adults, *Arterioscler. Throm.*, 13(4), 555–562, 1993.
89. T. K. Lee, T. P. Egbert, and D. R. Westenskow, Supraorbital artery as an alternative site for oscillometric blood pressure measurement, *J. Clin. Monit.*, 12(4), 293–297, 1996.
90. B. P. Imholz, G. J. Langenwouters, G. A. van Montrans, G. Parati, J. van Goudoever, K. H. Wesseling, W. Wieling, and G. Mancia, Feasibility of ambulatory, continuous 24-hour finger arterial pressure recording, *Hypertension*, 21(1), 65–73, 1993.

66

Blood Flow Measurements

Per Ask
Linkoping University

P. Åke Öberg
Linkoping University Hospital

66.1 Doppler Measurements

66.1.1 Ultrasound Doppler

The ultrasound Doppler measurements [1,2] are based on a principle discovered by the Austrian physicist Christian Doppler in 1842 [3], who theoretically predicted that a wave backscattered from a moving object will be shifted in frequency. The principle was verified by two hornblowers, one aboard a moving train and the other standing still, with the ability to assess the pitch of the sound.

To obtain a Doppler signal from a fluid, the fluid must contain scattering particles, which in the case of blood are the blood cells. The size of a red blood cell is about 2 × 7 μm, which means that the scatterers are much smaller than the wavelength of the ultrasound. Hence, a diffuse scattering of the ultrasound will occur (Rayleigh scattering). The scattering from tissues surrounding the heart and vessels usually gives a much larger signal (20–40 dB) than that from blood in motion. The velocity of tissue motion is usually much lower than that of blood. This contribution can therefore be suppressed by high-pass filtering. In recent years, ultrasound contrast agents (consisting of gas-filled shells) have been introduced to increase the blood flow signal. Figure 66.1 illustrates the ultrasound Doppler principle. An ultrasound beam is sent toward a moving object. The beam hits the object and returns to the receiver with a Doppler-shifted frequency carrying information about the velocity of the object.

The Doppler shift f_d of an ultrasound signal with the nominal frequency f_c is given by

$$f_d = 2f_c \cdot \frac{v}{c} \tag{66.1}$$

where

v is the velocity component in the direction of the ultrasound beam

c is the speed of sound in the medium that is in the range of 1500–1600 m s^{-1} in soft tissue and usually set to 1540 m s^{-1}

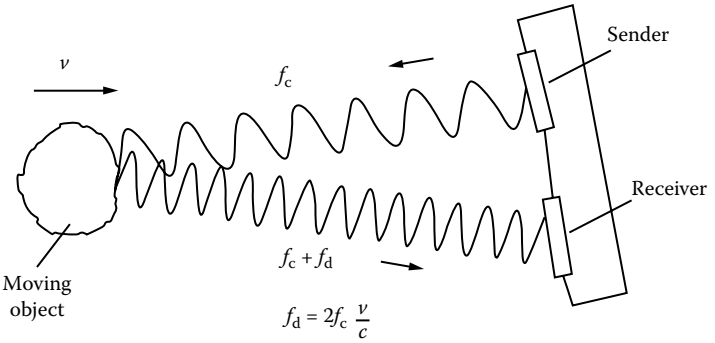


FIGURE 66.1 For Doppler ultrasound, the moving object shifts the received frequency.

The frequency f_c is in the range of 2–10 MHz, which gives a wavelength between 0.15 and 0.77 mm.

The peak systolic flow velocity in the heart and larger vessels is normally 0.5–1 m s^{-1}, resulting in a Doppler shift in the range of 1.3–13 kHz (depending on the ultrasound frequency f_c). The Doppler sound is therefore audible, which is helpful for the investigator in identifying vessels and phenomena of interest.

The most straightforward ultrasound investigation is to use a continuous wave system. The ultrasound beam is focused by suitable ultrasound transducer geometry and by a lens. By this arrangement, a narrow beam can be arranged and the backscattered information can come from any section along the beam.

In order to know from where along the beam the blood flow data are collected, a pulsed Doppler system has to be used. The transducer sends four to eight cycles of the ultrasound signal; at a specified time later a gate is opened and the transducer will act as a receiver. By using a preset time delay between the sending and the receiving signal, only flow information from a certain depth will be collected. The depth is obviously determined by the delay time and the propagation velocity in the tissue. Flow velocity information is obtained by spectral estimation of the Doppler signal.

By multirange gating, blood flow at various points along the ultrasound beam can be measured. By scanning with an ultrasound beam within a sector, a 2D velocity field can be presented. By color coding the information, a color Doppler flow image can be obtained. By scanning in one additional orthogonal plane, a 3D flow image can be created.

An important medical application of the ultrasound Doppler is in the study of heart valve flow when one might suspect stenosis or leaking flow in the valves. The pumping ability of the heart can be assessed from the general flow patterns. Regions with arteriosclerotic obstructions can be localized in the peripheral vessels.

66.1.2 Laser Doppler Flowmetry

The Doppler principle is also utilized in blood flow measurements in the microcirculatory bed in which laser Doppler flowmetry measures the blood perfusion. Laser light from a gas or semiconductor laser is launched into an optical fiber that leads the light to the tissue (Figure 66.2). Photons are reflected, scattered, and absorbed in the tissue matrix, and those that hit moving red blood cells become Doppler shifted, whereas those that are reflected in stationary structures are refracted without any change in frequency. A part of the photons returning to the fiber-optic system will be conducted by the pickup fiber to the detector. Shifted and nonshifted photons are mixed at the surface of a square-law photodetector. According to elementary wave mechanics, this type of mixing (coherent detection) results in the sum and difference frequencies. The frequency of the "difference wave" is proportional to the average velocity of the red blood cells. The amplitude of the same signal is proportional to the number of moving scatterers in the tissue volume. A velocity distribution will result in a Doppler spectrum, usually in the range of 30–12,000 Hz.

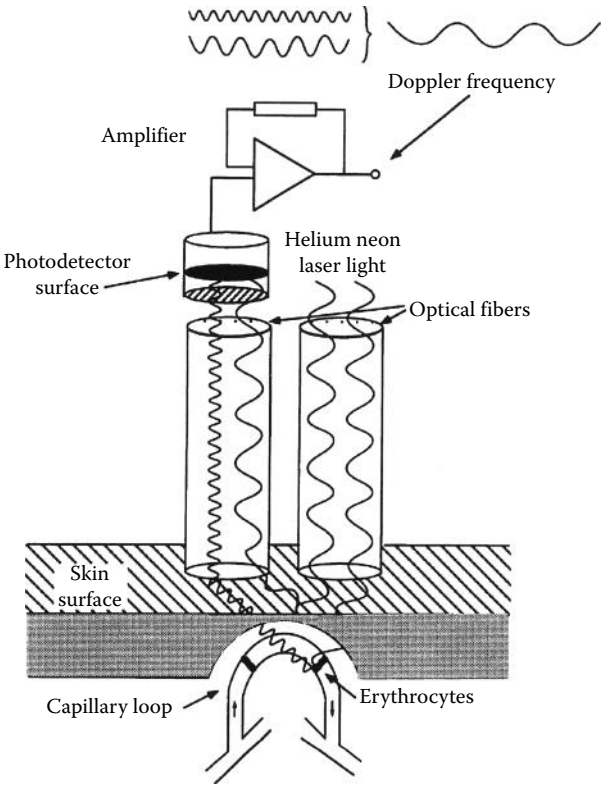


FIGURE 66.2 For laser Doppler flowmetry, the moving blood cells shift the received frequency.

Variants of the method utilizing fiber optics, airborne beams, microscope-based setups, and color-coded imaging scanners are described in the literature. For a review, see Shepherd and Öberg [4].

66.2 Indicator Dilution Methods

Cardiac output measurement is one of the most essential heart performance measures. The rest and exercise flows carry important diagnostic information. The monitoring of cardiac output is very important for the critically ill patient.

The principle is that an indicator is injected upstream in the circulation. Mixing with the circulating blood volume occurs and the indicator concentration is detected downstream. By knowing the added quantity of indicator and the time integral of the detected indicator concentration, the flow can be determined (Figure 66.3).

We can subdivide the indicator dilution methods into dye dilution and thermal dilution methods. See Webster [5] and Bronzino [6].

66.2.1 Dye Dilution Method

To determine the cardiac output (1/min), a known quantity (mass m) of a dye indicator such as Evans Blue is injected into the right heart, and a concentration of the indicator $c(t)$ is detected in the pulmonary artery. At the time t an indicator quantity Δm passes the detector at a time interval Δt. If F is the blood flow,

$$\Delta m = F\, c(t)\Delta t \tag{66.2}$$

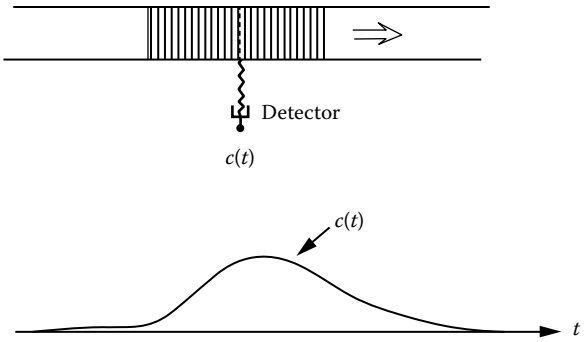


FIGURE 66.3 Indicator dilution measurements. The flow can be calculated if the amount of indicator and the time course of its concentration are known.

or by integration

$$F = \frac{m}{\int c(t)dt} \tag{66.3}$$

Thus, only the amount of indicator added and the time integral of the downstream concentration, assuming good mixing, have to be known to be able to calculate the flow, that is, the cardiac output.

66.2.2 Thermal Dilution Method

The thermal dilution method is a variant of the indicator dilution method family. A thermal dilution catheter is placed with the injection outlet in the right atrium of the heart and with a temperature sensor in the pulmonary artery (Bronzino [6] and Weissel et al. [7]).

A chilled solution of dextrose in water or saline solution is used for the injection and causes transient decrease in the pulmonary artery temperature $T(t)$. The blood flow (cardiac output) can be calculated from

$$F = \int T(t)dt\rho_b c_b = V(T_b - T_i)\rho_i c_i K \tag{66.4}$$

or

$$F = \frac{V(T_b - T_i)}{\int T(t)dt}\left(\frac{\rho_i c_i}{\rho_b c_b}\right)K \tag{66.5}$$

where

V is the injected volume
T_b is the undisturbed temperature in the pulmonary artery
T_i is the temperature of the indicator
ρ_i and ρ_b are the density of the indicator and the blood
c_i and c_b are the specific heat of the indicator and the blood
K is a correction factor that takes heat transfer along the catheter into account

The thermodilution technique is the standard technique for the monitoring of cardiac output in critically ill patients.

66.3 Plethysmography

The word *plethysmography* means methods for recording volume changes of an organ or a body part. Depending on the technique used, strain gage, impedance, and optical techniques can be used for the volume determination. See Webster [5] and Figure 66.4.

66.3.1 Strain Gage Plethysmography

The classical strain gage plethysmography is used to study circulation in the lower extremities from changes in the circumference of the legs.

Small-diameter silicone rubber tubes, filled with mercury or other types of conductive liquids, are placed around the circumference of the leg. Changes in the latter can be directly related to electric impedance changes of the silicone rubber tube. If a cylindrical cross section of the leg is assumed, volume changes should be proportional to the total circumference, times the change in impedance.

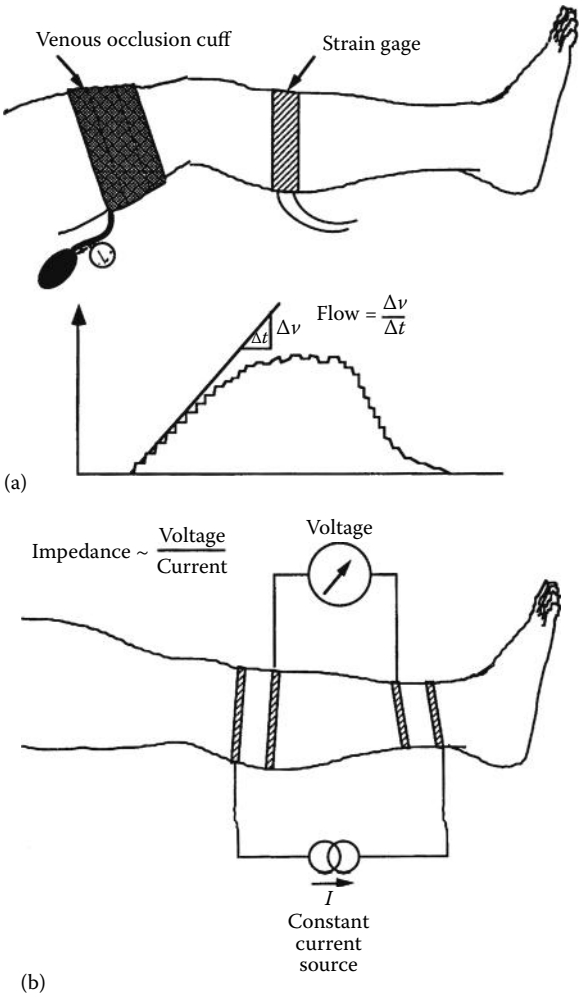


FIGURE 66.4 For occlusion plethysmography, increased volume stretches the strain gage (a) and decreases the impedance (b).

The strain gage tubes (Figure 66.4) are positioned around the lower part of the leg; an inflatable cuff is placed around the upper part (above the knee) and inflated to 40–50 mmHg, that is, above the venous pressure that will cause the outflow of blood to cease. The increased volume of the leg is therefore proportional to the arterial inflow. The latter is determined from the initial slope. Flow is volume increase per time unit.

When the cuff pressure is released, blood will flow out from the leg via the veins. The time course of the volume change will be related to venous function. If there is venous thrombosis, the decline in the volume curve will be slower.

66.3.2 Impedance Plethysmography

Bioelectric impedance measurements have a history dating back to the 1940s. The reason why impedance is useful for detection of volume changes is that different tissues in the body have different resistivity. Blood is one of the best conductors among the tissues of the body.

Impedance plethysmography has its most established applications in respiratory monitoring in newborn infants and for detection of venous thrombosis. Less-established applications are cardiac output measurements, peripheral blood flow studies, and body composition assessments.

A constant current with a frequency of 50–100 kHz and amplitude of 0.5–4 mA rms is applied via skin electrodes (Figure 66.4). Influences from skin impedance are eliminated by the use of the four-electrode technique.

The measurement object can be described with a conduction object with the constant impedance Z_0 in parallel with a time-varying impedance ΔZ. The impedance ΔZ is represented with a column of a conducting media with resistivity ρ and the length L. If the changes in ΔZ are small in comparison to those of Z_0, the volume changes can be obtained as

$$\Delta V \approx \rho \frac{L^2}{Z_0^2} \Delta Z \tag{66.6}$$

This technique can be used in the same way as the strain gage plethysmographic method to study circulation in the leg.

For the measurement of cardiac output, Equation 66.6 or a modified formula can be used [8]. The method seems to work rather well in normal persons for relative change, but for patients with cardiac disease, the cardiac output estimation might be poor.

66.3.3 Photoelectric Plethysmography

Hertzman and Spealman [9] and Hertzman [10] were the first to use the descriptive term *photoelectric plethysmography* (PPG). The first reports on the successful use of the principle were published in the middle of 1930 by Molitor and Kniazuk [11]. The principle on which PPG is based is simple, although the underlying detailed optical mechanisms remain unknown. A beam of light is directed toward the part of the tissue in which blood flow (or volume) is going to be measured (Figure 66.5). Reflected, transmitted, and scattered light leaving this volume is collected and focused on a photodetector. A signal modulated by the attenuation or scattering of light in the blood volume can be recorded. Two different components can be derived from the detector. One is pulsatile and synchronous with the heartbeat (the ac component), and the other is a constant voltage (the dc component). The physiological significance of the two signals is still under debate, but they reflected the blood volume and the orientation of erythrocytes during the cardiac cycle.

PPG has been used mainly for monitoring blood perfusion in skin, venous reflux conditions, and skin flaps during plastic surgery. Challoner [12], Roberts [13], and Bernstein [14] have reviewed the methodology and applications of PPG.

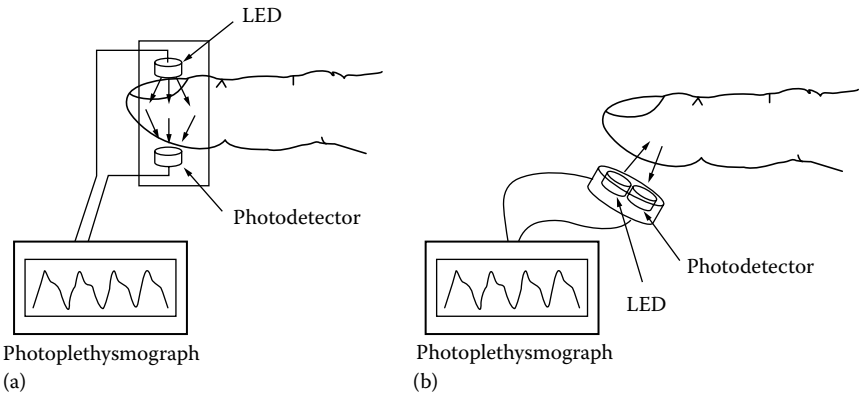


FIGURE 66.5 For photoplethysmography, increased blood decreases received light in (a) transmission mode and (b) reflection mode.

66.4 Radioisotopes

Kety [15] introduced the principle of tissue clearance of rapidly diffusing inert isotopes for blood flow measurements. An extensive theoretical treatment is given by Zierler [16]. In most applications, lipid-soluble gases like ^{133}Xe and ^{85}Kr have been used. These isotopes rapidly diffuse from blood to tissue and a rapid equilibration of the isotope concentration within a tissue volume takes place. Figure 66.6 illustrates the measurement principle. Their elimination from a microcirculatory bed is related to the blood flow rate. If the tissue is uniformly and constantly perfused, the activity of the isotope decays monoexponentially with time. The elimination of the isotope can be described by the equation

$$C(t) = C_0 \exp(-kt) \tag{66.7}$$

where

$C(t)$ and C_0 are the tissue concentrations at times t and at the onset of the injection C_0

k is the clearance constant related to the local blood flow by the relation

$$k = \frac{\ln 2}{t_{1/2}} \tag{66.8}$$

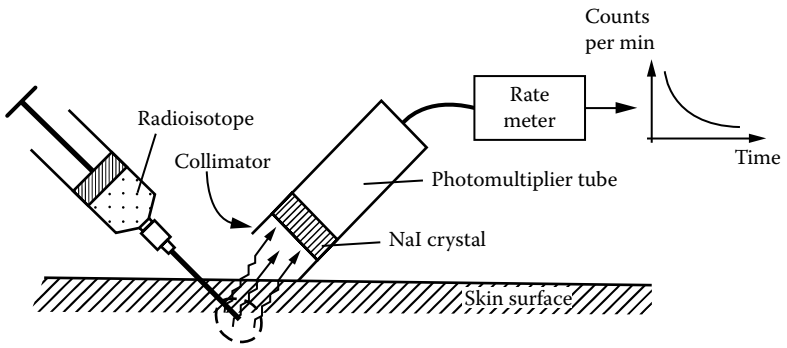


FIGURE 66.6 For isotope measurement of microcirculatory blood flow, the radioactivity decreases with washout.

in which $t_{1/2}$ is half the time of decay. Blood flow Q (in $k\ s\ 100$ mL $min^{-1} \cdot 100\ g^{-1}$) can then be derived from the formula

$$Q = k\ s\ 100 \tag{66.9}$$

where s denotes the tissue–blood partition coefficient. The indicator is administered via an injection into the tissue volume or through passive diffusion after deposition on the surface of the tissue volume under study.

The advantage of clearance methods is that they can be applied to the study of all kinds of tissue blood flow problems. One of the disadvantages is that the method does not give a continuous measurement of flow. In addition, the clearance curves are sometimes difficult to interpret. The trauma caused by injection of the isotope into the tissue seriously disturbs the flow, as shown by Holloway [17] and Sejrsen [18]. In spite of these shortcomings, the isotope clearance method has been applied extensively to the study of skin and tissue blood flow in experimental as well as clinical problem areas.

66.5 Thermal Convection Probes

Thermal convection probes have been developed specifically for skin blood flow measurements. Gibbs [19] pioneered the field by describing a probe in the form of a needle. Hensel and Bender [20] and van de Staak et al. [21] developed noninvasive variants by designing probes that can be positioned at the surface of the tissue.

All methods measure the rate of removal of heat from the tissue volume under the probe. A relation exists between the blood flow rate and the rate by which heat dissipates from the tissue volume under study. The sensing unit is usually designed around a central metal disk and a concentric outer ring between which a temperature difference is established (Figure 66.7). The two rings are thermally and electrically isolated from each other and both are in contact with the tissue. The temperature difference between the two rings is a measure of the blood flow under the probe.

A temperature difference of 2 °C–3 °C is usually established between the inner disk and the outer annulus. The central disk is heated with an electric current and kept at a constant temperature that only by 1°C or 2°C exceeds the resting temperature of the tissue under study. Thermal probes have not been extensively used because of their extreme nonlinear properties and the difficulties in their practical use, that is, the contact pressure sensitivity. Another difficulty is the highly variable thermal characteristics of the skin.

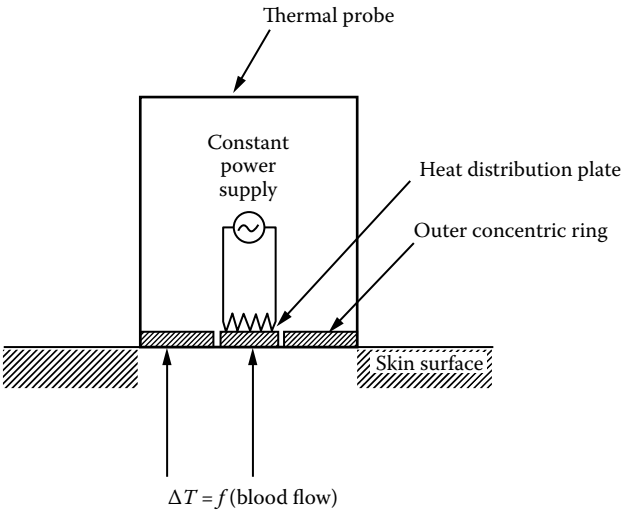


FIGURE 66.7 For thermal convection measurement of blood flow, heat dissipation increases with blood flow.

References

1. L. Hatle and B. Angelsen, Eds., *Doppler Ultrasound in Cardiology. Physical Principles and Clinical Applications*, Lea & Febiger, Philadelphia, PA, 1985.
2. S. Webb, Ed., *The Physics of Medical Imaging*, Adam Hilger, Philadelphia, PA, 1988.
3. C. Doppler, quoted by D. N. White, Johann Christian Doppler and his effects—A brief history, *Ultrasound Med. Biol.*, 8, 583–591, 1982.
4. A. P. Shepherd and P. Å. Öberg, Eds., *Laser Doppler Blood Flowmetry*, Kluwer Academic Publishers, Boston, MA, 1990.
5. J. G. Webster, Ed., *Medical Instrumentation. Application and Design*, 4th edn., John Wiley & Sons, New York, 2010, pp. 332–338.
6. J. D. Bronzino, Ed., *The Biomedical Engineering Handbook*, CRC Press, Boca Raton, FL, 1995, pp. 1212–1216.
7. R. D. Weissel, R. L. Berger, and H. B. Hechtman, The measurement of cardiac output by thermodilution, *N. Engl. J. Med.*, 292, 682, 1975.
8. R. P. Patterson, Fundamentals of impedance cardiology, *IEEE Eng. Med. Biol. Mag.*, 8(1), 35, 1989.
9. A. B. Hertzman and C. R. Spealman, Observations on the finger volume pulse recorded photoelectrically, *Am. J. Physiol.*, 119, 334, 1937.
10. A. B. Hertzman, Photoelectric plethysmography of the nasal septum in man, *Proc. Soc. Exp. Biol. Med.*, 124, 328, 1937.
11. H. Molitor and M. Kniazuk, A new bloodless method for continuous recording of peripheral circulatory changes, *J. Pharmacol. Exp. Ther.*, 57, 6, 1936.
12. A. V. J. Challoner, Photoelectric plethysmography for estimating cutaneous blood flow, in P. Rolfe, Ed., *Non-Invasive Physiological Measurements*, Vol. 1, Academic Press, London, U.K., 1979, p. 127.
13. V. C. Roberts, Photoplethysmography—Fundamental aspects of the optical properties of blood in motion, *Trans. Inst. M.C.*, 4, 101–106, 1982.
14. E. F. Bernstein, Ed., *Non-Invasive Diagnostic Techniques in Vascular Disease*, 2nd edn., C. V. Mosby, St Louis, MI, 1982.
15. S. S. Kety, Measurement of regional circulation by the local clearance of radioactive sodium, *Am. Heart J.*, 38, 321, 1949.
16. K. L. Zierler, Circulation times and the theory of indicator dilution methods for determining blood flow and volume, in Havillan and Dow, W. F. Hamilton (Ed.), *Handbook of Physiology*, Waerly Press, Baltimore, MD, Vol. 1, Section 2: Circulation, Chapter 18, p. 585, 1962.
17. G. A. Holloway, Jr., Laser Doppler measurements of cutaneous blood flow, in P. Rolfe, Ed., *Non-Invasive Physiological Measurements*, Vol. 2, Academic Press, London, U.K., 1983.
18. P. Sejrsen, Measurement of cutaneous blood flow by freely diffusible radioactive isotopes. Methodological studies on the washout of Krypton-85 and Xenon-133 from the cutaneous tissue in man, *Danish Med. Bull.*, 18(Suppl. 3), 9, 1971.
19. F. A. Gibbs, A thermoelectric blood flow recorder in the form of a needle, *Proc. Soc. Exp. Biol.*, New York, 31, 141, 1933.
20. H. Hensel and F. Bender, Fortlaufende Bestimmung der Hautdurchblutung am Menschen mit einem elektrischen Wärmeleitmesser, *Pflugers Arch.*, 263, 603, 1956.
21. W. J. B. M van de Staak, A. J. M. Brakkee, and H. E. de Rijke-Herweijer, Measurement of the thermal conductivity of the skin as an indication of skin blood flow, *J. Invest. Dermatol.*, 51, 149–154, 1968.

67

Medical Imaging

James T. Dobbins III
Duke University Medical Center

Sean M. Hames
Duke University Medical Center

Bruce H. Hasegawa
University of California, San Francisco

Timothy R. DeGrado
Duke University Medical Center

James A. Zagzebski
University of Wisconsin, Madison

Richard Frayne
University of Wisconsin, Madison

67.1 Introduction

Medical imaging has advanced considerably since the discovery of x-rays by Wilhelm Conrad Röntgen in 1895. Today, in addition to the continued use of x-rays for medical diagnosis, there are imaging methods that use sound (ultrasound), magnetic fields and radio waves (magnetic resonance imaging [MRI]), and radionuclides (nuclear medicine). Both projection imaging and cross-sectional imaging are routinely used clinically. This chapter will describe the principles behind the various imaging modalities currently in use and the various measurements routinely made with them.

67.2 Image Information Content

The vast majority of imaging procedures are qualitative in nature, where it is the visual presentation of anatomy that is the measurement outcome. There are also some quantitative measurements, which will be discussed in Section 67.5. However, since most imaging is concerned only with the qualitative nature of the image, a description of the salient features of image content follows.

There are three primary physical parameters of interest in image content: contrast, noise, and resolution. If these three features are known for a given image (or imaging system), then the entire

physical nature of the image has been characterized. There are also psychovisual effects, such as conspicuity [1], which affect the ability of the observer to detect a particular feature, but these issues are difficult to quantitate and are outside the scope of this handbook.

The first of the physical image features, contrast, is defined as the fraction of the total image signal occupied by a particular object:

$$C = \frac{B - S}{B} \tag{67.1}$$

where

S is the signal in the area of interest
B is the background signal

Contrast is determined by the properties of the object being imaged, the imaging modality, the properties of the image detector, postprocessing of the image (such as by digital processing), and the contrast of the display device.

Image noise is a measure of the stochastic nature of the image. All physical measurements, including medical images, contain a certain degree of uncertainty. In x-ray imaging, for example, the physics of x-ray production dictates that the number of x-rays incident on a unit area per unit time is random and given by a statistical distribution known as the Poisson distribution. The greater the image noise, the less likely it is that one will observe a given object. There is a relationship between the image noise, the contrast and area of an object, and its likelihood of being observed. This is summarized in the Rose model:

$$N = \frac{k^2}{C^2 A} \tag{67.2}$$

where N is the number of quanta (such as x-rays) per unit area needed to discern an object of contrast C and area A; assuming a signal to noise of k, Rose found that a signal-to-noise ratio of 5 is typically required to detect a visual object reliably [2].

Resolution is the ability of an imaging system to record faithfully the range of spatial detail in an object. Recording objects with finer spatial detail requires "sharper" imaging detectors. The resolving ability of a detector is largely determined by its point-spread function. The point-spread function describes how well the imaging apparatus can record an infinitesimal point object. No detector is perfectly sharp, and some spread of the infinitesimal dot occurs—the worse the spread, the less resolving the system.

67.2.1 Measurement of Imaging Performance

A linear-systems approach is typically used to quantify the performance of an imaging system. The relations among contrast, noise, and resolution of an imaging system are customarily described by two functions: the modulation transfer function (MTF) and the noise power spectrum (NPS), both of which are functions of spatial frequency. The MTF is the Fourier transform of the point-spread function and describes the inherent deterministic frequency response of the system. The NPS (also referred to as the Wiener spectrum) is proportional to the square of the Fourier transform amplitude at each frequency and represents the variance associated with noise in the system at each particular spatial frequency. The ratio of MTF and NPS, properly normalized, is the noise equivalent quanta (NEQ), which is the square of the maximum available signal to noise at each spatial frequency u:

$$\mathrm{NEQ}(u) \equiv \frac{(\text{Large area signal})^2 \cdot \mathrm{MTF}^2(u)}{\mathrm{NPS}(u)} \tag{67.3}$$

If the NEQ is divided by the number of incident quanta per area (e.g., the number of x-ray photons incident on the detector in x-ray imaging), the result is the detective quantum efficiency (DQE). The zero-frequency DQE is a measure of the fraction of incident quanta effectively used by the system. Alternatively, the DQE may be viewed as the efficiency with which the system utilizes the available signal to noise at each spatial frequency.

The actual measurement of MTF, NPS, and DQE is quite tedious and will be only briefly summarized here. The interested reader is encouraged to consult the suggested references for the appropriate detail on these measurements. Examples of these measurements will be given for x-ray imaging.

The MTF is typically measured by imaging either a very fine slit (typically 10–20 μm) [3–6] or an edge [7]. The profile across the slit image is called the line-spread function (LSF). The Fourier transform of the LSF gives the MTF in the direction perpendicular to the slit. The derivative of values along the edge-response function also gives the LSF. Detector response typically varies with energy so it is important to specify the conditions under which MTF is measured. With x-ray imaging it is typical to use a tube voltage of 70 kV with 0.5 mm Cu filtration placed in the beam to simulate the filtering of the x-ray spectrum expected from a patient, although other measurement techniques are also found in the literature.

The NPS is measured by taking an image of a flat field, where there is no structure in the image other than noise. Contemporary methods of NPS measurement on digital systems perform a 2D Fourier transform on the flat-field image [6], although when measuring the NPS of film, a scanning slit is used to generate a 1D NPS parallel to the direction of slit movement [8–10]. After appropriate scaling, the square of the amplitude of the 2D Fourier transform is the NPS. There are many details related to measuring the NPS properly, including eliminating background trends and the size of the region over which the Fourier transform is taken. These are all covered in detail in Refs. [6,8–18].

Measurement of imaging properties is easier on digital imaging systems than on film, since film must first be digitized at appropriately fine sampling intervals or else corrected for the use of 1D slits [19]. However, the effects of aliasing (fictitious frequency response in a digital system due to limited sampling) make the interpretation of MTF and NPS in digital systems more difficult than with film [11,12,20].

67.3 X-Ray Imaging

X-ray imaging requires an x-ray-generating apparatus (tube, high-voltage supply, and controls) and an appropriate x-ray detector. Typical x-ray detectors include photographic film (almost always used in concert with a fluorescent screen), image intensifiers, computed radiography phosphor plates, and newer dedicated digital detectors.

The x-ray generator is basically a high-voltage step-up transformer with appropriate rectification and control circuitry. Most contemporary generators are three-phase 12-pulse, full-wave rectified to give a very low-voltage ripple (3%–10%) [21]. For procedures requiring very fast pulses of several milliseconds or less (such as coronary angiography), a tetrode-based constant-potential generator is used. The operator selects the tube kilovoltage, tube current, and exposure time appropriate for the examination of interest.

X-ray tubes contain a heated filament (which serves as the cathode) and an anode made of a tungsten/rhenium combination for conventional use or molybdenum for mammography. With the exception of dental tubes, modern clinical x-ray tubes almost always contain a rotating anode to spread the heat out over a larger area, allowing for a greater tube output without damaging the anode. Many tubes contain two filaments, a large one and a small one, depending on tube output and resolution requirements of a particular exam. Measurements on x-ray tubes and generators involve calibrations to assure that kilovoltage, tube output, and exposure time are in good agreement with the control console settings [21]. Calibration of the high voltage is done by commercially available voltage dividers or by specially designed x-ray film cassettes with calibrating filters inside. Tube output is measured by ion chambers, and exposure time is measured either by a rotating-arm timer test tool placed over a film cassette during an x-ray exposure or by direct plotting of the exposure versus time output of an ion chamber.

67.3.1 X-Ray Imaging Detectors

The most common detector for x-ray imaging was film. X-ray film was typically placed in a sandwich between two fluorescent screens (or one screen in mammography for improved visibility of small detail). Contemporary screens were made of rare earth compounds such as Gd_2O_2S and served to convert the x-rays to visible light that exposed the film more efficiently than x-rays alone, thus reducing patient radiation dose. The response of these screen–film combinations had good contrast at intermediate exposure ranges (as given by the film γ, or contrast ratio), but poor contrast at low or high exposures. The contrast and latitude of films are described by the characteristic curve (often referred to as the Hurter–Driffield [HD] curve). Appropriate screen–film combinations were chosen based on the anatomy to be imaged, since screen–film combinations were designed with different contrast, latitude, and exposure sensitivity characteristics [22].

A second type of x-ray detector is the image intensifier, which is used with fluoroscopy. Fluoroscopy uses a low-exposure-rate x-ray output to image a patient continuously, typically to properly position the patient for a subsequent high-exposure film image. The image intensifier (Figure 67.1) comprises a cylindrical glass enclosure, inside of which are an input screen, photocathode, focusing electrodes, accelerating anode, and output screen [21]. The x-rays are absorbed in the input screen (typically CsI), giving off light that liberates electrons from the photocathode. The photoelectrons are then accelerated

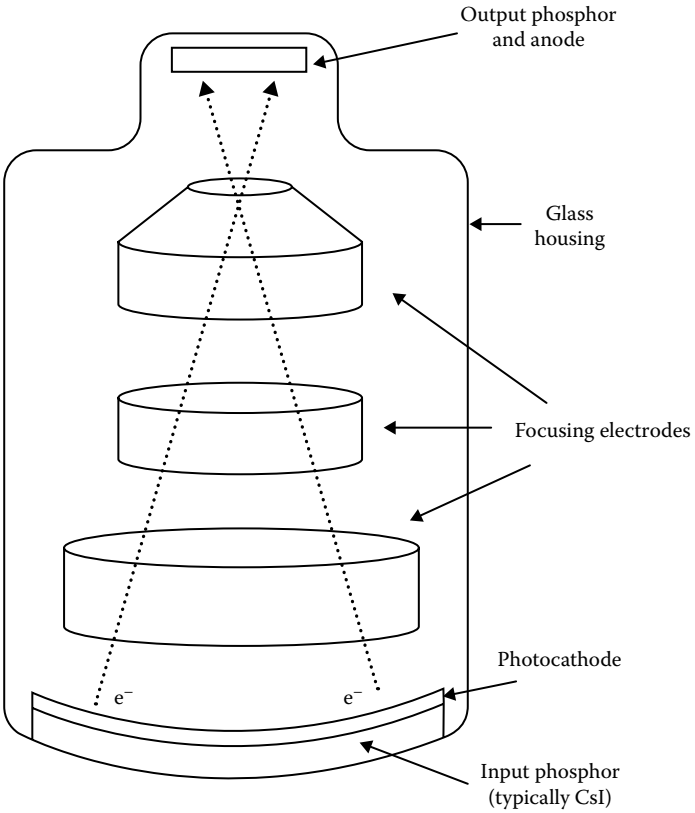


FIGURE 67.1 Schematic diagram of the major components of an image intensifier. The anode is typically at about 30 kV, and the three annular electrodes focus the beam and determine the usable area of the input surface for intensifiers having multiple formats.

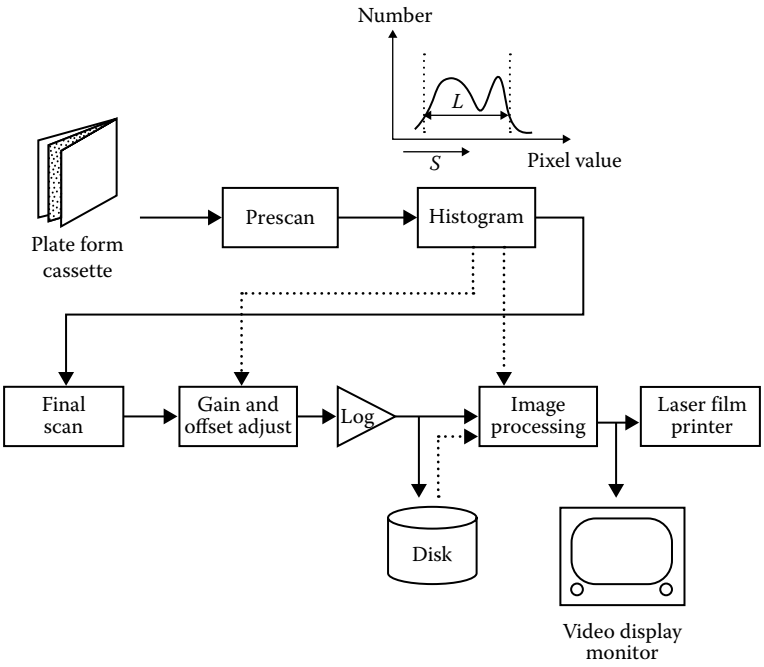


FIGURE 67.2 Schematic diagram of a typical photostimulable phosphor computed radiography system.

to the output screen where they strike the output phosphor screen with high energy (~30 keV), giving off a bright light, which is viewed by either a video camera or motion-picture (cineradiographic) camera.

A recently developed digital x-ray detector is the photostimulable phosphor, which is referred to commonly as computed radiography [23–25]. This detector uses a special type of phosphor that stores about half of the absorbed x-ray energy in metastable states, which are read out later by laser scanning. The laser light stimulates the phosphor to emit ultraviolet light in proportion to the original x-ray exposure. The photostimulated light is then detected with a photomultiplier tube (PMT) or solid-state photodetector and digitized. The clinical apparatus (Figure 67.2) first does a prescan of the imaging plate to adjust the input range of the analog-to-digital converter based on the image histogram; the digitized signal is then logarithmically transformed and stored, displayed on a video monitor, or printed on film following optional spatial filtering and contrast adjustment.

There are also currently available or in development a variety of other digital x-ray detectors, including selenium plate detectors [26], CCD-camera detectors with fluorescent screens [27], and flat-panel arrays with amorphous silicon [28] or amorphous selenium [29] detector elements. Computed radiography cassette scanners scan a storage phosphor screen mounted in a cassette of the same dimensions as a conventional x-ray cassette. This has replaced the use of film and permitted digital images for medical information systems.

67.4 Computed Tomography

A diagnostic computed tomography (CT) scanner comprises an x-ray tube with collimation to provide the slice thickness, a linear array of detector elements, and a reconstruction computer. The x-ray tube and the detectors typically rotate in a gantry. The number of detectors used depends on the generation of the scanner. First-generation scanners had only one detector that was translated across the patient with the tube for each projection, and then the entire assembly was rotated to acquire the next projection view.

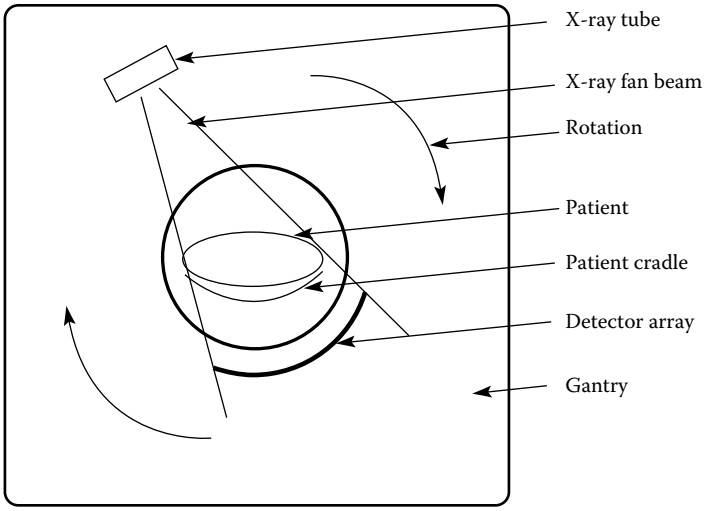


FIGURE 67.3 Orientation of components in a typical third-generation CT scanner.

To increase acquisition speed, second-generation scanners used several detectors in limited fan-beam geometry. Third-generation CT scanners, which are the most common in use today, utilize a large fan array of detectors (852 elements in a current scanner), which completely encompasses the patient and allows slice acquisition times of about 1 s [21]. The x-ray tube and the detector fan array are mechanically coupled and rotate together at high speed (Figure 67.3).

The implementation of electronic slip rings, which allow continuous electric contact, has removed the physical restriction imposed by the high-voltage cables of earlier scanners. Fourth- and fifth-generation scanners have a stationary, complete ring of detectors (typically 1200–4800 detectors). In fourth-generation scanners the x-ray tube is rotated alone, while the fifth-generation scanner design has a focused electron beam that traverses multiple target rings. Fifth-generation scanners can acquire a slice fast enough (50 ms per slice and 17 slices per second) to stop cardiac motion [30,31]. These last two generations are not in common use, primarily due to high cost. In recent years, a helical-scan adaptation of third-generation scanners, allowing continuous acquisition of data over a large patient volume, has become clinically popular [32,33].

While the x-ray tubes used for CT (tube potential range of 80–140 kV) are very similar to general radiographic tubes, the detectors are quite different from conventional radiographic detectors. Detectors used in CT are 1D photon counters that must be efficient and fast. Early CT devices used scintillation detectors, which converted the x-ray energy into light photons that were counted by PMTs. Originally, single-crystal NaI was used, but it proved to be insufficient in dynamic range and had too much afterglow of scintillation light. High-pressure (25 atm) xenon gas later replaced NaI as the detector. Currently, many CT scanners use scintillating ceramics (e.g., $CdWO_4$, $(Y,Gd)_2O_3$:Eu, and Gd_2O_2S:Pr,Ce) coupled to photodiodes, due to the high bulk density of the ceramics.

67.4.1 Reconstruction of an Object from Projections

CT is based on the image reconstruction theorem, which states that if one measures enough projections of an object, the 2D distribution of that object may be reconstructed from the projection data. In CT, the quantity of interest is the linear attenuation coefficient, μ, at each point in the object. The transmission of x-rays through an object of thickness x can be stated as

$$I(x) = I(0)e^{-\mu x} \tag{67.4}$$

where $I(0)$ is the incident intensity. Each ray from the focal spot of the tube to a discrete detector element is a measure of the line integral of the attenuation coefficient through the patient:

$$\lambda(x_r) = -\ln \frac{I(x_r)}{I(0)} = \int_{\ell} \mu(x_r, y_r) dy_r \tag{67.5}$$

where r represents the reference frame of one of the many projections through the patient.

Image reconstruction requires a method to invert Equation 67.5, in order to extract $\mu(x,y)$ of the object from the measured projection views, λ. The mathematical principles of image reconstruction from an infinite number of projections through an object were developed by Radon in 1917 [34]. An approximate solution to the Radon inversion, known as back projection, was later developed because of the need for rapid computation of images in clinical CT. Back projection involves smearing the data from each projection through the 2D space of the patient and summing overall projections. Simple back projection yields an estimate of the patient structures, but is plagued by artifacts due to the approximate nature of the reconstruction procedure. These artifacts are successfully removed, however, by prefiltering the projection data before back projecting. The 1D prefiltering is typically performed in frequency space by multiplying by a ramp function. This technique is known as filtered back projection and results in more accurate reconstructions of patient anatomy [35].

There are several conditions that can reduce the quality of image reconstruction. First, an insufficient number of angular projections or incomplete sampling of the object can lead to aliasing in the reconstructed view. Second, partial volume effects occur when the object is not of homogeneous composition in a particular voxel, causing the reconstructed pixel value (CT number) to be not representative of the tissue. Third, if the acquisition is not fast enough, patient motion leads to a ghosting artifact in the reconstructed image. Last, beam hardening occurs when a high-density structure, such as the skull, significantly changes the beam energy spectrum. The result is reduced intensity of adjacent structures. Beam hardening can be reduced by slightly altering the shape of the reconstruction filter to improve the reconstruction for a particular tissue type.

To present the reconstructed data in digital format, the CT number (also known as the Hounsfield unit [HU]) was developed.

$$\text{CT number} \equiv 1000 \times \frac{\mu_{\text{pixel}} - \mu_{\text{water}}}{\mu_{\text{water}}} \tag{67.6}$$

Using this normalization, pixel values in a CT image are stored as 12-bit integers between –1000 and 3095. A pixel containing only water would have a CT number equal to 0, while one containing bone or muscle would have a positive value, and one with only fat or air would have a negative value.

67.4.2 Clinical Measurements

The reconstructed images are only as accurate as the data input to the algorithm; therefore, a rigorous calibration and quality assurance program is vital to the performance of a diagnostic CT scanner. Calibration generates a baseline reference in air for the scanner and calibration values for every possible scan parameter. Detector channel variation and interaction along with x-ray tube focal spot size and position are quantified. A phantom is used to measure detector response for typical beam widths (1, 3, 5, 7, and 10 mm) and tube potentials (80, 100, 120, and 140 kV). The positioning accuracy of the scanner is also checked.

Quality assurance seeks to establish and maintain consistent image quality [36,37]. A specialized Plexiglas phantom is used to monitor the low-contrast detectability and high-contrast resolution and noise characteristics of the system. The low-contrast portion of the phantom consists of a set of holes of

different diameter in a thin polystyrene slab. The 0.75 mm thick polystyrene, when submerged in water and scanned with a 10 mm slice thickness, yields low contrast in the holes of about 1% (10 HU). The minimum detectable diameter is then found. The high-contrast resolution part of the phantom contains several repeating, equally sized bar/space patterns (spaces filled with water, contrast ~12% or 120 HU) with bar widths from 0.5 to 1.6 mm. The MTF is computed as a plot of the high-contrast frequency response [38]. The noise and uniformity of the scan are assessed with a homogeneous section of the phantom. An ROI is placed in the homogeneous area and the standard deviation is calculated, which should be approximately 3 HU.

67.5 Nuclear Medicine

Nuclear medicine techniques [39,40] use radiopharmaceuticals that are injected into the body to monitor or measure physiological function. Central to nuclear medicine is the role of the radiopharmaceutical as a tracer, that is, an agent with a predictable physiological action that is introduced without perturbing the function of the system. An external detector is used to record radioactivity emanating from the patient to determine the spatial distribution (and often temporal changes in concentration) of the radiopharmaceuticals in specific organs or tissues. Each radiopharmaceutical has an expected bio-distribution that a radiologist evaluates to diagnose the medical status of a patient. The radiopharmaceutical can be labeled either with positron-emitting radionuclides, which produce annihilation photons, or with "single-photon" radionuclides that emit γ-rays (or sometimes x-rays). This section considers only single-photon-emitting radionuclides, examples of which are given in Table 67.1.

TABLE 67.1 Examples of Tracers Used in Nuclear Medicine

Process	Tracer	Refs.
Blood flow		
Diffusible	^{133}Xe	[57]
	[^{99m}Tc]-HMPAO	[58]
Diffusible (trapped)	[^{123}I]IMP (brain)	[58]
	^{201}Tl (heart)	[44,59]
	[^{99m}Tc]-MIBI	[47]
Nondiffusible (trapped)	[^{99m}Tc]macroaggregated albumin, labeled microspheres	—
Effective renal plasma flow	[^{123}I]Hippuran	—
Blood volume		
Red blood cells (RBCs)	[^{99m}Tc]-RBC	[45,46]
Plasma	[^{125}I]-albumin	—
Transport and metabolism		
Free fatty acids	[^{123}I]-hexadecanoic acid	[44]
Bile	[^{99m}Tc]-HIDA	—
Osteoblastic activity	[^{99m}Tc]-MDP	[43]
Glomerular filtration rate	[^{99m}Tc]-DPTA	[60]
Molecular diffusion	[$^{99m}TcO_4$]	[58]
Receptor systems		
Dopaminergic	[^{123}I]-IBZM	[61]
Cholinergic	[^{123}I]-QNB	[62]
Adrenergic	[^{131}I]-MIBG	[63,64]
Somatostatin	[^{111}In]-octreotide	[65,66]

Source: Adapted from Sorenson, J.A. and Phelps, M.E., *Physics in Nuclear Medicine*, 2nd edn., Grune & Stratton, Orlando, FL, 1987.

The scintillation camera [41,42] is the most common device for imaging the distribution of single-photon-emitting radionuclides in vivo (Figure 67.4). The scintillation camera incorporates a large-field (e.g., 40 cm × 50 cm) position-sensitive photon detector with a collimator having a large number of small parallel holes (1–2 mm diameter, 4 cm length) so that only photons traveling perpendicular to the detector surface are recorded. Photons emitted by the patient and passing through the collimator are absorbed by a 1 cm thick sodium iodide scintillator coupled to an array of PMTs. The PMT signals are processed to generate signals proportional to the (*x,y*) coordinates of the interaction site of the photon in the crystal. In addition, the PMT signals are integrated to calculate the photon energy. Events falling within a specified range (typically ±7.5%) around the expected radionuclide photon energy are recorded, whereas those outside of this range are rejected as unwanted scatter or background events. An image is integrated from individual events at the calculated position and specified energy, representing detected photons emitted by the radiopharmaceutical. The camera acquires a planar projection image of the radiopharmaceutical distribution in the patient with a spatial resolution of about 1 cm. The image also can be acquired tomographically by rotating the scintillation camera around the axis of the patient.

This technique is called single-photon emission computed tomography (SPECT) and produces cross-sectional images representing the radiopharmaceutical concentration within the patient.

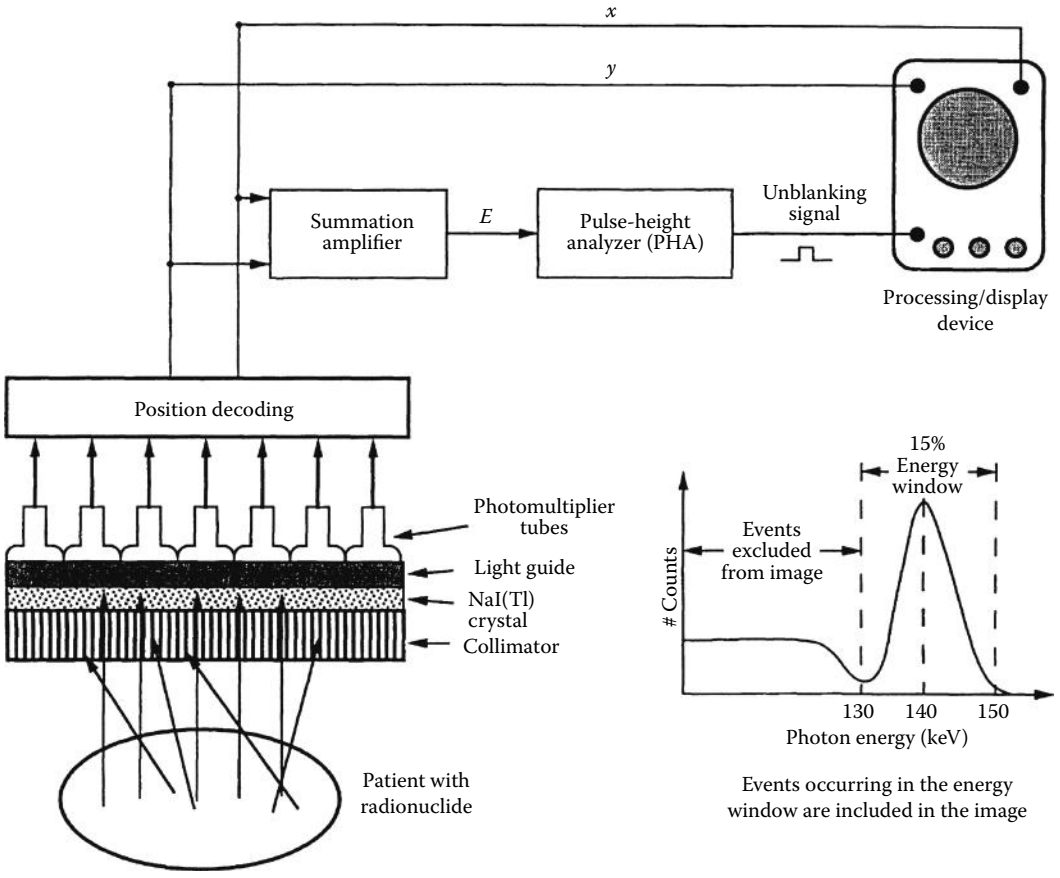


FIGURE 67.4 Scintillation camera incorporates collimator, scintillation crystal, PMTs, and electronic circuitry to generate position (*x,y*) and energy (*E*) of photons emitted by radiopharmaceutical distribution in patient. Only events falling within a specified energy window are recorded by the processing or display device to form the nuclear medicine image.

67.5.1 Measurement of Physiological Function

Radionuclide images can be interpreted visually or quantitatively. For example, the radiopharmaceutical [^{99m}Tc]-methylene diphosphonate (MDP) is incorporated into the bone matrix by osteoblastic activity [43]. A radiologist will inspect a nuclear medicine image for sites demonstrating focal uptake of ^{99m}Tc-MDP to determine the extent and degree of trauma, inflammation, degeneration, metastatic disease, or other skeletal disease processes. Typically, ^{99m}Tc-MDP images are interpreted visually but are not analyzed to determine the quantity of radiotracer incorporated into the skeleton.

Other nuclear medicine studies are assessed quantitatively in the sense that values extracted from the image represent the radioactivity concentration (and physiological function) in a specific organ or tissue region. Myocardial perfusion imaging with [^{99m}Tc]-hexatris-2-methoxyisobutylisonitrile (MIBI) is an example of one such "quantitative" nuclear medicine study for assessing a patient suspected of having coronary artery disease [44–46]. [^{99m}Tc]-MIBI is a lipophilic cation that accumulates in myocardial tissue roughly in proportion to blood flow [47]. Image data are acquired using SPECT to reconstruct tomograms of the myocardial concentration of [^{99m}Tc]-MIBI that are analyzed to assess regional myocardial blood flow. Although absolute measurements (μCi/g) of tissue activity are difficult (if not impossible) to obtain with SPECT, the images are interpreted "quantitatively" by extracting pixel values from the image to derive diagnostic information [48], rather than relying on "qualitative" visual interpretation of the images. Typically, ^{99m}Tc-MIBI is imaged in the "short-axis" view that presents the left ventricular myocardium in a series of annuli (or "doughnuts"). The image is analyzed using a "circumferential profile" representing the radionuclide concentration at 6° angular increments around each annular slice of the myocardium [49]. The extracted values are compared with standard values obtained from patients, in whom atherosclerotic disease has been excluded by coronary angiography, thereby allowing the nuclear cardiologist to assess both the presence as well as the regional extent of coronary artery disease.

67.5.2 Measurement of Technical Performance

Several parameters generally are measured to assess the performance of the scintillation camera [50–56]. *Spatial resolution* represents the precision with which the position of an event is localized and can be assessed from the full width at half maximum (FWHM) of a profile taken across the image of a point or linear radioactive object having small dimensions in comparison to the resolution of the system. *Spatial linearity* is quantified as the accuracy with which the position of an event is localized and represents the ability of a scintillation camera to produce a straight image of a straight object. Spatial linearity is measured as the deviation about the best-fit line in an image of a parallel line phantom or an orthogonal hole phantom, expressed as a percentage (ideally less than 1%) of the detector diameter. *Energy resolution* represents the precision with which the energy of a photon is recorded and generally is measured as the FWHM of the photopeak in an energy spectrum (number of detected photons recorded as a function of photon energy) of the radioactive source. *Flood field uniformity* assesses the ability of the camera to record a spatially uniform image when presented with a spatially uniform distribution of photons. An intrinsic measurement is performed by irradiating the uncollimated detector with the point source placed at a distance equal to at least five times the field of view of the detector. The system uniformity can be checked by irradiating the entire surface of a collimated detector with an extended source of uniform radioactivity. Sensitivity represents the number of photons recorded per unit of source radioactivity when the detector is operated either without (intrinsic sensitivity) or with (extrinsic sensitivity) a collimator. *Count-rate linearity* represents the ability of the camera to record a count rate proportional to the photon event rate received by the detector. At low event rates, the measured count rate increases linearly with the actual photon event rate. Because the scintillation camera acts as a paralyzable system, at higher event rates, the measured count rate is lower than that

predicted from linear response. At sufficiently high event rates, the measured count rate actually can decrease with increasing photon event rate and eventually can be extinguished when imaging radioactive sources of sufficiently high activities.

67.6 Positron Emission Tomography

Positron emission tomography (PET) involves a physiological administration of a positron-emitting radiopharmaceutical into the human body. The principal advantage of PET over single-photon imaging is the availability of a number of physiologically relevant radiotracers that are labeled with the short-lived positron-emitting radionuclides ^{11}C ($T_{1/2}$ = 20.4 min), ^{13}N (9.96 min), ^{15}O (2.04 min), and ^{18}F (109.8 min). A typical PET center consists of a cyclotron for on-site isotope production, a radiochemistry laboratory for synthetic incorporation of the isotopes into organic molecules, and a PET scanner. PET instrumentation is described in detail in several review articles [67,68].

67.6.1 Principle of Coincidence Detection

The proton-rich radioisotopes used with PET imaging undergo β-decay and emit positrons (anti-electrons). A positron travels a short distance and combines with an electron from the surrounding medium. The masses of the positron and electron are converted to electromagnetic radiation in the form of two γ-rays of energy 511 keV, which are emitted at nearly 180° to each other. The PET scanner utilizes multiple opposing γ-detectors that surround the positron emitter, each defining a linear volume of response between the detectors. Coincidence timing circuitry enables effective localization of the decay events occurring between detector pairs, rejecting events in each detector that originate from outside the volume of response. A typical modern PET scanner employs tens of thousands of small detectors (or analogous position-coded larger detectors), yielding as many as tens of millions of such volumes of response. The coincidence principle is also utilized to measure and correct for attenuation of photons within the body, allowing the measurement of radioactivity concentration in absolute terms (i.e., Bq/mL). In this case, a separate "transmission" measurement scan is performed, using an external positron-emitting source placed adjacent to the subject yet within the volume of response. A "blank" scan is similarly acquired but without the subject in the field of view. The ratio of coincident count rates in the blank/transmission scans multiplies the corresponding coincidence counts in the emission scan to correct for attenuation along each coincidence line of response.

67.6.2 Detector Composition

The choice of detector material for PET scanners is influenced by a number of considerations, including scanner geometry, detection efficiency (stopping power), output signal strength (energy resolution), signal decay time (count rate capability), physical stability (i.e., hygroscopicity), availability, and cost. Inorganic scintillators are best suited for detection of the 511 keV photons. The physical properties of the two most widely used scintillators, NaI(Tl) and bismuth germanate (BGO), are shown in Table 67.2. NaI(Tl) has found application in position-sensitive detector systems that utilize a small number of large crystals observed by multiple PMTs. NaI(Tl) offers the advantages of (1) good energy resolution for effective rejection of scattered radiation, (2) good timing resolution for minimizing the coincidence-resolving time window, (3) availability of large crystals, and (4) relatively low cost. The higher stopping power of BGO is advantageous for detector designs that use smaller crystals with one-to-one PMTs or position-encoded matrices of crystals [69]. The recently identified lutetium oxyorthosilicate (LSO) is a potential successor to BGO in detector block designs. LSO has a density of 7.4 g/mL, an effective atomic number of 59, a photofluorescent decay time of 40 ns, and light outputs that are 3/4 that of NaI(Tl) [70].

TABLE 67.2 Physical Properties of Scintillators Commonly Employed in PET Scanners

	NaI(Tl)	BGO
Density (g/cm^3)	3.67	7.13
Effective atomic number	51	75
Index of refraction	1.85	2.15
Relative emission intensity	100	15
Peak wavelength (nm)	410	480
Decay constant (ns)	230	300

67.6.3 PET Scanners

PET scanners use a number of different detector compositions and gantry configurations, each with its unique advantages and disadvantages [68]. Figure 67.5 shows two of the most common designs that are currently employed. At present, the majority of commercial designs employ a cylindrical geometry with individual BGO detector blocks arranged to form contiguous rings of detectors, each defining an image plane [71,72]. Most of these scanners have retractable lead (or tungsten) septa that are positioned between detector rings to attenuate photons that are emitted at angles not contained in the image plane. This minimizes the effect of out-of-plane scattered radiation, allowing accurate quantitation of the radioactivity distribution in each image plane by 2D (tomographic) image reconstruction. With the septa retracted, all axial angles are accepted, allowing true 3D volume imaging. Another scanner design uses large-area position-encoded NaI(Tl) detectors, allowing 3D volume imaging [73]. In all cases, computer-assisted image reconstruction is used to produce quantitative images of radiotracer concentration in the body.

The spatial resolution of the radioactivity distributions seen in the PET image is primarily determined by the size of the detector elements. In scanners employing cylindrical detector geometry, the in-plane spatial resolution is highest in the center of the field of view (typically 4–5 mm FWHM of the point source response for present state-of-the-art scanners). The spatial resolution slowly degrades as the radius increases due to inadequate stopping of photons within incident detectors for nonperpendicular entrance angles. Likewise, the resolution in the axial direction is determined by the axial dimension of the detector elements.

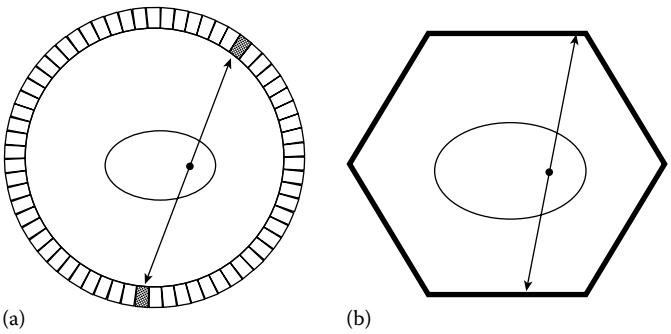


FIGURE 67.5 Representation of PET scanner geometries for typical scanners employing (a) multiple rings of small BGO scintillators and (b) six large NaI(Tl) position-sensitive planar detectors. The arrows represent positron annihilation photons that are emitted 180° from each other and detected in opposing detectors. (Courtesy of Dr. T. Turkington.)

67.7 Ultrasound Imaging

Ultrasound scanning provides a safe and noninvasive way to image the body. With this modality, brief pulses of sound are emitted by a transducer coupled to the skin surface. The sound pulse propagates through tissue at a fixed speed. Interfaces and other objects reflect portions of the acoustic energy back to the transducer, where they are detected as echoes. The ultrasound scanner forms 1D, or more commonly 2D, images of anatomic structures from the reflected echo patterns. In general imaging applications, ultrasound imaging uses frequencies in the 2–10 MHz range. Some newer ultrasound devices, for example, those used in emerging ophthalmology applications, use frequencies as high as 50 MHz [74,75].

67.7.1 Characteristics of Sound Waves in Tissue

The speed at which sound waves propagate through a medium depends on the density and compressibility of the medium. At 22 °C, the speed of sound in air is around 300 m/s, while in freshwater it is 1480 m/s. Human soft tissues behave somewhat like water, with speeds of sound ranging from 1460 m/s for fat to 1620 m/s for muscle. The average speed of sound in tissue is taken to be 1540 m/s (1.54 mm/μs) [76,77].

Any interface, large or small, can reflect a fraction of the ultrasound energy and produce an echo. The relative amount of energy reflected depends on the change in density and compressibility at the interface; the greater the change in these properties of the materials forming the interface, the greater the amplitude of an echo. Examples of reflectors include organ boundaries, blood vessels, and small scatterers distributed more or less randomly throughout most organs. The majority of the echo data displayed on images can be attributed to this scattering process [77]. Shung [78] has reviewed experimental work on ultrasonic scattering versus frequency in biological tissues.

As ultrasound pulses travel through tissue, they lose their strength due to attenuation. Attenuation is caused by scatter and reflection at interfaces and by absorption. For typical tissues, the amplitude of a 5 MHz beam decreases by about 50% for each centimeter traveled. The attenuation per unit distance is approximately proportional to the ultrasound frequency, so lower-frequency waves propagate greater distances through tissues than higher-frequency waves [79].

67.7.2 B-Mode Imagers

Figure 67.6 illustrates a typical configuration for an ultrasound imager. The operator places a handheld transducer on the skin surface of the patient. Early instruments utilized "single-element" transducers, but the majority of systems now use transducer arrays [80]. Acoustic pulses emitted by the transducer travel in well-defined beams. This beam can be "steered" in different directions, either mechanically with motors or electronically by using transducers arrays.

The same transducer detects echoes that arrive from interfaces in the body and applies them to the receiver, where they are amplified and processed for display. The instrument converts each echo signal into a dot on the display, the brightness of the dot being proportional to the echo amplitude at the transducer. The "scan converter" memory places dots in a location that corresponds to the reflector locations; information required to do this is the return time for each echo and the beam axis direction when the echo is detected.

The scanner constructs a cross-sectional image by sending out 100–200 such ultrasound beams, each in a slightly different direction, somewhat like a searchlight scanning the night sky. Echoes received from each beam direction are placed in the image memory using the scheme mentioned earlier. The entire image is updated at rates of 15–30 scans per second, producing a real-time image on the display monitor. This technique is referred to as B-mode imaging because echo signals simply modulate the intensity, or brightness, of the display at locations corresponding to their anatomic origin.

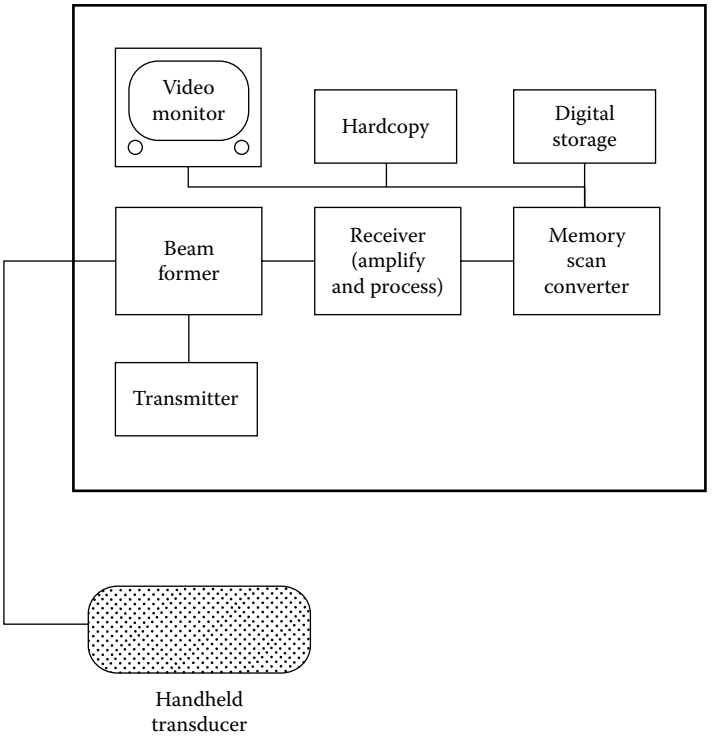


FIGURE 67.6 Components in a typical B-mode ultrasound device.

67.7.3 Doppler Techniques

Ultrasound instruments commonly provide Doppler records as well as B-mode images. Fundamentally, the Doppler effect is a change in the frequency of reflected waves when there is relative motion between the transducer and reflector. For motion directly toward or directly away from the transducer, the Doppler signal frequency f_d is given by

$$f_d = \frac{2 f_0 v}{c} \tag{67.7}$$

where

f_0 is the frequency of the transmitted ultrasound
v is the velocity of the reflector
c is the speed of sound

Thus, the Doppler signal frequency provides information on reflector velocity.

Continuous wave (CW) Doppler instruments consist of a transducer with separate transmitting and receiving elements, a transmitter–receiver unit, and a signal display. They extract a Doppler signal from the complex echo pattern, usually by heterodyning the echo signal with a signal that is coherent with the transmitted wave and then low-pass filtering. The most common applications are to detect and measure blood flow. With a 5 MHz ultrasound frequency and blood velocity of 50 cm/s, the Doppler signal frequency is 3.25 kHz, that is, in the audible frequency range. A simple loudspeaker may be all that is necessary for interpreting the Doppler signal, but very often a real-time spectral analyzer is available.

Pulsed Doppler instruments are a bit more complicated, but allow the operator to define precisely the distance from the transducer from which Doppler signals are selected. In pulsed Doppler, an acoustic

pulse is transmitted along a fixed beam line. Resultant echo signals are amplified and subjected to Doppler processing methods, similar to those outlined for the CW instrument. An operator-adjusted gate captures the waveform from the depth of interest, and a sample-hold device retains the value of this waveform until a subsequent pulse–echo sequence. Because the phase of the echo signal from moving reflectors changes from one pulse–echo sequence to the next, a Doppler signal can be constructed from the behavior over time of the sample-hold value.

67.7.4 Color Doppler Imaging

Color flow imagers may be thought of as extensions of pulsed Doppler machines. Rather than detecting Doppler signals from a single location, color flow imagers detect signals from all depths covered by the ultrasound beam and for many beam directions. Most instruments extract and display the mean Doppler signal frequency for each location throughout the scanned field [81,82]. A color Doppler image is almost always combined with a B-mode image to provide both anatomic and flow data from the scanned plane.

67.7.5 Measurement of Ultrasound Instrument Performance

Defining "image quality" in ultrasound, and specifying quantifiable factors that relate to optimal B-mode imaging, is controversial to say the least. Important factors that are considered include spatial and contrast resolution, sensitivity, penetration depth, and geometric accuracy.

High-quality ultrasound imagers interrogate the scanned field using a sufficient number of individual beam lines (more than 100) such that gaps between lines can be ignored in resolution considerations. An exception may be in color flow imaging, where sparse line densities are needed for sufficient frame rates [77]. Thus, spatial resolution is dictated by the volume of the ultrasound pulse propagating through the tissue. The dimension of this pulse volume in the direction the pulse travels, that is, the axial resolution, is determined by the duration of the pulse emitted by the transducer, while the dimension perpendicular to the beam axis or the "lateral resolution" is determined by the beam width. Although ultrasound beam energy is concentrated near the axis, it is the nature of beams from finite-sized apertures that the intensity falls off gradually with increasing distance from the beam axis. Finally, the size of the ultrasound beam perpendicular to the image plane determines the "slice thickness," the width of the volume of tissue contributing to the echo data viewed in the image plane.

A variety of methods have been used for determining in-plane resolution. The lateral and axial dimensions of a reflector whose size is small enough that it can be considered a pointlike object are frequently used [83]. For a 3.5 MHz transducer, this "spot size" can be as small as 0.7 mm in the axial dimension and 1–2 mm laterally. Smaller spot sizes are found with higher-frequency imagers, such as those using 10 MHz scan heads. Also, larger spot sizes are obtained with scanners that use fixed-focus, single-element transducers.

Slice thickness has been measured using a planar sheet of scatterers scanned with the ultrasound scanning plane intersecting the sheet at a 45° angle [83]. If the slice thickness were negligible, the image of the sheet in this projection would be a straight, horizontal line. The finite thickness of the scanned slice causes a thickening of the line; in fact, for the 45° orientation the vertical size of the image of the sheet corresponds to the slice thickness. For all ultrasound imaging systems, except annular array transducers, the slice thickness is the worst measure of spatial resolution, ranging from the 10 to 2 to 3 mm, depending on depth, for a 3.5 MHz transducer.

Physicians commonly use ultrasound imagers to detect cancerous tumors, for which the echoes are slightly stronger or weaker than the surrounding region. "Contrast-detail" tests [84,85] measure the smallest object that can be visualized at a fixed backscatter difference. Spherical mass detectability [86] assesses capabilities to visualize realistic focal lesions. Masses in the latter detection test are characteristic of actual tumors; furthermore, they are easily distributed throughout the scanning plane, assessing resolution at all depths.

Scanner sensitivity is an important performance feature, especially because spatial resolution can be enhanced with higher-frequency transducers. However, this is at the expense of increased ultrasound beam attenuation and poorer penetration. Although absolute measurements of sensitivity of scanners have been done [87], most centers rely upon clinically meaningful "maximum visualization distances" [83,88] for estimating and comparing sensitivity. Geometric accuracy also is important, as images frequently are used to determine structure dimensions, such as fetal head size when determining gestational age [89]. Calibration of distance measurements is done following standard protocols [83,90]; fortunately, modern scanners with digitally based image formation maintain their accuracy much better than previous systems, and many physicists maintain that tests for geometric accuracy are not crucial in routine performance assessments.

67.8 Magnetic Resonance Imaging

MRI is a new medical imaging modality that uses magnetic fields and radio-frequency (RF) energy to produce images of the body. This technique is based on nuclear magnetic resonance (NMR) [91], which is a quantum mechanical phenomenon exhibited by atoms having an odd number of either protons or neutrons. Such atoms have a nonzero nuclear magnetic moment, μ, and will precess (or rotate) about an external magnetic field ($\mathbf{B}_0$) with a frequency of $\omega_0 = \gamma \mathbf{B}_0$, where γ is the gyromagnetic ratio that for H is 42.57 MHz/T. A number of isotopes (including ^{1}H, ^{31}P, ^{23}Na, ^{2}H) exhibit the NMR phenomena; however, the majority of magnetic resonance (MR) scanners image ^{1}H. This is because, relative to other isotopes, ^{1}H has a high inherent sensitivity and abundance in tissue. Therefore, the following discussion is limited to ^{1}H MRI. When placed in a $\mathbf{B}_0$ field, ^{1}H nuclei align their spins either parallel or antiparallel to $\mathbf{B}_0$, with a slight excess in the lower-energy parallel state. At $T = 25$ °C and $|\mathbf{B}_0| = 1.5$ T, an excess of ~5 in 10^6 atoms is in the parallel state (this excess increases with $\mathbf{B}_0$ and T^{-1}). Because there are $\sim 10^{23}$ ^{1}H per milliliter of tissue, this excess, when summed over even a small volume, results in a net magnetization, $\mathbf{M} = \Sigma\mu$.

67.8.1 MR Imaging Techniques

The Bloch equations [91,92] are a set of phenomenological equations that succinctly describe the evolution of the net magnetization $\mathbf{M}(\mathbf{r},t)$ during an MRI experiment:

$$\frac{\partial \mathbf{M}(\mathbf{r},t)}{\partial t} \gamma \mathbf{M}(\mathbf{r},t) \times \mathbf{B}(\mathbf{r},t) - \frac{M_x(t)\hat{x} + M_y\hat{y}}{T2} + \frac{(M_z(t) - M_0)\hat{z}}{T1} \tag{67.8}$$

where $\mathbf{M}(\mathbf{r},t) = (M_x(t), M_y(t), M_z(t))$; M_0 is the initial (or equilibrium) magnetization, $\mathbf{B}(\mathbf{r},t) = \mathbf{B}_0 + \mathbf{G}(\mathbf{r},t)\ \mathbf{r} + \mathbf{B}_1(t)$ is the total applied magnetic field and includes terms representing the static field, $\mathbf{B}_0$, the field gradients, $\mathbf{G}(\mathbf{r},t)\ \mathbf{r}$, and the magnetic field component of any applied RF excitation, $\mathbf{B}_1(t)$; and $T1$ and $T2$ are the characteristic relaxation times of the tissues being imaged. The coordinate system is described in Figure 67.7. The $\mathbf{B}_0$ and $\mathbf{G}(\mathbf{r},t)\ \mathbf{r}$ fields are parallel to z, and the $\mathbf{B}_1$ field is orthogonal to z. Every MRI experiment consists of an excitation phase, in which the equilibrium magnetization is tipped away from z (the longitudinal axis) and into the transverse (x–y) plane. This is followed by a detection phase, in which the signal emitted by the excited spins is manipulated so that an echo forms. The echo time (TE) and the repetition time (TR) denote the time between excitation and echo formation and the time between successive excitation phases, respectively. Starting at equilibrium, RF energy at ω_0 is applied to create an oscillating $\mathbf{B}_1$(t) field that tips $\mathbf{M}$ toward the transverse plane, at which time it begins to precess about z (Figure 67.7a). Viewed in a frame of reference rotating at $-\omega_0$ about $z(x', y', z')$, the RF excitation corresponds to tipping $\mathbf{M}$ through an angle α from z (Figure 67.7b). Gradients, $\mathbf{G}(\mathbf{r},t)$, are then applied so that spatial information is encoded into the precessional frequency of $\mathbf{M}$, $\omega = \gamma\ [\mathbf{B}_0 + \mathbf{G}(\mathbf{r},t) \cdot \mathbf{r}]$. Additional gradients and/or RF-excitation pulses are used to refocus the magnetization into an echo (Figure 67.7c through f).

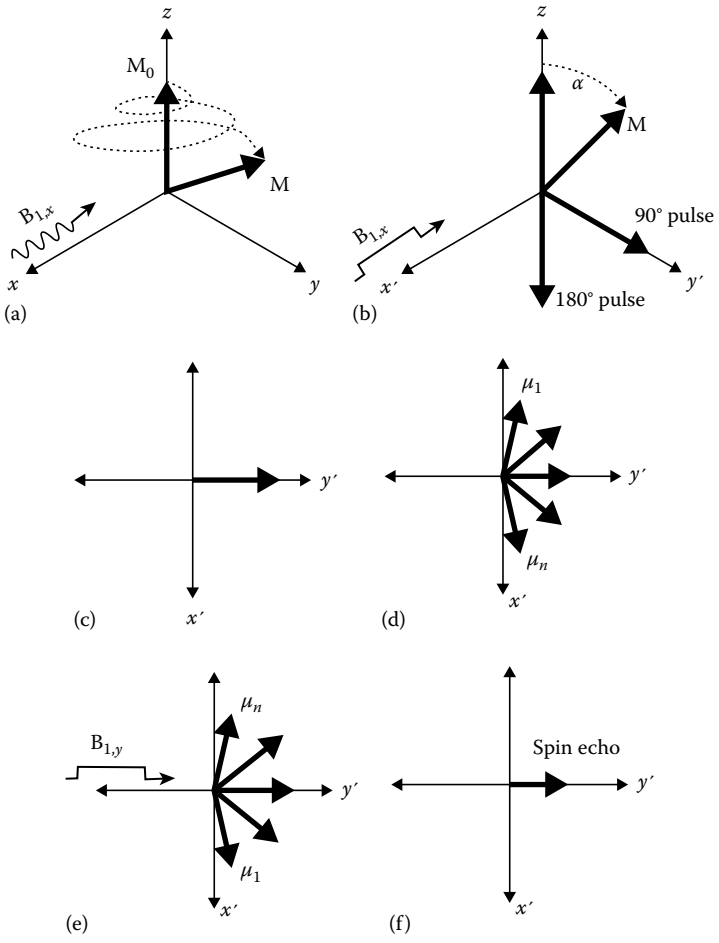


FIGURE 67.7 Graphical depiction of the MRI process showing the effect of the $\mathbf{B}_1$-excitation pulse (a–b) and formation of a spin-echo (c–f). In (a) an RF pulse, $\mathbf{B}_1$, is applied along the x-axis causing the net magnetization to tip away from equilibrium, $\mathbf{M}_0$, as it precesses about z. Viewed in the rotating frame of reference, this corresponds to a nutation by some angle α in the y'–z plane. After application of an $\alpha = 90°$ pulse, $\mathbf{M}$ lies in the x'–y' plane (c) and is subject to variations in the local magnetic field that cause the individual spins to precess at different frequencies, that is, μ_1 and μ_n in (d). The local field variations could be due to gradients or field inhomogeneity, the latter effect leading to $T2^*$ signal loss. By applying an $\alpha = 180°$ RF pulse along y', the spins are rotated about y' (e) and are refocused into an echo (f).

The precessing transverse component of the magnetization ($M_{xy} = M_x + jM_y$) induces a signal (EMF) in the receiver coil, and this is recorded during echo formation. The experiment is repeated many times (normally 128–256) using predetermined gradient strengths [93] so that a complete data set is collected. These data are in spatial-frequency space (known as **k**-space), and images are reconstructed after taking the multidimensional Fourier transform of **k**-space [94]. Gradients can be used to encode spatial information along all axes, so either 2D (planar) or 3D (volume) imaging is possible with MR.

67.8.2 MR Image Contrast

Excited spins undergo two principal relaxation processes: spin–lattice and spin–spin relaxation. Spin–lattice relaxation occurs when spins lose energy to surrounding molecules and return to the equilibrium position, M_0. Spin–spin relaxation is due to local interactions that cause spins to precess at different rates, resulting in

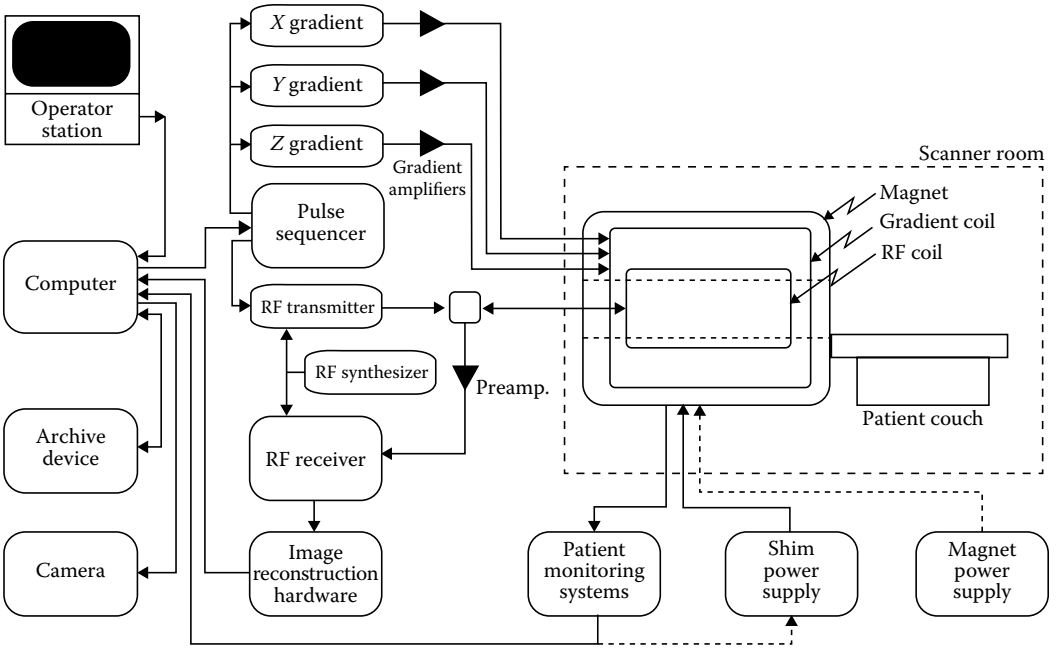


FIGURE 67.8 Overview of an MR scanner showing key components.

a reduction in $\mathbf{M}_{xy}$ (Figure 67.7d). Field inhomogeneity due to an imperfect magnet is reversible ($\delta\mathbf{B}(\mathbf{r})$), while inhomogeneity due to variations in local chemical structure is not. The relaxation times $T1$ and $T2$ characterize spin–lattice and spin–spin relaxation processes, respectively, and from the Bloch equation (Figure 67.8) the following equations can be derived for the relaxation of spins tipped by $\alpha = 90°$ at $t = 0$:

$$M_z(t) = M_0(1 - e^{-t/T1}) \quad \text{and} \quad M_{xy}(t) = M_{xy}(t = 0)e^{-t/T2} \tag{67.9}$$

Together with the differences in the inherent proton density between tissues, differences in $T1$ and $T2$ are the basis of the contrast seen in MR images. A spin-echo (SE) acquisition uses one or more additional RF excitations with $\alpha = 180°$ to refocus the reversible component of spin–spin relaxation so that one or more echoes are formed. Equation 67.9 shows that images with primarily $T1$ weighting, $T2$ weighting, or proton density weighting result when relative to $T1$ and $T2$: TR and TE are short, TR and TE are long, and TR is long and TE is short, respectively. In practice, images with a pure weighting are not obtainable because image contrast is due to a mixture of these contrast mechanisms. Gradient-recalled echo (GRE) imaging uses additional gradients, instead of an $\alpha = 180°$ pulse, to produce echoes and results in images with significantly shorter TR and TE than SE techniques. GRE images, however, do not refocus the reversible component of spin–spin relaxation and are thus susceptible to $T2^*$ signal loss, where $1/T2^* = 1/T2 + 2\pi\gamma|\delta\mathbf{B}(r)|$. This leads to lower signal to noise in GRE images compared with SE images. Since TR is short compared with $T1$, GRE images tend to have $T1$ weighting. If TE approaches $T2^*$, then the $T2^*$ weighting becomes significant.

67.8.3 MR Instrumentation

The key components of a modern MRI system include a magnet, a pulse sequencer, gradient and shim coils, and an RF transmitter/receiver—the function of which are controlled by a host computer (Figure 67.8). To obtain the $\mathbf{B}_0$-field, most commercial scanners use superconducting magnets, although some special-purpose (and often lower-cost) scanners may use resistive magnets. Superconducting magnets generally

have field strengths between 0.5 and 4.0 T, while resistive magnets normally have field strengths <0.3 T. The improved signal-to-noise ratio obtained with superconducting designs is offset by the need for periodic cryogen replacement. More modern magnets, however, minimize this cost by including a cryogen reliquefier. Two sets of auxiliary gradient coils are located within the main magnet to provide spatially varying fields, $\mathbf{G}(\mathbf{r},t)$, and to allow shimming of the $\mathbf{B}_0$-field. Current gradient coil hardware can generate maximum gradients of up to 40 mT/m with rise times of 120 μs and allow fields of view from 4 to 48 cm. Shim coils improve the homogeneity of the $\mathbf{B}_0$-field by decreasing $\delta\mathbf{B}(r)$ to a few parts per million in order to minimize $T2^*$ effects and spatial distortions. Modern scanners incorporate a digital RF subsystem that excites the spins and then records the emitted signals via one or more RF coils within the magnet. An RF synthesizer is coupled to both the RF transmitter and receiver, so that synchronous detection is possible. The RF system is connected either to separate transmit and receive coil(s) or to a combined transmit/receive coil(s).

To acquire an MR image, the host computer interacts with the operator who defines the imaging parameters (such as α, TR and TE, slice location, and field of view). The parameters are then translated into instructions that are executed on a synchronous state machine known as a pulse sequencer. This device provides real-time control of the gradient and RF waveforms as well as other control functions, such as unblanking the RF receiver and enabling the ADC during an echo. Data are collected and demodulated by the receiver, and then images are reconstructed using specialized hardware built normally around a fast-array processor. The images are sent to the host computer for operator station display, archival, or filming. In addition, many MR scanners incorporate devices for monitoring heart and respiration rate and allow these signals to trigger or gate image acquisition. Future MRI systems will probably include higher $\mathbf{B}_0$-field and gradients and faster data processing/reconstruction hardware. In addition to current imaging apparatus, it is likely that dedicated instruments will be increasingly used to study the heart and for performing neurofunctional imaging and MR-guided interventional procedures.

References

1. G. Revesz, H. L. Kundel, and M. A. Graber, The influence of structured noise on the detection of radiologic abnormalities, *Invest. Radiol.*, 9: 479–486, 1974.
2. A. Rose, *Vision: Human and Electronic*, New York: Plenum Press, 1973.
3. K. Doi, K. Strubler, and K. Rossmann, Truncation errors in calculating the MTF of radiographic screen-film systems from the line spread function, *Phys. Med. Biol.*, 17: 241–250, 1972.
4. ICRU, Modulation transfer function of screen-film systems, ICRU Report 41, Bethesda, MD, 1986.
5. H. Fujita, D.-Y. Tsai, T. Itoh, K. Doi, J. Morishita, K. Ueda, and A. Ohtsuka, A simple method for determining the modulation transfer function in digital radiography, *IEEE Trans. Med. Imag.*, 11: 34–39, 1992.
6. J. T. Dobbins III, D. L. Ergun, L. Rutz, H. Blume, D. A. Hinshaw, and D. C. Clark, DQE(f) of four generations of computed radiography acquisition devices, *Med. Phys.*, 22: 1581–1593, 1995.
7. J. M. Boone and J. A. Seibert, An analytical edge spread function model for computer fitting and subsequent calculation of the LSF and MTF, *Med. Phys.*, 21: 1541–1545, 1994.
8. J. M. Sandrik and R. F. Wagner, Absolute measures of physical image quality: Measurement and application to radiographic magnification, *Med. Phys.*, 9: 540–549, 1982.
9. R. C. Bunch, R. Shaw, and R. L. van Metter, Signal-to-noise measurements for a screen-film system, *Proc. SPIE*, 454: 154–163, 1984.
10. R. C. Bunch, K. E. Huff, and R. van Metter, Analysis of the detective quantum efficiency of a radiographic screen-film combination, *J. Opt. Soc. Am.*, 4: 902–909, 1987.
11. J. T. Dobbins III, Effects of undersampling on the proper interpretation of modulation transfer function, noise power spectra and noise equivalent quanta of digital imaging systems, *Med. Phys.*, 22: 171–181, 1995.

12. M. L. Giger, K. Doi, and C. E. Metz, Investigation of basic imaging properties in digital radiography. 2. Noise Wiener spectrum, *Med. Phys*., 11: 797–805, 1984.
13. R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra*, New York: Dover, 1958.
14. R. F. Wagner and J. M. Sandrik, An introduction to digital noise analysis, in A. G. Haus (Ed.), *The Physics of Medical Imaging: Recording System Measurements and Techniques*, New York: American Association of Physicists in Medicine, pp. 524–545, 1979.
15. G. T. Barnes, Radiographic mottle: A comprehensive theory, *Med. Phys*., 9: 656–667, 1982.
16. R. F. Wagner, Fast Fourier digital quantum mottle analysis with application to rare earth intensifying screen systems, *Med. Phys*., 4: 157–162, 1977.
17. H. H. Barrett and W. Swindell, *Radiological Imaging*, New York: Academic Press, 1981.
18. G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*, San Francisco, CA: Holden-Day, 1968.
19. K. Koedooder, J. Strackee, and H. W. Venema, A new method for microdensitometer slit length correction of radiographic noise power spectra, *Med. Phys*., 13: 469–473, 1986.
20. M. L. Giger and K. Doi, Investigation of basic imaging properties in digital radiography. 1. Modulation transfer function, *Med. Phys*., 11: 287–295, 1984.
21. J. T. Bushberg, J. A. Seibert, E. M. Leidholdt Jr., and J. M. Boone, *The Essential Physics of Medical Imaging*, Baltimore, MD: Williams & Wilkins, 1994.
22. M. Braun and B. C. Wilson, Comparative evaluation of several rare-earth film-screen systems, *Radiology*, 144: 915–919, 1982.
23. M. Sonoda, M. Takano, J. Miyahara, and H. Kato, Computed radiography utilizing scanning laser stimulated luminescence, *Radiology*, 148: 833–838, 1993.
24. W. Hillen, U. Schiebel, and T. Zaengel, Imaging performance of a digital storage phosphor system, *Med. Phys*., 14: 744–751, 1987.
25. H. Blume, Stimulable phosphor systems—Technical aspects, in W. W. Peppier and A. Alter (Eds.), *Proc. Chest Imaging Conf '87*, University of Wisconsin, Madison, WI, pp. 194–201, 1987.
26. U. Neitzel, I. Maack, and S. Günther-Kohfahl, Image quality of a digital chest radiography system based on a selenium detector, *Med. Phys*., 21: 509–516, 1994.
27. C. H. Slump, G. J. Laanstra, H. Kuipers, M. A. Boer, A. G. J. Nijmeijer, M. J. Bentum, R. Kemner, H. J. Meulenbrugge, and R. M. Snoeren, Image quality characteristic of a novel x-ray detector with multiple screen-CCD sensors for real-time diagnostic imaging. *Proc. SPIE*, 3032: 60–71, 1997.
28. L. E. Antonuk, Y. El-Mohri, J. H. Siewerdsen, J. Yorkston, W. Huang, V. E. Scarpine, and R. A. Street, Empirical investigation of the signal performance of a high-resolution, indirect detection, active matrix flat panel imager (AMFPI) for fluoroscopic and radiographic operation, *Med. Phys*., 24: 51–70, 1997.
29. D. L. Lee, L. K. Cheung, and L. S. Jeromin, A new digital detector for projection radiography, *Proc. SPIE*, 2432: 237–249,1995.
30. D. P. Boyd, R. G. Gould, J. R. Quinn, R. Sparks, J. H. Stanley, and W. B. Herrmannsfeldt, A proposed dynamic cardiac densitometer for early detection and evaluation of heart disease, *IEEE Trans. Nucl. Sci*., 26: 2724–2727, 1979.
31. C. H. McCollough and R. L. Morin, The technical design and performance of ultrafast computed tomography, *Radiol. Clin. North Am*., 32: 521–536, 1994.
32. P. M. Silverman, C. J. Cooper, D. I. Weltman, and R. K. Zeman, Helical CT: Practical considerations and potential pitfalls, *Radiographics*, 15(1): 25–36, 1995.
33. J. Hsieh, A general approach to the reconstruction of x-ray helical computed tomography, *Med. Phys*., 23: 221–229, 1996.
34. J. Radon, On the determination of functions from their integrals along certain manifolds, *Ber. Verh. Sächs. Akad. Wiss. Leipzig, Math. Phys. Kl*., 69: 262–277, 1917.
35. A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, New York: IEEE Press, 1988.

36. E. C. McCullough, J. T. Payne, H. L. Baker, Jr., R. R. Hattery, P. F. Sheedy, D. H. Stephens, and E. Gedgaudus, Performance evaluation and quality assurance of computed tomography scanner, with illustrations from the EMI, ACTA, and delta scanners, *Radiology*, 120: 173–188, 1976.
37. E. C. McCullough and J. T. Payne, X-ray-transmission computed tomography, *Med. Phys*., 4: 85–98, 1977.
38. R. T. Droege and R. L. Morin, A practical method to measure the MTF of CT scanners, *Med. Phys*., 9: 758–760, 1982.
39. H. N. Wagner, Jr., Z. Szabo, and J. W. Buchanan (Eds.), *Principles of Nuclear Medicine*, 2nd edn., Philadelphia, PA: Saunders, 1995.
40. M. P. Sandler (Ed.), *Diagnostic Nuclear Medicine*, 3rd edn., Baltimore, MD: Williams & Wilkins, 1996.
41. R. Chandra, *Introductory Physics of Nuclear Medicine*, 4th edn., Philadelphia, PA: Lea & Febiger, 1992.
42. J. A. Sorenson and M. E. Phelps, *Physics in Nuclear Medicine*, 2nd edn., Orlando, FL: Grune & Stratton, 1987.
43. J. L. Littlefield and T. G. Rudd, [^{99m}Tc]-hydroxymethylene diphosphonate and [^{99m}Tc]-methylene diphosphonate biologic and clinical comparison: Concise communication, *J. Nucl. Med*., 24: 643, 1983.
44. G. B. Saha, W. J. MacIntyre, R. C. Brunken, R. T. Go, S. Raja, C. O. Wong, and E. Q. Chen, Present assessment of myocardial viability by nuclear imaging, *Semin. Nucl. Med*., 26(4): 315–335, 1996.
45. T. R. Miller and J. W. Wallis, Cardiac nuclear medicine, *Curr. Probl. Diagn. Radiol*., 17(5): 157–193, 1988.
46. R. J. Boudreau and M. K. Loken, Functional imaging of the *heart, Semin. Nucl. Med*., 17(1): 28–38, 1987.
47. F. J. Wackers, D. S. Berman, J. Maddahi, D. D. Watson, G. A. Beller, H. W. Strauss, C. A. Boucher, M. Picard, B. L. Holman, and R. Fridrich, Technetium-99m hexakis 2-methoxyisobutyl isonitrile: Human biodistribution, dosimetry, safety, and preliminary comparison to thallium-201 for myocardial perfusion imaging, *J. Nucl. Med*., 30: 301–311, 1989.
48. B. M. W. Tsui, X. Zhao, E. C. Frey, and W. H. McCartney, Quantitative single-photon emission computed tomography: Basics and clinical considerations, *Semin. Nucl. Med*., 24: 38–65, 1994.
49. J. Caldwell, D. Williams, G. Harp, J. Stratton, and J. Ritchie, Quantitation of size of relative perfusion defect by single-photon emission computed tomography, *Circulation*, 70: 1048–1056, 1984.
50. NEMA, *Performance Measurements of Scintillation Cameras*, Standards Publication No. NUl-1980, Washington, DC: National Electrical Manufacturers Association, 1980.
51. G. Muehllehner, R. H. Wake, and R. Sano, Standards for performance measurements in scintillation cameras, *J. Nucl. Med*., 22: 72–77, 1981.
52. AAPM, *Scintillation Camera Acceptance Testing and Performance Evaluation*, AAPM Report No. 6, New York: American Association of Physicists in Medicine, 1980.
53. AAPM, *Computer-aided Scintillation Camera Testing and Performance Evaluation*, AAPM Report No. 9, New York: American Association of Physicists in Medicine, 1987.
54. NEMA, *Performance Measurements of Scintillation Cameras*, Standards Publication No. NU 1-1986, Washington, DC: National Electrical Manufacturers Association, 1986.
55. NEMA, *Guide to Revised Standards for Performance Measurements of Scintillation Cameras*, Washington, DC: National Electrical Manufacturers Association, 1986.
56. L. S. Graham, Quality assurance of Auger cameras, in D. V. Rao, R. Chandra, and M. C. Graham (Eds.), *Physics of Nuclear Medicine Recent Advances*, Monograph 10 of the American Association of Physicists in Medicine (AAPM), New York: American Institute of Physics, pp. 68–82, 1984.
57. L. A. O'Tuama and S. T. Treves, Brain single-photon emission computed tomography for behavior disorders in children, *Semin. Nucl. Med*., 23(3): 255–264, 1993.
58. G. B. Saha, W. J. MacIntyre, and R. T. Go, Radiopharmaceuticals for brain imaging, *Semin. Nucl. Med*., 24(4): 324–349, 1994.
59. S. Steien and J. Aaseth, Thallium-201 as an agent for myocardial imaging studies, *Analyst*, 120(3): 779–781, 1995.
60. A. Taylor, Jr. and J. V. Nally, Clinical applications of renal scintigraphy, *Am. J. Roentgenol*., 164: 31–41, 1995.

61. R. Schlosser and S. Schlegel, D2-receptor imaging with [^{123}I]IBZM and single photon emission tomography in psychiatry: A survey of current status, *J. Neural Transm. Gen. Sect*., 99(1–3): 173–185, 1995.
62. T. Sunderland, G. Esposito, S. E. Molchan, R. Coppola, D. W. Jones, J. Gorey, J. T. Little, M. Bahro, and D. R. Weinberger, Differential cholinergic regulation in Alzheimer's patients compared to controls following chronic blockade with scopolamine: A SPECT study, *Psychopharmacology*, 121(2): 231–241, 1995.
63. B. Shapiro, J. C. Sisson, B. L. Shulkin, M. D. Gross, and S. Zempel, The current status of meta-iodobenzylguanidine and related agents for the diagnosis of neuroendocrine tumors, *Q. J. Nucl. Med*., 39(4 Suppl. 1): 3–8, 1995.
64. B. Shapiro, Imaging of catecholamine-secreting tumours: Uses of MIBG in diagnosis and treatment, *Baillieres Clin. Endocrinol. Metabol*., 7(2): 491–507, 1993.
65. M. P. Stokkel, B. M. Kwa, and E. K. Pauwels, Imaging and staging of small-cell lung cancer: Is there a future role for octreotide scintigraphy? *Br. J. Clin. Pract*., 49: 235–238, 1995.
66. E. P. Krenning, D. J. Kwekkeboom, J. C. Reubi, P. M. Van Hagen, C. H. van Eijck, H. Y. Oei, and S. W. Lamberts, [^{111}In]-octreotide scintigraphy in oncology, *Metab. Clin. Exp*., 41(9 Suppl. 2): 83–86, 1992.
67. E. J. Hoffman and M. E. Phelps, Positron emission tomography: Principles and quantitation, in M. Phelps, J. Mazziotta, and H. Schelbert (Eds.), *Positron Emission Tomography and Autoradiography: Principles and Applications for the Brain and Heart*, New York: Raven Press, pp. 237–286, 1986.
68. R. A. Koeppe and G. D. Hutchins, Instrumentation for positron emission tomography: Tomograph and data processing and display systems, *Semin. Nucl. Med*., 22(3): 162–181, 1992.
69. M. E. Casey and R. Nutt, A multicrystal two dimensional BGO detector system for positron emission tomography, *IEEE Trans. Nucl. Sci*., NS-33: 460–463, 1986.
70. C. L. Melcher and J. S. Schweitzer, Cerium-doped lutetium oxyorthosilicate: A fast, efficient new scintillator, *IEEE Nuclear Science Symp*., Santa Fe, NM, November 1991 (Abstract).
71. K. Wienhard, M. Dahlbom, L. Eriksson, C. Michel, T. Bruckbauer, U. Pietrzyk, and W. D. Heiss, The ECAT EXACT HR: Performance of a new high resolution positron scanner, *J. Comput. Assist. Tomogr*., 18: 110–118, 1994.
72. T. R. DeGrado, T. G. Turkington, J. J. Williams, C. W. Steams, J. M. Hoffman, and R. E. Coleman, Performance characteristics of a whole body PET scanner, *J. Nucl. Med*., 35: 1398–1406, 1994.
73. J. A. Karp, G. Muehllehner, D. A. Mankoff, C. E. Ordonez, J. M. Ollinger, M. E. Daube-Witherspoon, A. T. Haigh, and D. J. Beerbohm, Continuous-slice PENN-PET: A positron tomograph with volume imaging capability, *J. Nucl. Med*., 31: 617–627, 1990.
74. C. J. Pavlin, K. Harasiewicz, M. D. Sherar, and F. S. Foster, Clinical use of ultrasound biomicroscopy, *Ophthalmology*, 98: 287–295, 1991.
75. D. H. Turnbull, High frequency ultrasound imaging, in L. Goldman and J. B. Fowlkes (Eds.), *Medical CT and Ultrasound, Current Technology and Applications*, Madison, WI: Advanced Medical Publishing, pp. 285–297, 1995.
76. P. N. T. Wells, Propagation of ultrasound waves through tissue, in G. D. Fullerton and J. A. Zagzebski (Eds.), *Medical Physics of CT and Ultrasound: Tissue Imaging and Characterization*, AAPM Monograph 6, College Park, MD: American Association of Physicists in Medicine, pp. 367–387, 1980.
77. J. A. Zagzebski, *Essentials of Ultrasonic Physics*, St. Louis, MO: Mosby, 1996.
78. K. Shung, In vitro experimental results on ultrasonic scattering in biological tissues, in K. Shung and G. Thieme (Eds.), *Ultrasonic Scattering in Biological Tissues*, Boca Raton, FL: CRC Press, pp. 291–312, 1993.
79. M. Insana, Sound attenuation in tissue, in L. Goldman and J. B. Fowlkes (Eds.), *Medical CT and Ultrasound, Current Technology and Applications*, Madison, WI: Advanced Medical Publishing, pp. 15–33, 1995.
80. D. H. Turnbull, Fundamentals of acoustic transduction, in L. Goldman and J. B. Fowlkes (Eds.), *Medical CT and Ultrasound, Current Technology and Applications*, Madison, WI: Advanced Medical Publishing, pp. 49–66, 1995.

81. K. Ferrara and G. DeAngelis, Color flow mapping, *Ultrasound Med. Biol.*, 23(3): 321–345, 1997.
82. K. J. W. Taylor, P. N. Burns, and P. N. T. Wells, *Clinical Application of Doppler Ultrasound*, 2nd edn., New York: Raven Press, pp. 19–34, 55–94, 1995.
83. AIUM, *Standard Methods for Measuring the Performance of Pulse Echo Ultrasound Imaging Equipment*, Laurel, MD: American Institute of Ultrasound in Medicine, 1991.
84. H. Lopez, M. H. Loew, P. F. Butler, M. C. Hill, and R. M. Allman, A clinical evaluation of contrast-detail analysis for ultrasound images, *Med. Phys.*, 17: 48–57, 1993.
85. AIUM, *Methods for Measuring Performance of Pulse-Echo Ultrasound Equipment—Part II: Digital Methods*, Laurel, MD: American Institute of Ultrasound in Medicine, 1995.
86. J. J. Rownd, E. L. Madsen, J. A. Zagzebski, G. R. Frank, and F. Dong, Phantoms and automated system for testing the resolution of ultrasound scanners, *Ultrasound Med. Biol.*, 23(2): 245–260, 1997.
87. AIUM, *Standard Specification of Echoscope Sensitivity and Noise Level Including Recommended Practice for such Measurements*, Laurel, MD: American Institute of Ultrasound in Medicine, 1991.
88. AIUM, *Quality Assurance Manual for Gray Scale Ultrasound Scanners*, Laurel, MD: American Institute of Ultrasound in Medicine, 1995.
89. C. M. Rumack, S. R. Wilson, and J. W. Charboneau, *Diagnostic Ultrasound*, 2nd edn., St. Louis, MO: Mosby, 1997.
90. P. L. Carson and M. M. Goodsitt, Pulsed echo acceptance and quality control testing, in L. Goldman and J. B. Fowlkes (Eds.), *Medical CT and Ultrasound, Current Technology and Applications*, Madison, WI: Advanced Medical Publishing, pp. 155–196, 1995.
91. C. P. Slichter, *Principles of Magnetic Resonance*, Berlin, Germany: Springer-Ver1ag, 1980.
92. F. Bloch, Nuclear induction, *Phys. Rev.*, 70: 460–485, 1946.
93. W. A. Edelstein, J. M. S. Hutchinson, G. Johnson, and T. Redpath, Spin-warp imaging and applications in whole body imaging, *Phys. Med. Biol.*, 25: 751–756, 1980.
94. W. S. Hinshaw and A. H. Lent, An introduction to NMR imaging: From Bloch equation to the imaging equation, *Proc. IEEE*, 71: 338–350, 1983.

Further Information

Bracewell R. N., *The Fourier Transform and Its Applications*, 2nd edn., New York: McGraw-Hill, 1978.
Brigham E. O., *The Fast Fourier Transform*, Englewood Cliffs, NJ: Prentice-Hall, 1974.
Brown M. A. and R. C. Semelka, *MRI: Basic Principles and Applications*, New York: Wiley-Liss, 1995.
Dainty J. C. and R. Shaw, *Image Science*, New York: Academic Press, 1974.
Fukushima E. and S. B. W. Roeder, *Experimental Pulse NMR. A Nuts and Bolts Approach*, Reading, MA: Addison-Wesley, 1981.
Huang H. K., *Elements of Digital Radiology*, Englewood Cliffs, NJ: Prentice-Hall, 1987.
Macovski A., *Medical Imaging Systems*, Englewood Cliffs, NJ: Prentice-Hall, 1983.
Sprawls P. and M. J. Bronskill (Eds.), *The Physics of MRI*: *1992 AAPM Summer School Proceedings*, Woodbury, NY: American Association of Physicists in Medicine, 1993. (An excellent overview of current MR instrumentation and clinical applications written by experts in the field.)

68

Blood Chemistry Measurement

Terry L. Rusch
Douglas Scientific

Ravi Sankar
University of South Florida

68.1 Introduction

The study of blood and its effects on sustaining life dates back to the start of understanding human anatomy. There is a diverse background associated with blood chemistry, and knowledge of the makeup and purpose of blood and its components is growing by leaps and bounds. The typical discussion and history of blood chemistry relate to oxygen transport and its effects. This chapter will deal mainly with the measurement of oxygen related to health and critical-care applications, with limited descriptions of noncritical health status measurements. The beginning of blood analysis and the measurement of a person's health status were not practically and conveniently performed until technology advanced. It has been within the last 50 years that routine monitoring began in critical care. Critical-care health monitoring is the focus of this discussion because it has been one of the main motivations for developing analytical measurements. However, most critical-care monitoring has led to diagnostic tools and functional assessment of blood and its components.

Current healthcare professionals use a variety of common analytical measurements to assess and maintain a person's health. The measurements of interest include hemoglobin, hematocrit (Hct), blood gases (O_2 and CO_2), pH, glucose, and concentrations of electrolytes. Blood gas analysis is divided into a wide variety of specific tests. The tests include arterial, mixed venous, and transcutaneous oxygen tension; carbon dioxide tension; oxygen and carbon dioxide concentration; and oxygen saturation. Each test has a specific purpose in diagnostics. Continuous real-time monitoring of oxygen saturation by pulse oximetry measurements is routinely used in anesthesia and critical-care settings. However, to determine heart and lung efficiency or true tissue oxygenation, other measurements are required. Different measurements are required because of the complex nature of the human body, and no one measurement gives a complete picture of overall status. For example, arterial oxygen saturation does not give true cardiac output efficiency nor does it give a true tissue

oxygenation state. Measurements can also be very misleading. During normal conditions, the measured values correspond very well. For example, under a given temperature and pH, oxygen saturation corresponds to oxygen tension. A fixed relationship occurs under these conditions. This does not mean the same relationship applies to all circumstances. A change in temperature or pH will shift the oxygen tension to a saturation relationship.

68.2 Background

Most blood chemistry measurements have a long history. In the late 1800s, the hydrogen (H^+) electrode was used to measure pH and, in 1925, John Peters at Yale showed a relationship between pH and CO_2 content. The measurement of blood gas started in the early 1920s with *in vitro* analysis of oxygen saturation. It was not until much later, however, that common patient care measurements were accepted. For example, it was not until 1952 in Copenhagen that pH monitoring was performed on artificially ventilated patients. The monitoring of pH was the start of the clinical use of blood gas monitoring. Until that time, physiological measurements were performed only in the laboratory setting. Electrochemical measurements were more common and performed with a Clark electrode or similar electrode arrangement. Leland Clark made a working model of an oxygen electrode around 1950, which introduced a practical method of measuring oxygen tension. The probe was originally developed to aid in heart bypass instruments. Not long after the Clark electrode was developed, optical transmission analysis tools were developed. The measurement of oxygen saturation, called oximetry, was one of the early optical measurements. Oximetry has a long history dating back more than 70 years to the 1930s. Matthes and Millikan recorded the earliest noninvasive reading around 1935, and Glen Millikan introduced the name *oximetry* in 1942. A real research effort was started during World War II as a part of military aviation. It took almost 70 years for pulse oximetry to become a standard piece of equipment in operating rooms, critical-care units, and emergency health care. On January 1, 1990, the American Society of Anesthesiologist (ASA) made intraoperative monitoring with pulse oximetry a standard [1]. Despite the acceptance of pulse oximetry and its benefits, there still has been little change in pulse oximetry techniques. A good historic perspective on most blood gas analysis is written in a series of articles in the *Journal of Clinical Monitoring* and in *International Anesthesiology Clinics* [2–7].

68.2.1 Oxygen Delivery

As Severinghaus and Astrup [6] note in the history of blood gas analysis, nothing is more important than oxygen supply for life. Therefore, it can be argued that oxygen content and consumption should be one of the most important things to monitor in critical and unconscious patients. It is important to look at the mechanisms for oxygen transport. All human tissue needs oxygen to function, and it is therefore critical to understand how oxygen is delivered. *Hemoglobin* is the binding agent for oxygen in blood, but it is necessary to oxygenate the blood and move it to the tissue that uses the oxygen. A simple explanation is as follows. Oxygen is diffused into blood via the lungs. Blood is then circulated by a system of *arteries, veins*, and *capillaries* where the oxygen is diffused to tissue. The tissue also diffuses carbon dioxide to blood in the capillaries, which transports back to the lungs, where the carbon dioxide is diffused to the lungs for expiration. Figure 68.1 shows a diagram of the general makeup of the components of interest. This is a very simplistic view of blood and oxygen transport. There are many other regulatory factors, such as sodium, potassium, calcium, pH, and glucose, but this simplistic view is a good starting point to understand the measurements that follow. It is also important to note that several measurement options and locations exist for gathering the same or similar data. This makes understanding the interaction and importance of different measurements difficult. For example, arterial and venous blood oxygen tensions are interdependent but have different readings. Different physical states and health problems cause varying interdependence. A simple requirement to identify whether or not cell tissue is

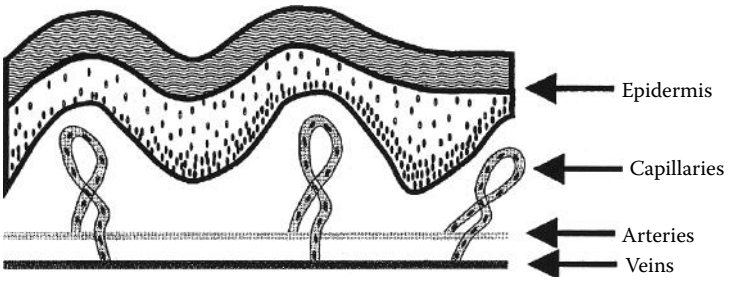


FIGURE 68.1 Components involved in oxygen transport. Oxygen is delivered to the epidermis by arteries and exchanged through the capillaries. Veins return the by-product.

being properly oxygenated becomes very complex. Oxygen delivery to tissue is determined by the differential in partial pressure between cell tissue and capillaries. The partial pressure difference determines the rate and efficiency of diffusion. However, the rate is also dependent on blood flow, concentration of hemoglobin (CtHb), and the saturation of oxygen. No individual parameter or measurement gives a complete picture of cell oxygen consumption.

68.2.2 Blood Hemoglobin

Blood contains many components, but hemoglobin is the main transporter of oxygen (O_2) to tissue. In addition, blood is responsible for removing the by-product of oxygen expenditure, carbon dioxide (CO_2). Hemoglobin is a good binding agent for oxygen, producing oxyhemoglobin (O_2Hb), also called oxygenated hemoglobin. Hemoglobin is a combination of four peptide chains, each containing several hundred amino acids. Hemoglobin is composed of two elements, heme and globin. Heme is the iron pigment of the electrolytes, and globin is a simple protein. Hemoglobin contains about 6% heme and 94% globin. Hemoglobin is the principal component of red blood cells. Unbound hemoglobin is called deoxyhemoglobin (RHb) or reduced hemoglobin. Hemoglobin also binds to produce additional components, such as carboxyhemoglobin (COHb) and methemoglobin (MetHb). MetHb is a result of oxidation of ferrous iron in hemoglobin [8]. There are other compounds that are found in blood, but they are usually in very low concentrations and are of interest in more specialized cases.

68.2.3 Measurement Methods

Most blood chemistry analysis measurements are derived from one of two methods: electrochemical and optical. There are other specific measurement techniques, such as gas chromatography, but they are more of a specialized measurement and will not be addressed. Electrochemical measurements are based on the Clark electrode with the blood component or electrolyte of interest using a different ion-specific electrode. The Clark electrode consists of an electrode in a medium. The first electrode for oxygen tension measurements was a platinum electrode and a silver anode. The current or voltage generated at the electrode is measured and is proportional to the amount of that component. The relationship is usually a calculated concentration derived from empirical data. For optical measurements, there are many approaches, but all are based on absorption, reflection, scattering, and fluorescence techniques. One optical probe has been called an optode. The optode is a cross measurement of optical fluorescence and electrochemical measurement. Figure 68.2 shows a diagram of one configuration for an optode arrangement. In an optode, an ion-selective membrane is used to diffuse an ion or compound into the fluorescent dye measurement chamber. The ion is associated with a fluorescent dye, excited with a light source, and the emitted fluorescent light is measured. Many measurements can be obtained, such as oxygen tension, carbon dioxide tension, pH, potassium concentration, calcium concentration, and chloride concentration.

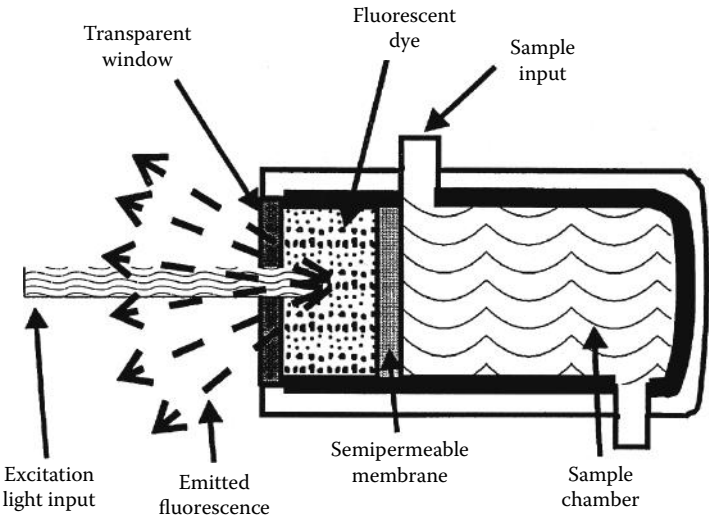


FIGURE 68.2 Chemical fluorescent optode for oxygen tension measurements. Light input through a window excites encased dye. The sample interacts with the dye through a semipermeable membrane. The sample is introduced to the dye via a flow-through chamber.

Measurements are listed in one of four categories, (1) invasive or (2) noninvasive measurements and (3) continuous or (4) periodic measurements. *Invasive* measurements can be continuous or periodic and can be performed *in vivo* or *in vitro*. Invasive *catheter* sensors can record real-time data continuously with no loss of blood sample. *Intra-arterial* measurements can be drawn periodically for minimally invasive measurements with no sample loss, or samples can be withdrawn and discarded. Periodic samples can also be drawn for analysis and measured in a physically different location at a later time. Noninvasive measurements can also be continuous or periodic but are generally on patient, real-time measurements. Noninvasive measurements are always preferred over invasive measurements, because invasive measurements increase risk of infection and usually mean some delay between the time of acquisition and the time results are available. However, noninvasive measurements are not always as accurate as invasive measurements.

68.3 Measurements and Techniques

There are typically several measurement techniques that yield the same or similar data. Each technique will give information about different physical states, such as cardiac efficiency or tissue oxygenation state. For example, oxygen saturation and partial pressure of oxygen are related. However, the exact relationship is dependent on pH and temperature. A text by Kenneth McClatchey on *Clinical Laboratory Medicine*, the International Federation of Clinical Chemistry (IFCC), and instrument operator's manuals are good references for laboratory protocols relating to invasively drawn samples [9,10]. Sample preparation and handling protocols, as well as standard measurement levels and physical states relating to excess or deficient readings, are detailed. The range, price, and suppliers for instrumentation are extensive. A standard midrange laboratory blood gas analyzer can cost from $5,000 to $50,000. The more common measurements are described in the following text.

68.3.1 Hemoglobin and Hematocrit Concentration

The total CtHb or Hct indicates the oxygen-carrying potential of blood. The combination of amount of hemoglobin, partial pressure of oxygen or percent oxygen saturation, and rate of flow

of hemoglobin determines the amount and efficiency of tissue oxygenation. Invasively drawn samples are typically used to measure hemoglobin and Hct values.

68.3.1.1 Hemoglobin

The total hemoglobin concentration is not the same as the red blood cell count, because red blood cells have different amounts of hemoglobin. The total hemoglobin concentration is measured optically by absorptive intensity at the isosbestic point. The Beers–Lambert law, also referred to as Beer's law, shown in Equation 68.1 regulates the absorptive property of a substance:

$$I_t = I_0 e^{-DC\alpha_e} \tag{68.1}$$

where

I_t is the transmitted intensity
I_0 is the incident intensity on the sample
D is the distance light travels through the substance
C is the concentration of the solution
α_e is the extinction coefficient at a specified wavelength

The isosbestic point is the crossover point in extinction curves for oxyhemoglobin and RHb. The wavelength of 805 nm is an isosbestic point, and at this wavelength the absorption is independent of hemoglobin type. If blood is assumed to be composed of only O_2Hb and RHb, then the absorbance at the isosbestic point determines the total concentration of Hb. The most common error in assuming only O_2Hb and RHb is during elevated COHb. COHb, however, generally is optically indistinguishable from O_2Hb for absorption. An alternative approach to isosbestic measurement is to calculate CtHb. Assuming only O_2Hb and RHb, the total Hb concentration is simply the sum of the two concentrations. A more accurate sum is shown in Equation 68.2:

$$\text{CtHb} = \text{O}_2\text{Hb} + \text{RHb} + \text{COHb} + \text{MetHb} \tag{68.2}$$

Here, the four most common Hb derivatives are used to calculate a more accurate CtHb. The individual Hb derivatives are measured as outlined later.

68.3.1.2 Hematocrit

Hct is the volumetric fraction occupied by red blood cells and is generally measured by conductivity. Hct is also referred to as the packed cell volume (PVC). Hct can be determined by conductivity based on the plasma ion content. Hct does not contribute to the conductivity and therefore is inversely proportional to the conductivity. Hct can also be determined optically in various ways. One way is through optical density measurements, where the total optical density is the sum of optical absorbance and optical scattering density. The total optical density is linearly proportional to Hct, and, at clinically relevant Hct levels of 20%–40%, scattering is dominant over absorption. One optimization study showed that an optimum wavelength of 624 nm, at a measured angle of 90° from the incident light, gave an inverse linear intensity to the Hct level [11].

68.3.2 Oxygen Tension

Oxygen tension or the partial pressure of oxygen (PO_2) is a common measure of oxygenation states. The partial pressure of oxygen in hemoglobin determines how well oxygen is delivered to the cell tissue of the body. The more common partial pressure reference is arterial partial pressure of oxygen (PaO_2). If the partial pressure of oxygen is higher than the surrounding tissue, oxygen is diffused to the tissue.

If the partial pressure is lower than the tissue partial pressure, no oxygen is diffused to the tissue, and tissue damage can start to occur. Equation 68.3 shows Henry's law:

$$C = \alpha_s * PO_2 \tag{68.3}$$

where
C is the concentration of oxygen (O_2)
α_s is the solubility coefficient
PO_2 is the partial pressure of oxygen

In diffusion, a partial pressure between the tissue and blood supply is trying to maintain equilibrium. Diffusion will occur until the two partial pressures are equal. Equation 68.4 shows diffusion equilibrium:

$$PO_2 = \frac{C_1}{\alpha_{s1}} = \frac{C_2}{\alpha_{s2}} \tag{68.4}$$

The rate of oxygen diffusion is dependent on the difference in partial pressure. The larger the difference in partial pressure, the faster the diffusion rate.

The partial pressure difference is partially determined by the oxygen delivery (D_o) rate. Oxygen delivery can be identified as the concentration of arterial oxygen minus the concentration of venous oxygen times the rate of blood flow (R), as shown in Equation 68.5:

$$D_o = (C_aO_2 - C_vO_2) * R \tag{68.5}$$

where
C_aO_2 is the arterial oxygen concentration
C_vO_2 is the venous oxygen concentration

The oxygen delivered, however, is not the oxygen delivered to the tissue. Some oxygen is transpired through the skin. The transpired oxygen allows an alternative approximation measurement of arterial oxygen tension. The more common oxygen tension measurement is arterial oxygen tension. Oxygen tension can be measured electrochemically, transcutaneously, or optically. The techniques are listed in the following text.

68.3.2.1 Electrochemical PO_2 Measurements

Electrochemical measurements of PO_2 are obtained using a basic Clark electrode with a platinum electrode and a silver/silver chloride reference electrode. Figure 68.3 shows a diagram of one type of Clark electrode. The electrode provides a path for the reduction in Equation 68.6:

$$O_2 + 4e^- \rightarrow 2O^- \tag{68.6}$$

The platinum electrode has an affinity with oxygen. If a reference PO_2 at the electrode is known, preferably zero, then the current depends only on the oxygen tension variations of the sample. Early electrodes were consumption measurements where the oxygen is removed and the sample is altered. Oxygen was attracted to the electrode and a current proportional to the oxygen content was observed. Improved electrodes use a semipermeable, constant diffusion membrane. An electrolyte is used with the membrane to improve response, longevity, and stability. An exposed electrode will become coated in whole blood

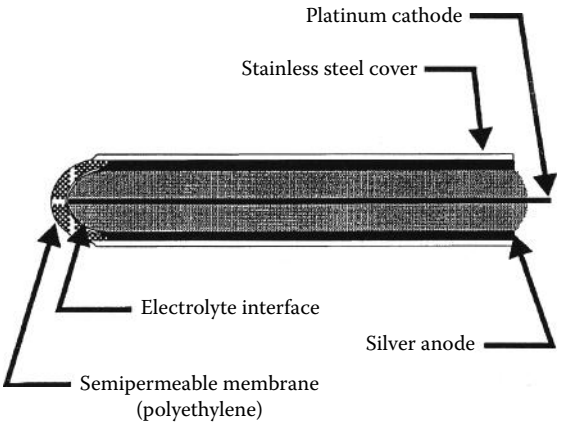


FIGURE 68.3 Clark-type needle electrodes for PO_2 measurements. Oxygen ion interaction through the membrane tip allows the ion concentration to be measured by dissimilar metals.

and the sensitivity will degrade. The membrane and electrolyte allow the diffusion of oxygen without the sample directly contacting the electrode. In addition, a properly designed membrane reduces the stirring effect noted by Clark and others using an exposed electrode. The size and location of electrodes have been extensively studied. The exact setup for the electrode can be varied for optimum performance in different areas. For example, a less permeable diffusion layer reduces the stirring effect and makes the probe more stable, but the result is a slower response time [3,6,12,13].

68.3.2.2 Transcutaneous Partial Pressure of Oxygen

Partial pressure and oxygen saturation are not always true indicators of actual tissue oxygen consumption. It has been known since 1851 that oxygen is respired from living tissue. A combination of heat and optical means can be used to determine the amount of oxygen expired and correlated to arterial partial pressure under controlled conditions. This method of measuring arterial oxygen partial pressure is called transcutaneous oxygen partial pressure ($P_{tc}O_2$). There are, however, many complications in the measurement of $P_{tc}O_2$. Many studies have tried to correlate arterial oxygen saturation with the measured expired oxygen. The most accurate measurements are made by maintaining a constant temperature as high as 45 °C, to assure perfusion. This causes the complication of burns to skin, and sensors need to be moved on a regular basis.

Measurements can be performed polarographically using Clark electrodes or alternatively by mass spectroscopy. Because of the nature of $P_{tc}O_2$, adult measurements are not common, but premature infant hemoglobin and skin are more responsive to $P_{tc}O_2$ measurements, which have had a place in monitoring neonatal oxygenation [7,14].

68.3.2.3 Optical-Based PO_2

There are many variations on optical-based blood gas measurements. The basic optical measurement for PO_2 is a *fluorescent* measurement. A fiber-optic cable is used to excite a fluorescent dye remotely. The excited dye emits a higher-wavelength signal, known as Stokes shift. The emitted signal is measured and is correlated to the PO_2 value. Figure 68.4 shows a configuration for a fluorescent sensor. Oxygen is used as a fluorescence quencher. Oxygen has a fluorescence-quenching property and attenuates the fluorescent intensity. A zero state can be measured by filling the sample chamber with a zero oxygen concentration mixture. The measured intensity is compared with the incident light source and is used to calculate the concentration of the sample. For oxygen tension, the measurement is done by a fluorescent electrochromic dye, such as pyrenebutyric acid, with an ion-selective membrane such as silicone rubber.

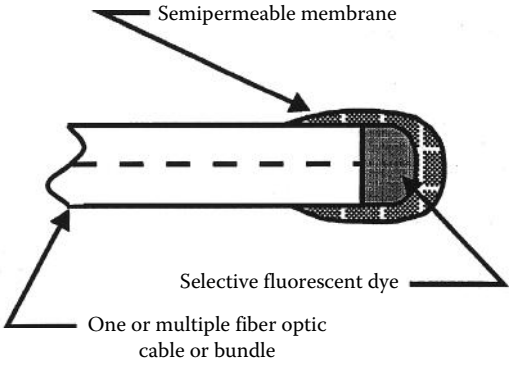


FIGURE 68.4 Chemical fluorescent fiber-optic probes. Dye is encased on a fiber tip by a membrane. Light enters the dye through the fiber and the emitted fluorescence is returned via the fiber. The amount of fluorescence is an indication of the interaction between the dye and the substance in which the fiber tip is placed.

The oxygen tension is then calculated using the Stern–Volmer quenching formula, shown in Equation 68.7, and the PO_2 value is calculated empirically:

$$\frac{I_0}{I} = 1 + K * PO_2 \quad \text{or} \quad I(PO_2) = \frac{I_0}{(1 + K * PO_2)} \tag{68.7}$$

where

$I(PO_2)$ and I_0 are the relative fluorescence intensities in the presence of oxygen and in the absence of oxygen, respectively

K is the overall quenching constant and is given in Equation 68.8:

$$K = k^+ * \alpha_O * \tau_0 \tag{68.8}$$

where

k^+ is the collisional quenching constant

α_O is the oxygen solubility coefficient

τ_0 is the mean lifetime of the excited state at zero PO_2

The ratio of I to I_0 is then plotted against the PO_2 value. This plot is used to calibrate the sensor empirically for different PO_2 values. $1/I_0$ is the intercept, and the slope is K/I_0. The system is empirically derived and errors can occur with varying excitation intensity. One way to correct for intensity variations is to use a reference fluorescence dye that is not affected by oxygen quenching. The same procedure can be used to calculate the percent oxygen concentration [15,16].

68.3.3 Oxygen Saturation

There are several readings that are interrelated regarding oxygen saturation. The most common measurements are arterial blood saturation (S_aO_2), mixed venous blood saturation (S_vO_2), and *photoplethysmogram* arterial blood saturation (S_pO_2). S_pO_2 measurements, also called pulse oximetry measurements, are arterial measurements but are only related to true arterial saturation. Care must be taken in knowing which measurements are actually taken, because of inaccuracies in various measurements. Saturation measurements can be performed invasively using reflectance or fluorescence oximetry or noninvasively by transmission or reflectance photoplethysmogram readings.

There is an important relationship between oxygen saturation and partial pressure. A plot shown in Figure 68.5 of arterial oxygen saturation vs. arterial partial pressure is called the oxygen–hemoglobin

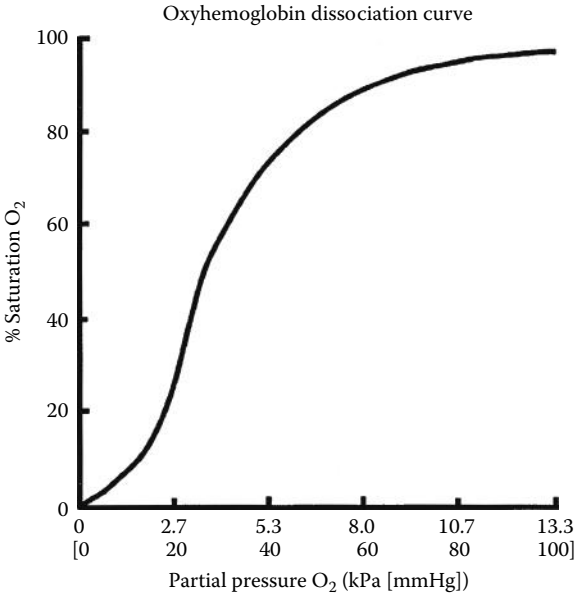


FIGURE 68.5 Oxygen saturation versus PO_2 oxygen–hemoglobin dissociation curve. The partial pressure shows a sigmoidal relationship to the percent oxygen saturation.

dissociation curve. Partial pressure falls at a linear rate and is a good indicator of changing oxygen delivery to cell tissue. The percent oxygen saturation is nonlinear with respect to partial pressure in a sigmoidal relationship. A partial pressure of 13.3 kPa (100 mmHg) is effectively 100% saturation. An increase in partial pressure above 13.3 kPa (100 mmHg) indicates free oxygen not bound to hemoglobin and in that state does not contribute much to tissue oxygenation. This means the percent saturation gives a delayed notice of desaturation. In a patient with healthy lungs, partial pressure values of greater than 24.0 kPa (180 mmHg) can be measured. A 50% decrease in partial pressure could go undetected using the percent oxygen saturation. The oxygen dissociation curve, as mentioned previously, is dependent on pH and temperature. Under normal conditions, the relationship between partial pressure and saturation is maintained, but under altered conditions, such as low perfusion, the relationship can deviate.

Roughton and Severinghaus [18] describe a computation to approximate the dissociation curve based on the Hill equation. The Hill equation is given in Equation 68.9:

$$Y = \frac{(PO_2/P_{50})^n}{1 + (PO_2/P_{50})^n} \tag{68.9}$$

where

Y is the oxygen saturation
PO_2 is the oxygen partial pressure
P_{50} is the partial pressure at 6.7 kPa (50 mmHg)

The P_{50} point is determined by the temperature and pH of the sample. Temperature and pH correction equations are also presented in Refs. [5,7,17–19].

68.3.3.1 Transmission Oximetry

Continuous S_aO_2 and S_vO_2 measurements are used to help evaluate whether oxygen is being adequately delivered to tissue. Oxygen is either consumed or expired. If oxygen is delivered to tissue and not consumed, then tissue oxygenation is not adequate. Early oximeters used invasively drawn samples and

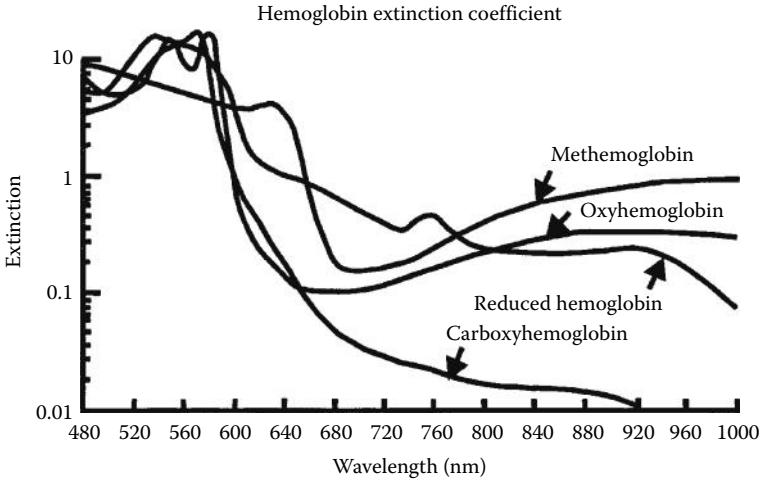


FIGURE 68.6 Hemoglobin extinction coefficients plotted by wavelength. The light extinction coefficient varies by wavelength and type of hemoglobin.

light transmission to measure oxygenation. The basic concept in oximetry is to transmit light through a blood sample, and blood absorbs a determined amount of light according to Beer's law, shown in Equation 68.1. For oximetry applications, hemoglobin is assumed to be composed of only two substances, oxygenated hemoglobin or oxyhemoglobin (O_2Hb) and deoxygenated or RHb. This is a very simplistic approach, but it is the basis for most oximetry measurements. For *in vivo* measurements, the path length for the light is constant and known. Both the O_2Hb and RHb can be measured simultaneously by using two separate wavelengths. Flash lamps and filter wheels were originally used to illuminate blood samples, and the transmitted signals were measured. Current technology uses light-emitting diodes (LEDs) and alternate on and off cycles.

If two light wavelengths are used and the two substances have different extinction coefficients (α_e), or equivalently attenuation coefficients (σ), then the percentage of each substance can be calculated. The extinction coefficients of hemoglobin are well documented. Figure 68.6 is a plot of the extinction coefficients for oxyhemoglobin, RHb, COHb, and MetHb. The two wavelengths of light yielding good results for oximetry are the red (660 nm) and infrared (940 nm) wavelengths. Red has the largest difference between the two extinction curves, and infrared has maximal difference after the isosbestic point where the two extinction curves cross. The functional arterial oxygen saturation (S_aO_2) is calculated using Equation 68.10 given the concentration of O_2Hb and RHb:

$$\text{Functional } S_aO_2 = \frac{C_o}{(C_o + C_r)} * 100 \tag{68.10}$$

The concentration of oxyhemoglobin is given by C_o and the concentration of deoxygenated hemoglobin is given by C_r. This formula must be adjusted for other contents in blood that influence measurements if present, such as COHb and MetHb. The more accurate fractional saturation is given as the total hemoglobin concentration (tHb) in Equation 68.11:

$$\text{Fractional } S_aO_2 = \frac{C_o}{(\text{tHb})} * 100 = \frac{C_o}{(C_o + C_r + C_c + C_m)} * 100 \tag{68.11}$$

where

C_c is the COHb saturation

C_m is the MetHb concentration

To account for COHb and MetHb accurately, four transmission wavelengths would be required, or for accurate S_aO_2 readings the total Hb concentration needs to be known or measured [5,7,20].

68.3.3.2 Pulse Oximetry

Pulse oximetry is based on the transmission, absorption, and dispersion of light as it passes through hemoglobin. Beer's law, as stated in Equation 68.1, determines the transmission of light through a substance. For pulse oximetry, the light illuminates both arterial and venous blood and the light must traverse all tissue between light source and receiver. Figure 68.7 represents the light path and indicates a variable (ac) path length as well as a constant (dc) path length.

To calculate the pulse oximetry digital photoplethysmogram (DPP) oxygen saturation (S_pO_2), two equations are used. An example of a DPP signal is shown in Figure 68.8 and it indicates the ac and dc DPP components. The first step is to use the red and infrared time signal to calculate an R value.

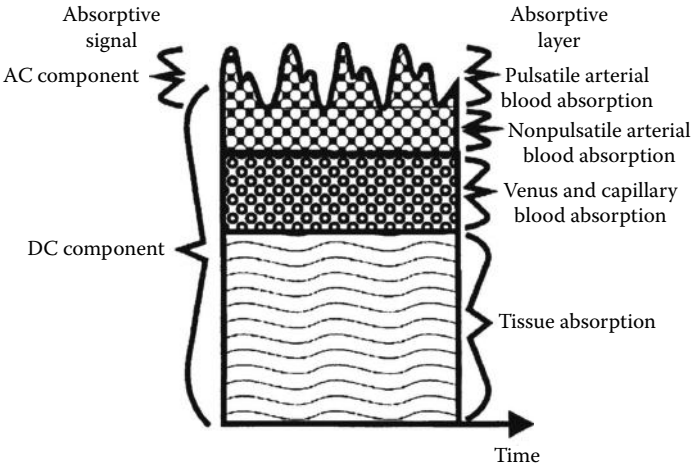


FIGURE 68.7 Absorption components encountered during transmission oximetry. The arteries, veins, capillaries, and tissue absorb light. The total light absorbed has a steady state (dc) and a varying (ac) signal component. The varying signal is due to the pulsatile arterial volume change.

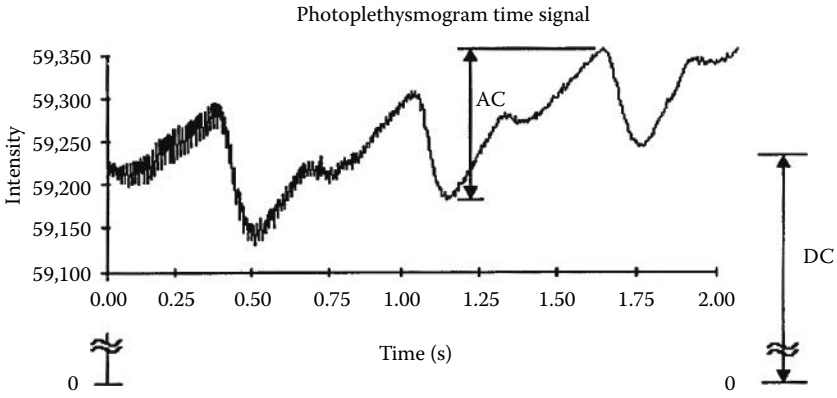


FIGURE 68.8 Photoplethysmogram with ac and dc components labeled. The light intensity transmitted through the finger varies with time. The signal contains a constant (dc) component and a varying signal (ac) component. The varying signal is due to the pulsatile arterial volume change.

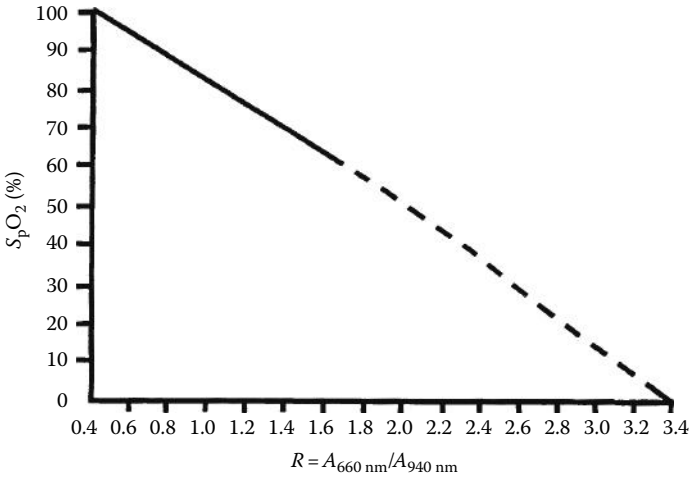


FIGURE 68.9 Measured oxygen saturation vs. *R* value. Empirically derived 660–940 nm ratio oxygen saturation. The values are used to formulate a first-order equation to calculate the pulse oximetry saturation.

The *R* value is the normalized ratio of the red to infrared transmitted light intensity and is shown in Equation 68.12:

$$R = \frac{(ac_r/dc_r)}{(ac_i/dc_i)} \tag{68.12}$$

The *R* value for two specific light wavelengths can be plotted against a measured S_aO_2 value, as shown in Figure 68.9. A linear approximation can then be used to calculate a S_pO_2 value. The empirical linear approximation for Figure 68.9 is listed in Equation 68.13:

$$S_pO_2 = 110 - 25R \tag{68.13}$$

The empirical approximation is used to correct for errors in the measured values. Pulse oximeters on the market traditionally used weighted moving average (WMA) techniques to identify the transmitted ac and dc DPP components. The dc component is the averaged signal intensity. The ac signal is computed using the WMA as a band-pass filter to single out only the ac cardiac signal.

There are many areas where pulse oximetry has limitations. One limitation was already mentioned, and that is the assumption of only two substances in hemoglobin. For most measurements, the percentage of these substances is small enough not to affect the pulse oximetry measurements. There are cases in emergency care when these substances may be present and do affect readings. At present, the user must be aware of the limitations and not use the pulse oximetry reading if the person has any of the additional substances. The solution to this problem is to use additional light sources to calculate the additional substance. The first pulse oximeters tested used a multiple-wavelength light source and could easily be implemented in new instruments.

A second limitation of pulse oximetry is background light. Because pulse oximetry currently uses transmitted light, the photodiode receiver is susceptible to ambient light. Ambient light can be from indoor lighting, sunlight, or phototherapy lights. To correct for this problem, a third light measurement can be collected with no light source and subtracted from the transmitted intensity. Using a digital microprocessor, the subtraction can be easily performed.

A third limitation is called low *perfusion*. Pulse oximetry is based on having a pulsatile signal. If the pulsatile signal is small compared with the dc signal, usually 1 count in 1000, the *R* value calculation becomes inaccurate. There are two reasons for the inaccuracy, roundoff error and resolution. The roundoff error can be compensated for by using more precision during the *R* value calculations.

Another limitation, which is difficult to eliminate, is motion artifact and autonomic nervous system response. Whenever there is an autonomic nervous system action, there is a transition or movement in the pulse oximetry signal. The most common technique to correct for motion artifact is averaging of consecutive measurements. Averaging works, but it slows the response time and lengthens the processing time. Another improvement would be to identify and eliminate inaccurate readings. This approach, however, is difficult to implement [7,21–23].

68.3.3.3 Reflectance Oximetry

An alternative technique to consider for DPP is reflectance probe measurements, as shown in Figure 68.10. Reflectance pulse oximetry has been addressed for *in vivo* studies by Mendelson and Ochs [24]. Because reflectance probes can be placed flat on the measurement area, they provide better shielding than a probe placed across a finger. However, reflectance measurements have less intensity than transmission measurements. Reflectance probe location and temperature effects have been analyzed. Increased temperature results in perfusion and increased signal intensity.

Schmidt et al. [25] reported an integrated circuit-based optical sensor for *in vivo* surface measurements. A set of equations was derived to estimate the amount of reflectance at a given intensity. The solution is a three-wavelength reflectance probe, with a red LED wavelength emission of 660 nm and a near and far infrared LED. The reflectance depth was calculated to be within 2 mm of the surface of the probe. The third wavelength is used to eliminate errors in measurement due to additional hemoglobin derivatives. The third wavelength was used to calculate a Hct value. The saturation was calculated using Equation 68.14:

$$SO_2 = A - B\left(\frac{R_{805}}{R_{660}}\right) \tag{68.14}$$

where A and B are constants empirically derived for a specific light source and detector under specific physiological conditions.

A current method to measure S_vO_2 is by reflectance oximetry methods. A bundled fiber-optic cable is placed through a catheter. Light enters by a fiber or fibers and the reflected light is analyzed for oxygen content using spectrophotometry. The probe is similar to the fluorescence probe shown in Figure 68.4 with no chemistry envelope. The reflectance is used to determine the oxygen concentration empirically [24–26].

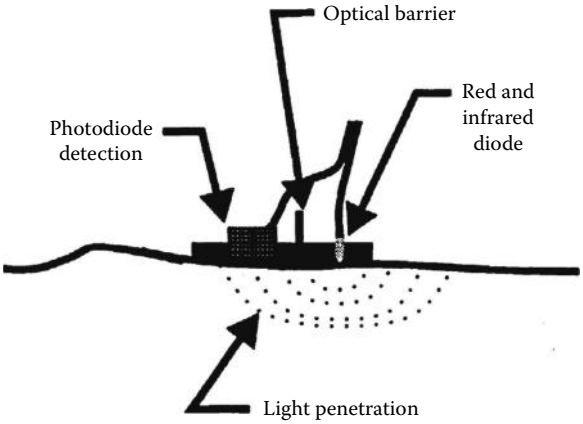


FIGURE 68.10 Reflectance oximetry probe. Red and infrared diodes illuminate the surface alternately and the reflected light is detected in close proximity by a photodiode. An optical barrier is required to eliminate direct light detection.

68.3.4 pH Measurements

Blood pH is very important in sustaining life. There is a small range of values allowable to maintain life, and pH is one of the most tightly regulated parameters in the body. Typical pH measurements are taken invasively either through a catheter probe or as a blood sample measurement. pH is noted as the negative decade logarithm of the molal activity of hydrogen ions. The acid and alkaline hydrogen reactions are given in Equations 68.15 and 68.16, respectively:

$$O_2 + 4H^+ + 4e^- \rightarrow 2H_2O \tag{68.15}$$

$$O_2 + 2H_2O + 4e^- \rightarrow 4OH^- \tag{68.16}$$

The pH level can be determined by electrochemical means. The original pH measurements were performed using hydrogen electrodes after the oxygen (O_2) was eliminated. Current electrochemical measurements use a glass electrode sensitive to hydrogen and a concentrated KCl bridge.

Optical measurements of pH are performed fluorescently. It was determined that pH corresponds to CO_2 tension. Both PCO_2 and pH can be measured using a dye sensitive to hydrogen ions and using a variant of the Henderson–Hasselbalch equation, shown in Equation 68.17:

$$\frac{I_0}{I} = 1 + K * H^+ \tag{68.17}$$

where
- I is the relative fluorescence intensity
- H^+ is the hydrogen ion concentration
- I_0 and K are calibration constants

pH fluorescence measurements can be obtained using a pH-sensitive dye, such as phenol red. The basic form of phenol red is green absorbing, and the acidic form is blue absorbing. Exciting the phenol red buffer with a green 560 nm and red 600 nm light, the ratio of intensities can be used to calculate the pH according to Equation 68.18:

$$R = k * 10^{[-C/(10^{-\delta}+1)]} \tag{68.18}$$

where
- k is the system optical constant
- C is the green base form intensity
- δ is the difference between the pH and pK of the dye

Intensity shifts are accounted for by measuring the difference in intensity of two different excitation wavelengths. An alternative two-wavelength approach using hydroxypyrene trisulfonic acid (HTA) can be used. The basic form of HTA has a maximum excitation at a wavelength of 460 nm and an acidic maximum excitation wavelength of 410 nm. The ratio of the two fluorescent intensities at 520 nm is used to calculate the pH value [2,6,16,27].

68.3.5 Glucose Measurement

A common glucose measurement technique is by enzymatic amperometric measurements. Through selective binding, glucose and other compounds can be measured. The bound glucose changes the conductance and the glucose concentration can be determined.

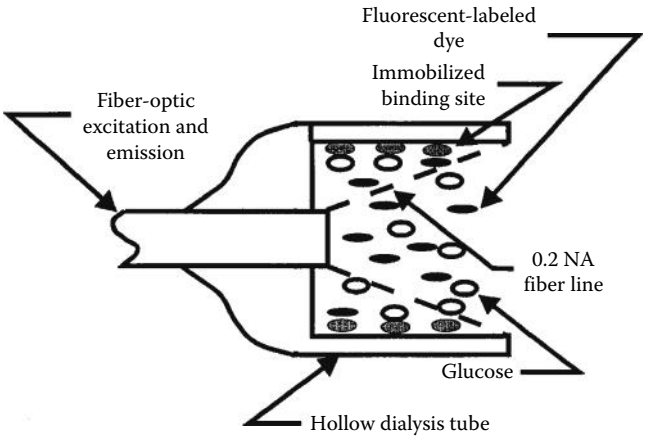


FIGURE 68.11 Optical glucose probe. A hollow tube is used to localize fluorescent dye outside the numerical aperture of the fiber. The fiber is used both to introduce excitation light and to collect emitted fluorescent levels. Only dye in the fiber view is illuminated. Dye bound to the hollow tube is not illuminated and is not detected.

Another more recent and interesting measurement technique is enzymatic optical measurements. Competitive binding between glucose and a fluorescein-labeled analog is used to determine glucose concentrations. An immobilized binding site is used to bind a fluorescent material, such as fluorescein-labeled dextran. A diagram of a glucose sensor is shown in Figure 68.11. The binding sights are fixed outside the excitation illumination. As the glucose concentration increases, the bound fluorescent material is released into the optical path. As the concentration of glucose increases, the viewed fluorescence increases and the reflected intensity increases. The reflected intensity will follow the concentration of glucose, and the actual response is measured empirically. The probe operation is reversible, allowing the probe to be reused. Reversibility is a requirement for implantable probes, which is the intended application for the optical glucose probe [27,28].

68.3.6 Electrolyte Concentration Measurements

In the clinical setting, it is important to monitor the concentration of various *electrolytes*. Typical measurements include sodium (Na^+), calcium (Ca^{++}), potassium (K^+), chloride (Cl^-), and magnesium (Mg^{++}). The concentrations are typically invasively drawn samples measured in a clinical laboratory. The measurement is performed using an ion-specific electrode, which is similar to the Clark electrode described previously. The clinical significance of different electrolytes is outlined in the text by McClatchey, *Clinical Laboratory Medicine* [9]. Optical measurement techniques have been performed but are not typically used in the clinical setting and are not covered here.

68.4 Combined Analysis Techniques

There are many combinations of instruments available. As instruments become more accurate and smaller, the trend is to simplify measurements and increase usefulness of equipment. Several selected instruments are described that have combined blood chemistry measurements into one unit.

68.4.1 CO-Oximetry

CO-oximetry has been noted for some time as the gold standard for oximetry readings. The more typical CO-oximeter does not measure just oxygen saturation but several hemoglobin concentrations.

For example, the Instrumentation Laboratory, Inc., IL 282 CO-oximeter uses an invasively drawn sample and measures total hemoglobin, oxyhemoglobin, COHb, MetHb, and oxygen content in blood. The measurement technique is optical absorption. Four wavelengths at 535, 585.2, 594.5, and 626.6 nm are generated using a hollow cathode lamp. The extinction coefficients are used to calculate the concentration of RHb, O_2Hb, COHb, and MetHb. The total hemoglobin (tHb) concentration is calculated as the sum of the individual hemoglobin groups. The oxygen saturation (S_aO_2) is calculated as the concentration of oxyhemoglobin divided by the total CtHb. The total hemoglobin concentration, however, is usually noted as only the oxyhemoglobin concentration divided by the concentration of RHb and oxyhemoglobin. Therefore, care must be used in comparisons, because variations in S_aO_2 calculations can occur. The oxygen content is calculated as 1.39 times the concentration of oxyhemoglobin [29].

68.4.2 Intra-Arterial Probes

Intra-arterial probes have drawn much interest. They allow continuous or periodic measurements in the clinical setting in real time. These probes also offer minimally invasive measurements. The intent is to use intra-arterial measurements on patients who require catheterization. The goal is a small probe that can be inserted into a catheter and not interfere with administering therapies or blood sampling. An optode-type probe is described by Shapiro et al. [30,31]. A fiber-optic probe is described by Gehrich et al. [15] for blood gas monitoring. The probe is designed to measure pH, PCO_2, and PO_2. In addition, a thermocouple is used to adjust for temperature variations, and values are normalized to standard temperatures. The sample drawing of the probe is shown in Figure 68.12. A set of chemical optical fluorescent probes is used.

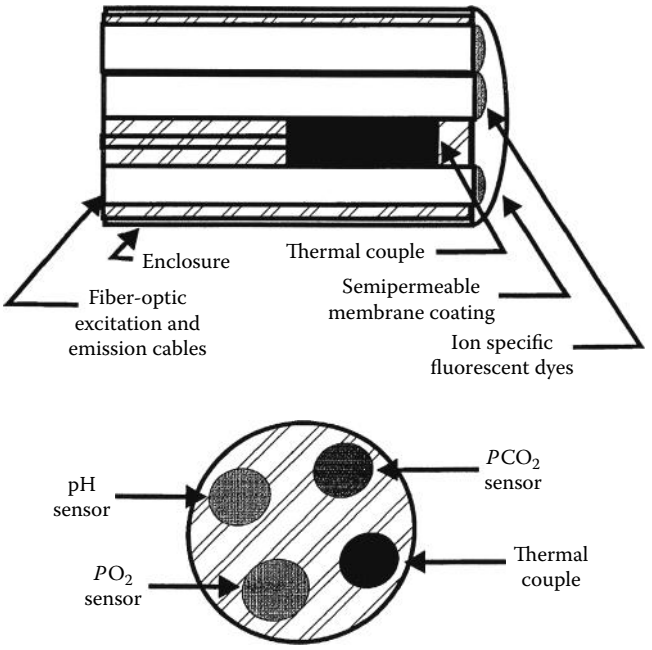


FIGURE 68.12 Combined fiber-optic sensor for pH, PCO_2, PO_2, and temperature correction measurements. Each sensor is bundled into a single enclosure and has separate detection areas.

68.5 Evaluation of Technology

There have been numerous measurement types and techniques used in developing an understanding of blood chemistry. The diverse background has resulted in very good indicators of how blood is used to sustain life. Despite the diverse background, many changes and advances are occurring at an impressive rate. Immediate improvements are being tested and implemented on a regular basis, and novel approaches will be seen in the near future. In addition, there is an explosion of studies being performed to understand how the body works and is put together. A brief overview of some improvements that can be expected follows.

Some basic engineering changes could be performed to enhance accuracy. Oximetry measurements could easily see engineering improvements. New techniques have been tested, but approval and acceptance are slow. There has been progress made in size, with new oximeters that can clip on the finger. The accuracy has seen only moderate improvement. Alternative probes and algorithms have been tested and even patented but have not been marketed. One example of a patent is for Fourier analysis of pulse oximetry signals for arterial saturation measurements. The spectral analysis would allow the possibility to implement some novel filter and analysis routines. This would improve stability from motion artifact. Hardware changes could also bring changes. Resolution could be improved by using more accurate photodiodes and A/D converters or by using laser diodes for more intense and accurate light illumination. Reflectance probe advances would improve versatility.

The future of blood chemistry should see continued miniaturization. Fiber optics have been a major research direction. Fiber optics allow remote sensing and have reduced the possibility of electrical and magnetic complications for patient monitoring. Miniature fiber-optic sensors are being developed for numerous applications, and the improvement and developments can be applied to biological readings. Miniaturization not only reduces size but can also improve stability. Small probes also allow multiple readings in one sensor. This trend has already started with blood gas probes being used for intravascular and extravascular measurements. The next trend envisioned is nanoprobes. Implantable glucose probes are already being tested. Sensor and light sources can be implemented on an integrated circuit and continue to shrink. Microsensors are already being used in biomedical application, such as blood pressure sensors and single-cell neural stimulators. Miniature chemistry and thermal-cycling laboratories are being developed into single integrated chips. The same chemical and optical arrangements can be performed for blood chemistry analysis. A complete laboratory measurement system could be at the patient location for continuous monitoring [28–30,32].

The other major anticipated change is the trend for additional and more accurate noninvasive readings. Invasive measurements can mean delays and noncontinuous monitoring, as well as increased possibility of infection. Given equal measurement accuracy, noninvasive measurements are an obvious choice over invasive readings.

It is important to note that measurement techniques will continue to improve, but the rate of growth and utilization will not grow as fast. Acceptance and utilization are dependent on cost and the effort involved in using the instruments. Cost can slow or even eliminate instrumentation from being used. However, with the increasing cost of medical care, there is interest in providing fast diagnosis and recovery for patients. This will provide incentive for improving measurement technique and accuracy.

Defining Terms

Artery: Vessel used to carry blood away from the heart.

Capillary: Semipermeable membrane used to exchange oxygen and other substances with tissue.

Catheter: Surgical instrument inserted into the body for drawing or administering fluids.

Electrolyte: A substance that dissolves into ions, becoming capable of conducting electricity.

Fluorescence: The property of emitting light when exposed to an excitation light.

Glucose (dextrose): Substance found in normal blood, which is the main source of energy in living tissue.

Hematocrit: The total CtHb, or packed red cells, in blood.

Hemoglobin: Oxygen-carrying pigment found in blood, which has the property of reversible oxygenation.

Intra-arterial: Within the artery.

Invasive: Pertaining to the insertion of an instrument into the body.

Oximetry: Photoelectric determination of arterial blood oxygen saturation (S_aO_2).

Perfusion: Pertaining to the passage of fluid through vessels. Low perfusion refers to reduced blood flow.

Photoplethysmogram: Chart of the volume change in an organ or limb. In oximetry it is the chart of the light passed through an ear, nose, or digit. The light intensity varies with the arterial pulse, which is related to the arterial volume.

Transcutaneous: Pertaining to entering through the skin.
Vein: Vessel used to return blood to the heart.

References

1. F. W. Cheney, The ASA closed claims study after the pulse oximeter, *ASA Newslett*. 54: 10–11, 1990.
2. J. W. Severinghaus and P. B. Astrup, History of blood gas analysis II: pH and acid-base measurements, *J. Clin. Monitoring*, 1: 259–277, 1985.
3. J. W. Severinghaus and P. B. Astrup, History of blood gas analysis IV: Leland Clark's oxygen electrode, *J. Clin. Monitoring*, 2: 125–139, 1986.
4. J. W. Severinghaus and P. B. Astrup, History of blood gas analysis V: Oxygen measurement, *J. Clin. Monitoring*, 2: 175–189, 1986.
5. J. W. Severinghaus and P. B. Astrup, History of blood gas analysis VI: Oximetry, *J. Clin. Monitoring*, 2: 270–288, 1986.
6. J. W. Severinghaus and P. B. Astrup, History of blood gas analysis, *Int. Anesth. Clin.*, 25(4): 1–214, 1987.
7. K. K. Tremper and S. J. Barker, Advances in oxygen monitoring, *Int. Anesth. Clin.*, 25(3): 1–96, 113–208, 1987.
8. *Churchhill's Illustrated Medical Dictionary*, Churchhill Livingston, New York, 1989.
9. K. D. McClatchey, *Clinical Laboratory Medicine*, Williams & Wilkins, Baltimore, MD, 1994.
10. R. W. Burnett, A. K. Covington, N. Fogh-Andersen, W. R. Kulpmann, A. H. Maas, O. Muller-Plathe, O. Siggaard-Andersen, A. L. Van Kessel, P. D. Wimberly, and W. G. Zijlstra, International federation of clinical chemistry, *Eur. J. Clin. Chem. Clin. Biochem.*, 33: 247–253, 399–404, 1995.
11. J. W. Gilbert, F. P. Holladay, and H. C. Weiser, Hematocrit monitor, *Crit. Care Med.*, 17(9): 929–966, 1989.
12. L. C. Clark, Measurement of oxygen tension, *Crit. Care Med.*, 9(10): 690–693, 1981.
13. D. W. Lubbers, Oxygen measurement in blood and tissues and their significance: Methods of measuring oxygen tension of blood and organ surfaces, *Int. Anesth. Clin.*, 4(1): 103–122, 1966.
14. D. W. Lubbers, F. Hannebauer, and N. Opitz, Continuous transcutaneous blood gas monitoring: PCO_2-optode, fluorescence fotometric device to measure the transcutaneous PCO_2, *Birth Defects Orig. Art. Ser.*, XV(4): 123–126, 1979.
15. J. L. Gehrich, D. W. Lubbers, N. Opitz, D. R. Hansmann, W. W. Miller, J. K. Tusa, and M. Yafuso, Optical fluorescence and its application to an intravascular blood gas monitoring system, *IEEE Trans. Biomed. Eng.*, BME-33: 117–131, 1986.

16. R. Narayanaswamy, Current developments in optical biochemical sensors, *Biosens. Bioelectron*., 6(6): 467–475, 1991.
17. J. F. O'Riordan, T. K. Goldstick, J. Ditzel, and J. T. Ernst, Characterization of oxygen-hemoglobin equilibrium curves using nonlinear regression of the Hill equation: Parameter values for normal human adults, in H.I. Bicher and D.F. Bruley (Eds.), *Oxygen Transport to Tissue—IV*, Plenum Press, New York, pp. 435–444, 1983.
18. F. J. W. Roughton and J. W. Severinghaus, Accurate determination of O_2 dissociation curve of the human blood above 98.7% saturation with data on O_2 solubility in unmodified human blood from 0° to 36°C, *J. Appl. Physiol*., 35: 861–869, 1973.
19. J. W. Severinghaus, Simple accurate equations for human blood O_2 dissociation computations, *J. Appl. Physiol. Respir., Environ. Exercise Physiol*., 46: 599–602, 1979.
20. E. H. Wood and J. E. Geraci, Photoelectric determination of arterial oxygen saturation in man, *J. Lab. Clin. Med*., 34: 387–401, 1949.
21. J. W. Severinghaus and J. F. Kelleher, Recent developments in pulse oximetry, *Anesthesiology*, 76: 1018–1038, 1992.
22. J. A. H. Bos, W. Schelter, W. Gumbrecht, B. Montag, E. P. Eijking, S. Armbruster, W. Erdmann, and B. Lachmann, Development of a micro transmission cell for in vivo measurement of S_aO_2 and Hb, in *Oxygen Transport to Tissue XII*, Plenum Press, New York, pp. 47–52, 1990.
23. Y. Mendelson and J. C. Kent, An in vitro tissue model for evaluating the effect of carboxyhemoglobin concentration on pulse oximetry, *IEEE Trans. Biomed. Eng*., BME-36: 625–627, 1989.
24. Y. Mendelson and B. D. Ochs, Noninvasive pulse oximetry utilizing skin reflectance photoplethysmography, *IEEE Trans. Biomed. Eng*., BME-35: 798–805, 1988.
25. J. M. Schmitt, J. D. Meindl, and F. G. Mihm, An integrated circuit-based optical sensor for in vivo measurement of blood oxygenation, *IEEE Trans. Biomed. Eng*., BME-33: 89–107, 1986.
26. W. Cui, L. E. Ostrander, and B. Y. Lee, In vivo reflectance of blood tissue as a function of light wavelength, *IEEE Trans. Biomed. Eng*., BME-37: 623–639, 1990.
27. J. I. Peterson and G. G. Vurek, Fiber-optic sensor for biomedical applications, *Science*, 224(4645): 123–127, 1984.
28. J. S. Schultz, S. Mansouri, and I. J. Goldstein, Affinity sensor: A new technique for developing implantable sensors for glucose and other metabolites, *Diabetes Care*, 5(3): 245–253, 1982.
29. L. J. Brown, A new instrument for the simultaneous measurement of total hemoglobin, % oxyhemoglobin, % carboxyhemoglobin, % methemoglobin, and oxygen content in whole blood, *IEEE Trans. Biomed. Eng*., BME-27: 132–138, 1980.
30. B. A. Shapiro, R. D. Cane, C. M. Chomka, L. E. Bandala, and W. T. Peruzzi, Preliminary evaluation of an intra-arterial blood gas system in dogs and humans, *Crit. Care Med*., 17(5): 455–460, 1989.
31. B. A. Shapiro, Clinical and economic performance criteria for intra-arterial and extra-arterial blood gas monitoring with comparison with in vitro testing, *Am. J. Clin. Pathol*., 104(4 Suppl. 1): S100–S105, 1995.
32. C. Ajluni, Microsensors move into biomedical applications, *Electron. Design*, 44(11): 75–84, 1996.
33. J. G. Webster (Ed.), *Design of Pulse Oximeters*, IOP Publishing, Bristol, U.K., 1997.

69

Human Occupancy Detectors

Jacob Fraden
Fraden Corporation

69.1 Introduction

September 11 has changed the way people think about the airport, aviation, and security in general. The threat is expanding interest in more reliable systems to detect presence of people within the protected perimeters. All sensors that respond to a human body are divided on the occupancy and motion detectors. The *occupancy* detectors detect the presence of people (and sometimes animals) in a monitored area. The *motion detectors* respond only to moving objects. A distinction between the two is that the occupancy detectors produce signals whenever an object is stationary or not, while the motion detectors are selectively sensitive to the moving objects. The applications of these sensors include security, surveillance, energy management (electric lights control), personal safety, friendly home appliances, point-of-sale advertisements, interactive toys, and novelty products. Depending on the applications, presence of humans may be detected through any means that is associated with some kind of a human body's property or body's actions [1]. For instance, a detector may be sensitive to body weight, heat, sounds, dielectric constant, etc. The following is a nonexhaustive list of the types of detectors that are presently used for the occupancy and motion sensing of people:

1. *Air pressure sensors*—detect miniscule changes in air pressure resulted from opening doors and windows or movement of an intruder
2. *Capacitive*—detectors of human body capacitance
3. *Acoustic*—detectors of sound produced by people
4. *Photoelectric*—interruption of light beams by moving objects
5. *Optoelectric*—detection of variations in illumination or optical contrast in the protected area

6. *Pressure mat switches*—pressure-sensitive long strips or pressure sensors imbedded into floors to detect weight of an intruder
7. *Stress detectors*—strain gauges imbedded into floor beams, staircases, and other structural components
8. *Switch sensors*—electrical contacts connected to doors and windows
9. *Magnetic switches*—a noncontact version of switch sensors
10. *Vibration detectors*—react to the vibration of walls or other building structures. Also, may be attached to doors or windows
11. *Glass breakage detectors*—sensors reacting to specific vibrations produced by shattered glass
12. *Infrared (IR) motion detectors*—devices sensitive to heat waves emanated from warm or cold moving bodies
13. *Microwave detectors*—active sensors radar-type devices responsive to microwave electromagnetic signals reflected from objects
14. *Ultrasonic detectors* (USDs)—devices similar to microwave detectors except that instead of electromagnetic radiation, ultrasonic waves are used
15. *Video motion detectors*—video equipment that compares a stationary reference image with the current image of a protected area. May operate in visible or IR spectral ranges
16. *Facial recognition*—the image analyzers that compare facial features with a database
17. *Laser system detectors*—similar to photoelectric detectors, except that they use narrow light beams and combinations of reflectors
18. *Triboelectric detectors*—sensors capable of detecting static electric charges carried by moving objects

One of the major aggravations in detecting occupancy or intrusion is a *false-positive* detection. The term "false positive" means that the system indicates an intrusion when there is none. In some noncritical applications where false-positive detections occur once in a while, for instance, in a toy or switch controlling electric lights in a room, this may be not a serious problem: the lights will be erroneously turned on for a short time, which unlikely do any harm.* In other systems, especially used for the security and military purposes, false-positive detections, while generally not as dangerous as false negative ones (missing an intrusion), may become a serious problem. While selecting a sensor for critical applications, considerations should be given to its reliability, selectivity, and noise immunity. It's often a good practice to form a multiple sensor arrangement with the symmetrical interface circuits. It may dramatically improve a reliability of a system, especially in the presence of the external transmitted noise. Another efficient way of reducing the erroneous detections is to use in concert the sensors operating on different physical principles [2], for instance, combining the capacitive and IR detectors is an efficient combination as they are receptive to different kinds of transmitted noise. In the following, we proved overview of the most popular or promising detectors of people.

69.2 Ultrasonic Detectors (USD)

These devices are based on transmission to the object and receiving the reflected acoustic waves. A transmission and reception of the ultrasonic energy is the basis for very popular ultrasonic range meters and velocity detectors. The USDs for humans are employed for the automatic door openers and security sensors. Ultrasonic waves are mechanical compressions and rarefaction of air covering the frequency range well beyond the capabilities of human ears, that is, over 20 kHz. However, these frequencies may be quite perceptive by smaller animals, like dogs, cats, rodents, and insects. Indeed, the USDs are the biological ranging devices for bats and dolphins.

* Perhaps just steering up some suspicion about living in a haunted house.

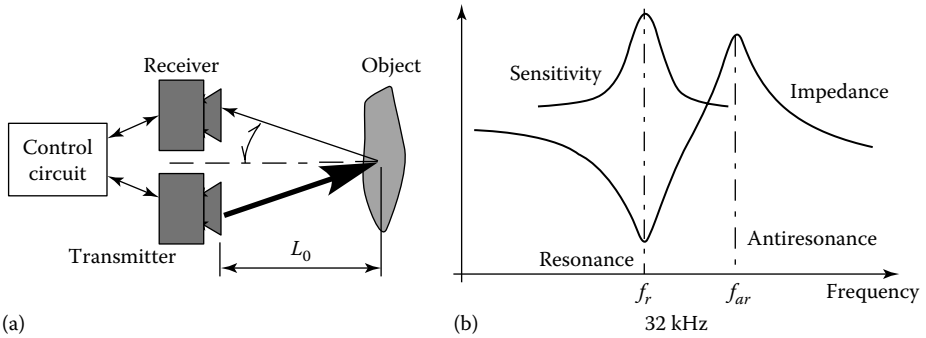


FIGURE 69.1 Ultrasonic distance measurement: (a) basic arrangement and (b) impedance characteristic of a piezoelectric transducer.

When the waves are incident on an object, part of their energy is reflected. In many practical cases, the ultrasonic energy is reflected in a diffuse manner. That is, regardless of the direction where the energy comes from, it is reflected almost uniformly within a wide solid angle, which may approach 180°. If an object moves, the frequency of the reflected wavelength will differ from the transmitted waves. This is called the Doppler effect.

A distance L_0 to the object can be calculated through the speed v of the ultrasonic waves in the media and the angle Θ (Figure 69.1a):

$$L_0 = \frac{vt\cos\Theta}{2} \tag{69.1}$$

where t is the time for the ultrasonic waves to travel to the object and back to the receiver (thus 2 in the denominator).

If transmitter and receiver are positioned close to each other when compared with the distance to the object (which usually is the case), then $\cos\Theta \approx 1$. The ultrasonic waves have an obvious advantage over the microwaves: they propagate with the speed of sound, which is much slower than the speed of light at what the microwaves propagate. Thus, time t is much longer and its measurement can be accomplished easier and cheaper.

To generate any mechanical waves, including ultrasonic, the movement of a surface is required. This movement creates compression and expansion of medium (air). The most common type of the excitation device that can generate surface movement in the ultrasonic range is a piezoelectric transducer operating in the so-called *motor* mode. The name implies that the piezoelectric device directly converts electrical energy into mechanical energy.

Figure 69.2a shows that the input voltage applied to the ceramic element causes it to flex and transmit ultrasonic waves in air. Because piezoelectricity is a reversible phenomenon, the ceramic generates voltage when incoming ultrasonic waves flex it. In other words, the element may work as both the sonic transmitter and receiver (a microphone). A typical operating frequency of the transmitting piezoelectric element is near 32 kHz. For better efficiency, frequency of the driving oscillator should be adjusted to the resonant frequency f_r of the piezoelectric ceramic (Figure 69.1b) where the sensitivity and efficiency of the element is the best. When the measurement circuit operates in a pulsed mode, the same piezoelectric element is used for both transmission and receiving. If the system requires continuous transmission of ultrasonic waves, separate piezoelectric elements are employed for the transmitter and receiver. Typical design of air-operating sensors is shown in Figures 69.2b and 69.3b. A directional sensitivity diagram (Figure 69.3b) is important for a particular application. The narrower the diagram is the more sensitive the transducer.

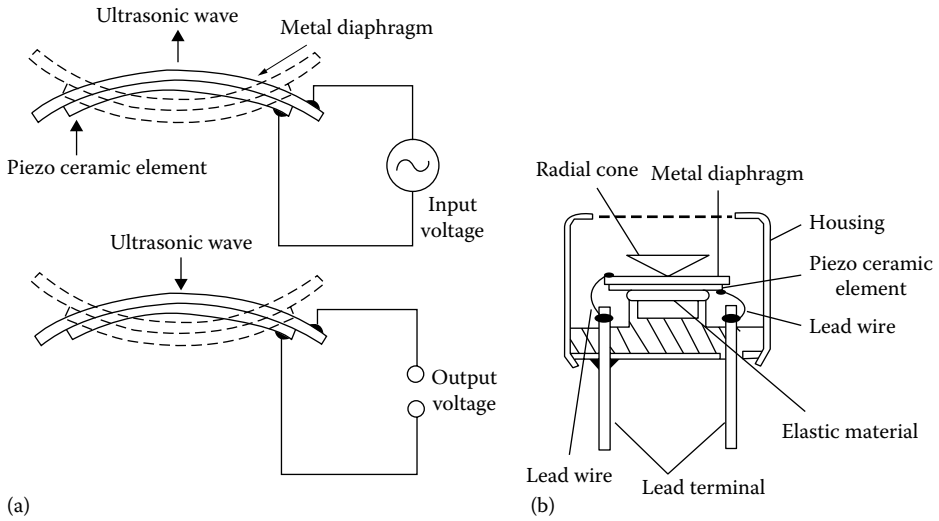


FIGURE 69.2 Piezoelectric ultrasonic transducer: (a) input voltage flexes the element and transmits ultrasonic waves, while incoming waves produce output voltage and (b) open aperture type of ultrasonic transducer for operation in air. (Courtesy of Nippon Ceramic, Japan.)

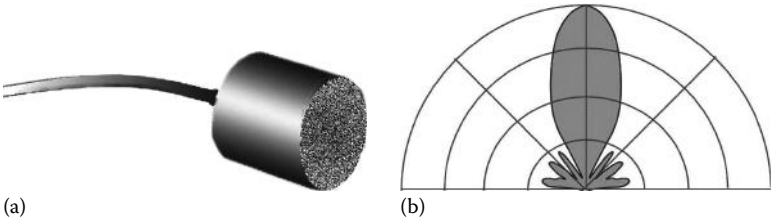


FIGURE 69.3 (a) Ultrasonic transducer for air and (b) directional diagram.

69.3 Microwave Motion Detectors

The microwave detectors offer an attractive alternative to other detectors when it is required to cover large areas and to operate over an extended temperature range under the influence of strong interferences, such as wind, acoustic noise, fog, dust, and moisture. Just like the USDs, these detectors belong to the active sensors as they generate an excitation signal. That is, they emit pulses of the electromagnetic energy and can operate at day or night. The operating principle of a microwave detector is based on radiation of electromagnetic radio-frequency (RF) waves toward a protected area. The electromagnetic waves backscattered (reflected) from objects whose sizes are comparable with or larger than the wavelength of the excitation signal. The reflected waves are received, amplified, and analyzed. A time delay between the sent (pilot) signal and received reflected signal is used to measure distance to the object, while the frequency shift is used to measure speed of motion of the object.

The microwave detectors belong to the class of devices known as *radars*. Radar is an acronym for *radio detection and ranging*. The radar frequencies are shown in Table 69.1.

The name *microwave* is arbitrarily assigned to the wavelengths shorter than 4 cm (K_a, K, and X bands). They are long enough ($\lambda \approx 3$ cm at X band) to pass freely through most contaminants, such as fog and airborne dust, and short enough for being reflected by humans.

The microwave part of the detector consists of a Gunn oscillator, an antenna, and a mixer diode. The Gunn oscillator is a diode mounted in a small precision cavity that, upon application of power, oscillates at microwave frequencies. The oscillator produces electromagnetic waves (frequency f_0), part of which

TABLE 69.1 Radar Frequencies Various Bands

Band	Frequency Range (GHz), f	Wavelength Range (cm), λ
K_a	26.0–40.0	0.8–1.1
K	18.0–26.5	1.1–1.67
X	8.0–12.5	2.4–3.75
C	4.0–8.0	3.75–7.50
S	2.0–4.0	7.5–15
L	1.0–2.0	15–30
P	0.3–1.0	30–100

is directed through an iris into a waveguide and focusing antenna that directs the radiation toward the object. Focusing characteristics of the antenna are determined by the application. As a general rule, the narrower the directional diagram of the antenna, the more sensitive it is (the antenna has a higher gain). Another general rule is that a narrow-beam antenna is much larger, whereas a wide-angle antenna can be quite small. The typical radiated power of the transmitter is 10–20 mW. A Gunn oscillator is sensitive to the stability of applied dc voltage and, therefore, must be powered by a good quality voltage regulator. The oscillator may run continuously, or it can be pulsed, which reduces the power consumption from the power supply.

A smaller part of the microwave oscillations is coupled to the Schottky mixing diode and serves as a reference signal (Figure 69.4a). In many cases, the transmitter and the receiver are contained in one module called a transceiver. The target reflects some waves back toward the antenna, which directs the received radiation toward the mixing diode whose current contains a harmonic with a phase differential between the transmitted and reflected waves. The phase difference is in a direct relationship to the distance to the target. The phase-sensitive detector is useful mostly for detecting the distance to an object. However, in a motion sensor, movement, not distance, should be detected. Thus, for the occupancy and motion detector, the Doppler effect is the basis for operation of microwave and USDs. It should be noted that the Doppler effect device is a true motion detector because it is responsive only to moving targets. Here is how it works.

An antenna transmits the frequency f_0 that is defined by the wavelength λ_0 as

$$f_0 = \frac{c_0}{\lambda_0} \tag{69.2}$$

where c_0 is the speed of light in air.

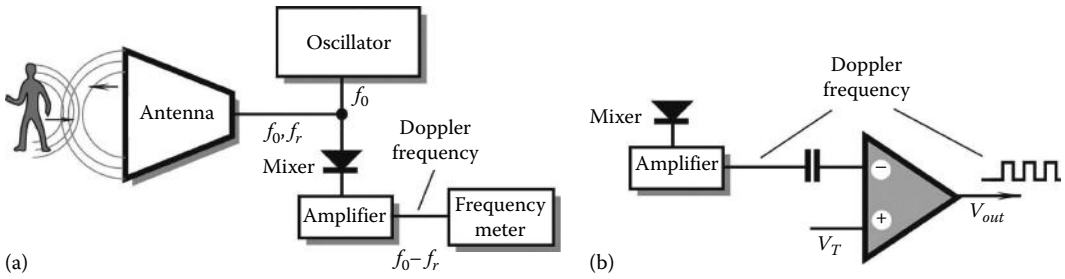


FIGURE 69.4 Microwave occupancy detector: (a) a circuit for measuring Doppler frequency and (b) circuit with a threshold detector.

When the target moves toward or away from the transmitting antenna, the frequency of the reflected radiation will change. Thus, if the target is moving away with velocity v, the reflected frequency will decrease and it will increase for the approaching targets. This is called the Doppler effect, after the Austrian scientist Christian Johann Doppler (1803–1853).* Although the effect first was discovered for sound, it is applicable to electromagnetic radiation as well. In contrast to sound waves that may propagate with the velocities dependent on movement of the source of sound, electromagnetic waves propagate with the speed of light, which is an absolute constant, independent of the light source movement. The frequency of reflected electromagnetic waves can be predicted by Einstein's special theory of relativity as

$$f_r = f_0 \frac{\sqrt{1-(v/c_0)^2}}{1+(v/c_0)} \tag{69.3}$$

For practical purposes when detecting a relatively slow moving objects, the quantity $(v/c_0)^2$ is very small as compared with unity; hence, it can be ignored. Then, the equation for the frequency of the reflected waves becomes identical to that for the acoustic waves:

$$f_r = f_0 \frac{1}{1+(v/c_0)} \tag{69.4}$$

As follows from Equation 69.4, due to a Doppler effect, the reflected waves have a different frequency f_r. The mixing diode combines the radiated (reference) and reflected frequencies and, being a nonlinear device, produces a signal that contains multiple harmonics of both frequencies. The electric current through the diode may be represented by a polynomial:

$$i = i_0 + \sum_{k=1}^{n} a_k (U_1 \cos 2\pi f_0 t + U_2 \cos 2\pi f_r t)^k \tag{69.5}$$

where

i_0 is a dc component
a_k are the harmonic coefficients that depend on a diode operating point
U_1 and U_2 are amplitudes of the reference and received signals, respectively
t is time

The current through a diode contains an infinite number of harmonics, among which there is a harmonic of a differential frequency: $\Delta f = a_2 U_1 U_2 \cos 2\pi (f_0 - f_r)t$, which is called a Doppler frequency.

The Doppler frequency in the mixer can be found from Equation 69.4:

$$\Delta f = f_0 - f_r = f_0 \frac{1}{1+(c_0/v)}, \tag{69.6}$$

and since $c_0/v \gg 1$, the following holds after substituting Equation 69.2:

$$\Delta f \approx \frac{v}{\lambda_0} \tag{69.7}$$

* During Doppler times, the acoustical instruments for precision measurements were not available yet. To prove his theory, Doppler placed trumpeters on a railroad flatcar and musicians with a sense of absolute pitch near the tracks. A locomotive engine pulled the flatcar back and forth at different speeds for two days. The musicians on the ground "recorded" the trumpet notes as the train approached and receded. The equations held up.

Therefore, the signal frequency at the output of the mixer is proportional to the velocity of a moving target. For instance, if a person walks toward the detectors with a velocity of 0.6 m/s, a Doppler frequency for the *X*-band detector is $\Delta f = 0.6/0.03 = 20$ Hz.

Equation 69.6 holds true only for movements in the normal direction. When the target moves at angles Θ with respect to the detector, the Doppler frequency is

$$\Delta f \approx \frac{v}{\lambda_0}\cos\Theta \tag{69.8}$$

This implies that Doppler detectors theoretically become insensitive when a target moves at angles approaching 90°.

In the velocity meters, to determine velocity of a target, it is required to measure the Doppler frequency and phase to determine direction of the movement (Figure 69.5a). This method is used in police radars. For the supermarket door openers and security alarms, instead of measuring frequency, a threshold comparator is used to indicate the presence of a moving target (Figure 69.5b). It should be noted that though Equation 69.8 predicts that the Doppler frequency is near zero for targets moving at angles $\Theta = 90°$, the entering of a target into a protected area at any angle results in an abrupt change in the received signal amplitude, and the output voltage from the mixer changes accordingly. Usually, this is sufficient to trigger the response of a threshold detector.

A signal from the mixer is in the range from μV to mV, so amplification is needed for the signal processing. Because the Doppler frequency is in the audio range, the amplifier is relatively simple. However, it generally must be accompanied by the so-called notch filters that reject a power line frequency and the main harmonic from full-wave rectifiers and fluorescent light fixtures: 60 and 120 Hz (or 50 and 100 Hz). For the normal operation, the received power must be sufficiently high. It depends on several factors, including the antenna aperture area A, target area a, and distance to the target r:

$$P_r = \rho\frac{P_0 A^2 a}{4\pi\lambda^2 r^4} \tag{69.9}$$

where P_0 is the transmitted power.

For effective operation, the target's cross-sectional area a must be relatively large, because for $\lambda^2 \le a$, the received signal is drastically reduced. Further, the reflectivity ρ of a target in the operating wavelength is also very important for the magnitude of the received signal. Generally, conductive materials and objects with high dielectric constants are good reflectors of electromagnetic radiation, whereas

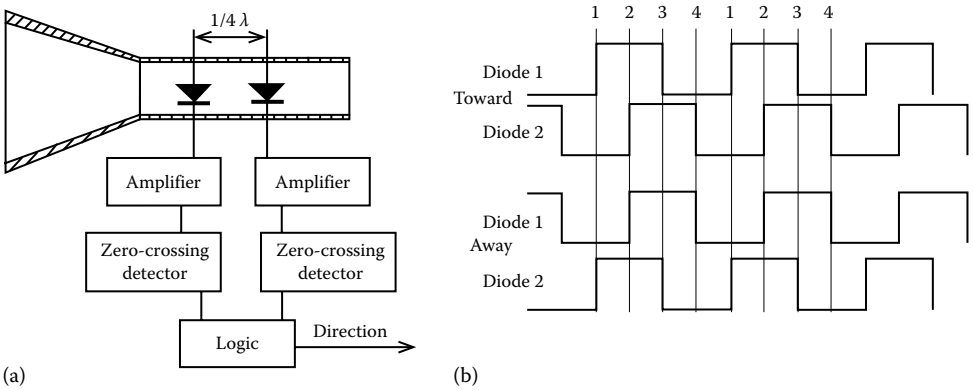


FIGURE 69.5 Block diagram (a) and timing diagrams (b) of a microwave Doppler motion detector with directional sensitivity.

many dielectrics absorb energy and reflect very little. Plastics and ceramics are quite transmissive and can be used as windows in the microwave detectors.

The best target for a microwave detector is a smooth, flat conductive plate positioned normally toward the detector (doesn't look like a human, does it?). A flat conductive surface makes a very good reflector; however, it may render the detector inoperable at angles other than 0°. Thus, an angle of $\Theta = 45°$ can completely divert a reflective signal from the receiving antenna. This method of diversion, along with other techniques, was used quite effectively in designs of the stealth bombers and fighters that are virtually invisible on radar screens. Nevertheless, at relatively short distances, signals reflected from the human bodies are still sufficiently strong for many practical purposes because they are scattered rather than reflected.

To detect whether a target moves toward or away from the antenna, the Doppler concept can be extended by adding another mixing diode to the transceiver module. The second diode is located in the waveguide in such a manner that the Doppler signals from both diodes differ in phase by one-quarter of wavelength or by 90° (Figure 69.5a). These outputs are amplified separately and converted into square pulses that can be analyzed by a logic circuit. The circuit is a digital phase discriminator that determines the direction of motion (Figure 69.5b). The door openers and traffic control devices are two major applications for this type of a module. Both applications need the ability to acquire a great deal of information about the target for discrimination before enabling a response. In door openers, limiting the field of view and transmitted power may substantially reduce the number of false-positive detections. Although discrimination is optional for the door openers direction, for traffic control it is a necessity to reject signals from the vehicles moving away. If the module is used for intrusion detection, the vibration of building structures may cause a large number of false-positive detections. A direction discriminator will respond to vibration with an alternate signal, and it will respond to an intruder with a steady logic signal. Hence, the direction discriminator is an efficient way to improve reliability of the detection.

Generally, the transmission and reception are alternated in time. That is, the receiver is disabled during the transmission; otherwise, a strong transmitted energy not only will saturate the receiving circuitry but may damage its sensitive components. In Nature, bats use ultrasonic ranging to catch their small prey. The bats become deaf for the short time when the ultrasonic burst of energy is transmitted. This temporary blinding of the receiver is the main reason why radars and acoustic rangers are not effective for short distances—it's just not enough time to disable and enable the receiver.

Whenever a microwave detector is used in the United States, it must comply with the strict requirements (e.g., MSM20100) imposed by the Federal Communication Commission. Similar regulations are enforced in many other countries. Also, emission of the transmitter must be below 10 mW/cm^2 as averaged over any 0.1 h period, as specified by OSHA 1910.97 for the frequency range from 100 MHz to 100 GHz.

69.4 Capacitive Occupancy Detectors

Being a conductive medium with a high dielectric constant, a human body develops a coupling capacitance to its surroundings.* This capacitance greatly depends on such factors as body size, clothing, materials, type of surrounding objects, and weather. However wide the coupling range is, the capacitance may vary from few pF to several nF. When a person moves, the coupling capacitance changes, thus making it possible to discriminating static objects from moving objects.

All objects form some degree of a capacitive coupling with respect to one another. If a human moves into vicinity of the objects whose coupling capacitance with each other has been previously established, a new capacitive value arises between the objects as a result of presence of an intruding body. Figure 69.6 shows that the capacitance between a test plate and earth† is equal to C_1. When a person moves into

* At 40 MHz, the dielectric constant of muscle, skin, and blood is about 97. For fat and bone, it is near 15.

† Here, by "earth" we mean any large object, such as the earth, lake, metal fence, car, ship, and airplane.

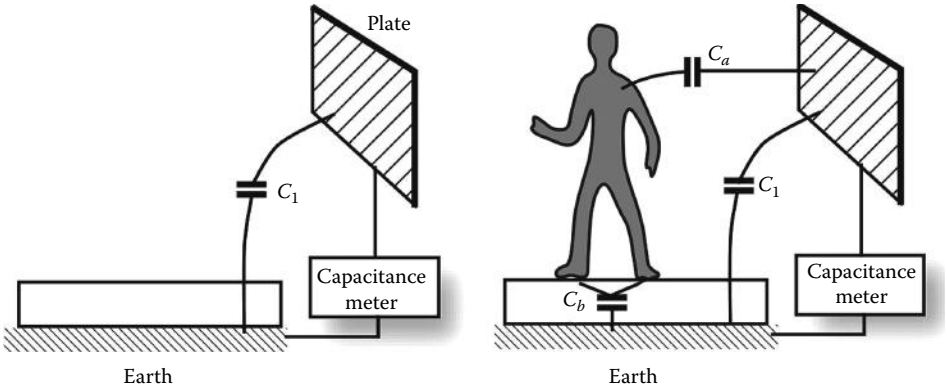


FIGURE 69.6 An intruder brings in an additional capacitance to a detection circuit.

vicinity of the plate, it forms two additional capacitors: one between the plate and its own body, C_a, and the other between the body and the earth, C_b. Then, the resulting capacitance C between the plate and the earth becomes larger by the incremental capacitance ΔC:

$$C = C_1 + \Delta C = C_1 + \frac{C_a C_b}{C_a + C_b} \tag{69.10}$$

With the appropriate apparatus, this phenomenon can be used for occupancy detection. What is required is to measure capacitance between a test plate (the probe) and a reference plate (the earth).

Figure 69.7 illustrates a capacitive security system for an automobile [3]. A sensing probe is imbedded into a car seat. It can be fabricated as a metal plate, metal net, a conductive fabric, etc. The probe forms one plate of a capacitor C_p. The other plate of the capacitor is formed either by a body of an automobile or by a separate plate positioned under a floor mat. A reference capacitor C_x is composed of a simple fixed or trimming capacitor that should be placed close to the seat probe. The probe plate and the reference capacitor are, respectively, connected to two inputs of a charge detector (resistors R_1 and R_2). The conductors preferably should be twisted to reduce the introduction of spurious signals as much as possible. For instance, strips of a Twinflex cabling were found quite adequate. A differential charge detector is

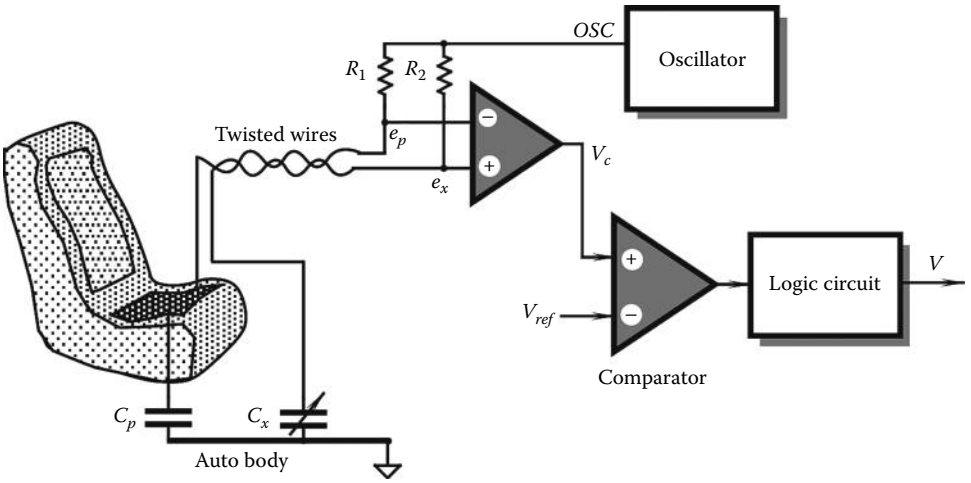


FIGURE 69.7 Capacitive intrusion detector for automotive applications.

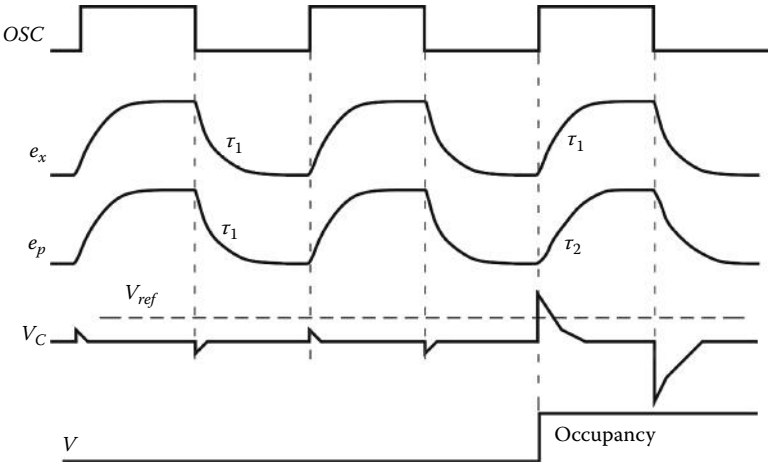


FIGURE 69.8 Timing diagrams for a capacitive intrusion detector.

controlled by an oscillator that produces square pulses (Figure 69.8). Under a no-seat-occupied condition, the reference capacitor is adjusted to be approximately equal to C_p. Resistors and the corresponding capacitors define time constants of the networks. Both *RC* circuits have nearly equal time constants τ_1. Voltages across the resistors are fed into the inputs of a differential amplifier, whose output voltage V_c is near zero. Small spikes at the output are the result of some unavoidable imbalance. When a person is positioned on the seat, her body forms an additional capacitance in parallel with C_p, thus increasing a time constant of the R_1C_p network from τ_1 to τ_2. This is indicated by the increased spike amplitudes at the output of a differential amplifier. The comparator compares V_c with a predetermined threshold voltage V_{ref}. When the spikes exceed the threshold, the comparator sends an indication signal to the logic circuit that generates signal V manifesting the car occupancy.

When a capacitive occupancy (proximity) sensor is used near or on a metal device, its sensitivity may be severely reduced due to a capacitive coupling between the electrode and the device's metallic parts. An effective way to reduce that stray capacitance is to use driven shields. Figure 69.9a shows a robot with a metal arm. The arm moves near people and other potentially conductive objects with which it could collide if the robot's control computer is not provided with advance information on proximity to the obstacles. An object, while approaching the arm, forms a capacitive coupling with it, which is equal to C_{so}. An arm is covered with an electrically isolated conductive sheath that is called an *electrode*. The nearby massive metal arm (Figure 69.9b) forms a much stronger capacitive coupling

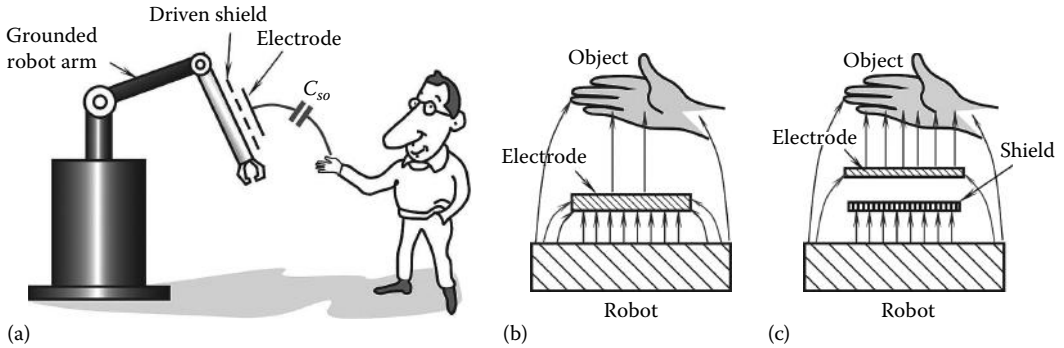


FIGURE 69.9 Capacitive proximity sensor: a driven shield is positioned on the metal arm of a grounded robot (a). Without the shield, the electric field is distributed between the electrode and the robot (b), while a driven shield directs electric field from the electrode toward the obstacle (c).

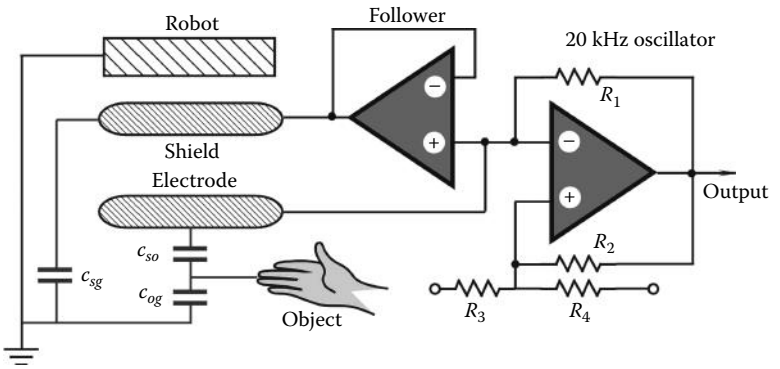


FIGURE 69.10 Simplified circuit diagram of a frequency modulator controlled by the input capacitances.

with the electrode that drugs the electric field from the object. An elegant solution is to shield the electrode from the arm by an intermediate shield as shown in Figure 69.9c. The sensor's assembly is a multilayer cover for the robotic arm, where the bottom layer is an insulator, then there is a large electrically conductive shield, then another layer of insulation, and on the top is a narrower sheet of the electrode. To reduce a capacitive coupling between the electrode and the arm, the shield must be at the same potential as the electrode, that is, its voltage must be driven by the electrode voltage (hence the name is a *driven shield*). As a result, no electric field is formed between them. Yet, a strong electric field will exist between the shield and the arm. The electric field is squeezed out from beneath the electrode and distributed toward the object.

Figure 69.10 shows a simplified circuit diagram of a square-wave oscillator whose frequency depends on the net input capacitance, comprised of C_{sg} (sensor to ground), C_{so} (sensor to object), and C_{og} (object to ground). The electrode is connected to the shield through a voltage follower. A frequency-modulated signal is fed into the robot's computer for controlling the arm movement. This arrangement allows detecting proximity to conductive objects over the range of 30 cm.

69.5 Triboelectric Detectors

Any object can accumulate, on its surface, static electricity. These naturally occurring charges arise from the triboelectric effect, which is a process of charge separation due to object movements, friction of clothing fibers, air turbulence, atmosphere electricity, etc. Usually, air contains either positive or negative ions that can be attracted to the human body, thus modifying its charge. Under the idealized static conditions, an object is not charged—its bulk charge is equal to zero. In reality, any object that at least temporarily is isolated from the ground can exhibit some degree of its bulk charge imbalance. In other words, it becomes a carrier of electric charges.

Any electronic circuit is made of conductors and dielectrics. If a circuit is not shielded, all its components exhibit a certain capacitive coupling to the surrounding objects. A pickup *electrode* can be added to the circuit's input to increase its coupling to the environment, very much like in the capacitive detectors. The electrode can be fabricated in form of a conductive surface that is well isolated from the ground. The difference between the triboelectric and capacitive sensors is that in the former no capacitance is being measured but rather an electric charge that is accumulated on the capacitance.

Electric field is established between the surrounding objects and the electrode whenever at least one of them carries electric charges. In other words, all distributed capacitors formed between the electrode and the environmental objects are charged by static or slow-changing electric fields resulted from a triboelectric effect. Under a no-occupancy condition, electric field in the electrode vicinity is either constant or changes relatively slow.

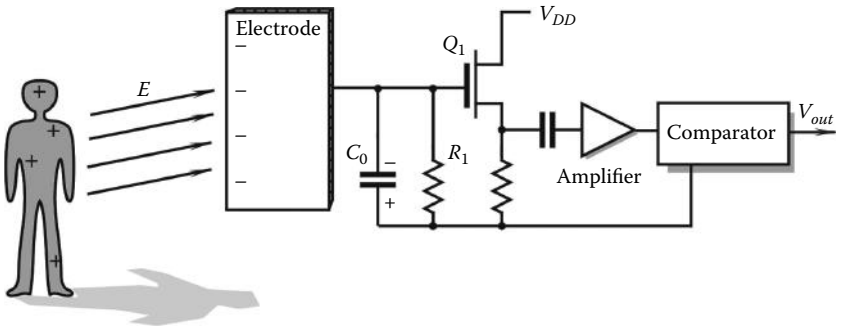


FIGURE 69.11 Monopolar triboelectric motion detector.

If a charge carrier (an intruder) changes its position, moves away, or a new charge-carrying object enters into the vicinity of the electrode, a static electric field is disturbed. This results in a redistribution of charges between the coupling capacitors, including those that are formed between the input electrode and the surroundings. The charge depends on the atmospheric conditions and nature of the objects. For instance, a person in dry manmade clothes walking along a carpet carries a million times stronger charge than a wet intruder who has come from the rain. An electronic circuit can be adapted to sense these variable charges at its input. In other words, it can be made capable of converting the induced variable charges into electric signals that may be amplified and further processed. Thus, static electricity that is a naturally occurring phenomenon can be utilized to generate alternating signals in the electronic circuit to indicate the movement of objects.

Figure 69.11 shows a monopolar triboelectric motion detector [4]. It is comprised of a conductive electrode connected to an analog impedance converter made with a MOS transistor Q_1, a bias resistor R_1, an input capacitance C_0, a gain stage, and a window comparator [5]. While the rest of the electronic circuit may be shielded, the sensing electrode is exposed to the environment and forms a coupling capacitor C_p with the surrounding objects. Being a charge carrier, the intruder becomes a source of an electric field, having intensity E. The field induces a charge of the opposite sign in the electrode. Under the static conditions, when the person doesn't move, the field intensity is constant and the input capacitance C_0 is discharged through a bias resistor R_1. To make the circuit sensitive to relatively slow motions, the resistor R_1 should be selected of a very high value: on the order of 10^{10} Ω or higher. When a person moves, the intensity, E, of the electric field changes. This induces a corresponding variable electric charge in the input capacitor C_0 and results in appearance of a variable electric voltage across the bias resistor. That voltage is fed through the coupling capacitor into the gain stage whose output signal is applied to a window comparator. Note that contrary to a capacitive motion detector that is an active sensor, a triboelectric detector is passive. That is, it does not generate or transmit any signal that makes it more difficult to detect. This detector may be hidden in or behind nonmetallic objects such as wood and bricks.

There are several possible sources of interferences that may cause spurious detections by the triboelectric detectors. The detector may be subjected to a transmitted noise resulting in false-positive detection. Among the noise sources are 60 or 50 Hz power line signals, electromagnetic fields generated by radio stations, power electric equipment, lightnings, etc. Most of these interferences generate electric fields that are distributed around the detector quite uniformly and can be compensated for by employing a symmetrical input circuit with a significant common mode rejection ratio.

69.6 Optoelectronic Motion Detectors

By far, the most popular intrusion sensors are the optoelectronic motion detectors. They rely on electromagnetic radiation in the optical range, specifically having wavelengths from 0.4 to 20 μm. This covers the visible, near, and part of the far-IR spectral ranges. The detectors are primarily used for indication of

movement of people. They operate over distances ranging up to several hundred meters and, depending on the particular need, may have either a narrow or wide field of view.

The operating principle of the optical motion detectors is based on detection of light (either visible or IR) reflected or passively emanated from surface of a moving object into the surrounding space. Such radiation may be originated either by an external light source and then reflected by the object, just like the microwave or ultrasonic waves, or it may be produced by the object itself in form of a natural IR emission. The former case is classified as an active detector and the latter—a passive. Hence, an active detector requires an additional light source, for instance, daylight, electric lamp, IR LED projector, and laser. The passive detectors detect mid- and far-IR natural emission from objects having temperatures that are different from the surroundings.

First, we shall consider limitations of the optoelectronic detectors as opposed to such devices as microwave, ultrasonic devices, and laser detectors. Presently, optoelectronic detectors are used almost exclusively to detect the presence or absence of movement qualitatively rather than quantitatively. In other words, the optoelectronic detectors are very useful to indicate whether an object moves or not, while they cannot distinguish one moving object from another and they can't be utilized to accurately measure distance to a moving object or its velocity. The major application areas for the optoelectronic motion detectors are in security systems (to detect intruders), in energy management (to turn lights on and off), and in the so-called "smart" homes where they can control various appliances, such as air conditioners, cooling fans, and stereo players. They also may be used in robots, toys, point-of-sale advertisements, and novelty products. The most important advantage of an optoelectronic motion detector is simplicity and low cost.

69.6.1 Sensor Structures

A general structure of an optoelectronic motion detector is shown in Figure 69.12a. Regardless of what kind of sensing element is employed, the following components are essential: a focusing device (a lens or curved mirror), a light detecting element, and a threshold comparator. An optoelectronic motion detector resembles a photographic camera. Its focusing component creates on a focal plane an image of the field of view. Instead of an imaging sensor, a "single-pixel" light-sensitive element is used. The element converts the focused light into an electric signal.*

Let's assume that the motion detector is mounted in a room. A focusing lens creates an image of the room on a focal plane where the light-sensitive element is positioned. If the room is unoccupied, the image is static and the output signal from the element is steady/stable. When an "intruder" enters the room and keeps moving, her image on the focal plane also moves. In a certain moment, the intruder's body is displaced for an angle α and the image overlaps with the light sensing element. This is an important point to understand—the detection is produced only at the moment when the object's image either coincides with the detector's surface or clears it. That is, no crossing—no detection. Assuming that the intruder's body creates an image whose photon flux is different from that of the static surroundings, the light-sensitive element responds with a deflecting voltage V. To cause detection, a moving image shall have a certain degree of an optical (or thermal) contrast with its surroundings.

Figure 69.12b shows that the output signal is compared with two thresholds in a window comparator. The purpose of the comparator is to convert the analog signal V into two logic levels: 0—no motion detected and 1—motion is detected. In most cases, signal V from the element first must be amplified and conditioned before it becomes suitable for the threshold comparison. The window comparator contains both the positive and negative thresholds, while signal V is positioned in between. Whenever an image of a moving object overlaps with the light-sensitive element, voltage V deflects from its steady-state position and crosses one of two thresholds. The comparator generates a positive voltage (1), thus indicating

* In a differential sensor, as described in the following text, two "pixels" are employed.

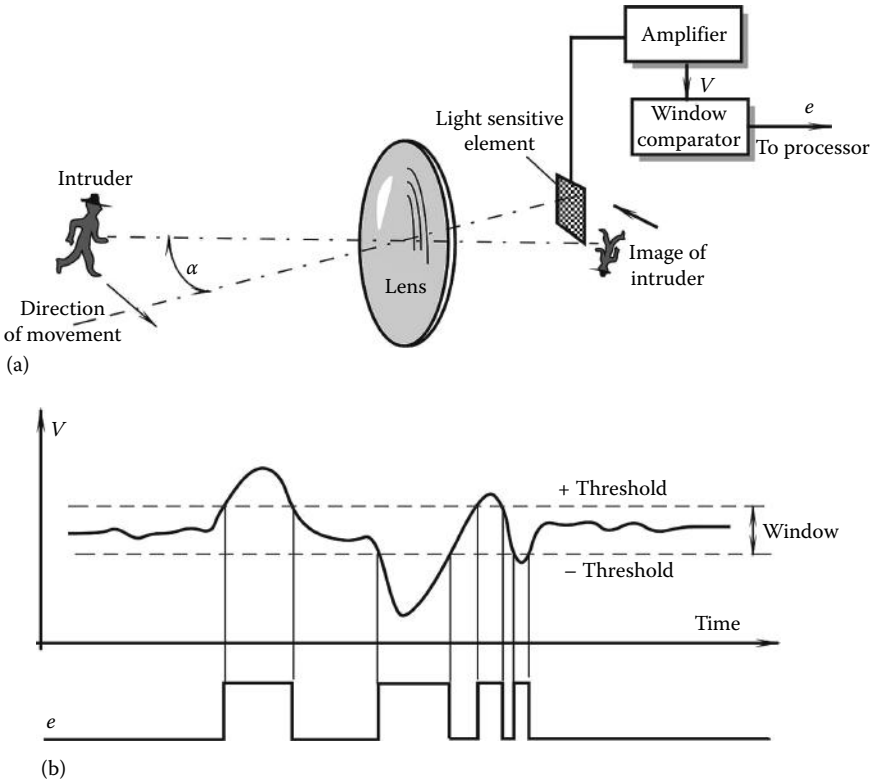


FIGURE 69.12 General arrangement of an optoelectronic motion detector. A lens forms an image of a moving person (intruder). When the image crosses the sensor's optical axis, it covers the sensitive element (a). The element responds with a signal that is amplified and compared with a window threshold in a comparator (b).

detection of movement in the field of view. The operation of this circuit is identical to the threshold circuits described earlier for other types of occupancy detectors.

Note from Figure 69.12 that the detector has quite a narrow field of view: if the intruder keeps moving, her image will overlap with the sensor only once. After that the window comparator output will produce steady zero. This is the result of a small area of the sensing element. In some instances, when a narrow field of view is required, it is quite all right; however, in the majority of cases, a much wider field of view is desirable. This can be achieved by several methods described later.

69.6.2 Multiple Detecting Elements

An array of detecting elements (multiple pixels) may be placed in the focal plane of a wide-angle focusing mirror or lens. Each individual element covers a narrow portion of the field of view, while in combination they encompass a larger area. All detectors in the array either shall be multiplexed or otherwise interconnected to produce a composite detection signal.

69.6.3 Complex Sensor Shapes

If the detecting element's surface area is sufficiently large to cover the entire angle of view, the sensor's surface may be optically broken into smaller elements, thus creating an equivalent of a multiple detector array. To break up the surface area into several parts, one may shape the sensing (detecting) element in an odd pattern like that shown in Figure 69.13a. Each part of the element acts as a separate light detector.

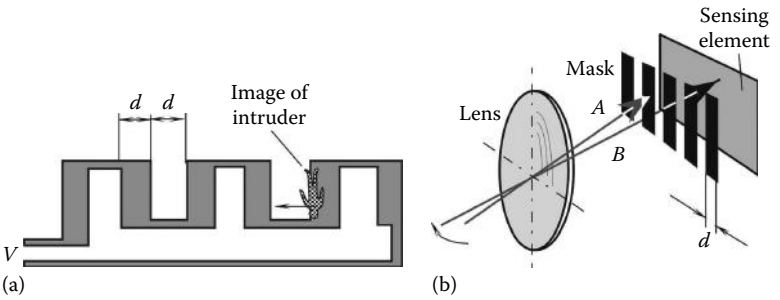


FIGURE 69.13 (a) Complex shape of a sensing element and (b) image distortion mask.

All such detectors are electrically connected either in parallel or in series, being arranged in a serpentine pattern. The parallel or serially connected detectors generate a combined output signal, for instance, voltage *V*, when image of the object moves along the element surface crossing alternatively sensitive and nonsensitive areas. This results in an alternative signal *v* at the detector terminals. For a better sensitivity, such sensitive and nonsensitive areas should be sufficiently large to overlap with most of the object's image.

69.6.4 Image Distortion

Instead of making the detector in a complex shape, an image of an entire field of view may be broken into several parts. This can be done by placing a distortion mask [6] in front of the detector having a sufficiently large area as it is depicted in Figure 69.13b. The mask is opaque and allows formation of an image on the detector's surface only within its clearings. The mask operation is analogous to the complex sensor's shape as described earlier.

69.6.5 Facet Focusing Device

A more popular way of broadening the field of view while employing a small area light detector is to use multiple focusing devices. A focusing mirror or a lens may be divided into arrays of smaller mirrors or lenses called *facets*, resembling an eye of an insect. Each facet works as an individual lens (mirror) thus creating its own image on a common focal plane. The optical axis of each facet is shifted with regard to its neighbor to project an image of an adjacent area of the field of view. All facets together form multiple images as shown in Figure 69.14a. When the object moves, the multiple images also move across the

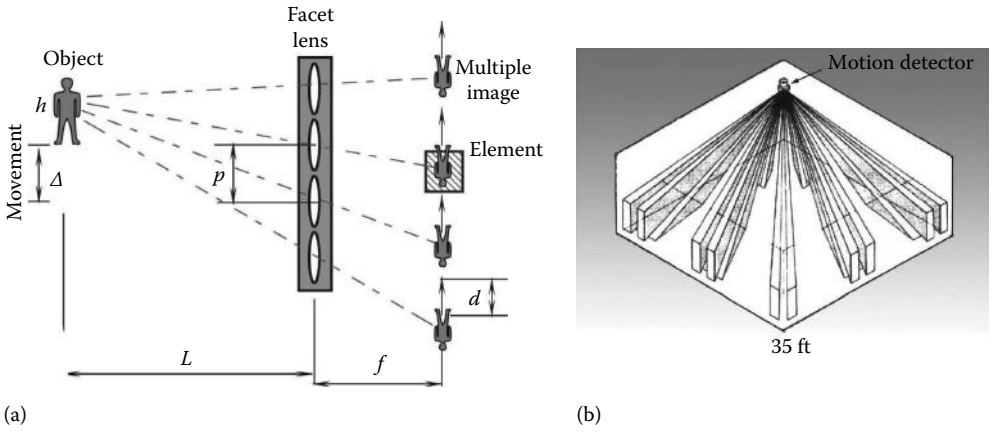


FIGURE 69.14 (a) A facet lens creates multiple images near the sensing element and (b) sensitive zones created by a complex facet lens.

FIGURE 69.15 Example of curved IR faceted Fresnel lens molded of HDPE.

sensing element causing it to produce an alternative signal every time at least one of the images crosses the sensing element. By combining multiple facets, it is possible to shape any desirable detecting pattern in the field of view, in both horizontal and vertical planes. Positioning of the facet lens, focal distances, number, and a pitch of the facets (a distance between the optical axes of two adjacent facets) may be calculated in every case by applying rules of geometrical optics. The following practical formulas may be applied to find the focal length of a facet:

$$f = \frac{Ld}{\Delta} \tag{69.11}$$

and the facet pitch (a distance between their optical axes) is

$$p = 2nd \tag{69.12}$$

where

L is distance to the object
d is width of the sensing element
n is number of the sensing elements (evenly spaced)
Δ is the object's minimum displacement that must result in detection

By combining facets, one may design a lens that covers a large field of view (Figure 69.14b)—up to 180°—where each facet creates a relatively narrow angle sensitive zone. Currently, the facet lenses are primarily used in the mid- and far-IR spectral ranges. These lenses are molded of high-density polyethylene, HDPE (Figure 69.15), and quite inexpensive.

69.6.6 Visible and Near-IR Light Motion Detectors

Most of objects (apart from very hot) radiate electromagnetic waves only in the mid- and far-IR spectral ranges. Hence, visible and near-IR light motion detectors have to rely on additional sources of light that illuminate objects. The light is reflected by the object's surface toward the focusing device for a subsequent detection. The light sources may be the sun, lamps, or invisible near-IR light-emitting diodes (LEDs). The use of visible light for detecting moving objects goes back to 1932 [6] when in the pre-radar era, inventors were looking for ways of detecting moving cars and flying airplanes. In one invention, an airplane detector was built in form of a photographic camera where the focusing lens made of glass was aimed at the sky. A moving plane's image was focused on a selenium photodetector that reacted to a changing contrast in the sky image. Naturally, such a detector could operate only at a daytime to detect planes flying below clouds. Obviously, those detectors were not too practical.

Another version of a visible range light motion detector was patented and produced for less demanding applications: controlling lights in a room [5] and to make interactive toys [7].

To turn the lights off in a nonoccupied room, the visible range motion detector was combined with a timer and a solid-state relay. The detector was activated when the room is illuminated. Visible light photons carry a relatively high energy and may be detected by quantum photovoltaic or photoconductive cells whose detectivity is quite high. Thus, the optical system may be substantially simplified. In the motion switch, the focusing device was built in form of a pinhole lens (Figures 69.16b and 69.17a). Such a lens is just a tiny hole in an opaque foil. To avoid a light diffraction, a hole diameter must be substantially larger than the longest detectable wavelength (red). The motion switch had a three-facet pinhole lens with apertures of 0.2 mm in diameter (Figure 69.16c). A pinhole lens has a theoretically infinite depth of focusing range; hence, a photodetector can be positioned at any distance from the lens. For practical reasons, that distance was calculated for the object's displacement, view angle, and the photoresistor dimensions used in the design. The CdS (cadmium sulfide) photoresistor was selected with a serpentine pattern of a sensing element (Figure 69.16a) and connected into a low frequency ac amplifier. When the

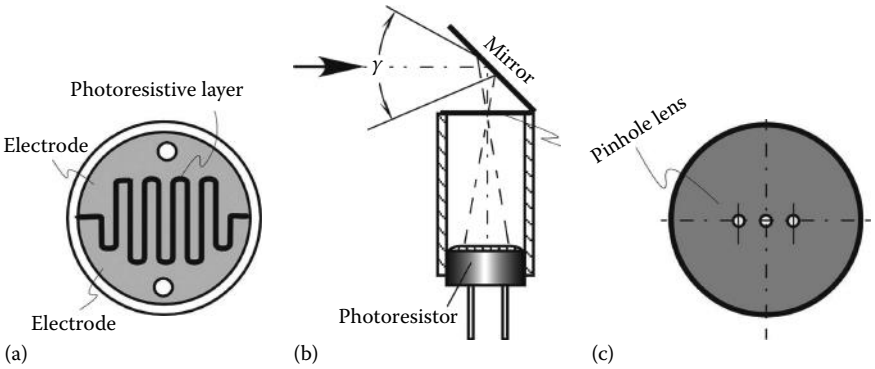


FIGURE 69.16 Simple optical motion detector for a light switch and toys: (a) a sensitive surface of a photoresistor forms a complex sensing element, (b) a flat mirror and a pinhole lens form an image on a surface of the photoresistor, and (c) a triple-pinhole lens.

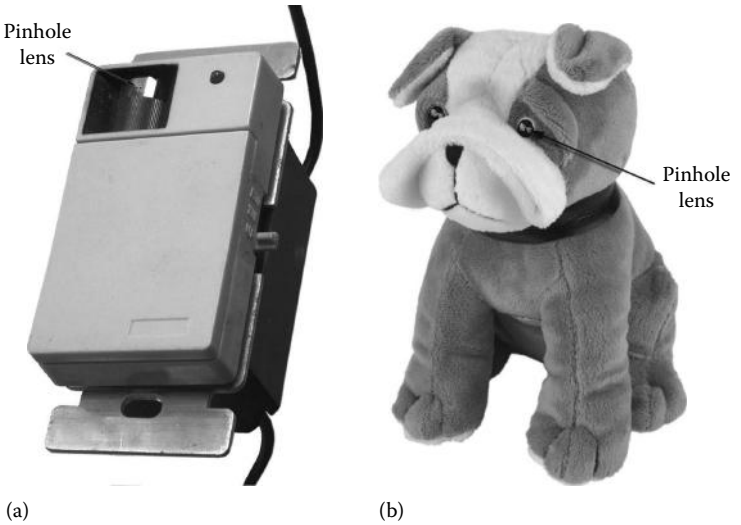


FIGURE 69.17 Motion-controlled light switch with a photoresistor and pinhole lens (a). Interactive toy (b) that reacts to a child's movement—the dog barks when motion is detected.

room was illuminated, the motion sensor acted as a miniature photographic camera: an image from the lens' field of view was created on a surface of the photoresistor. Moving people in the room caused the image to change in such a way as the optical contrast changed across the serpentine pattern of the photoresistor (see Figure 69.13a). In turn, its resistive value was changing, resulting in modulation of the electric current passing through the photoresistor. This signal was further amplified and compared with a predetermined threshold. Upon crossing that threshold, the comparator generated electric pulses that reset a 15 min timer. If no motion was detected within 15 min from the last movement, the timer disabled the solid-state relay to turn lights off. Then, light could be turned on back only manually, because this motion detector can't function in darkness. Thanks to its very low cost, this type of a motion sensor was used in interactive toys that react to movement of children [7]. An example of such a toy is shown in Figure 69.17b where a pinhole lens was built into an eye of a mechanized barking dog. Normally the dog was sitting quietly but when motion in its vicinity was detected, the dog started moving and barking. If you pet it on the back, the barking stopped and the dog wagged the tail (a tactile sensor was installed on the back under the coat).

69.6.7 Far Infrared Motion Detectors

Another version of a motion detector operates in the optical range of thermal radiation, the other name for which is mid- and far IR. Such detectors are responsive to radiative heat exchange between the sensing element and moving object [9–11].

The principle of thermal motion detection is based on the physical theory of natural emission of electromagnetic radiation from any object whose temperature is above absolute zero.

For motion detection, it is essential that surface temperature of an object be different from that of the surrounding objects, so a thermal contrast would exist, just as a visible contrast in the sensors described earlier. All objects emanate thermal radiation from their surfaces and intensity of that radiation is governed by the Stefan–Boltzmann law. If the object is warmer than the surroundings, its thermal radiation is shifted toward shorter wavelengths and the intensity becomes stronger. Most objects whose movement is to be detected are nonmetals; hence, they radiate thermal energy quite uniformly within a hemisphere. Moreover, the dielectric objects generally have high emissivity of thermal radiation. Human skin is a quite good emitter of thermal radiation. Its emissivity is well over 90%. Most of natural and synthetic fabrics also have high emissivities between 0.74 and 0.95.

There are two types of the IR motion detectors: passive and active. Here we briefly overview the passive passive infrared (PIR) detectors as most popular for the security and energy management systems. The PIR sensing element is responsive to mid- and far-IR radiation within a spectral range from approximately 4 to 20 μm where most of the thermal power emanated by humans is concentrated (body surface temperatures ranging from about 28°C and up). There are three types of the sensing elements that are potentially useful for that detector: bolometers, thermopiles, and pyroelectrics; however, the pyroelectric elements are used almost exclusively for motion detection thanks to their simplicity, low cost, high responsivity, and broad dynamic range. Let's see how a pyroelectric may be employed in a practical motion sensor design.

A pyroelectric material generates an electric charge in response to a thermal energy flow through its body. In a very simplified way, it may be described as a secondary effect of a thermal expansion (Figure 69.18). Since all pyroelectrics are also piezoelectrics, absorbed heat causes the front side of the sensing element to expand. This miniscule thermally induced stress leads to development of a piezoelectric charge on the element electrodes. The charge is manifested as voltage across the electrodes deposited on the opposite sides of the material. Unfortunately, the piezoelectric properties of the element have also a negative effect. If the sensor is subjected to a minute mechanical stress due to any external force, like sounds or structural vibrations, it also generates a charge that in many cases is indistinguishable from that caused by the IR heat waves. This would cause a false-positive detection.

To separate thermally induced charges from the piezoelectrically induced charges, a pyroelectric sensor is usually fabricated in a symmetrical form (Figure 69.19a), as a two-pixel element. Two identical

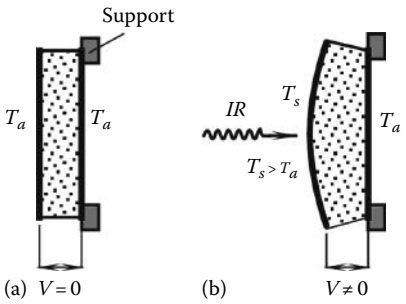


FIGURE 69.18 Simplified model of a pyroelectric effect as a secondary effect of piezoelectricity. Initially, the element has a uniform temperature (a); upon exposure to thermal radiation, its front side warms up and expands, causing a stress-induced charge (b).

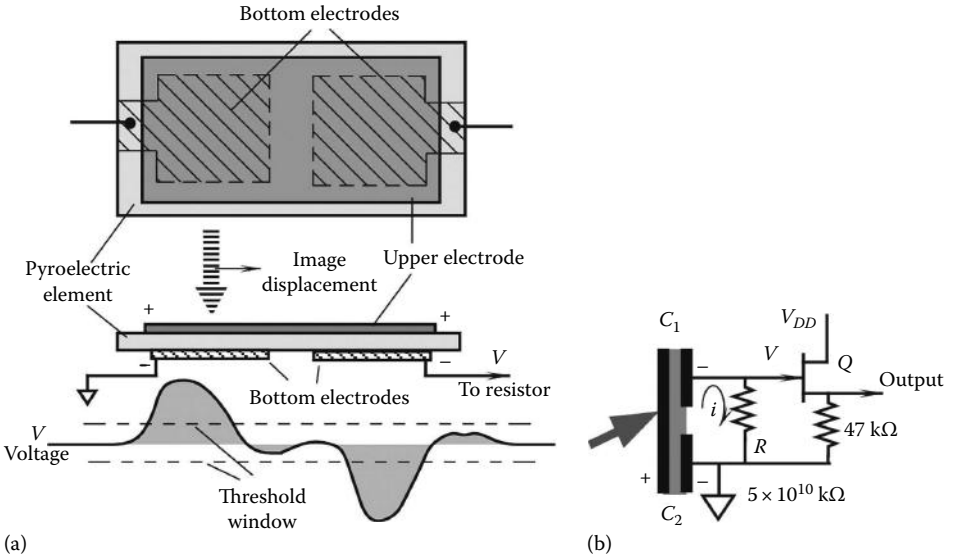


FIGURE 69.19 Dual pyroelectric sensor: sensing element with a front (upper) electrode and two bottom electrodes deposited on a common crystalline substrate (a). A moving thermal image travels from left part of the sensor to the right generating an alternate voltage across the bias resistor, *R* (b).

elements are positioned inside the sensor's housing. The IR radiation is focused on the element by a faceted Fresnel lens, such as depicted in Figure 69.15. The pixels are connected to the electronic circuit in such a manner as to produce the out-of-phase signals when subjected to the same in-phase inputs. The idea is that interferences that are produced by the piezoelectric effect or spurious signals are applied to both electrodes equally (in phase) and thus cancelled at the input of the electronic circuit, while the variable thermal radiation thanks to a correctly designed faceted lens will be absorbed by only one element at a time, thus avoiding a cancellation. This arrangement is called a differential sensor.

One way of fabricating a differential sensor is to deposit two pairs of electrodes on both sides of the pyroelectric element. Each pair forms a capacitor that may be electrically charged either by heat or by a mechanical stress. The electrodes on the upper side of the sensor are connected together forming one continuous electrode, while the two bottom electrodes are separated, thus creating the opposite serially connected capacitors. Depending on the side where the electrodes are positioned, the output signal will have either a positive or negative polarity for a thermal influx. In same applications, a more complex pattern of the sensing electrodes may be required (for instance, to form predetermined detection zones), so that more than one pair of the electrodes are needed. In such a case, for better rejection

of the in-phase signals (common mode rejection), the sensor still should have an even number of pairs where positions of the pairs alternate for a better geometrical symmetry. Sometimes, such an alternating connection is called an interdigitized electrode [8].

A symmetrical sensing element should be mounted in a way assuring that both parts of the element generate the same in-phase signals if subjected to the same external factors. At any moment, the optical component (e.g., a Fresnel lens) must focus a thermal image of a target on the surface of one part of the sensor only, otherwise signals from the thermal image will be cancelled. The element generates a charge only across the electrode pair that is subjected to a heat flux. When a thermal image moves from one electrode to another, the current i flowing from the sensing element to the bias resistor R (Figure 69.19b) changes from zero, to positive, then back to zero, then to negative, and again back to zero. A field-effect transistor Q is used as an impedance converter. Resistor R value must be very high. For example, a typical alternate current generated by the element in response to a moving person is on the order of 1 pA (10^{-12} A). If a desirable output voltage for a specific distance is v = 50 mV, according to Ohm's law, the resistor value should be $R = v/I$ = 50 GΩ (5×10^{10} Ω). Such a resistor cannot be directly connected to a regular electronic circuit; hence, transistor Q serves as a voltage follower (the gain is close to unity). Its typical output impedance is on the order of several kilohm.

69.7 Optical Presence Sensors

A reflection of light is the optical phenomenon that is used quite extensively in detecting not only motion but a mere presence of a human in a monitored area. The operating principle is very simple. The sensor contains two key components: a source of light (usually a near-IR LED) and a photodetector. The LED emits a light beam that illuminates surroundings in the field of view of a photodetector. First, the background reflection from the surrounding objects is established in terms of the output voltage from the photodetector V_0. A new object that appears in a foreground either absorbs more light or reflects more. In most cases it alters the background signal by the increment ΔV that can be detected by a threshold detector in the electronic processor. This sensor is a reflectance monitor. It will not measure a distance to the object because the value of ΔV depends on many factors, such as size of the object, its shape, material, surface finish, and distance to the sensor. The sensor is merely a presence detector, yet in many practical cases it's just what's needed. An example of the sensor application is a presence detector for a bathroom faucet that is used to control flow of water when hands are placed under the water outlet [11,12]. Placement of hands under the faucet controls the actuator of water flow (a solenoid-valve assembly). A similar detector is frequently employed in hand dryers, toilet tanks, light switches, robotic vacuum cleaners, and many other products.

Figure 69.20a and b show two possible locations of the sensor in a water fixture. One location is on the spout while in the other the sensor is built in directly into the faucet. It's important to make sure that the

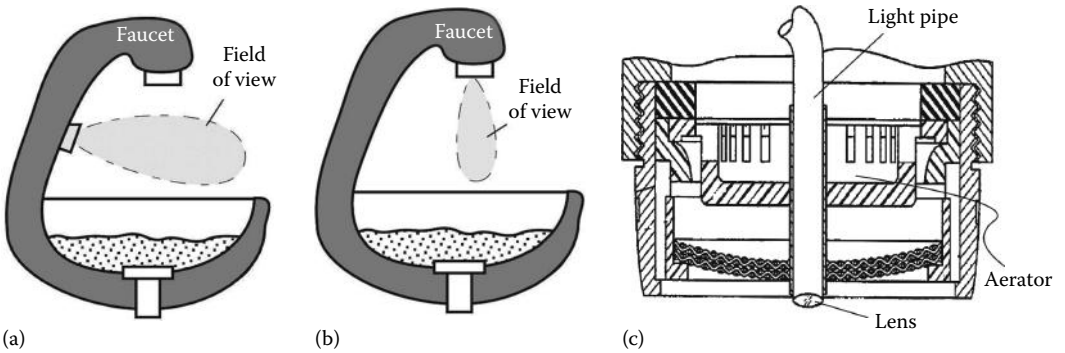


FIGURE 69.20 Installation of the optical presence detector into a spout (a) and faucet (b). Cross-sectional view (c) of the faucet with a light pipe. (Adapted from Parsons, N.E. et al. Passive sensors for automatic faucets and bathroom flushers, U.S. Patent No. 7,396,000, 2008.)

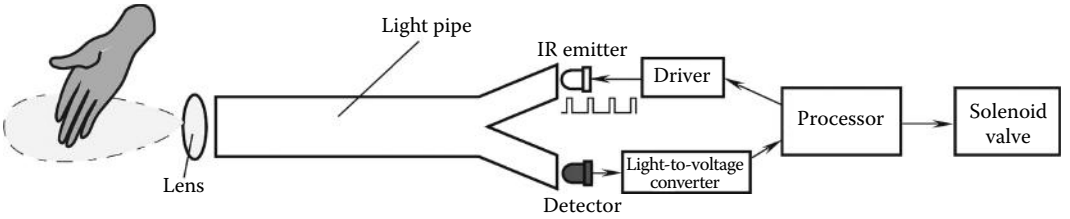


FIGURE 69.21 Block diagram of a water flow controller with the optical presence detector.

detection area is situated where the hands are normally being placed. Figure 69.20c illustrated the faucet having a light pipe and the parts that are normally needed for dispensing water. Figure 69.21 shows a block diagram of the water flow control system. The light pipe can be a bundle of the optical fibers or a solid translucent rod molded of polycarbonate resin. Usually, the emitted light is modulated by pulses. This helps to separate the reflected light signals into a background (ambient) component and that controlled by the pulsing LED driver, since the ambient light is a dc or slow-changing signal. Note that the background component also may contain a pulsing signal since some light will be constantly reflected from it, for example, from a sink surface. This can be taken care of by both the sensitivity adjustment and by detecting only a variable component of signal ΔV. It is possible to eliminate the pilot light LED and entirely rely on the ambient illumination by independent light sources [13]; however, this design may be not as robust as the one with a modulated LED.

69.8 Pressure-Gradient Sensors

An efficient sensor can be designed to detect intrusion into a closed room by monitoring small variations in the atmospheric air pressure resulting from opening doors and windows or movement of people. In general, variations in air pressure can be monitored by a conventional air pressure sensor. However, this is not an efficient solution. A conventional air pressure sensor is designed for a relatively large span of the input pressures, while the peak amplitudes of the pressure variations that are associated with intrusion are very small—over three orders of magnitude smaller than a conventional pressure sensor's span. In fact, these variations approach the noise floor of stress-sensitive membrane pressure sensors. Besides, such a sensor is just not sensitive enough for them. Appending the sensor with a high-gain amplifier is not a solution because noise will be amplified as well. The solution would be to design a sensor having a narrow pressure span but high sensitivity. It is also desirable to make the sensor responsive only to the pressure changes rather than to an absolute value of pressure. Preferably the sensor should produce a signal similar to a first derivative of the air pressure. Since the only purpose of the sensor is to detect intrusion and not measure the actual air pressure, accuracy requirements can be significantly relaxed. This sensor produces a qualitative rather than quantitative output.

High sensitivity can be achieved by making the sensing membrane very thin with a relatively large area. An example of the intrusion air pressure sensor design [14] is shown in Figure 69.22. The main part of the sensor is an enclosed chamber. The left wall of the chamber is covered with a thin stretched membrane made of a plastic or metal foil having thickness on the order of 25 μm. The membrane area should be relatively large—about 200 mm^2 or larger. The right side of the chamber is a rigid backplate with a small venting hole whose purpose is to equalize static air pressures inside and outside of the chamber. The distance d from the membrane to the backplate is monitored by a built-in displacement detector (sensor). All exterior surfaces of the sensor are exposed to ambient air. When all doors and window in the monitored room are closed, the ambient air pressure is either static or changes slowly. Thanks to the vent in the backplate, pressures P_h inside and outside of the sensor's chamber are equal. When the door or window opens, the ambient air pressure changes slightly but rapidly. Because the vent is very narrow and air has a finite viscosity, air pressure P_h inside the chamber can't change instantly; thus, any changes inside the chamber will lag behind the outside changes. The phase lag creates a temporary air pressure

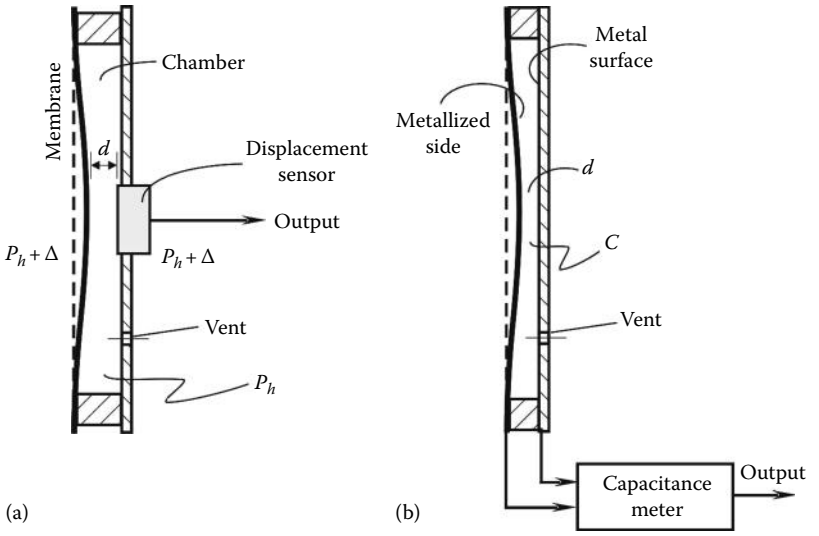


FIGURE 69.22 Air pressure-gradient sensor (a) and the sensor with a capacitive displacement detector (b).

differential across the membrane that deflects it to or from the backplate in relation to the differential amplitude and sign. The distance *d* from the membrane to the backplate is measured by a displacement sensor and is used as an indication of the intrusion. When the differential pressure is small, the membrane remains substantially flat and the distance *d* is at its base level. Figure 69.23 illustrates the timing diagrams of air pressures inside and outside of the sensing chamber and the differential pressure across the membrane. Note that the signal representative of the displacement *d* is compared with a threshold to detect an intrusion.

There are numerous ways of designing a displacement senor for monitoring the membrane deflection [15]. As an example, Figure 69.22b illustrates a capacitive displacement sensor, where the sensing chamber was built in form of a flat capacitor with two plates. The first plate of a capacitor is a metal foil (or metalized plastic membrane) and the other plate is a metal layer on the backplate. The baseline gap *d* between the membrane and backplate should be rather small—a few millimeters maximum. A value

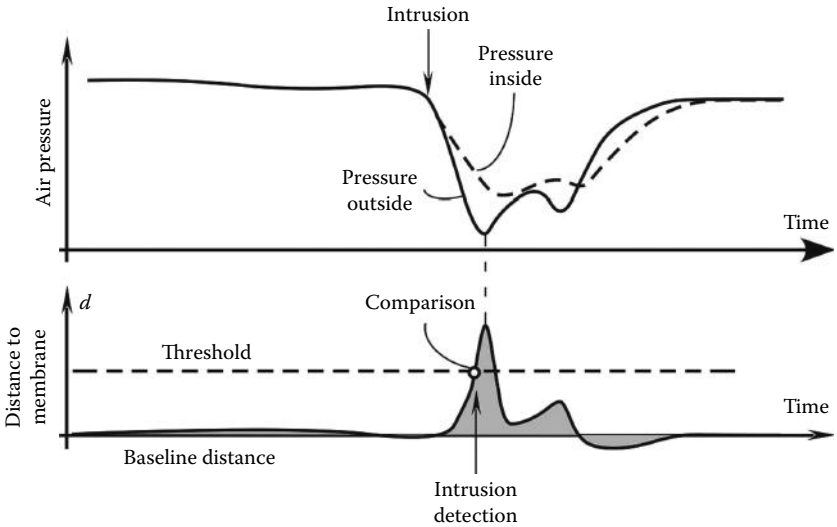


FIGURE 69.23 Timing diagrams for the pressure-gradient detector.

of the capacitance C will change when distance d varies according to the air pressure differential. The capacitance variations are measured and converted into a useful signal.

An alternative design of the earlier approach may include a solid-state MEMS sensor with a thermo-anemometer as a flow sensor [16]. Such sensors can detect pressure gradients as low as a few Pa (pascal) or mm of H_2O. This is a sufficient sensitivity to detect minute air movements in a room. However, unlike the sensor of Figure 69.22 that outputs the signal proportional to the rate of change in differential pressure, to detect changes in air pressure, the flow sensor would require a differentiator as part of the signal processing.

References

1. Blumenkrantz, S. (1989) *Personal and Organizational Security Handbook*. Government Data Publications, Washington, DC.
2. Ryser, P. and Pfister, G. (1991) Optical fire and security technology: Sensor principles and detection intelligence. In: *Transducers'91. International Conference on Solid-State Sensors and Actuators. Digest of Technical Papers*, pp. 579–583, ©IEEE, San Francisco, CA.
3. Long, D.J. (1975) Occupancy detector apparatus for automotive safety system. U.S. Patent No. 3,898,472
4. Fraden, J. (1991) Apparatus and method for detecting movement of an object. U.S. Patent No. 5,019,804.
5. Fraden, J. (1984) Motion discontinuance detection system and method. U.S. Patent No. 4,450,351.
6. Fitz Gerald, A. S. (1932) Photo-electric system. U.S. Patent No. 2,016,032.
7. Fraden, J. (1984) Toy including motion-detecting means for activating same. U.S. Patent No. 4,479,329.
8. Fraden, J. (1988) Motion detector. U.S. Patent No. 4,769,545.
9. Fraden, J. (1990) Active infrared motion detector and method for detecting movement. U.S. Patent No. 4,896,039.
10. Fraden, J. (1992) Active far infrared detectors. In: *Temperature, Its Measurement and Control in Science and Industry*. Vol. 6, Part 2, pp. 831–836, ©AIP, New York.
11. Parsons, N. E. et al. (1999) Object-sensor-based flow-control system employing fiber-optic signal transmission. U.S. Patent No. 5,984,262.
12. Parsons, N. E. et al. (2003) Automatic flow controller employing energy-conservation mode. U.S. Patent No. 6,619,614.
13. Parsons, N. E. et al. (2008) Passive sensors for automatic faucets and bathroom flushers. U.S. Patent No. 7,396,000.
14. Fraden, J. (2007) Alarm system with air pressure detector. U.S. Patent Publication No. US 2008/0055079.
15. Fraden, J. (2010) *Handbook of Modern Sensors*. 4th Edn., Chapter 6, Springer Verlag, New York.
16. Fraden, J. (2009) Detector of low levels of gas pressure and flow. U.S. Patent No. 7,490,512.

VIII

Environmental

70

GPR Principles and Applications

Jan Francke
International Groundradar Consulting, Inc.

70.1 Introduction

Ground-penetrating radar, also known as surface-penetrating radar, geo-radar, or more commonly by the abbreviation GPR, has seen significantly increased acceptance as a viable near-surface geophysical technique in recent years. The practice of employing radio waves to image the subsurface dates to work conducted on glaciers in the Austrian Alps in the 1920s (Stern, 1929). Research on the ice-penetrating capabilities of radio waves languished until the late 1950s, when it was noted that newly installed radar altimeters on aircraft could penetrate through the Greenland icecap and display the aircraft's height above the underlying bedrock, which led to a number of mishaps (Waite and Schmidt, 1961). Although the ability to penetrate ice was a limitation for the reliance on radar altimetry for low-altitude flying in Polar regions, agencies such as NASA took an interest in the ability of radar to penetrate the ground in suitable environments. A rudimentary GPR system was launched onboard Apollo 17 in an attempt to determine surface conditions prior to a manned mission to the moon (Simmons et al., 1972).

Stemming from NASA's requirements for instrument portability and miniaturization and research conducted by the Atomic Energy Commission, the technology was commercialized in the mid-1970s, initially for archaeological investigations (Morey, 1974). However, GPR's potential to locate discrete targets in the ground, such as buried pipes, cables, tunnels, or land mines, was quickly realized. It was also found that low-frequency radar waves (10–100 MHz) could image geological contacts in some environments to depths of dozens of meters (Clough, 1976; Thierbach, 1974). Sedimentologists, geologists, groundwater hydrologists, and the mineral exploration sector all experimented with the technology in various environments. This research was assisted by work conducted to better understand the

electrical properties of geological materials at radio frequencies as well as the relationship between electrical conductivity and dielectric polarization of such media (Olhoeft, 1975, 1987).

By the late 1980s, the major restrictions on the further use of GPR as a geophysical tool were the bulkiness and power requirements of the instrumentation. Furthermore, data were generally recorded on electrostatic plotters, limiting the geophysicist's ability to employ advanced seismic processing techniques, which were keeping pace with the rapidly developing power of desktop computing.

The advent of portable computing and fast A/D microchips in the late 1980s led to a significant reduction in power requirements, more portable instrumentation, and the ability to digitally capture radar data for subsequent processing.

Since the early 1990s, the possible applications for GPR have burgeoned to include the mapping of pavement thickness and cracks at highway speeds (Saarenketo and Scullion, 2000; Spagnolini and Rampa, 1999), network-wide mapping of utilities in full 3D imaging (Eide and Hjelmstad, 2002), and the detection of movement in collapsed buildings (Sachs et al., 2008; Zaikov et al., 2008). GPR has also led to the development of subtechnologies such as wall-penetrating radar (Nag et al., 2001), which is used for the detection of live targets through nonmetallic walls by law enforcement of military personnel, and pipe-penetrating radar (Ariaratnum et al., 2005), which is employed within pipes to detect variations in wall thickness and adjacent voids.

Regardless of the application, all penetrating-radar technologies rely on electromagnetic (EM) fields to probe lossy dielectric materials to image variations in material properties. In most geological material, as well as manmade media such as concrete and asphalt, EM fields will penetrate some distance before being attenuated. This depth could range from many kilometers through Antarctic ice sheets to a few centimeters in some concretes.

70.2 Radar Wave Propagation in the Subsurface

The radar signals are emitted by a transmitter and penetrate into the subsurface as nondispersive waves. When the signals encounter a variation in the subsurface material's impedance, the signals are scattered or reflected, with some of the energy returning back to the surface, where it is captured and digitized by a receiver. Due to the nondispersive nature of radar, the returned reflections are similar to the transmitted signal, making signal recognition and interpretations simple. The frequency range of GPR signals is generally between 1 MHz and 2 GHz. Below the MHz range, the EM fields become dispersive, whereas higher frequencies are limited in their practical application because several factors increase attenuation to a degree where penetration would be limited to a few millimeters.

Maxwell's equations mathematically describe the physics of the EM theory and thus GPR energy propagation as a 3D polarized vector field. In geological media, the waves travel at velocities lower than the speed of light and are scattered by variations in the electrical and magnetic properties of the subsurface. The subsurface media through which GPR energy must penetrate are referred to as semiconductors, or dielectrics, and are characterized by three EM properties: electrical conductivity, electrical permittivity, and magnetic permeability.

Electrical conductivity, σ, is the measure of a material's ability to transmit a direct current, which results in energy dissipation through the conversion of electrical energy to heat. Dielectric permittivity, ε, refers to the degree to which a geological medium resists the flow of electrical charge divided by the degree to which free spaces resists the same charge. The dielectric permittivity is thus defined as the ratio of the electric displacement to the electric field strength and is an important quantity for GPR. Most often, the terms *relative permittivity* or *dielectric constant* are used, which may be defined as the ratio of $\varepsilon/\varepsilon_0$, where ε_0 is the permittivity of a vacuum. The velocity of an EM wave propagating through a medium is the reciprocal of the square root of the dielectric permittivity.

Magnetic permeability, μ, is the result of electron spin and motion in atomic orbits and also results in energy loss and storage. In most GPR applications, the effect of magnetic permeability is negligible and is often excluded from calculations.

GPR is most effective in environments with low electrical loss media. If the media through which GPR is imaging had $\sigma = 0$, the radar system would be able to image great depths, limited only by the power of the radar transmitter and the sensitivity of the receiver. However, neither geological material nor manmade media are perfect insulators, and their semiconductive nature is highly variable and often unpredictable.

One of the greatest determinant factors of the GPR range in most media is the concentration of both pore and bound water. Generally, bulk minerals are good GPR targets with low pore water content and thus exhibit low dielectric permittivities. Conversely, sands and soils have pore spaces filled with air, water, and other minerals. Furthermore, the water within the pores can often contain ions from dissolved minerals, with the ionic mobility within the water being the dominant contributor to bulk electrical conductivity.

A number of factors control the maximum range of a radar system. As a radar wave field encounters a discontinuity in impedance, a portion of the energy is returned, while the remainder continues on to greater depths. In an environment with a large number of impedance changes with depth, there is a decrease in the portion of the original imparted energy that is still available to continue deeper. Since the radar energy path is generally two-way, these losses, known as reflection losses, are doubled due to the return journey to the surface. Although this describes a simple layered model, the effects on penetration of which may be modeled easily, most GPR environments also include scattering effects from unwanted targets such as gravels and boulders. Scattering losses cause the incident radar energy to reflect in multiple directions, with only a portion aimed back toward the receiving antenna.

The electrical conductivity of the media being propagated through leads to material losses, which have a significant effect on the maximum range of a GPR system. With increasing conductivity, the radar energy is attenuated more quickly, thereby restricting the effective penetration depths.

The final common type of loss is due to the geometric spreading of the radar energy. Although generally difficult to quantify in real-world complex geologies, these losses are exponentially proportional to the range being imaged. The maximum range of a GPR system can be generally estimated by summing these losses along with the losses encountered within the instrument itself.

70.3 GPR Instrumentation

A generalized radar block diagram is shown in Figure 70.1. The transmitted energy may be a waveform modulated in frequency, amplitude, or phase, as well as random noise. The amount of power, the pulse rate, and the bandwidth will depend on the lossiness of the subsurface media and the target dimensions. The transmitting and receiving antennas are usually matched and are also designed with target dimensions and signal polarization taken into account, as well as practical size and weight limitations. The receiver must be designed to capture the returned signals quickly and at a sufficient dynamic range to overcome the losses encountered by the radar signals on their two-way journey from the surface to the target and back to the surface.

Of the possible modulation techniques, amplitude modulation, commonly referred to as impulse radar, is most commonly employed for GPR instrumentation. Presently, there are over a dozen manufacturers of commercial GPR systems worldwide, with all but two employing impulse transmitters. In such systems, a train of impulses are applied to a transmitting antenna. The impulse sequence may have amplitudes ranging from 20 to 400 V, a width of a few nanoseconds, and a pulse repetition frequency (PRF) in the range of 100–800 MHz. The signals returned from the ground are applied to a flash A/D converter or, more commonly, to a sequential-sampling receiver with a high-speed sample-and-hold circuit. The requirement for ultrafast sampling generally restricts impulse radars to the range of 10 MHz to 1 GHz.

For radar frequencies greater than 1 GHz, frequency-modulated continuous wave (FMCW) radar architectures become more suitable. Thus, such systems are generally employed for shallow,

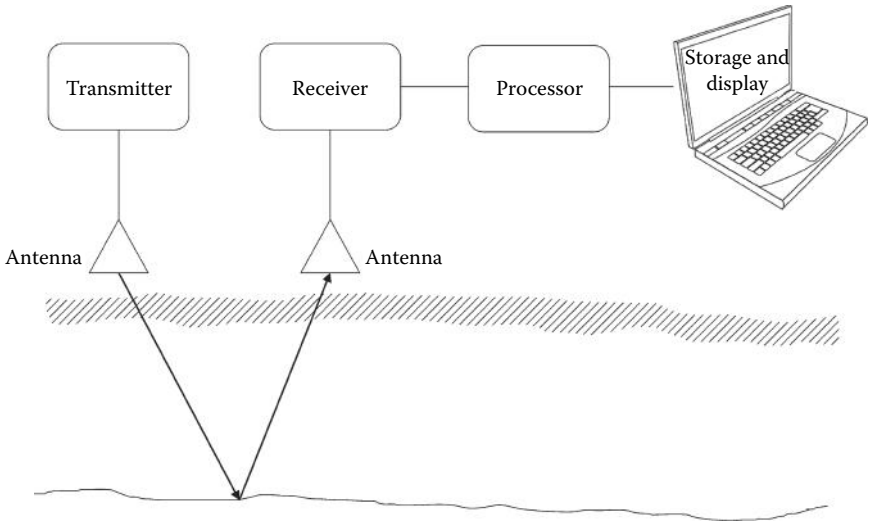


FIGURE 70.1 General radar block diagram.

high-resolution applications such as pavement analysis and landmine detection. In an FMCW system, a continuously changing carrier frequency is repeatedly transmitted using a voltage-controlled oscillator over a range of frequencies. The returned signal is combined with a sample of the original transmitted waveform, which produces a difference frequency related to the phase of the received signal and thereby a different time delay and range in relation to the target reflector. To produce a radar signal similar in the time domain, the difference frequency must be derived from an *I*/*Q* mixer pair (Daniels, 2004).

70.4 Commercial Radar Systems

As of 2012, it is estimated that there are between 10,000 and 20,000 GPR systems in existence worldwide. Although the majority of these are used on rare occasion for research purposes, many thousand systems are employed regularly for commercial surveys. The first commercial manufacturer of GPR systems, Geophysical Survey Systems, Inc., of Salem, NH, United States (www.geophysical.com), remains the world leader in terms of sales volumes, followed closely by Sensors and Software, Inc., of Mississauga, ON, Canada (www.sensoft.ca), and Malå Geosience of Malå, Sweden (www.malags.se). In addition, over the last decade, Ingegneria dei Sistemi SpA of Pisa, Italy, has become known for their multichannel GPR systems. More recently, researchers from the former Eastern Bloc countries and from China and South Korea have also begun to commercialize GPR systems.

Each of these manufacturers tends to focus on the lucrative civil infrastructure market, which includes rebar detection in concrete, road, and railbed analysis, utility mapping, and void identification. Although the instrumentation is packaged differently and the robustness and flexibility of the design vary greatly, all systems generally employ impulse transmitters with shielded bowtie antennas and sequential-sampling receivers. Figure 70.2 shows a typical commercial cart-mounted GPR system employed for utilities detection.

For low-frequency radar applications (10–200 MHz), manufacturers rely on resistively loaded dipoles for the transmitter and receiver antennas, as depicted in Figure 70.3. Such antennas can be unwieldy in difficult terrain, such as in forests, where cut lines up to 4 m wide are required. Alternatively, dipole antennas may be placed collinearly, allowing the radar system to be housed within a long "snake" and towed easily along a trail. The Malå Geoscience's RTA system and the UltraGPR from Groundradar, Inc., in Vancouver, Canada (www.groundradar.com), both employ such a design with real-time radar receivers for long-range imaging of mineral deposits (Figure 70.4). Such a collinear orientation of

FIGURE 70.2 Typical cart-mounted utility-detection GPR.

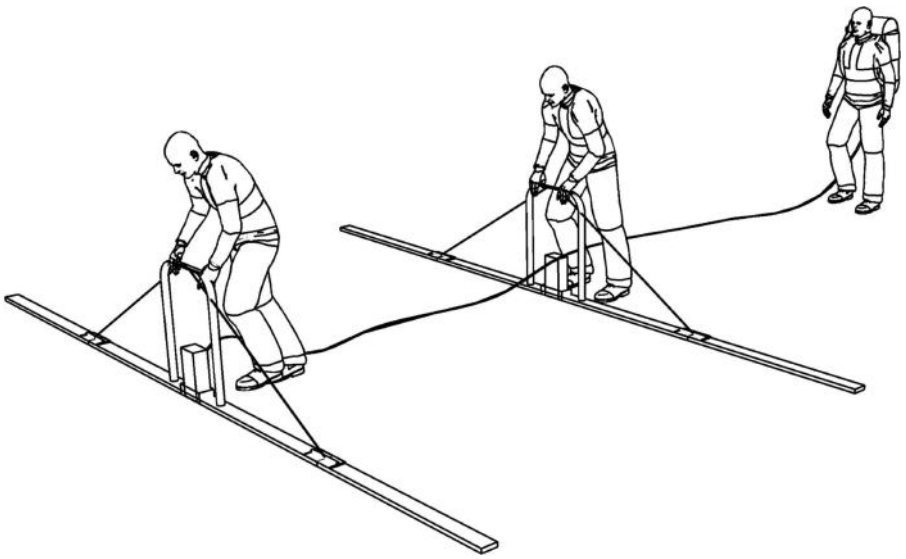


FIGURE 70.3 Low-frequency dipole antennas.

radar antennas often produces clutter patterns on the received radar profiles due to the end fire from the antennas reflecting off of aboveground targets such as trees.

In recent years, the availability of large array radar systems for 3D mapping of wide swaths has grown significantly (Figure 70.5). One of the pioneers in the field, 3D-Radar AS of Trondheim, Norway (www.3d-radar.com), has commercialized large radar arrays for pavement and utility mapping using FMCW radar architecture. Radar Portal, Pty., of Brisbane, Australia (www.radarportal.com.au), specializes in integrating large radar arrays using noise-modulated FMCW radars with complementary sensors such a high-speed falling weight deflectometers, panoptic cameras, and laser profilometers to comprehensively map road surfaces and structures at highway speeds.

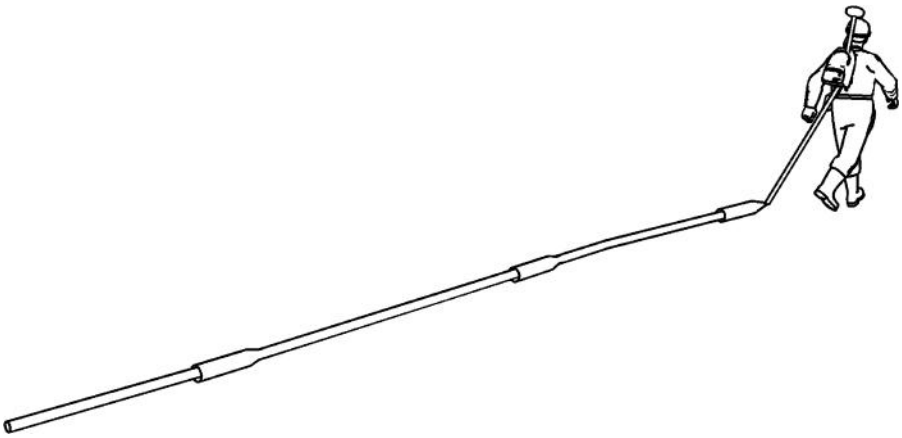


FIGURE 70.4 Low-frequency collinear towed dipole antennas.

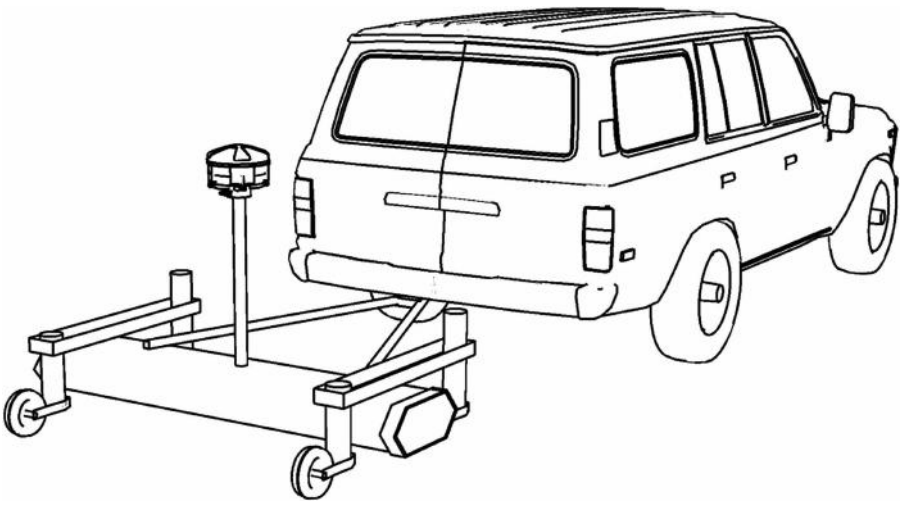


FIGURE 70.5 Array of GPR antennas for pavement and utility surveys.

70.5 Typical Commercial Applications

A review of GPR manufacturers' websites would suggest that GPR is a panacea to all near-surface imaging problems. What are often neglected or downplayed are the inherent limitations of GPR in many, if not most, environments. These limitations are governed by the laws of physics, and no degree of advanced radar design, virtual-reality 3D software, or high instrument price can circumvent them. Listed in the following text are the most common applications for GPR technology for commercial surveying. A myriad of other applications are possible, ranging from tree root imaging and the detection of wood rot and landmine detection.

70.5.1 Utilities Detection

Perhaps the first substantial commercial application of GPR technology was the mapping of buried utilities. With many urban areas, "as-planned" maps are often unrelated to "as-is" conditions underground. The severing of a sewer or natural gas line can have disastrous consequences. Traditional methods such

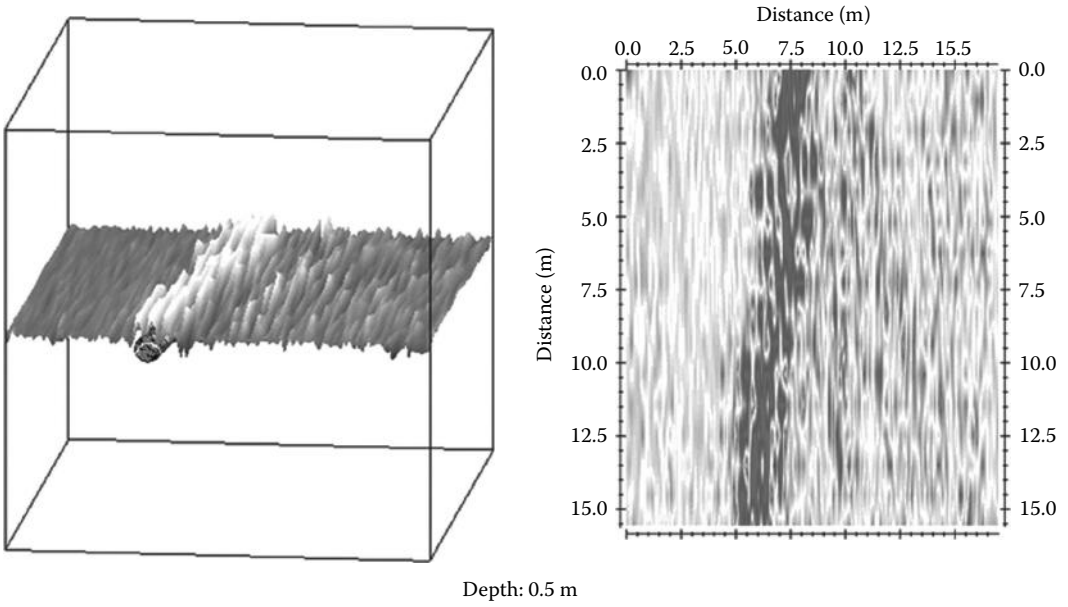


FIGURE 70.6 3D and plan view representation of radar-detected pipe. (Courtesy of Malå Geoscience.)

as radiodetection can yield ambiguous depths or inaccurate results in clustered pipes. While disruptive trenching has given way to trenchless technology such as horizontal directional drilling, which enables far more rapid installations in urban settings, engineers still rely on a degree of chance to avoid collateral damage to preexisting plants. As existing plants age and cities expand, the need for a tool to rapidly and accurately map existing utilities has become critical.

Although seemingly an ideal solution, particularly for metal pipes and cables, GPR suffers from some limitations. The most common is due to signal attenuation in the subsurface. Many cities built along floodplains or in coastal zones are situated atop clay units, which significantly restrict radar suitability to as shallow as a few decimeters. It has been estimated that in the United Kingdom, only 60%–70% of the total land area may be suitable for GPR surveying (Daniels, 2004). A further limitation is that in a situation where radiodetection is ambiguous and GPR is most useful, such as clusters of pipes and cables, radar is generally ineffective without extreme care in sensor positioning and complete high-density 3D imaging by surveying the region in perpendicular directions. Such surveys are time-consuming with single antennas, often require positioning accuracies not practical in urban settings, and generally still produce data volumes subject to ambiguous interpretations.

Nevertheless, a significant portion of the commercial GPR market is aimed at utilities detection. In many environments, such as parking lots and open fields with uncluttered pipe layouts, GPR can be a highly effective tool at mapping utilities. Figure 70.6 shows a typical radar cross section of a series of utilities, each shown a characteristic hyperbolic reflection signature.

70.5.2 Pavement and Railbed Analysis

Traditionally, the health of a road or highway has been measured using methods such as falling weight deflectometers to determine dynamic moduli and coring to measure the thickness of various construction layers. Together with other sensors such as surface roughness, profilometers, and data on traffic flow, estimates of the residual life of the road can be determined. These approaches are time-consuming and costly and also interrupt traffic flow during surveys. Furthermore, coring disturbs the road surface and provides information only at disparate points.

This was one of the first commercial applications for GPR, as radar can be used to determine the thickness of asphalt layers and concrete slabs, the spacing and location of slab reinforcement, their depth of cover, dowel bars, joining slabs, and cracking. Compared to traditional methods, GPR can provide continuous coverage at a fraction of the cost (Noureldin et al., 2003). Modern systems operate at highway speeds, eliminating the disruption of traffic flow (Maser, 2001). Although depth measurements generally have a 90%–95% accuracy rate, variations in radar velocity due to differing moisture contents both across a lane and along a stretch of road can lead to inaccuracies in layer depths (Saarenketo, 1997). In addition, the construction layers of the roadbed must have sufficient variation in impedance to create radar reflections.

A similar application of radar in the area of transport infrastructure is the rapid surveying of rail ballast conditions. GPR is regularly used on railway tracks to provide continuous measurements of the track substructure to assist in maintenance decision making. The track substructure consists of the ballast, the subballast, and subgrade layers, each having a significant influence on track performance. Designed to facilitate water drainage and distribute load, a deterioration of these layers or an accumulation of moisture forces trains to slow down or, at worst, can cause derailments. A significant portion of operating expenses for a rail company is the correction of rough track caused by shifting or mixing of these layers under repeated heavy loads. Standard monitoring methods include trenching, cone penetrometer tests, and core drilling, all of which cause traffic delays and further disrupt the integrity of the layers.

GPR has been applied to provide a continuous profile of the track substructure to characterize layer thicknesses and degree of mixing as well as accumulations of moisture (Al-Nuaimy, 2004a; Gallagher et al., 1999). Often employing radar antennas in a range of 800 MHz to 1.5 GHz, there is generally sufficient variation in the impedance between the substructure layers to produce radar reflections. While layer thicknesses can be mapped by tracking these reflections, the amplitude of the reflections can provide some indication of variations in moisture content (Olhoeft and Selig, 2002). In regions where water is trapped in the substructure, the higher dielectric contrast produces stronger radar reflections. The texture of the radar data can also be analyzed to produce a relative indication on the coarseness of the ballast.

Antennas may be either ground coupled or air coupled. Ground-coupled antennas on roads are designed to handle rough road conditions, whereas such antennas on railbeds must be able to cope with frequent obstructions higher than the sleepers (Figure 70.9). High-frequency air-coupled horn antennas have been used for both road and rail applications (Selig et al., 2008). Often, an array of antennas is used to provide maximum coverage across a lane or railbed (Al-Nuaimy, 2004b).

70.5.3 Concrete and NDT

The advent of high sampling rate A/D converters for high-frequency radar receivers and the further miniaturization of electronics have led to the proliferation of relatively low-cost concrete radar systems. Radar now forms an important role in the nondestructive testing (NDT) of in situ concrete. In many jurisdictions, standard practices for GPR use in concrete NDT have been published (American Concrete Institute, 1998; Concrete Society, 1997; Texas Transportation Institute, 2001). Products such as the Hilti PS 1000 X-Scan (www.hilti.com) have moved concrete GPR from the exclusive realm of geophysics to a consumer product on par with inspection cameras. The acquisition of accurate 3D grids for rebar mapping is now a trivial task using adhesive templates (Figure 70.7). Applications for these systems include the location of rebar and posttensioning cables as well as metal and plastic conduits and voids, with the former being the most common. The optimization of antenna orientation to maximize the effect of signal polarization is important for both the imaging of rebar and for attempting to image through them to deeper targets (Annan et al., 2002).

Recent research in this field has focused on determining the size and shape of voids (Pollock et al., 2008), the location of honeycombing and cracking (Beutel et al., 2008), and the estimation of rebar size and corrosion (Lai et al., 2010). Each of these subjects has been studied extensively through numerical modeling and laboratory measurements, although practical applications have remained elusive.

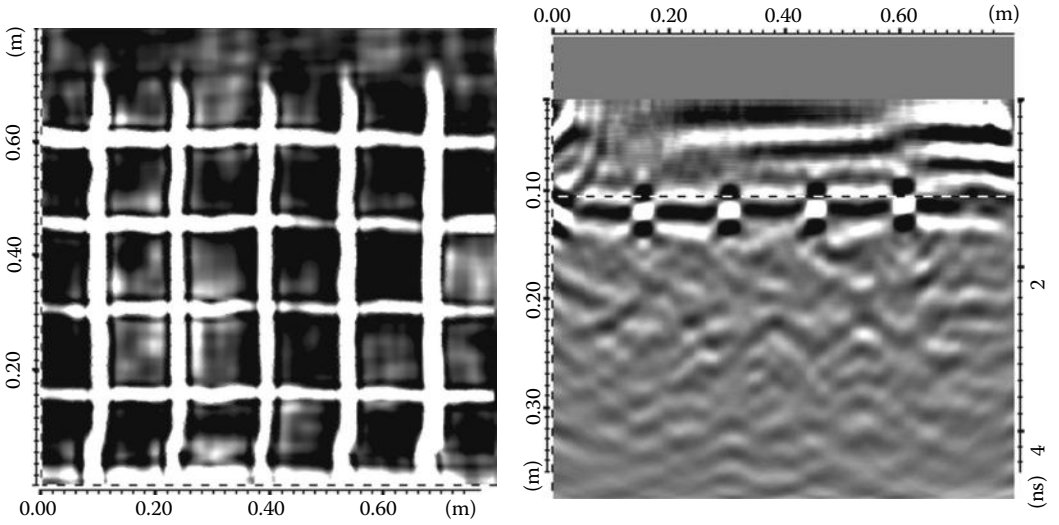


FIGURE 70.7 Plan view and radar cross section of rebar in concrete. (Courtesy of Malå Geoscience.)

70.5.4 Snow and Ice

The original applications of GPR, snow and ice, provide the ideal environment for radar penetration. Extremely low conductivity, low dielectric permittivity, and no magnetic relaxation processes combine to permit penetration to many kilometers. Spurred by an interest in climate change, ground-based, as well as increasingly airborne, GPR surveys are being regularly conducted by researchers over glaciers with impressive results (Arcone, 1996; Arcone et al., 2000; Maijala at al., 1998). Such data are of use in the examination of ice structures and paleoclimatic studies.

Common commercial applications include the delineation of ice thickness for ice roads and runways as well as the creation of snow thickness maps on ski slopes. Recent discoveries of major oil and diamond resources in the Arctic have created the need for an extensive network of seasonal ice roads and runways. Ice thickness is traditionally controlled by manual auger coring, a process that is both slow and inherently inferential due to practical limitations on hole spacing. Furthermore, as weather conditions change, it is critical the ice roads and runways be monitored regularly in order to ensure that minimum ice thicknesses are maintained. High-frequency antennas (500 MHz–1 GHz) are often used for this application. Water lenses in the ice may cause significant distortion in the radar profiles or inhibit further radar penetration altogether (Galley et al., 2009). The ideal radar velocity of ice is approximately 0.17 m/ns, while water has a velocity of 0.03 m/ns. Not compensating for this significant variation in velocity will lead to inaccurate depth maps.

Recent developments have included the use of multichannel radar systems designed for roadbed surveys to automatically collect radar velocity information. While allowing for more precise ice depth maps, the dielectric permittivity variations, calculated from the radar velocities, may also yield information on the load-bearing characteristics of the ice (Proskin et al., 2011).

GPR is now a common tool used for monitoring snow depth at ski resorts (Heilig, 2008). The optimal distribution of natural snow and the creation of manmade snow are reliant on accurate and dense snow depth data, which have been traditionally measured by probing. GPR systems fitted with GPS can be easily mounted on hand-towed sleds to rapidly provide snow maps with 3D depth. When mounted on groomers, GPR can provide a real-time display of snow thickness, allowing the operator to redistribute the snow evenly over a run. Such radar systems eliminate the need for an expert interpretation of a GPR profile by automatically interpreting the depth of snow and storing only these depths and GPS locations for subsequent mapping.

70.5.5 Archaeology and Forensics

Archaeologists were among the first to realize the potential benefits of the rapid subsurface imaging that GPR made possible in the 1970s. Early surveys were conducted to search for buried foundations and stone walls (Bevan and Kenyon, 1975). In addition, extensive surveys were conducted during the 1980s in Japan for burial mounds and ancient pit dwellings (Imai et al., 1987).

Although hundreds of successful surveys have been performed and the technique has been well established as a viable tool for both archaeology and forensics, these surveys are often conducted by untrained users in environments unsuitable for GPR. As the intended targets for such GPR investigations are often discrete objects buried among other radar reflections from geological features, the results of a "radar search" often produce ambiguous anomalies. Prior to the ability to integrate high-precision GPS positioning to track the radar antennas, surveys only relied on widely spaced 2D cross sections to map targets. Modern GPR systems with integrated RTK-DGPS allow the user to traverse a search site rapidly in a random semigrid pattern, often with lines spaced decimeters apart (Nuzzo, 2002). These data may then be input into a gridding algorithm to produce a 3D voxel cube, which may subsequently be sliced both vertically and horizontally to discriminate a potential investigation dig site from background clutter.

Regardless of these advancements in the 3D visualization of archaeological GPR data, the inherent ambiguity of the technique in all but the most ideal environments suggests that GPR is best considered in tandem with another sensing tool such as magnetics and/or EMs.

70.5.6 Tunnel Detection

Concurrent with the development of radar for lunar exploration and monitoring nuclear test explosions, the US military experimented with early radar systems to detect tunnels beneath the DMZ along the 38th parallel (Kim and Ra, 1993). Tunnel detection, whether for security or geotechnical applications, has always been an important application of GPR.

While manufacturers often display images of clear radar profiles across large highway tunnels through solid bedrock, the most practical applications for tunnel detection are situated in areas of significant subsurface and aboveground clutter and involve relatively small tunnels at significant depths. Research on the use of GPR to monitor tunneling beneath sensitive borders such as the US–Mexico frontier and the West Bank and the Gaza Strip has yielded moderate results due to the generally conductive nature of the soils in these locations (Cechak, 2007; Mahrer and List, 1995). Furthermore, due to the ambiguous nature of radar targets, security personnel are often reluctant to rely solely on the technique due to the costly implications of false anomalies (Pappalardo, 2009).

Similarly, in the geotechnical community, GPR has found mixed success at locating abandoned mine workings (Duff, 1983; Fenner, 1995; Leggo and Leech, 1983). Perhaps the most crucial need for the remote sensing of tunnels is that of abandoned coal mines. Often, urbanization has encroached on historical mining areas, raising the risk of ground subsidence or collapse. Such applications generally require shielded antennas to minimize the effects of interference from aboveground reflectors and other EM signal sources. Due to size limitations, the lowest frequency shielded antenna commercially available is 100 MHz, which restricts the practical search depths to less than 8 m in most environments. Recently, a partially shielded 40 MHz system has become commercially available for tunnel detection (Francke, 2012).

70.5.7 Sedimentary Geology

Geologists have long employed radar imaging to past depositional environments and the nature of sedimentary processes in a variety of settings, particularly in describing unconsolidated recent and Quaternary fluvial deposits both above and below the water table. Numerous studies dating to the early 1990s provide excellent examples of fluvial stratiforms (Jol and Smith, 1991, 1992; Vandenberghe and van Overmeeren, 1999).

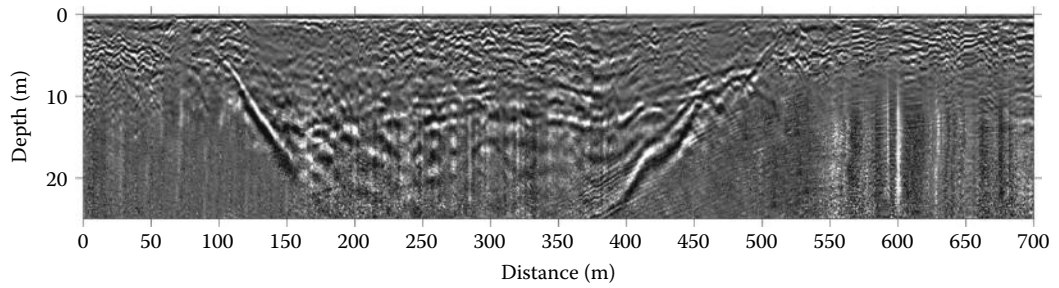


FIGURE 70.8 GPR cross section of a paleochannel in Brazil.

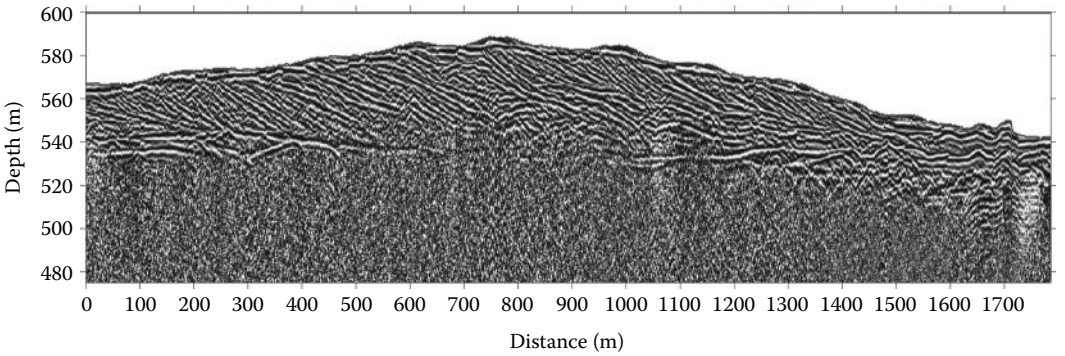


FIGURE 70.9 GPR profile through a Saharan sand dune in Libya.

An associated commercial application of sedimentary geology studies is the mapping of paleochannels for alluvial gold and diamond exploration. A complete understanding of the paleofluvial environment is critical to resource exploration and project economics. Exploration is normally conducted using drill holes and trenching, with seismic or EM geophysical surveying occasionally employed. GPR is increasingly being trialled in suitable environments where the in-filled sediments consist of course-grained sands and gravels in order to map the ore-rich basal gravels. Figure 70.8 shows an example of a radar-imaged paleochannel from an alluvial project in Guyana.

Given the low electrical conductivity and dry nature of the environment, desert sand dunes are often ideal radar settings. Studies of the internal structure of dunes are useful for studying desertification and climate change processes as well as for constructing petroleum reservoir analogies. While studies of sand dunes with GPR have been conducted over the last two decades (Bristow and Jol, 2003; Bristow et al., 1996, 2000), recent advances in rapid long-range radar have enabled the deep profiling and 3D surveying of sand dunes (Tatum and Francke, 2012). The example shown in Figure 70.9 demonstrates the suitability of radar in mapping the complex sedimentary bedding features of large duneforms.

70.5.8 Tropical Weathering Environments

One of the most common applications for GPR technology in mineral exploration is in tropical weathering environments with residual enrichments of either nickel (nickel laterites) or aluminum oxide (bauxites) (Francke, 2012). Both environments pose particularly challenging problems for traditional resource definitions that use drill holes alone, due to the extreme lateral variability in weathering thicknesses. Radar penetration depths of up to 70 m have been recorded in these weathering environments, with many tropical types of clay exhibiting unexpectedly low electrical resistivities. Although the clay-rich limonite and saprolite contain high bound water and thus are characterized by relatively slow radar

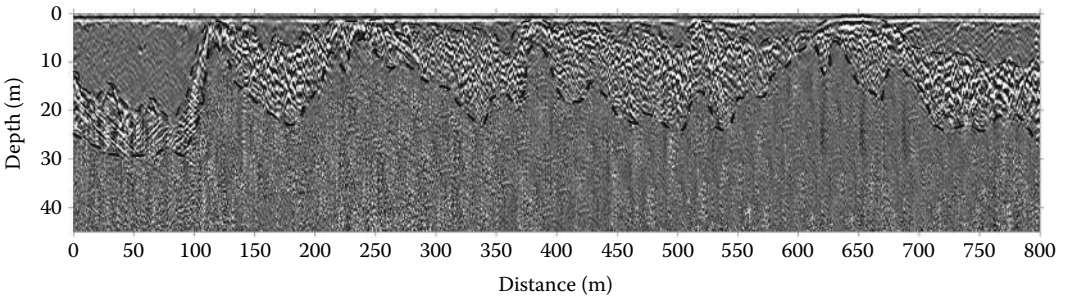


FIGURE 70.10 GPR profile of a lateritic weathering sequence in Indonesia, with dashed lines indicating the top of saprolite (upper line) and bedrock (lower line).

velocities, the lateritization process, with its repeated seasonal groundwater fluctuation, effectively leaches most of the conductive mineralogy from the profile.

The primary objective of nickel laterite and bauxite exploration is the mapping of the contact between weathered materials with the underlying bedrock, which commonly defines the limit of mining. Radar reflections from this interface are due to the dramatic change in water saturation between the moist weathered zone and the dry solid parent rock beneath. Similarly, the boundary between weathered soil and the interstitial rocky saprolite is well imaged with radar as a region of superimposed hyperbolic reflections. Figure 70.10 shows an example of GPR data acquired over a nickel laterite deposit.

In some lateritic bauxite deposits, particular challenges are faced due to the thin ore zone and the undulating nature of the underlying ferricrete. Seasonal fluctuations in the water table result in regions of diffuse bauxite to ironstone transitional zones, whereas a static water table produces a more abrupt transition. Resource exploitation consists of a large earth-moving operation, wherein the avoidance of bauxite ore contamination by the silica in the underlying ferricrete is critical while still maximizing the tonnage of extracted bauxite. GPR has been used to accurately map the depth of bauxite and the texture of the underlying horizon with high accuracy, with the intention of eventually using a GPR-generated digital surface model to form a "mine to" surface for semiautomated excavators (Francke and Yelf, 2003).

70.6 Future Advancements in GPR

70.6.1 Depth Range

With the ongoing development of faster processors and A/D converters, the future of long-range radar is seemingly limitless. However, both the laws of physics and practical limitations will significantly inhibit deeper radar systems in the future. In the realm of impulse radars, to double the penetration of an existing GPR transmitter in a given geological environment, an increase of 30 dB in system performance would be required, which in turn requires an increase in peak transmitter voltage of 1000 times. Generating hundreds of kilovolts in a portable instrument is not practical. Furthermore, the PRF of the transmitter would need to be lowered substantially to avoid saturation of the components.

An alternative approach to increasing penetration in environments where the radar range is limited by the noise floor is through stacking or averaging of radar signals at colocated positions, thereby increasing the signal-to-noise ratio. Stacking 1000 times would theoretically double the penetration over a single-stack system. The sequential sampling used in most radar systems limits the practical number of stacks to approximately 32. Technologies such as UltraGPR employ real-time sampling to effectively stack 32,000 times in the same amount of time that it takes a sequentially sampled system to stack 32 times. In addition, the latest GPR systems employ completely wireless designs and use PDAs or mobile phones as data acquisition computers.

Even further depth penetration may be achieved by using pulse compression techniques, which refer to the wave-shaping process that is produced as a propagating waveform is modified by the properties

of the transmission system (Utsi, 2007). Pulse compression combines the high energy of a long pulse width with the high resolution of a short pulse width. Since each part of the pulse has a unique frequency, the resultant returns can be completely separated. A GPR system developed for extreme penetration in suitable radar environments using pulse compression has imaged to 200 m through Libyan sand dunes. This penetration through dunes demonstrates the possible applicability of radar techniques to seismic static correction problems in deserts. A thorough understanding of the variations in the shallow low seismic velocity (LVL) zones in these environments is critical for seismic processing. Accepted approaches of up-hole surveys and drilling are costly and time-consuming and provide only point data over a large concession. GPR is showing promise of being able to penetrate to sufficient depths to image the base of the LVL.

70.6.2 Rapid 3D Surveys

The proliferation of multichannel array radar systems over the past 5 years foretells the future of GPR surveying for shallow civil infrastructure surveys through the use of instruments that can survey wide swaths of ground at once and often at high speeds. These systems are linked to RTK-DGPS or robotic theodolites to provide accurate positions for each of the antennas. The latest radar processing software, Reflex (www.sandmeier-geo.de) and GPR-Slice (www.gpr-survey.com), are capable of importing and manipulating these large data volumes to visualize buried targets. However, such systems are an order of magnitude more costly than single-channel radar units, suggesting that most future GPR surveys will be conducted by large companies capable of the required capital investment.

70.7 Conclusion

Four decades after commercialization, GPR has gained acceptance as a mature geophysical method. While the rapid growth in the number of manufacturers has yet to lower equipment prices, the increased competition in commercial systems appears to have split the market into two approaches to instrumentation. Systems intended for small-scale utilities detection, pavement studies, and concrete NDT are no longer designed as high-level geophysical instrumentation but as standard job site tools to be used by untrained workers. Conversely, large-scale surveys conducted by multichannel radar arrays and long-range GPR systems designed for mineral exploration have become increasingly expensive and complex, relegating their use to the largest and most experienced service providers. This bifurcation of the radar instrumentation market is likely to continue for the foreseeable future, with more manufacturers being spun off from university research laboratories. Eventually, competition should drive down the cost of single-channel units.

Although new uses for GPR appear in the literature each year, they are generally ancillary to those that have been well established over the past decades. What appears to be the future is the incorporation of GPR as a complementary tool in a suite of sensors, such as those used for roadbed analysis. EM sensors, magnetometers, precision positioning and guidance sensors, and LIDAR (light detection and ranging) all have applicability when fused with subsurface radar for use in many different scenarios.

Partial List of GPR Manufacturers and Suppliers

Geoscanners AB, Sweden
http://geoscanners.com

Ingegneria Dei Sistemi, Italy
http://www.idscompany.it

International Groundradar Consulting, Inc., Canada
www.groundradar.com

Koden Electronics Co., Ltd., Japan
http://www.koden-electronics.co.jp/eng/

Mala Geoscience, Sweden
http://www.malags.com

Non-Intrusive Inspection Technology, Inc., United States
http://www.niitek.com

PipeHawk, Plc., United Kingdom
http://www.pipehawk.com/

Radar Systems, Inc., Latvia
http://www.radsys.lv

3D-Radar, Norway
http://www.3d-radar.com
Geophysical Survey Systems, Inc., United States
http://www.geophysical.com

Radarteam Sweden AB, Sweden
http://www.radarteam.se

Sensors and Software, Inc., Canada
http://www.sensoft.ca/

Subsurface Imaging Systems, United States
http://www.usradar.com/

Transient Technologies Company, Ukraine
http://viy.com.ua/e

UTSI Electronics, Ltd., United Kingdom
http://www.utsielectronics.co.uk

References

Al-Nuaimy, W., Eriksen, A., and Gascoyne, J., Improved productivity and reliability of ballast inspection using road-rail multichannel GPR, *International Railway Engineering Conference*, London, U.K., 2004a.

Al-Nuaimy, W., Eriksen, A., and Gascoyne, J., Train-mounted GPR for high-speed rail trackbed inspection, *Proceedings 10th International Conference on Ground Penetrating Radar,* Delft, the Netherlands, pp. 631–634, 2004b.

American Concrete Institute, *Non-Destructive Test Methods for Evaluation of Concrete in Structures*, 2R-98, Farmington Hills, MI, p. 228, 1998.

Annan, A. P., Cosway, S. W., and DeSouza, T., Application of GPR to map concrete to delineate embedded structural elements and defects, *Proceedings of 9th International Conference on Ground Penetrating Radar*, Santa Barbara, CA, pp. 359–364, 2002.

Arcone, S. A., High resolution of glacial ice stratigraphy: A ground penetrating radar study of Pegasus Runway, McMurdo Station, Antarctica. *Geophysics*, 61, 1653–1663, 1996.

Arcone, S. A., Lawson, D. E., Delaney, A. J., and Moran, M. L., 12–100 MHz depth and stratigraphic profiles of temperate glaciers. *Proceedings of 8th International Conference on Ground Penetrating Radar,* Gold Coast, Australia, pp. 377–382, 2000.

Ariaratnum, S., Webb, R., and Conroy, A., Utilizing SSET with GPR for assessing large diameter lined concrete sewers, *Proceedings of No-Dig Show,* Orlando, FL, 2005.

Beutl, R., Reinhardt, H., Grosse, C., Glaubitt, A., Krause, M., Maierhofer, C., Algernon, D., Wiggenhauser, H., and Schickert, M., Comparative performance tests and validation of NDT methods for concrete testing, *J. Nondestr. Eval.*, 27, 59–65, 2008.

Bevan, B. and Kenyon, J., Ground-probing radar for historical archaeology. MASCA, University of Pennsylvania, *Newsletter*, 11, 2–7, 1975.

Bristow, C. S. and Jol, H. M., *Ground Penetrating Radar in Sediments*. Geological Society Special Publication, London, U.K., p. 211, 2003.

Bristow, C. S., Bailey, S. D., and Lancaster, N., The sedimentary structure of linear sand dunes, *Nature*, 406, 56–59, 2000.

Bristow, C. S., Pugh, J., and Goodall, T., Internal structure of aeolian dunes in Abu Dhabi revealed using ground penetrating radar, *Sedimentology*, 43, 995–1003, 1996.

Cechak, J., Detection of the under-soil intruder activity, *Proceedings of the International Society for Optical Engineering*, 6562, Orlando, FL, 2007, 656210.

Clough, J. W., Electromagnetic lateral waves observed by earth sounding radars. *Geophysics*, 41, 1126–1128, 1976.

Concrete Society, Guidance on radar testing of concrete structures, Technical Report 48, Camberley, U.K., 1997.

Daniels, D. J., *Ground Penetrating Radar*, 2nd edn., ser. IEE Radar, Sonar, Navigation, and Avionics Series 15, Institution of Electrical Engineers, London, U.K., 2005.

Duff, B., A review of electromagnetic methods used for detection of underground tunnels and cavities, *Proceedings of the IEEE International Symposium on Antennas and Propagation*, Houston, TX, 21, pp. 634–638, 1983.

Eide, E. S. and Hjelmstad, J. F., 3D utility mapping using electronically scanned antenna array, *Proceedings of the 9th International Conference on Ground Penetrating Radar*, Santa Barbara, CA, SPIE 4758, pp. 192–196, 2002.

Fenner, T., Ground penetrating radar for identification of mine tunnels and abandoned mine stopes, *Mining Eng*., 47(3), 280–284, 1995.

Francke, J., A review of selected ground penetrating radar applications to mineral resource evaluations, *J. Appl. Geophys*., 81, 29–37, 2012.

Francke, J., Utsi, V., and Singh, K., The design of an intrinsically safe GPR for underground coal mines, *IEEE Proceedings of 14th International Conference on Ground Penetrating Radar*, Shanghai, China, pp. 125–130, 2012, CFP12538.

Francke, J. and Yelf, R., Applications of GPR for surface mining, in *Second International Workshop on Advanced Ground Penetrating Radar*, A. Yarovoy (ed.), Delft, the Netherlands, Delft University of Technology, pp. 115–119, 2003.

Gallagher, G. P., Leiper, Q., Williamson, R., Clark, M. R., and Forde, M. C., The application of time domain ground penetrating radar to evaluate railway track ballast, *NDT & E Int*., 32(8), 463–468, 1999.

Galley, R. J., Trachtenberg, M., Langlois, A., Barber, D. G., and Shafai, L., Observations of geophysical and dielectric properties and ground penetrating radar signatures for discrimination of snow, sea ice and freshwater ice thickness. *Cold Reg. Sci. Technol*., 57, 29–38, 2009.

Heilig, A., Schober, M., Schneebeli, M., and Fellin, W., Next level for snow pack monitoring in real-time using ground-penetrating radar (GPR) technology, *Proceedings of International Snow Science Workshop*, Whistler, BC, Canada, 2008.

Imai, T., Sakayama, T., and Kanemori, T., Use of ground-probing radar and resistivity surveys for archaeological investigations. *Geophysics*, 52, 137–150, 1987.

Jol, H. M. and Smith, D. G., Ground penetrating radar of northern lacustrine deltas, *Can. J. Earth Sci*., 2, 1939–1947, 1991.

Jol, J. M. and Smith, D. G., Geometry and structure of deltas in large lakes: A ground penetrating radar overview, *Proceedings of the 4th International Conference on Ground Penetrating Radar*, Rovaniemi, Finland, pp. 159–168, 1992.

Kim, S.-Y. and Ra, J.-W. The role of cross borehole radar in the discovery of Fourth tunnel at Korea DMZ, *4th Tunnel Detection Symposium on Subsurface Exploration Technology*, Fort Belvoir, VA, Combat Engineering Directorate, U.S. Army Belvoir Research, Development and Engineering Center, pp. 253–257, 1993.

Lai, W. L., Kind, T., and Wiggenhauser, H., Detection of accelerated reinforcement corrosion in concrete by ground penetrating radar, *Proceedings of the 13th International Conference on Ground Penetrating Radar*, Lecce, Italy, 2010.

Leggo, P. J. and Leech, C., Sub-surface investigation for shallow mine workings and cavities by the ground impulse radar technique, *Ground Eng*., 16, 2–203, 1983.

Mahrer, K. D. and List, D. F., Radio frequency electromagnetic tunnel detection and delineation at the Otay Mesa site, *Geophysics*, 60(2), 1995, 413–422.

Maijala, P., Moore, J. C., Hjelt, S. E., Palli, A., and Sinisalo, A., GPR investigations of glaciers and sea ice in the Scandinavian arctic, *Proceedings of 7th International Conference on Ground Penetrating Radar*, Lawrence, KS, pp. 143–148, 1998.

Maser, K. R., Highway speed radar for pavement and bridge deck evaluation, *Proceedings of Infrastructure Planning and Management Conference, American Society of Civil Engineers*, Denver, CO, p. 6–10, 1993.

Morey, R. M., Detection of subsurface cavities by ground penetrating radar, *Proceedings Highway Geological Symposium*, Chicago, IL, 27, pp. 28–30, 1974.

Nag, S., Fluhler, H., and Barnes, H., Preliminary interferometric images of moving targets obtained using a time-modulated ultra-wide band through-wall penetration radar, Fluher and Barnes, Atlanta, GA, *IEEE Radar Conference*, pp. 64–69, 2001.

Noureldin, A. S., Zhu, K., Li, S., and Harris, D., Network pavement evaluation with falling-weight deflectometer and ground-penetrating radar, *J. Transp. Res. Board*, 1860, 90–99, 2003.

Nuzzo, L., Leucci, G., Negri, S., Carrozzo, M., and Quarta, T., Application of 3D visualization techniques in the analysis of GPR data for archaeology. *Ann. Geophys*., 45(2), 321–337, 2009.

Olhoeft, G. R., The electrical properties of permafrost, PhD thesis, University of Toronto, Ontario, Canada, 1975.

Olhoeft, G. R., Electrical properties from 10 − 3 to 10 + 9 Hz—Physics and chemistry: Proceedings of 2nd International Symposium Physical Chemistry Porous Media, *American Institute of Physics Conference Proceedings*, 154, 281–298, 1987.

Olhoeft, G. R. and Selig, E. T., Ground penetrating radar evaluation of railroad track substructure conditions, *Proceedings of 9th International Conference on Ground Penetrating Radar*, Santa Barbara, CA, SPIE 4758, pp. 48–53, 2002.

Pappalardo, J., High-tech border patrol: 5 new tricks to find smuggler tunnels, *Popular Mechanics*, New York, October 1, 2009.

Pollock, D., Dupuis, K., LaCour, B., and Olsen, K., Detection of voids in prestressed concrete bridges using thermal imaging and ground-penetrating radar, *Research Report for Federal Highway Administration (FHWA) Project DTFH61-05-C-00008* (*Task No. 8*), Washington State Transportation Center, Department of Civil & Environmental Engineering, Washington State University, Washington, DC, pp. 115–119, 2008.

Proskin, S. A., Parry, N. S., and Finlay, P., Applying GPR in assessing the ice bridges, ice roads and ice platforms, *Proceedings of 16th Workshop on River Ice, CGU HS Committee on River Ice Processes and the Environment*, Winnipeg, Manitoba, Canada, 2011.

Saarenketo, T., Using ground-penetrating radar and dielectric probe measurements in pavement density quality control, *J. Transp. Res. Board*, 1575, 34–41, 1997.

Saarenketo, T. and Scullion, T., Road evaluation with ground penetrating radar, *J. App. Geoph*., 43(2), 119–138, 2000.

Sachs, J., Aftanas, M., Crabbe, S., Drutarovsky, M., Klukas, R., Kocur, D., Nguyen, T. T., Peyerl, P., Rovnakova, J., and Zaikov, E., Detection and tracking of moving or trapped people hidden by obstacles using ultra-wideband pseudonoise radar, *European Radar Conference*, Amsterdam, the Netherlands, pp. 408–411, 2008.

Selig, E. T., Hyslip, J. P., Olhoeft, G. R., and Smith, S., Ground penetrating radar for track substructure condition assessment, *Proceedings of Implementation Heavy Haul Technology Network Efficiency, IHHA/REMSA, World Rail Expo*, Dallas, TX, pp. 6.27–6.33, 2008.

Simmons, G. D., Strangway, D. W., Bannister, R., Baker, D., Cubley, G., La Torraca, G., and Watts, R., The surface electrical properties experiment, in *Lunar Geophysics: Proceedings of Conference Lunar Science Institution*, Houston, TX, October 18–21, 1971, Z. Zopal and D. W. Strangway (ed.), Dordrecht, the Netherlands, D. Reidel Publishing, pp. 258–271, 1972.

Spagnolini, U., and Rampa, V., Multitarget detection/tracking for monostatic ground penetrating radar: Application to pavement profiling, *IEEE Trans. Geosci. Remote Sens.*, 37, 383–394, 1999.

Stern, W., Versuch einer elektrodynamischen Dickenmessung von Gletschereis. *Gerl. Beitz. zur Geophys.*, 23, 292–333, 1929.

Tatum, D. and Francke, J., Constructing hydrocarbon reservoir analogues of aeolian systems using ground penetrating radar, *J. App. Geophys.*, 81, 21–28, 2012.

Texas Transportation Institute, Specification Tx DOT-845-49-62, 2001.

Thierbach, R., Electromagnetic reflections in salt deposits. *J. Geophys*. 40, 633–637, 1974.

Utsi, V., Design of a GPR for deep investigations, in *4th International Workshop Advanced Ground Penetrating Radar*, Washington, DC, IEEE Computer Society Press, pp. 222–225, 2007.

Vandenberghe, J. and van Overmeeren, R. A., Ground penetrating radar images of selected fluvial deposits in the Netherlands, *Sediment. Geol.*, 128(3), 245–270, 1999.

Waite, A. H. and Schmidt, S. J., Gross errors in height indication from pulsed radar altimeters operating over thick ice or snow, *IRE International Convention Record*, Part 5, pp. 34–54, 1961.

Zaikov, E., Sachs, J., Aftanas, M., and Rovnakova, J., 2008, Detection of trapped people by UWB radar, *German Proceedings of Microwave Conference (GeMIC)*, Hamburg, Germany, pp. 1–4, 2008.

71

Airborne Radiometers to Measure Electromagnetic Radiation in the Earth's Atmosphere: Mature and Emerging Technologies

Manfred Wendisch
University of Leipzig

71.1 Introduction

71.1.1 Importance of Airborne Measurements of Electromagnetic Radiation in the Atmosphere

Electromagnetic (EM) radiation is important for many atmospheric and oceanic processes. The solar radiative energy reaching the top of atmosphere (TOA) is the major source of energy feeding weather and climate phenomena on Earth (Sellers, 1965). All biological life depends on the energy provided by solar radiation. Almost every motion of air in the atmosphere or of water in the oceans and numerous chemical processes within the atmosphere are powered by solar radiative energy. If the Sun would hypothetically be switched off, then plants would die within several weeks, and animals would lose food and would die shortly after plants have disappeared. Within about one year, no more higher life would be possible on Earth, and ice and snow would cover the continents. After roughly 15 years, the Earth would be a snowball.

Within the atmosphere and at the Earth's surface, solar radiation is scattered (redistributed in direction) and absorbed (transformed into other forms of energy, mainly heat or chemical energy). As a result of absorption, the temperatures of the atmospheric components (gases, aerosol particles, clouds) and the Earth's surface rise.

Earth's terrestrial radiation is emitted corresponding to the atmospheric and Earth's surface temperatures and respective emissivities. The radiation emitted by the Earth's surface is partly reabsorbed and reemitted by the atmospheric components whereby the atmospheric reemission takes place at mostly lower temperatures compared to the Earth's surface temperature, which causes the atmospheric greenhouse effect; a phenomenon of fundamental importance for life on Earth.

Globally and averaged over long time periods, the climate system is in energetic equilibrium (balanced net radiation, i.e., downward minus upward radiative energy at TOA equals zero) (see Trenberth and Fasullo, 2012). However, on a local and short time scale, there is no such energetic equilibrium. As a result, thermal meridional gradients develop, which are causing compensating dynamic processes, such as atmospheric air motions or oceanic water circulations in the vertical or horizontal directions, thermodynamic processes like evaporation or condensation initiating clouds and precipitation, and chemical reactions. Thus, all processes modifying solar and terrestrial radiation within the atmosphere and at the Earth's surface are crucial for the energy budget of the global climate systems. For this purpose, airborne measurements of atmospheric EM radiation are of highest importance.

Additionally, airborne observations of EM radiation are fundamental to develop, test, and apply remote sensing techniques to retrieve atmospheric and Earth's surface properties. In order to characterize the global climate system, satellites comprise an essential tool of observations. All satellite techniques are based on measuring radiative energy in different wavelength ranges. Therefore, the transfer of atmospheric radiative energy through the atmosphere and its interactions with the Earth's surface needs to be measured by airborne techniques and modeled by radiative transfer simulations in order to interpret the radiation data received by the satellite instruments.

71.1.2 Quantitative Description of Electromagnetic Radiation

The following definitions of quantities to describe the atmospheric radiation field follow Wendisch and Yang (2012) and Wendisch et al. (2013).

71.1.2.1 Angular Coordinates

In order to describe direction, spherical polar coordinates are applied. The zenith angle θ represents the planar angle from local vertical; it is measured in units of radian (rad) or degrees (°). Overhead means $\theta = 0$ rad or 0° (zenith direction), and the horizon corresponds to $\theta = \pi/2$ rad or 90°. $\theta = \pi$ rad or 180° denotes the nadir direction. Often instead of the zenith angle, the quantity $\mu = \cos\theta$ is used. The azimuth angle φ measures the horizontal angle in units of rad or °. φ is counted counterclockwise from a reference point such that 0 rad < ϕ < 2π rad. Often, the projection of the direction of the Sun with respect to the horizontal plane is defined as reference. For the direction of the Sun, the subscript "0" is applied.

Furthermore, the solid angle Ω is introduced, which is something like "square degrees" and carries the unit of "steradian" (sr). If an observer is situated at the center of a sphere of unit radius, then the total surface area of the sphere is 4π square radius units. Thus, the total solid angle is defined as 4π sr. The surface area of one-half of the sphere is 2π square radius units; the angle is 2π sr. This holds for the upper and lower hemispheres separately. For the assumed unit sphere with radius of 1 m, the solid angle of 1 sr is bordered by a surface of 1 m^2 on the unit sphere. The incremental solid angle dΩ is defined by:

$$\mathrm{d}\Omega := \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi = \mathrm{d}\mu\;\mathrm{d}\varphi = \frac{\mathrm{d}A}{a^2}, \tag{71.1}$$

with dA the enclosed area on the surface of the sphere with radius a. Integration over all directions yields 4π sr:

$$\iint_{4\pi}\mathrm{d}\Omega = \int_0^{2\pi}\int_0^{\pi}\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi = 2\pi\ \mathrm{rad}\int_0^{\pi}\sin\theta\,\mathrm{d}\theta = 4\pi\,\mathrm{sr}. \tag{71.2}$$

71.1.2.2 Vertical Coordinate: Optical Depth

We define the vertical spectral optical depth $\tau(\lambda, z)$ at the given wavelength λ as a transformed vertical coordinate:

$$\tau(\lambda, z) = \int_{z}^{z_{\mathrm{TOA}}} b_{\mathrm{ext}}(\lambda, z')\,\mathrm{d}z'. \tag{71.3}$$

The volume extinction coefficient, $b_{\mathrm{ext}}(\lambda, z)$, quantifies the extinction (sum of absorption and scattering) by atmospheric constituents (gases, aerosol particles, clouds), see Wendisch and Yang (2012). If τ is used to characterize the extinction properties of a layer, it is called the optical thickness instead of depth.

71.1.2.3 Radiant Energy Quantities

The radiant energy E_{rad} in units of Joule (J) is the basic quantity to describe the EM radiation field in the atmosphere. It can be measured using first principles. From the differential radiant energy E_{rad} per time increment $\mathrm{d}t$, the radiant energy flux (or radiant power) ϕ in units of Watt (W) is derived:

$$\phi(\tau) = \frac{\mathrm{d}E_{\mathrm{rad}}}{\mathrm{d}t} \tag{71.4}$$

The radiant energy flux density F (often wrongly quoted as flux) in units of W m^{-2} is given by the incremental radiant flux $\mathrm{d}\phi$ per incremental area element $\mathrm{d}A$ (with no preferred orientation):

$$F(\tau) = \frac{\mathrm{d}\phi}{\mathrm{d}A} = \frac{\mathrm{d}E_{\mathrm{rad}}}{\mathrm{d}t\,\mathrm{d}A}. \tag{71.5}$$

F is a measure of total radiant flux per unit area transported by EM radiation through or deposited on a planar surface A. Finally, the radiance I in units of W m^{-2} sr^{-1} is defined as:

$$I(\tau,\theta,\varphi) = \frac{\mathrm{d}\phi}{\mathrm{d}A\,\mathrm{d}\Omega} = \frac{\mathrm{d}\phi}{\cos\theta\ \mathrm{d}A\,\mathrm{d}\Omega} = \frac{1}{\cos\theta}\frac{\mathrm{d}F}{\mathrm{d}\Omega}, \tag{71.6}$$

with $\mathrm{d}A_{\perp}$ the differential area oriented perpendicularly to the direction of propagation of the EM radiation. The integration of Equation 71.6 over the hemisphere yields the relationship between irradiance and radiance:

$$F(\tau) = \int_{2\pi} I(\tau,\theta,\varphi)\cdot\cos\theta\,\mathrm{d}\Omega. \tag{71.7}$$

In atmospheric applications, the reference unit area is usually defined as horizontal. Therefore, upward $F^{\uparrow}$ and downward $F^{\downarrow}$ irradiances are obtained from radiance I by applying Equations 71.7 and 71.1:

$$F^{\uparrow}(\tau) = \int_{0}^{2\pi}\int_{0}^{\pi/2} I(\tau,\theta,\varphi)\cdot\cos\theta\sin\theta\,\mathrm{d}\theta\ \mathrm{d}\varphi, \tag{71.8}$$

and

$$F^{\downarrow}(\tau) = \int_{0}^{2\pi}\int_{\pi}^{\pi/2} I(\tau,\theta,\varphi)\cdot\cos\theta\sin\theta\,\mathrm{d}\theta\ \mathrm{d}\varphi, \tag{71.9}$$

The dimensionless albedo α is the ratio of upward to downward irradiance:

$$\alpha(\tau) = \frac{F^{\uparrow}}{F^{\downarrow}}. \tag{71.10}$$

Actinic flux density F_{act}, sometimes called average intensity, is the integral of radiance over solid angle:

$$F_{act}(\tau) = \int_0^{2\pi}\int_0^{\pi} I(\tau,\theta,\varphi)\cdot\sin\theta\,\mathrm{d}\theta\;\mathrm{d}\varphi. \tag{71.11}$$

Like irradiance, actinic flux density has units of W m^{-2}. It represents the energy flux on a unit sphere, normalized by the cross section of the sphere, and therefore is related to flux divergence.

Irradiance (F), radiance (I), and actinic flux density (F_{act}) can be either spectral or band-integrated (broadband) quantities. For example, the spectral flux density F_λ is the irradiance per unit wavelength interval indicated by the subscript λ (λ itself is the symbol for wavelength). As a result, the units of a spectral radiant energy quantity contain an additional term nm^{-1}. The band-integrated flux density $F(\lambda_1, \lambda_2)$ includes radiant energy contributions from wavelengths within an interval (λ_1, λ_2). Here, we concentrate on spectral irradiance, radiance, and actinic flux density measurements from aircraft.

Irradiance and radiance may be divided into the contributions from scattering (diffuse) and direct transmission (direct) indicated by the subscripts "dif" and "dir," respectively (Wendisch et al., 2013). The sum of both corresponds to the total irradiance or radiance with the subscript "tot." For example, for the total irradiance, we obtain:

$$F_{tot} = F_{dir} + F_{dif}. \tag{71.12}$$

71.1.3 Solar and Terrestrial Spectra

The important spectral range for the application of airborne measurements of EM radiation in energy budget and remote sensing studies covers wavelengths between 0.3 and 100 μm. Generally, two subranges may be distinguished: The solar spectrum spans from 0.2 to 5 μm, and the terrestrial wavelength range covers wavelengths larger than 5 μm. The thermal infrared (IR) usually refers to wavelengths between 5 and 50 μm. Furthermore, the ultraviolet (UV, 10–370 nm wavelength) and microwave (MW, 0.3 mm–30 cm) spectral ranges are specified. The MW range corresponds to a frequency range of 1000–1 GHz. Frequency ν or wavenumber $\tilde{\nu}$ is sometimes used alternatively to wavelength λ. They are defined by:

$$\lambda = \frac{c}{\nu}. \tag{71.13}$$

With the speed of light in a vacuum, c is given by:

$$c = 2.997925\times10^{8}\ \mathrm{m\ s^{-1}}. \tag{71.14}$$

The wavenumber is defined as:

$$\tilde{\nu} = \frac{1}{\lambda}. \tag{71.15}$$

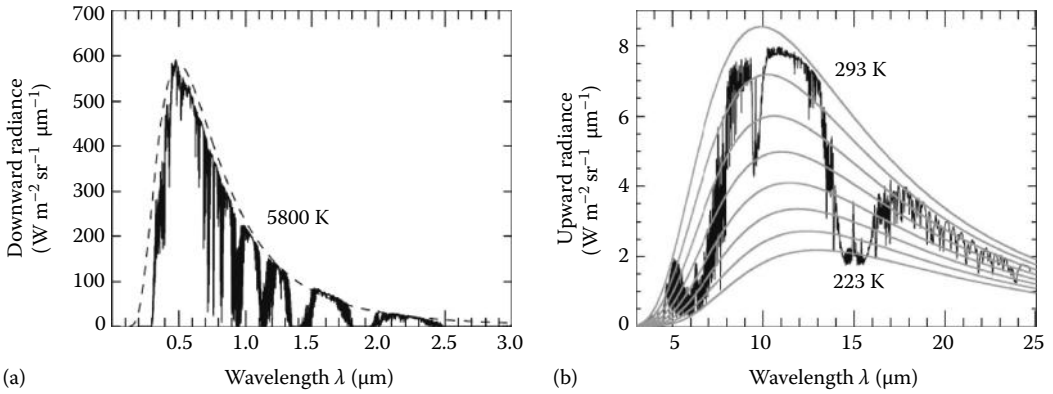


FIGURE 71.1 Black solid lines represent the spectral distribution of simulated solar downward radiances at the surface (a) and terrestrial upward radiances emitted at the TOA (b). "Tropical" profiles of atmospheric parameters have been applied from Anderson et al. (1986). The simulations were performed with MODTRAN4 (Version 3 Revision 1, MOD4v3r1) (see Berk et al. 1999), with the Sun in the zenith. The dashed line in (a) represents the Planck's function for 5800 K. The gray lines in (b) indicate Planck's functions for 293 K (upper curve) and 223 K (lower curve) in 10 K increments. (Partly adapted from Wendisch, M. and Yang, P.: *Theory of Atmospheric Radiative Transfer—A Comprehensive Introduction*. Weinheim, Germany. 2012. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reprinted with permission.)

Figure 71.1 shows typical simulated spectra for major parts of the solar (Figure 71.1a) and the thermal IR (Figure 71.1b) spectral ranges. At the TOA, the extraterrestrial solar radiation can roughly be approximated by black-body radiation with a temperature of about 5800 K (see dashed line of Figure 71.1a). From the downward radiance reaching the ground (see the solid line of Figure 71.1a), it is obvious that the deeper the radiation penetrates into the atmosphere, the more radiation is absorbed by atmospheric gases such as ozone (O_3): <0.3 μm; oxygen (O_2): 0.76 μm; water vapor (H_2O): 0.72, 0.82, 0.94, 1.1, 1.38, 1.87, and 2.7 μm; and carbon dioxide (CO_2): 1.4, 2.0, and 2.7 μm.

The upward radiance at TOA is illustrated by the black solid line in Figure 71.1b. In selected wavelength regions, for example, the atmospheric window (8–12 μm), the upward thermal IR radiance emitted at the TOA can be approximated in most parts by black-body radiation corresponding to a temperature of about 293 K, similar to the Earth's surface temperature (see Figure 71.1b). Thus, the atmospheric window region can be used to retrieve Earth's surface parameters, for example, temperature, by remote sensing techniques. Several absorption/emission bands appear in the spectrum. For example, the wavelength of 15 μm is in the center of a strong CO_2 absorption band. Thus, the signal measured by a satellite sensor near 15 μm stems from the cold stratosphere and even higher altitudes. A narrow peak is observed in the spectra at 15 μm at the very center of the CO_2 absorption. This spike occurs where absorption (and emission) is by far the most intense of any wavelength in the thermal IR. Emission from this wavelength originates almost entirely from the relative warm stratopause. Keep in mind that the CO_2 is very well mixed in the atmosphere up to an altitude of 100 km. Further major absorption/emission bands in the TOA upward radiance correspond to carbon dioxide (CO_2): 4.3 and 15 μm; water vapor (H_2O): 6.3 μm; and ozone (O_3): 9.6 μm (within the atmospheric window).

71.2 Airborne Radiometers

In general, radiometers measure radiant energy E_{rad} directly or quantities derived from E_{rad}, such as radiance or irradiance. These are called radiometric quantities in the following text. A comprehensive overview of airborne radiometers including issues and future challenges is given in Chapter 7 of Wendisch and Brenguier (2013).

71.2.1 Instruments for Broadband Solar Radiation Measurements (Broadband Radiometers)

Instruments that measure broadband solar irradiance are often called pyranometers or broadband radiometers. Mostly, the total irradiance is measured with pyranometers; however, there are also versions that may separate direct and diffuse irradiance (Valero et al., 1989; Pilewskie and Valero, 1993). Examples of pyranometers widely used in airborne applications are those manufactured by Kipp & Zonen (http://www.kippzonen.com/) and Eppley (the Eppley Laboratory, Inc., http://www.eppleylab.com/). Here, we describe the pyranometer manufactured by Kipp & Zonen and the total ultraviolet radiometer (TUVR) from Eppley. Furthermore, pyroelectric radiometers are in use.

The pyranometer is the most common type of instrument to measure broadband solar irradiances from an aircraft, for example, Wendisch and Keil (1999). It has a circular multijunction thermopile of the plated (copper–constantan, hot and cold junction) wire wound type (Coulson, 1975). The sensitivity of the thermopile is in the range of 10 μV/(W m^{-2}); therefore, the signal needs electronic magnification in order to be logged. The temperature sensed by the detector is nearly linear with the flux density of incident radiation. The pyranometer manufactured by Kipp & Zonen is supplied with a pair of concentric hemispheres of Schott optical glass (see Figure 71.2). The inner dome is transparent to wavelengths in the range of 0.285–2.8 μm. The outer dome can be replaced by Schott glass hemispherical filters, which transmit within specified bandwidths. The outer dome is for mechanical protection of the inner dome. Furthermore, the air between the domes assures thermal insulation of the inner dome, which is then not heated by incoming solar radiation preventing emission due to solar heating. The response time of the pyranometer is in the range of 1 s. One of the advantages of the pyranometer is its mechanical robustness, which makes it well suited for airborne use. The largest problem of the pyranometer is the fact that temperature changes of the surrounding air also influence the housing and thus the temperature of the cold thermopile sensor. To mitigate this problem, thermocompensation is implemented in the pyranometer.

The Eppley TUVR measures broadband irradiance within the wavelength range from 0.295 to 0.385 μm (broadband UV) (see Wendisch et al., 1996; Wendisch and Keil, 1999). It utilizes a hermetically sealed selenium barrier-layer cell (photocell), which is protected by a quartz window. A specially designed Teflon diffuser reduces the radiant flux to acceptable levels and assures close adherence to the Lambert cosine law corresponding to Equation 71.7. An encapsulated narrow band-pass (interference) filter limits the spectral response of the photocell to the UV.

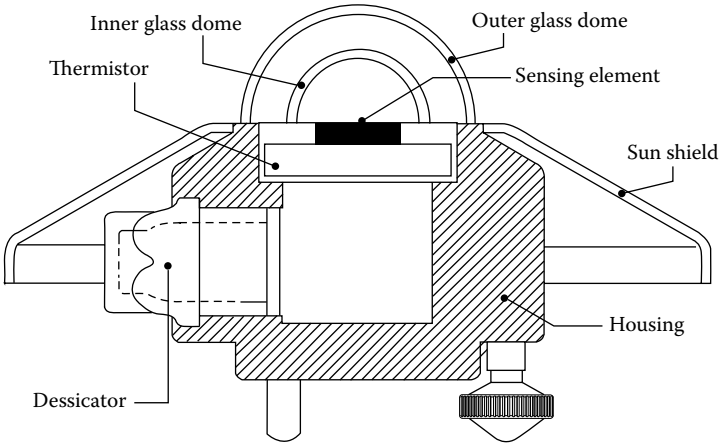


FIGURE 71.2 The Kipp & Zonen pyranometer. (Reprinted with permission from Kipp & Zonen, *CMP Series Pyranometer Instruction Manual*, The Netherlands, 2010.)

Pyroelectric solar radiometers use the temperature dependence of electric polarization in a pyroelectric crystal (Geist and Blevin, 1973; Valero et al., 1982; Hengstberger, 1989). If the temperature is changed, then the lattice spacing in the crystal also changes, which causes a change in the electric polarization of the crystal and consequently leads to a charging of the crystal surface generating a current that is used as a measure of the radiant energy. Thus, these sensors need a change in temperature to generate a signal. Therefore, they need to be equipped with a chopper, which, however, causes mechanical problems in aircraft installation. The major advantage of this type of radiometer is its fast response time (>100 Hz).

71.2.2 Spectroradiometers to Measure Spectral Solar Radiation

Spectral radiometers (also called spectroradiometers) measure radiometric quantities in narrow spectral bands as function of wavelength. In this way, spectroradiometers allow more detailed insights into the nature of solar radiation as can be achieved by broadband radiometers.

71.2.2.1 Spectral Solar Irradiance and Radiance (Spectroradiometers)

Typically, solar spectroradiometers consist of four components (Shetter and Müller, 1999; Wendisch et al., 2001; Pilewskie et al., 2003): the optical inlet, fiber optics, spectrometer (including detector), and electronics with data acquisition.

The optical inlet (fore optic) comprises either a hemispheric light collector for irradiance or a telescope for radiance observations. The hemispheric light collector is a diffuse element that transmits or reflects radiant energy flux (power) proportional to the cosine of the angle of the incident light, corresponding to Equation 71.7. For the diffusively scattering hemispheric light collector based on transmission, mostly, materials like flashed opal, Teflon, or Delrin are applied. Because scattering is wavelength dependent and mostly drops in the near IR, instead of transmitting hemispheric light collectors, an integrating sphere (using highly reflective Spectralon) can be applied (Crowther, 1997; Kindel et al., 2010).

Optical fibers transmit the light from the optical inlet (either hemispheric light collector or telescope) into the entrance slit of a spectrometer (monochromator including the detector). The wavelength dispersion within the spectrometer is often realized by gratings; sometimes also prisms or circular filter wheels are applied. The detector elements include photovoltaic and photoconductive devices, photomultiplier tubes (PMTs), photodiode, or charge-coupled device (CCD) arrays.

Solar spectroradiometers require careful calibration with regard to radiant power, as well as angular and spectral response. More details can be found in Wendisch and Brenguier (2013), Chapter 7.

71.2.2.2 Spectral Actinic Radiation

The major purpose of measuring spectral actinic flux density from aircraft is to determine the photolysis frequencies J by integrating (over wavelength) the product of spectral actinic flux density, absorption cross section of the molecular species to be photolyzed, and the quantum yield of the photoproduct. J in units of s^{-1} is given by:

$$J(\mathrm{A} \rightarrow \mathrm{B}) = \int_0^\infty \sigma_\mathrm{A}(\lambda, T, p) \cdot \phi_\mathrm{B}(\lambda, T, p) \cdot F'_{\mathrm{act},\lambda}\, \mathrm{d}\lambda, \tag{71.16}$$

where

σ_A is the absorption cross section of the molecular species A in units of cm^2

ϕ_B is the quantum yield of the photoproduct B (dimensionless)

T is the air temperature

p is the atmospheric pressure

Note that in Equation 71.16, the quantity F' (spectral photon actinic flux density) is introduced instead of the spectral actinic flux density; $F_{act,\lambda}$ is defined in Equation 71.11. $F_{act,\lambda}$ must be converted to the corresponding spectral photon actinic flux density F' applying the following relation:

$$F'_{act,\lambda} = \frac{\lambda}{h \cdot c} \times 10^{-13} F_{act,\lambda}, \tag{71.17}$$

where

λ is given in units of nm
$F_{act,\lambda}$ is given in units of W m^{-2} nm^{-1}
$F'_{act,\lambda}$ is given in units of cm^{-2} nm^{-1} s^{-1}

Two types of instruments are applied: actinic spectroradiometers and filter radiometers. Both types are described by Hofzumahaus (2006).

71.2.2.2.1 Actinic Spectroradiometers

The main setup of the spectroradiometers for actinic radiation measurements is similar to that of a spectroradiometer for irradiance measurements, except that the optical inlet has an isotropic angular response. Mainly two separate 2π optical inlets are used (one for the upper, the other for the lower hemisphere). There are two kinds of spectroradiometer for actinic radiation measurements: (a) scanning double monochromators (Shetter and Müller, 1999) and (b) single monochromators (Jäkel et al., 2005):

a. Scanning double monochromators select one wavelength band by a fixed grating position, while a second moving diffraction grating is scanning the fixed wavelength range with high spectral resolution. The spectral radiation is directed onto the detector (PMT) by moving the diffraction grating. The scanning double monochromators have the advantages of low stray light contribution and high sensitivity in the UV spectral range. However, the disadvantages in particular for airborne operation are that the wavelength scanning is time-consuming (in the order of several seconds up to minutes) and that the mechanical stability during aircraft operation is not guaranteed.
b. Therefore, in recent years, more and more single-monochromator spectroradiometers have been applied for airborne research to measure spectral actinic flux density. The multichannel spectrometer (MCS) has a photodiode array (PDA) or CCD detectors. The advantages of the single-monochromator approach are the fast time resolution of the spectra (one spectrum in less than 1 s) and the mechanical stability because of the fixed grating (no moving grating). The disadvantages are the relatively high stray light contribution (problems in the UV) and the lower sensitivity in the UV compared to scanning-monochromators with PMT detector.

71.2.2.2.2 Filter Radiometers for Actinic Radiation

Filter radiometers are designed such that the spectral response of the interference filters is close to the spectral shape of the product of the quantum yield φ_A of a specific species A and the absorption cross section σ_B. Widely used are filter radiometers for the photolysis of NO_2 and ozone (Junkermann, 1994; Volz-Thomas et al., 1996; Früh et al., 2000). The input optics (optical inlet) consists of a series of frosted hemispherical quartz domes that act as a diffuser with nearly isotropic response similar to the inlets of the actinic spectroradiometers. The photons are transferred by a pair of optical filters and detected by a photodiode. Ideally, the signal is directly proportional to J; in reality, however, the spectral response matches not exactly the spectral relative shape of the product $\varphi_A \cdot \sigma_B$. Actually, the accuracy of filter radiometers is temperature-dependent and sensitive to changes of the solar zenith angle and the total ozone column.

71.2.2.3 Direct Solar Radiation

Airborne measurements of direct solar radiation are performed to derive the optical thickness of major atmospheric constituents, mostly aerosol particles (see Equation 71.3). The optical thickness of aerosol

particles comprises one of the important input parameters for radiative transfer simulations. In general, Sun photometers are applied for this type of measurements. The direct portion of solar irradiance is measured by restricting the field of view of the instrument to the solar disc, that is, to 0.5° angular width of the Sun using a cylindrical tube (Gershun tube). Interference filters are applied for wavelength separation and photodiodes for detection. An automated drive mechanism finds and tracks the Sun (Matsumoto et al., 1987; Russell et al., 1993). Sometimes handheld Sun photometers are being used; however, the data quality in this case is not always sufficient (Porter et al., 2007). For calibration, the so-called Langley method is employed where the extraterrestrial spectral irradiance is obtained by extrapolating irradiance measurements during rising or setting Sun to a zero air mass, that is, to TOA (Langley, 1903).

71.2.3 Leveling Issues

A typical airborne irradiance sensor (broadband or spectral) is fixed with the fuselage of the aircraft. Therefore, it moves with respect to the Earth-fixed coordinate systems due to attitude changes of the aircraft during flight. However, atmospheric applications of irradiance (i.e., flux density) strictly refer to a horizontal receiving plane. If the sensor receiving plane is out of horizontal level, serious deviations compared to measurements with a horizontal sensor plane may arise (Wendisch et al., 2001). To correct for misalignments, either a software postprocessing procedure or an active, mechanical leveling of the sensor heads during the flight is required.

71.2.3.1 Software Postprocessing Correction

The following technique holds for the direct portion of solar radiation only. Although this software, past-flight procedure is not optimal, it is widely used to correct for horizontal misalignment of irradiance sensors, because the active leveling technique is technically challenging and expensive. To correct the irradiance measurements for airplane attitude deviations from the horizon, the following equation is applied (Bannehr and Schwiesow, 1993):

$$k = \frac{\sin h_0}{\cos h_0 \ \sin\phi \sin(\varphi_0 - \psi) - \cos h_0 \ \sin\Theta \cos(\varphi_0 - \psi) + \sin h_0 \ \cos\Theta \cos\phi}, \tag{71.18}$$

with the solar altitude angle $h_0 = \pi - \theta_0$; θ_0 and φ_0 are the solar zenith distance angle and the solar azimuth angles, respectively. The aircraft attitude angles (no index) are given as ψ, the true heading of the aircraft with respect to the Earth-fixed coordinate system, Θ being the aircraft pitch angle and ϕ representing the aircraft roll angle. This formula does not consider for the effects, in case the radiometer is not level with the aircraft-fixed horizontal plane.

71.2.3.2 Active Horizontal Stabilization of Sensor Heads

An active and fast horizontal stabilization technique has been developed by Wendisch et al. (2001) and Bucholtz et al. (2008). The technical challenge is to precisely measure the roll and pitch angles on an accelerated platform (aircraft) and to use these measurements to compensate for aircraft attitude changes in real time with desired final adjustment accuracy better than ±0.2°. To achieve these demands, an active horizontal stabilization involves two parts: (a) an accurate aircraft roll and pitch angle measurement unit and (b) an active horizontal mechanical adjustment system:

a. Usual inclination sensors measure the tilt of a coordinate system with respect to the Earth's gravity vector. On accelerated platforms like an aircraft, this technique does not work because the sensor cannot distinguish between the Earth's gravity and the acceleration vector of the moving platform. In this case, a so-called artificial horizon (AHZ) has to be applied. The AHZ consists of three linear servo-acceleration sensors, which deliver the aircraft velocity and position with respect to the inertial Earth-fixed coordinate system by integration of the lateral acceleration measurements and three fiber optical gyros, which measure angular rates also with respect to the inertial Earth–fixed frame. From these data, the aircraft attitude angles (i.e., roll and pitch) are derived. However, the measurements

of the acceleration sensors as well as of the fiber optical gyros are affected by temporal drifts due to sensor errors and electronic noise. These drifts are compensated by using supporting information from a Global Positioning System (GPS). The combined AHZ and GPS data are processed, and the resulting accurate position and attitude data are stored on a first data acquisition system.

b. The processor unit transfers analog output signals of the roll and pitch angles to a personal computer. This drives (after amplification) two separate 2D tilt stages, which are connected with the optical inlet systems of the radiation sensors, that is, cardianically mounted. Each of the two separate 2D tilt stages consists of two servomotors (servos), which realize the horizontal adjustment of the optical inlets. Additionally, the measured attitude data from the combined AHZ–GPS system and the signals to drive the servos are stored on a second data acquisition system.

71.2.4 Techniques to Measure Terrestrial Radiation

71.2.4.1 Broadband Thermal IR Irradiance

Terrestrial radiation can be measured with an IR thermometer or a pyrgeometer, for example, Eppley precision IR radiometer (PIR) or Kipp & Zonen pyrgeometers. The measurement principle for the pyrgeometer is based on the exchange of energy between hotter and cooler objects (Foot, 1986; Philipona et al., 1995). Assuming that the instrument is one of the objects, the instrument absorbs or emits energy if it is cooler or warmer than the object that is being sensed. A pyrgeometer is used to measure, hemispherically, the exchange of terrestrial radiation between a horizontal blackened surface (the detector) and the target viewed (i.e., the sky or the ground). The PIR has the same circular multijunction thermopile detector as the pyranometer, which is used as a method of transducing the terrestrial radiation flux into an electric response. The inner silicon dome is coated with a vacuum-deposited interference filter. Radiation at wavelengths from 0.3 to 0.4 μm to approximately 50 μm is transmitted into the sensor. The amount of radiation in the visible spectrum 0.3–4.0 μm is not significant; absorption and reemission effects are small and are compensated for in the calibration.

71.2.4.2 Spectral Thermal IR Radiance

These instruments are composed of three major parts: the optical system, a detector, and a signal processing system (electronics and data acquisition). Furthermore, calibration targets are required that are usually a hot and cold black-body target. An optical window transmits the radiation from the outside atmosphere into the instrument's optics and to the detector. Zinc selenide (ZnSe) is a common material that has a fairly flat spectral response and transmits radiation across a wide spectral range. Then a filter selects wavelengths; it consists of a narrow band-pass highly transmissive material. Some instruments have individual filters in front of a detector; others use a filter wheel to rotate a selection of filters in front of the detector. Then the light hits the detector (thermal or quantum). Thermal detectors convert the absorbed radiative energy into thermal one. Most thermal detectors do not require cooling but have slow response times and poor detection capability. Quantum detectors usually are faster but often require cooling.

71.2.5 Microwave Radiometers

The MW spectral range of terrestrial EM radiation is characterized by a typically high transmissivity of the atmosphere (>80% below 40 GHz), which allows the retrieval of surface properties by downward looking MW radiometers from aircraft. Because wavelengths are large compared to atmospheric particles, scattering of EM radiation can be ignored for most applications in the MW spectral range. This simplifies the respective equations of the radiative transfers significantly.

Some examples of scanning airborne MW radiometers are given here. The conical scanning millimeter-wave imaging radiometer (CoSSIR) (Wang et al., 2007) has a fixed angle of incidence and a vertical

axis of rotation. The polarimetric scanning radiometer (PSR) (Stankov et al., 2008) allows both cross-track and conical scanning. Similarly, the microwave airborne radiometer scanning system (MARSS) (McGrath and Hewison, 2001) is operated in an external pod.

The beam width should be limited to a few degrees. Therefore, the antenna size becomes relatively large (up to 0.5 m for <10 GHz), which often proves to be a limiting factor for the operation on a smaller aircraft. Flight altitude and antenna beam width determine the spatial resolution of the observations of the Earth's surface. For example, for a flight altitude of 6 km and a beam width of 8°, the spatial resolution at the surface of the Earth is about 2 km.

MW radiometers need to significantly amplify the signal received at the antenna. Low-noise amplifiers (LNAs) are available for frequencies up to about 100 GHz.

71.3 Examples of Applications

71.3.1 Airborne Surface Albedo Measurements

The albedo α in general is defined as the ratio of upward and downward irradiance (see Equation 71.10) at any altitude in the atmosphere. It is a nonlinear function of the vertical coordinate τ; it is wavelength-dependent and a function of the underlying atmospheric and ground properties. The spectral surface albedo at the ground is the spectral albedo at the Earth's surface, which constitutes a major input function for radiative transfer simulations. Surface albedo cannot be measured representatively at the ground. Also, it cannot be directly derived from aircraft measurements of α, because scattering and absorption processes in the atmospheric layer between the aircraft and the Earth's surface also influence the signal measured at the aircraft altitude. To correct for this so-called atmospheric masking, a nonlinear retrieval procedure has been developed by Wendisch et al. (2004) and applied over different terrains. Figure 71.3 shows some exemplary results of surface albedo measurements over water, desert, ice, and vegetation corrected for atmospheric masking by using the method of Wendisch et al. (2004). Different spectral features become obvious from Figure 71.3. Generally, low values of surface albedo are observed over water; ice surfaces exhibit high surface albedo values. The typical vegetation step around 750 nm wavelengths dominates the surface albedo spectra over planted surfaces.

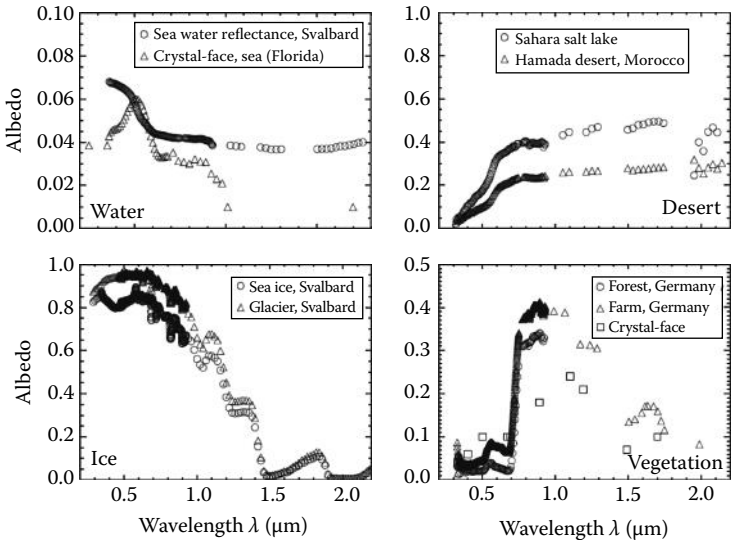


FIGURE 71.3 Different surface albedo spectra for various surface conditions. (Adapted from Wendisch, M. and Yang, P.: *Theory of Atmospheric Radiative Transfer—A Comprehensive Introduction*. Weinheim, Germany. 2012. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reprinted with permission; Courtesy of E. Bierwirth.)

71.3.2 Impact of Crystal Shape in Cirrus on Thermal IR Irradiance

Examples of airborne pyrgeometer measurements along straight, horizontal flight paths are plotted in Figure 71.4 as open squares with horizontal lines indicating the standard deviation during the flight leg. The curve notation indicating the assumed crystal shape is given in Table 71.1.

Figure 71.4a and c show that the cirrus decreases the upward broadband thermal IR irradiance above its base height, compared to the cloudless atmosphere (greenhouse effect of cirrus). The strength

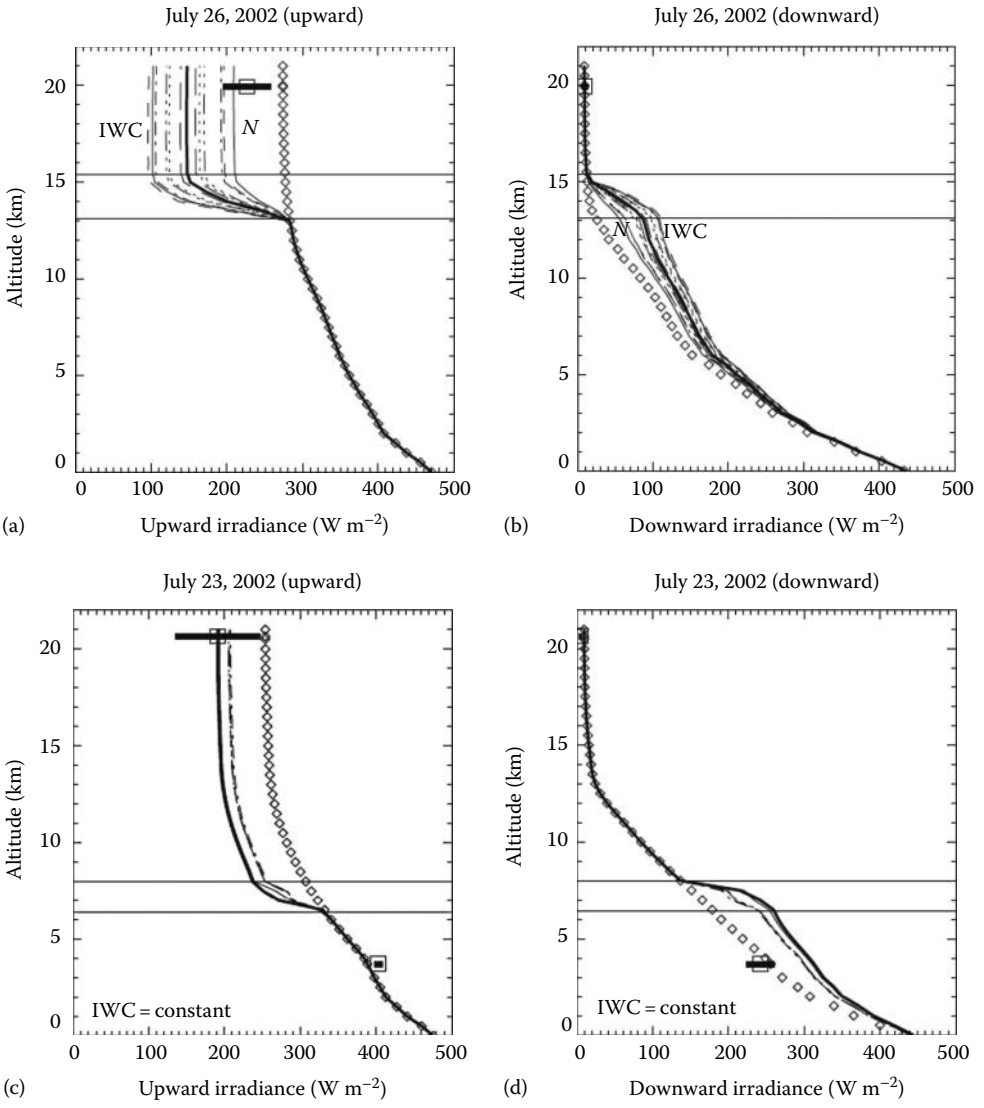


FIGURE 71.4 Vertical profiles of simulated broadband thermal IR upward and downward irradiance. The horizontal lines mark the cirrus top and bottom heights. The open diamonds indicate the simulations without cirrus (cloudless conditions). The open squares show the average of the measured irradiance, with horizontal bars representing the two-sigma standard deviations along the flight track. The first scenario is labeled by *N* (or *N* = constant), and the second scenario is marked by *IWC* (or *IWC* = constant), respectively. For details, see Wendisch et al. (2007). The curve notation is specified in Table 71.1. (Adapted from Wendisch, M., Yang, P., and Pilewskie, P., Effects of ice crystal habit on thermal infrared radiative properties and forcing of cirrus, *J. Geophys. Res.*, 112, D03202. Copyright 2007 American Geophysical Union.)

TABLE 71.1 Curve Notation Depending on Crystal Shape

Ice Crystal Shape	Curve Notation
Sphere	Thick solid lines
Droxtal	Long dashes
Solid hexagonal column	Dotted lines
Hollow hexagonal column	Dash-dot-dot-dotted lines
Solid hexagonal plate	Dash-dotted lines
Bullet rosette	Dashed lines
Aggregate	Thin solid lines

of this decrease depends on the altitude of the cirrus (i.e., its temperature) and its optical thickness. The height (temperature) dependence is because the cirrus emits less thermal IR radiation in colder environment (higher altitude). Below the cloud, there are no effects of the assumed crystal habit on the upward thermal IR irradiances. Above the cloud base, there are significant shape-related effects on upward irradiance in the cirrus case of small optical thickness. An outgoing longwave radiation (OLR) bias is seen as the difference in calculated and measured upward terrestrial irradiance in Figures 71.4a. Most probably this OLR bias is caused by the neglect of horizontal cirrus inhomogeneities. The effect is most significant for semitransparent cirrus (low optical thickness), which is located in the cold upper troposphere. For the low (warm), optically thick cirrus observed on July 23, 2002, there is no such OLR bias.

Figure 71.4b and d demonstrate that the cirrus generally increases the downward broadband thermal IR irradiance below its top height, compared to the cloudless atmosphere. This is because (due to its emission) the cirrus constitutes an additional source of thermal IR radiation, which adds to the cloud-free atmosphere emission. The cirrus has almost no impact on the downward broadband thermal IR irradiance reaching the surface; the emission source (cirrus) is too far away from the surface. The impact of crystal shape on atmospheric radiant energy budget simulations is concluded to be important for the high, optically thin cirrus only.

References

Anderson, G., Clough, S., Kneizys, F., Chetwynd, J., and Shettle, E.: *AFGL Atmospheric Constituent Profiles (0–120 km)*, Tech. Rep. AFGL-TR-86-0110, AFGL (OPI), Hanscom AFB, Bedford, MA, 1986.

Bannehr, L. and Schwiesow, R.: A technique to account for the misalignment of pyranometers installed on aircraft, *J. Atmos. Oceanic Technol.*, 10, 774–777, 1993.

Berk, A., Anderson, G., Bernstein, L., Acharya, P., Dothe, H., Matthew, M., Adler-Golden, S. et al.: MODTRAN4 radiative transfer modeling for atmospheric correction, in *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III*, Vol. 3756, pp. 348–353, SPIE, 1999.

Bucholtz, A., Bluth, R., Kelly, B., Taylor, S., Batson, K., Sarto, A., Tooman, T., and McCoy Jr., R.: The stabilized radiometer platform (STRAP)—An actively stabilized horizontally level platform for improved aircraft irradiance measurements, *J. Atmos. Oceanic Technol.*, 25, 2161–2175, 2008, doi:10.1175/2008JTECHA1085.1.

Coulson, K.: *Solar and Terrestrial Radiation–Methods and Measurements*, Academic Press, New York, 1975.

Crowther, B.: The design, construction, and calibration of a spectral diffuse/global irradiance meter, PhD thesis, University of Arizona, Tucson, AZ, 1997.

Foot, J.: A new pyrgeometer, *J. Atmos. Oceanic Technol.*, 3, 363–370, 1986.

Früh, B., Trautmann, T., Wendisch, M., and Keil, A.: Comparison of observed and simulated NO_2 photodissociation frequencies in a cloudless atmosphere and in continental boundary layer clouds, *J. Geophys. Res.*, 105, 9843–9857, 2000.

Geist, J. and Blevin, W.: Chopper–stabilized null radiometer based upon an electrically calibrated pyroelectric detector, *Appl. Opt*., 12, 2532–2535, 1973.

Hengstberger, F.: Thermal detectors of optical radiation, in *Absolute Radiometry*, ed. F. Hengstberger, Academic Press, Inc., San Diego, CA, 1989.

Hofzumahaus, A.: Measurement of photolysis frequencies in the atmosphere, in *Analytical Techniques for Atmospheric Measurement*, D. E. Heard (ed.), Blackwell Publishing, Oxford, U.K., 2006, pp. 406–500.

Jäkel, E., Wendisch, M., Kniffka, A., and Trautmann, T.: Airborne system for fast measurements of upwelling and downwelling spectral actinic flux densities, *Appl. Opt*., 44, 434–444, 2005.

Junkermann, W.: Measurement of the J(O^1D) actinic flux within and above stratiform clouds and above snow surfaces, *Geophys. Res. Lett*., 21, 793–796, 1994.

Kindel, B., Schmidt, K., Pilewskie, P., Baum, B., Yang, P., and Platnick, S.: Observations and modeling of ice cloud shortwave spectral albedo during the tropical composition, cloud and climate coupling experiment, *J. Geophys. Res*., 115, D00J18, 2010, doi: 10. 1029/2009JD013127.

Kipp & Zonen, *CMP Series Pyranometer Instruction Manual*, The Netherlands, 2010.

Langley, S.: The "Solar Constant" and related problems, *Astrophys. J*., 17, 89–99, 1903.

Matsumoto, T., Russell, P., Mina, C., van Ark, W., and Banta, V.: Airborne tracking sunphotometer, *J. Atmos. Oceanic Technol*., 4, 336–339, 1987.

McGrath, A. and Hewison, T.: Measuring the accuracy of MARSS—An airborne microwave radiometer, *J. Atmos. Oceanic Technol*., 18, 2003–2012, 2001.

Philipona, R., Fröhlich, C., and Betz, C.: Characterization of pyrgeometers and the accuracy of atmospheric long-wave radiation measurements, *Appl. Opt*., 34, 1598–1605, 1995.

Pilewskie, P., Pommier, J., Bergstrom, R., Gore, W., Howard, S., Rabbette, M., Schmid, B., Hobbs, P. V., and Tsay, S. C.: Solar spectral radiative forcing during the Southern African Regional Science Initiative, *J. Geophys. Res*., 108, 8486, 2003, doi: 10. 1029/2002JD002411.

Pilewskie, P. and Valero, F.: Optical depths and haze particle sizes during AGASP III, *Atmos. Environ*., 27A, 2895–2899, 1993.

Porter, J., Clarke, A., Reid, J., Shaw, G., Maring, H., Reid, E., and Kress, D.: Handheld Sun photometer measurements from light aircraft, *J. Atmos. Oceanic Technol*., 24, 1588–1597, 2007.

Russell, P., Livingston, J., Dutton, E., Pueschel, R., Reagan, J., DeFoor, T., Box, M. et al.: Pinatubo and pre-Pinatubo optical-depth spectra: Mauna Loa measurements, comparisons, inferred particle size distributions, radiative effects, and relationship to Lidar data, *J. Geophys. Res*., 98, 22969–22985, 1993.

Sellers, W.: *Physical Climatology*, The University of Chicago Press, Chicago, IL, 1965.

Shetter, R. and Müller, M.: Photolysis frequency measurements using actinic flux spectroradiometry during the PEM-Tropics mission: Instrumentation description and some results, *J. Geophys. Res*., 104, 5647–5661, 1999.

Stankov, B., Cline, D., Weber, B., Gasiewski, A., and Wick, G.: High-resolution airborne polarimetric microwave imaging of snow cover during the NASA cold land processes experiment, *IEEE T. Geosci. Remote*, 8, 1635–1671, 2008.

Trenberth, K. E. and Fasullo, J. T.: Tracking Earth's energy: From El Niño to global warming, *Surv. Geophys*., 33, 413–426, 2012, doi:10.1007/s10712-011-9150-2.

Valero, F., Ackerman, T., and Gore, W.: The effects of the Arctic haze as determined from airborne radiometric measurements during AGASP II, *J. Atmos. Chem*., 9, 225–244, 1989.

Valero, F., Gore, W., and Giver, L.: Radiative flux measurements in the troposphere, *Appl. Opt*., 21, 831–838, 1982.

Volz-Thomas, A., Lerner, A., Paetz, H.-W., Schultz, M., McKenna, D., Schmitt, R., Madronich, S., and Roeth, E.: Airborne measurements of the photolysis of NO_2, *J. Geophys. Res*., 101, 18613–18627, 1996.

Wang, J., Racette, P., Piepmeier, J., Monosmith, B., and Manning, W.: Airborne CoSMIR observations between 50 and 183 GHz over snow-covered Sierra mountains, *IEEE Trans. Geosci. Remote*, 45, 55–61, 2007, doi:10.1109/TGRS.2006.885410.

Wendisch, M. and Brenguier, J.-L. (Eds.): *Airborne Measurements for Environmental Research—Methods and Instruments*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2013.

Wendisch, M. and Keil, A.: Discrepancies between measured and modeled solar and UV radiation within polluted boundary layer clouds, *J. Geophys. Res*., 104, 27373–27385, 1999.

Wendisch, M., Mertes, S., Ruggaber, A., and Nakajima, T.: Vertical profiles of aerosol and radiation and the influence of a temperature inversion: Measurements and radiative transfer calculations, *J. Appl. Meteorol*., 35, 1703–1715, 1996.

Wendisch, M., Müller, D., Schell, D., and Heintzenberg, J.: An airborne spectral albedometer with active horizontal stabilization, *J. Atmos. Oceanic Technol*., 18, 1856–1866, 2001.

Wendisch, M., Pilewskie, P., Bohn, B., Bucholtz, A., Crewell, S., Harlow, C., Jäkel, E., Schmidt, K. S., Shetter, R., Taylor, J., Turner, D. D., and Zöger, M.: Atmospheric radiation measurements, in Wendisch M. and Brenguier J.-L. (eds.): *Airborne Measurements for Environmental Research—Methods and Instruments*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2013.

Wendisch, M., Pilewskie, P., Jäkel, E., Schmidt, S., Pommier, J., Howard, S., Jonsson, H., Guan, H., Schröder, M., and Mayer, B.: Airborne measurements of areal spectral surface albedo over different sea and land surfaces, *J. Geophys. Res*., 109, 2004, doi: 10. 1029/2003JD004392.

Wendisch, M. and Yang, P.: *Theory of Atmospheric Radiative Transfer—A Comprehensive Introduction*, Wiley–VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2012.

Wendisch, M., Yang, P., and Pilewskie, P.: Effects of ice crystal habit on thermal infrared radiative properties and forcing of cirrus, *J. Geophys. Res*., 112, D03202, 2007.

72

In Situ Sensors for Monitoring the Marine Environment

Graham A. Mills
University of Portsmouth

Gary R. Fones
University of Portsmouth

Silke Kröger
Centre for Environment, Fisheries and Aquaculture Science

72.1 Introduction

The marine environment has been affected by growing population and industrial pressures in the coastal zone. Activities such as dumping of waste, construction of harbors, dredging, and extraction processes all contribute to changes in environmental quality. In order to assess the impacts of these activities and instigate appropriate remedial actions where needed, it is necessary to monitor a wide range of properties (i.e., biological, chemical, and physical) of the marine environment. Recently, such monitoring activities are becoming harmonized within regulatory frameworks and international agreements (e.g., within the Europe-specific initiatives under the Helsinki and OSPAR Commissions, the European Union's Water Framework Directive [200/60/EC], and more recently Marine Strategy Framework Directive [2008/56/EC]). The latter requires for a thematic strategy for the protection and conservation of the marine environment and mandates for the establishment of clear environmental targets and monitoring programs. Most work on developing monitoring methods has been focused on freshwater (e.g., ground and surface water) applications, but now there is an increasing demand to adapt these methods for use in the marine sector. Monitoring the marine environment, however, presents a number of problems that are not encountered with freshwaters. For example, oceans cover very large areas, and there are complex currents and, in coastal systems, tides that can cause large shifts in salinity. Large dilutions of pollutants occur, and so many contaminants can appear at very low, but often biologically significant, concentrations. Boats and ships are usually needed to reach monitoring sites, with samples of water, sediment, and biota transported back to land-based laboratories for analysis or in some cases analyzed onboard ship. The use of sampling coupled with classical biological or chemical analysis is expensive and time-consuming and often fails to provide representative information over changes in time (short term over a tidal cycle and longer term over seasons, years, and decades). An additional problem is that delays between sampling and analysis can compromise sample integrity (especially for

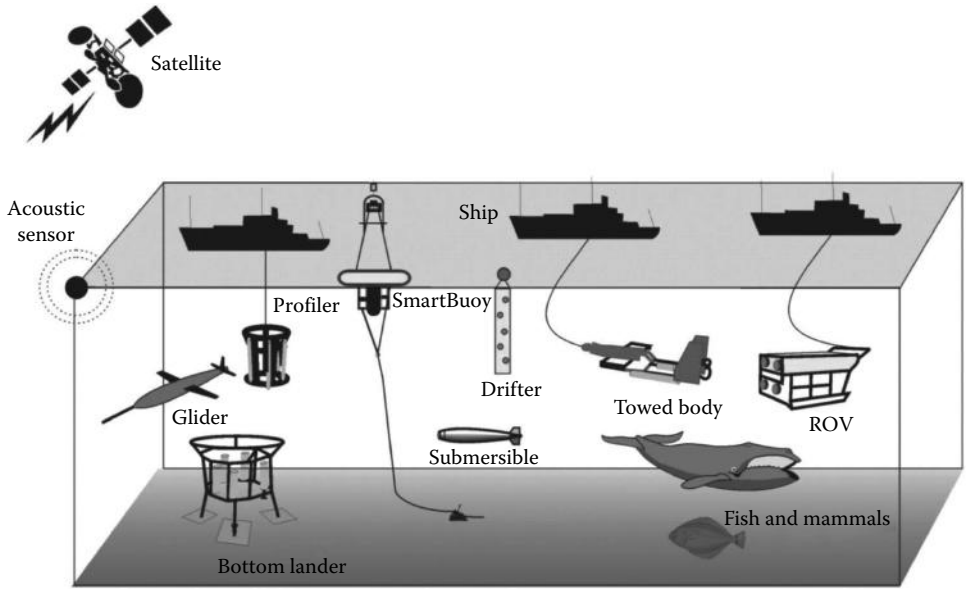


FIGURE 72.1 Schematic overview of marine observational strategies. Each of the elements shown is or could be equipped with a range of sensors.

labile compounds), while changes in temperature and pressure can also be an issue. A solution to some of the problems outlined earlier is the use of sensors that are particularly suited for making *in situ* measurements in the marine environment and can provide continuous or semicontinuous observations and transmit the information telemetrically to shore-based data-gathering facilities. Their use can overcome some of the problems of the undersampling (in both space and time) of coastal waters and the ocean. Many engineering platforms are available on which these sensors can be deployed in the environment and include APEX floats, autonomous underwater vehicles (AUVs), gliders, benthic landers, moorings, remotely operated vehicles (ROVs), nodes attached to cable networks, and even tagged live stock and wildlife (Figure 72.1). This approach thus eliminates the need for expensive (>$15–20,000 day^{-1}) sampling from ships, and there can be real-time acquisition of information where any strategic decisions can be made in a short time scale.

The first *in situ* analyzer used in the marine environment was for determining salinity on the basis of measurements of conductivity and temperature. This was reported in 1964 and is considered to represent a turning point in marine chemistry (Wangersky 2005). Today, many *in situ* techniques are used for monitoring hydrological and physicochemical variables in the marine environment. For example, a system of observational buoys (such as the Cefas "SmartBuoys" [Figure 72.2], http://www.cefas.defra.gov.uk/our-services/monitoring-and-mapping/autonomous-monitoring/smartbuoys.aspx) is in operation in European waters, and these devices house a range of measuring instruments (e.g., for monitoring pH, pO_2, turbidity, conductivity, and concentrations of nutrients such as phosphate and nitrate) (Figure 72.3). The European Marine Ecosystem Observatory (EMECO, http://www.emecogroup.org) is a consortium of European Marine Institutes that aim to integrate marine environmental monitoring, ecosystem modeling, and coastal and ocean research. But there are still significant gaps to fill, for example, there are fewer sensors in use for monitoring concentrations of heavy metals, organic pollutants, and algal toxins. Like any technology deployed in nutrient-rich environments, all these systems suffer from the problem of biofouling, which can be a factor in limiting overall deployment times. This chapter aims to provide information on the current developments in the area of *in situ* monitoring of the chemical quality of marine waters. It does not address remote sensing observations of, for example, phytoplankton blooms,

FIGURE 72.2 Cefas SmartBuoy being deployed in UK coastal waters. (Photo from The Centre for Environment, Fisheries and Aquaculture Science, http://www.cefas.co.uk.)

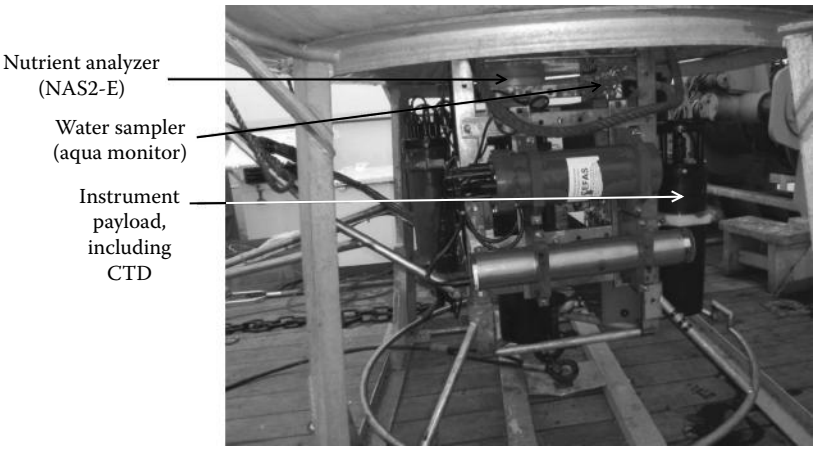


FIGURE 72.3 Instrument packages on the Cefas SmartBuoy. (Photo from The Centre for Environment, Fisheries and Aquaculture Science, http://www.cefas.co.uk.)

using satellite-based technologies. The topics covered here include physical and chemical sensors, spectroscopic and electrochemical devices, and biosensors, together with a view on the future developments and applications of these technologies to this sector. It is not the authors' aim to describe every device available or that could be potentially useful; the reader is referred for further information to other reviews cited in the bibliography and the web links.

72.2 Physical- and Chemical-Based Sensors

A sensor is defined as a device that produces a response to a change in a physical condition, such as temperature or thermal conductivity, or to a change in chemical concentration. Sensors, therefore, are particularly suited for making *in situ* measurements in the marine environment. Their use can overcome some of the problems of the undersampling (in both space and time) of coastal waters and the ocean. Currently, there is a significant research effort being undertaken to push forward the use of *in situ* sensors. An example is as follows: this is the ongoing work at the Centre for Marine Microsystems at the University of Southampton, United Kingdom (http://www.southampton.ac.uk/cmm/), for the

development of miniaturized multiparametric biogeochemical sensors. Here, they are developing a range of chemical sensors for nitrate/nitrite, phosphate, ammonia, oxygen, methane, pH, alkalinity, and iron/manganese using lab-on-a-chip technologies (Sieben et al. 2010).

In situ sensors have been used for a number of years to measure physical-based parameters such as conductivity, oxygen, pH, and carbon dioxide in seawater. The majority of these are available commercially (Table 72.1), particularly in sensor packages such as conductivity, temperature, and depth (CTD) profiling instruments. The term CTD is often used today to describe a package that includes the actual CTD as well as auxiliary sensors to measure parameters such as dissolved oxygen, pH, turbidity, chlorophyll-*a*, rhodamine, blue-green algae, ammonia, and nitrate. Often, arrays of sensors are used so that multiple parameters can be measured simultaneously, and these systems can be fully automated. These can either be deployed from a ship in the profiling mode or can be deployed on a mooring for long-term monitoring as part of an observation system.

Commercially available conductivity sensors use an ac voltage applied to nickel electrodes or graphite electrodes. The challenges are to maintain the calibration parameters during long-term deployments (Owens and Wong 2009). Oxygen sensors are also available commercially for a number of applications, such as profiling sediment–water interfaces. These can be either electrochemically or optically based systems (Glud et al. 2000). Most oxygen sensors are based on the original Clark polarographic electrode (Reimers 2007); however, there are now optical dissolved oxygen sensors that utilize luminescent technology (Glud et al. 2000; Reimers 2007), such as the AADI oxygen optode, which has been used in a number of marine applications (Körtzinger et al. 2004; Tengberg et al. 2006; Hydes et al. 2009; Queste et al. 2012).

There are numerous sensors available both commercially and as research devices for the measurement of pH and pCO_2. These parameters are important in understanding climate change and therefore need to be measured accurately in monitoring programs, ideally by *in situ* measurements. The requirements for making accurate measurements of oceanic carbonate chemistry including pH have been documented by Dickson at al. (2007). The sensors include electrochemical, potentiometric, photometric, and fiber-optic-based instruments (Cai and Reimers 2000). There are two primary techniques for measuring *in situ* pH. The first is electrochemical (potentiometric) using a hydrogen electrode and reference electrode and is the standard method. However, there is a major problem with these devices when ionic strength and composition change, such as in estuarine environments. The second approach is the photometric/fiber-optic measurement of pH (Cai and Reimers 2000). This uses an indicator dye that is added directly to the test solution or is confined within a sensing cell or matrix. These techniques have excellent precision (0.001 pH units) and accuracy (0.004 pH units) but ideally are used in flow-through systems and are therefore best suited for operation onboard research ships (Bellerby et al. 2002). Instruments for measuring pCO_2 have been used mainly as flow-through systems onboard ships (Hardman-Mountford et al. 2008); however, a pCO_2 system for deployment on ocean moorings has been developed (DeGrandpre et al. 1997). This consists of a fiber-optic pH sensor in a flow-through cell connected to a CO_2-permeable membrane equilibrator.

72.3 *In Situ* Automatic Analyzers

Another approach to achieving *in situ* measurements has been the miniaturization of existing analytical methods, such as colorimetric and spectroscopic methods. In the former, some quantitative chemical reaction that generates a colored product that can be measured spectroscopically is used; in the latter, analytes are measured directly in the water. There is considerable experience in using these methods in the laboratory and more recently online and in the field. Developments in optical fiber technology; solid-state electronics; long-life, stable power supplies; pumps; and valves have enabled the production of robust, miniaturized instruments that are portable for use at sea. An early example is the system developed to be used on research vessels to measure concentrations of hydrogen sulfide in seawater (Sakamoto-Arnold et al. 1986). The subsequent work to improve the reliability of these miniaturized

TABLE 72.1 Examples of Commercially Available Sensors for Applications in Marine Environments

Target Analytes (Group)	Sensor Principle	Sensor	Company, Location	Link
Physicochemical				
Conductivity	Conductance	SBE 4C	Sea-Bird Electronics, United States	www.seabird.com
CTD	Conductance, thermistor, transducer	YSI CastAway™ CTD	YSI, United States	www.ysi.com
CTD	Conductance, thermistor, transducer	DataSonde 5X (DS5X)	Hydrolab	www.hachhydromet.com
Temperature	Thermistor	SBE-39	Sea-Bird Electronics, United States	www.seabird.com
Pressure	Transducer	SBE 53 BPR	Sea-Bird Electronics, United States	www.seabird.com
Turbidity	Fiber optic	Analite390 series	McVan Instruments, Australia	www.mcvan.com
Turbidity	Optical	6136 Turbidity Sensor	YSI, United States	www.ysi.com
Photosynthetically active radiation (PAR)	Photodiode	PAR 600m	Satlantic, Canada	www.satlantic.com
Nutrients, oxygen, pH, and sulfide				
Nitrate, nitrite	Optical	ProPs UV	TriOS, Germany	www.trios.de
Nitrate, nitrite	Optical	ISUS V3	Satlantic, Canada	www.satlantic.com
Nitrate	Optical	SUNA (Submersible UV Nitrate Analyzer)	Satlantic, Canada	www.satlantic.com
Nitrate	Ca reduction/diazotization	9600 Nitrate Monitor	YSI, United States	www.ysi.com
Nitrate, nitrite, phosphate, ammonia	Chemical/optical	APNA II	SubChem Systems, Inc., United States	www.subchem.com
Ammonia	Chemical/optical	MARCHEM	SubChem Systems, Inc., Unites States	www.subchem.com
Nitrate, nitrite, ammonia, phosphate, silicate, iron(II), iron(III)	Chemical/optical	SubChemPak Analyzer (with SubChemPak reagent delivery module and ChemStar absorption detector by WET Labs, Inc.)	SubChem Systems, Inc., United States	www.subchem.com
Phosphate	Microfluidics/optical	CYCLE-PO_4	Wet Labs, United States	www.wetlabs.com
O_2, H_2, H_2S, N_2O, NO, pH, redox, NOx	Electrochemical and biosensor	Range of microsensors	Unisense, Denmark	www.unisense.com
O_2, pH, and CO_2	Optical	Range of sensor probes	PreSens, Germany	www.presens.de
O_2	Optode	Oxygen optode 4330F	AADI, Norway	www.aadi.no
pH	Ion selective	SeaFET Ocean pH Sensor	Satlantic, Canada	http://www.satlantic.com
Sulfide	Amperometric	Submersible sulfide/H_2S probe	Sea & Sun, Germany	www.sea-sun-tech.com

(*continued*)

TABLE 72.1 (continued) Examples of Commercially Available Sensors for Applications in Marine Environments

Target Analytes (Group)	Sensor Principle	Sensor	Company, Location	Link
Pollutants				
Hydrocarbons, PAH	Optical	enviroFlu-HC	TriOS, Germany	www.trios.de
Hydrocarbons	Optical	Hydrocarbon fluorometer	Sea & Sun, Germany	www.sea-sun-tech.com
Humic acids, amino acids, BTXE, PAH	Optical	HydroC™/PAH	CONTROS, Germany	www.contros.eu
Hydrocarbons	Optical	UViLux and UV AQUAtracka	Chelsea Technologies Group, United Kingdom	www.chelsea.co.uk
Crude oil	Optical	Cyclops-7 Submersible Sensors	Turner Designs, United States	www.turnerdesigns.com
Environmental effect monitoring (focus oil and gas)	Sensor array including mussel-based biosensor	Biota Guard AS	Biota Guard, Norway	www.biotaguard.no
PAHs	Antibody-based biosensor	KinExA Inline Sensor	Sapidyne Instruments, United States	www.sapidyne.com
Trace metals				
Trace metals Cu(II), Pb(II), Cd(II) Zn(II); Mn(II), Fe(II)	Electrochemical/voltammetric	VIP system	IDRONAUT, Italy	www.idronaut.it
Organisms				
Microalgae species composition	Optical	Multi-Exciter—*in vivo* multiwavelength excitation fluorescence	JFE ALEC Co. Ltd., Japan	www.jfe-alec.co.jp
Microalgae class composition/total Chl analysis	Optical	bbe FluoroProbe bbe AlgaeGuard	bbe moldaenke GmbH, Germany	www.bbe-moldaenke.de
Chlorophyll	Optical	MicroFlu-chl	TriOS, Germany	www.trios.de
Phytocyanin	Optical	MicroFlu-blue	TriOS, Germany	www.trios.de
Total chlorophyll/cyanobacteria	Optical	AlgaTorch	bbe moldaenke GmbH, Germany	www.bbe-moldaenke.de
Chlorophyll	Optical	ECO BB2F	WET Labs, United States	www.wetlabs.com
Chlorophyll, rhodamine, fluorescein, phycocyanin, phycoerythrin	Optical	AQUAtracka III and UniLux/TriLux series	Chelsea Technologies Group, United Kingdom	www.chelsea.co.uk
Chlorophyll-*a in vivo*, phycocyanin, phycoerythrin, cyanobacteria	Optical	Cyclops-7 Submersible Sensors	Turner Designs, United States	www.turnerdesigns.com
Plankton species composition, particles	Molecular biology/optical	Environmental Sample Processor	Spyglass	www.spyglassbio.com
Toxic algae and microbial species (*Vibrio, Escherichia coli, Salmonella*, etc.)	Flow cytometry/optical	CytoBuoy, CytoSense, CytoSub	CytoBuoy	www.cytobuoy.com

Source: Modified from Zielinski, O. et al., *Ocean Sci.*, 5, 329, 2009. With kind permission.

systems that are based on flow-injection analysis (FIA) methods was accompanied by developments in the design of instrument housings for deployments in robust environments such as fast-flowing water or deep water.

72.3.1 Flow-Injection Methods

Analyzers based on FIA methods for use *in situ* comprise a pump, detector, and narrow bore tube manifold (Vuillemin et al. 2009). The sample is pumped through a tube manifold where it is mixed with appropriate reagents to form a product that can be detected by a spectroscopic method. In some cases, the reaction product is colored and can be detected by visible wavelength spectroscopy; in others, it is detected by fluorimetry. A system of valves allows sample to be replaced by standard solutions from time to time to achieve an *in situ* calibration of the instrument and provide quality control information. When using such systems in the marine environment, there can be problems caused by large changes in pressure, temperature, and salinity (e.g., over a tidal cycle in inshore waters) that can lead to a significant bias in results. For example, when measuring nitrate and nitrite concentrations, the rate of formation of the colored reaction product was influenced by changes in salinity and differences in temperature with depth also affected the analysis by changing both reaction rates and the viscosity of sample and reagents (and hence flow rates and mixing times) (Daniel et al. 1995).

A number of miniaturized systems have been developed for the measurement of silicic acid (silicate), nitrate and nitrite, iron(II), total iron, phosphate, and total sulfide in the marine environment. One of the key driving forces is the need for information on the chemical quality of water in relation to the distribution of biota in the deep ocean, particularly in the region of hydrothermal vents. Some of these systems have incorporated limited temperature regulation to improve the reliability of the data. Among these is a system based on a colorimetric method (using the production of ß-silicomolybdic acid that is reduced [by tin{II}] into an intensely colored molybdenum blue) for the measurement of silicic acid (silicate) in seawater at a detection limit of 29 μg L^{-1} (Floch et al. 1998). An FIA analyzer for measuring ammonium in fresh and saline waters was developed by Aminot et al. (2001) and was based on a fluorometric method (using an orthophthaldialdehyde–sulfite reagent) since this was robust to variations in refractive index of the sample over a wide range of environmental conditions. The method gave a limit of detection in the range of 0.4–0.5 μg L^{-1}. A system with three analyzers was developed by Thouron et al. (2003) to measure the nutrients: nitrate, phosphate, and silicate. This system had a sampling rate of 2 h^{-1} and was capable of 400 autonomous measurements. The analyzers had detection limits of 6, 10, and 10 μg L^{-1} for nitrate, silicate, and phosphate, respectively. A nitrite microfluidic colorimetric sensor has been developed by Beaton et al. (2011). The low-power system is low cost, portable, and with a limit of detection of 15 nM nitrite. The device was tested in seawater over 57 h, and 375 discrete measurements were taken. Future developments include the use of on-chip valves and pumps to decrease size and operation power demands of the device.

A FIA analyzer (ALCHIMIST) was developed by IFREMER (France) to measure nitrate plus nitrite and total sulfide in deep marine locations near hydrothermal vents (Le Bris et al. 2000). The maximum sampling rate was 22 h^{-1}, and the device could take into account the effects of varying temperature and pressure. This system had detection limits of 26 and 27 μg L^{-1} for total sulfide and nitrite plus nitrate, respectively. More recently, Vuillemin et al. (2009) developed an *in situ* analyzer (the CHEMINI system) for measuring analytes at even greater depths, which demonstrated some of the latest developments in this type of technology. In this system, the reagents (up to 2 L) and standards (250 mL) are contained in flexible bags and enable 400 analyses per deployment. Prototypes of the CHEMINI have been deployed in deep-sea (1650–1750 m) hydrothermal environments for the measurement of dissolved iron and total sulfide. In this difficult environment, the analyzer gave useful spatial and temporal data by taking up to 8 measurements per day over a 6-month period. The combination of miniaturization; the availability of better protective deployment housings; stable, long-life power supplies; and advances in communications capabilities have enabled these FIA analyzer systems to be buoy based, located on robotic vehicles,

or towed by ships. However, the use of reagents sets limits on the time that instruments can be deployed and adds significantly to the bulk and weight of the analytical system. More recently, there has been a move to develop reagentless methods where possible (Lacombe et al. 2008).

72.3.2 Spectroscopic Methods

A number of spectroscopic- and spectrometric-based *in situ* devices have been developed for use in marine applications for measuring dissolved gases, nutrients, organic chemicals, and trace metals (Moore et al. 2009a). These are usually miniaturized versions of conventional laboratory analytical equipment that have been fitted with robust protective housing so as to withstand high pressures. A range of optical detectors can be used and include absorbance, reflectance, luminescence, fluorescence, refractive index, and light scattering. Underwater mass spectrometric detectors can also be employed, usually with a membrane inlet sample introduction system (Prien 2007; Moore et al. 2009b). The following provides a brief discussion of some of different optically and mass-based technologies available.

One of the simplest optical sensing systems measures changes in refractive index. These detectors can be used to measure bulk particulate loads (e.g., amounts of organic and inorganic particles in the water column). These devices work on the back scattering of light principle and can be useful in estimating phytoplankton loads and degree of sediment resuspension. Probably the most widely used method is optical absorption spectroscopy. This method is often used analytically in conjunction with FIA as described earlier. Absorption spectra (both visible and ultraviolet [UV] regions) in seawater are more complex than in pure water due to the influence of salts. The absorption peaks are broader, and this can cause overlap of spectra between different compounds making identification harder. Another interesting example is the use of flow cytometric techniques with optical sensing for the autonomous measurement of harmful algal blooms. Here, an increase in cell abundance correlated well with subsequent blooms (Campbell et al. 2008).

The ISUS is an example of a high-resolution *in situ* UV spectrophotometer that can take rapid (of the order of 1 s) direct measurements (Johnson et al. 2006). This was used to measure nitrate, bisulfide, and bromine as these species have distinctive UV absorption spectra. The device has the potential to operate at depths of 2 km and can be deployed for over 3 months without any degradation of performance due to biofouling or instrumental drift. Such *in situ* UV/visible spectrometers are frequently used to monitor surface, waste, and drinking water treatment processes. Here, light attenuation measurements are made over the whole spectrum (e.g., 200–735 nm), and over this region, it is possible to measure nitrate, chemical oxygen demand (COD), and total suspended solids (turbidity) directly in the water column (van den Broeke et al. 2006). Another deep-sea (up to 6 km) *in situ* UV spectroscopic device operating over 190–360 nm has been developed by TriOS Optical Sensors, Germany (http://www.trios.de/). The instrument was originally designed for nitrate measurements but can also be used for estimating biological oxygen demand (BOD), COD, and total organic carbon (TOC), humic acids, and some phenols. The device has "nanocoated" optical window that resists biofouling on its measurement surface.

An *in situ* dissolved manganese analyzer (Statham et al. 2003) based on the spectrophotometric detection of the 1-(2-pyridylazo)-2-naphthol (PAN)–manganese complex at 560 nm in a continuous flow system was interfaced with Autosub, which is a 7 m long AUV that is capable of carrying a wide range of instruments. The detection limit of the system was 25 nmol L^{-1}, and using 10 s averaged data, the along-track resolution of the system is about 10 m. The analyzer was successfully used with the Autosub (Figure 72.4) in Loch Etive, Scotland, to define the 3D distribution of dissolved manganese in subsurface waters (Statham et al. 2005). An *in situ* continuous flow analyzer for soluble Fe(II) species in seawater was developed by Meyer et al. (2012) and uses a colorimetric method with ferene as the spectrophotometric reagent. The method is based on the measurement of absorbance of the $[Fe^{II}(fer)_3]^{4-}$ complex at 594 nm. The method was applied *in situ*, during a cruise with the R/V "Heincke" in the central Baltic Sea and verified by concurrent analyses of discrete samples by ICP-OES. A detection limit of 20 nM L^{-1} was obtained (http://www.io-warnemuende.de/che-ag-in-situ-sensors.html).

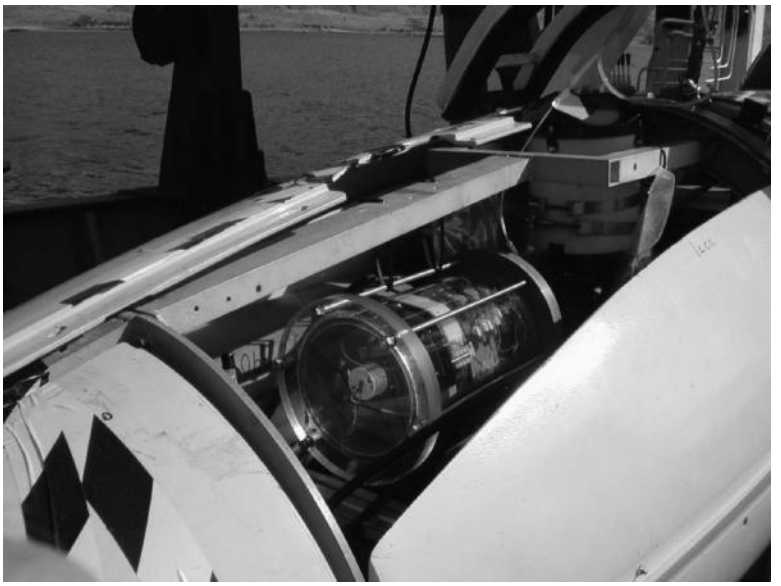


FIGURE 72.4 Manganese analyzer fitted in Autosub. (Photo by Professor Peter Statham, University of Southampton, Southampton, U.K. http://www.bodc.ac.uk/projects/uk/autosub/science_missions/mission2.html. *Source:* Copyright-Professor Peter Statham, University of Southampton, available at: http://www.bodc.ac.uk/projects/uk/autosub/science_missions/mission2.html)

In situ fluorescence spectroscopy can also be used to measure compounds that naturally exhibit this characteristic. Fluorescence spectroscopy is very sensitive; however, the drawback is the intrinsic lack of selectivity in the choice of excitation and emission wavelengths to be used. The broad character of fluorescence peaks usually prohibits finding an explicit excitation/emission wavelength combination that is unique to the analyte being measured. Due to their chemical structure, polynuclear aromatic hydrocarbons (PAHs) are inherently fluorescent. A submersible UV fluorometer has been designed for the measurement of PAHs in water and is commercially available. The enviroFlu-HC (TriOS Optical Sensors, Germany) is designed to excite compounds (using a xenon lamp) at $\lambda \approx 254 \pm 12.5$ nm and detect the emitted light at $\lambda \approx 360 \pm 25$ nm. The instrument weighs about 2 kg and is rated to 500 m depth and is hence suitable for deployment on a number of available systems (Tedetti et al. 2010). *In situ* measurements of chlorophyll are possible using fluorometers (e.g., Seapoint fluorometers used on SmartBuoys [Greenwood et al. 2010]) and can be used in the deduction of primary production rates (Kröger et al. 2009). More sophisticated, so-called fast repetition rate, PhytoFlash, and pulse-amplitude-modulation fluorometers are on the market (Table 72.1), and these can give an indication about the health of the algal population and their photosynthetic yield. With the use of the so-called saturation pulse method, the quantum yield of photosynthetic energy conversion can be derived (Kröger et al. 2009). These methods, however, still need careful calibration and checking with accepted carbon incorporation methods, but good matches between conventional and fluorescence-based production estimates have been found (Kromkamp et al. 2008). More recently, Fischer et al. (2012) developed an online, *in situ* optical fiber–based sensor to measure the growth of biofilm aquatic environments. The device is based on the detection of the natural fluorescence of microorganisms constituting the biofilm. Basically, the intrinsic fluorescence of the amino acid tryptophan is excited at a wavelength of $\lambda = 280$ nm and detected at $\lambda = 350$ nm utilizing a numerically optimized sensor head equipped with a UV–LED light source and optical fiber bundles for efficient fluorescence light collection. The sensor has been used in the Baltic Sea to monitor the growth of biofilm layers and could have other uses such as the mass measurement of biomass and cell activity using tryptophan and adenosine triphosphate fluorescence.

Optodes, sometimes referred to as optrodes, can also work using fluorescence spectroscopy principles. Here, indicator dyes (e.g., ruthenium or platinum complexes) are employed, and these are embedded or immobilized into sol–gel matrices of thin polymer layers. The presence of a test analyte in the light path modifies the properties of the indicator (e.g., emission of fluorescent light), and this forms the basis of operation these sensors. These devices are frequently used to measure oxygen (Prien 2007) but can also be used for pH and pCO_2. Recently, Clark et al. (2012) used an oxygen optode 4330F sensor (AADI, Norway) attached to an AUV to obtain a high-resolution 3D map of dissolved oxygen concentration in Hopavågen Bay, Norway. This powerful, 3D, mapping procedure offers much potential and could be used, for example, in estimating the quantity of dissolved oxygen per volume and has the potential to become a reliable test for the health of an underwater ecosystem (e.g., near fish farming activities).

Infrared devices can also be used, and this area was reviewed by Mizaikoff (2003). The application of these devices is mainly for the measurement of volatile chlorinated and aromatic hydrocarbons, and this relies on mid-infrared range optical systems. This wavelength region is particularly attractive for optical sensing, since molecule-specific information is provided by stimulation of the ground vibrational modes of organic substances. The use of Fourier transform methods enables simultaneous measurements of individual chemicals in multicomponent mixtures at concentrations in the $\mu g\ L^{-1}$ range. Interferences to the signal from changes in salinity and water turbidity also need to be taken into account. An *in situ* miniaturized modular Fourier transform-IR device with a fiber-optic sensor head was developed as part of the EU framework project SOFIE (Kraft and Mizaikoff 2000). The device can operate at depths of 300 m being deployed either on a tow frame or a remotely operated vehicle. The long-term stability and resistance to fouling of the device were good. The concentration of suspended particulate matter (turbidity) in a water body can be measured using the optical back scatter of infrared beams. An understanding of the concentration of suspended particulate matter is important when assessing the degree of penetration of sunlight as this can affect the growth of phytoplankton and this may vary hourly, daily, or seasonally.

Laser Raman spectroscopy is a type of vibrational spectroscopy based on Raman scattering (inelastic scattering) that is capable of nondestructive molecular identifications of solids, liquids, and gases. It can measure multiple species simultaneously and requires no reagents or consumables, making it ideal for *in situ* long-term deployments. Raman measurements are fast (2–10 s), and this makes them useful rapid sensing systems for applications such as profiling. *In situ* devices based on Raman spectroscopy have been developed. Brewer et al. (2004) were one of the first groups to develop a deep-ocean Raman *in situ* spectrometer (DORISS). The instrument uses a Nd:YAG green laser operating at 532 nm to provide the Raman excitation, with a holographic grating as a filter and a charged coupled device array to detect the weak Raman signals that are produced. The DORISS system is capable of being deployed from a remotely operated vehicle and can work to a depth of 3.6 km and at water temperatures of 2°C. *In situ* Raman spectra have been obtained from seawater at various depths and from various gases, liquids, and solid mineral samples. Another Raman-based device has been developed during the EU that funded SOFIE and MISPEC projects (Schmidt et al. 2004). *In situ* Raman sensors do suffer from a lack of sensitivity; however, this may in part be overcome by the use of surface-enhanced Raman scattering (SERS) techniques. Using SERS, Raman intensities can be enhanced by the use of chemical and electromagnetic mechanisms when molecules are adsorbed onto specially prepared metal surfaces such as electrodes or metal colloid particles. Proposed applications include the measurement of low concentrations of hydrocarbons such as dissolved methane (Boulart et al. 2010) and polycyclic aromatic hydrocarbons (Pfannkuche et al. 2012). Although both Raman spectroscopy and SERS show potential, these systems are still very much research instruments, and further work is needed to overcome issues such robustness, fluorescent interferences, and sensitivity issues.

Laser-induced breakdown spectroscopy is an emerging technology with potential *in situ* monitoring applications. This uses a laser pulse to create a microplasma in the sample; then, a spectrophotometer captures the transient light to identify and quantify the analytes. It is a relatively low-cost method and

can be easily adapted for use with automatic or robotic systems. Measurements can be taken rapidly (less than 1 s), and sensitivity (mg L^{-1} or better) is similar for most chemical elements. Michel et al. (2007) demonstrated, in laboratory, the potential of this approach for the detection of calcium, potassium, lithium, manganese, and sodium in seawater. Changes in pressures, temperature, and salinity had little effect on the signal. Preliminary results suggest that this technique could in the future be suitable for the *in situ* chemical analysis of the marine environment.

Membrane inlet mass spectrometers have been used for the *in situ* measurements of dissolved gases and volatile organic compounds at different ocean depths. Devices can be deployed on wide variety of platforms. As they use no reagents for their operation, the spectrometers can be deployed for extended periods. Using mass spectrometers underwater is still challenging since the devices operate under a high vacuum. Analytes must be transported from the pressurized aqueous environment into a vacuum system, and this can place severe demands on pumps, particularly for long-term deployments. It has also been necessary to give consideration to the power requirements of the equipment. Two designs (linear quadrupole or ion trap) of mass spectrometer are used with electron impact ionization (Short et al. 2001). Both have similar detection limits (about 1 mg L^{-1}); however, ion trap systems do have a higher mass limit (650 Da) compared with quadrupole detectors (100 Da). The most common sample introduction method is a high-pressure polydimethylsiloxane membrane, and this can be used to depths of 2 km. These membranes allow the selective permeation of gases and nonpolar, volatile organic compounds. The thickness of the membrane affects the response time and sensitivity of the instrument. Although thin membranes are often preferred, these are mechanically weak and not able to withstand harsh field conditions. Mass spectrometers can be used for the simultaneous detection of multiple species. Most underwater applications have been the *in situ* detection and quantification of dissolved gases (e.g., CO_2, O_2, methane) and small volatile organic compounds (e.g., dimethylsulfide, toluene) at different depths (Camilli and Hemond 2004; Short et al. 2006; Schlüter and Gentz 2008; Camilli and Duryea 2009). Much of this work is still at a preliminary stage. Larger molecular weight (i.e., above 200 Da) compounds are not detected as the permeability of the inlet membrane to such molecules is very low. By the use of global positioning technology, that is, time matched to the concurrent mass spectral data, it is possible to create high-resolution 3D maps of the spatial distribution of chemicals and locate pollutant plumes. Such systems offer potential for the future in monitoring the distribution of specific chemicals. Further developments in microfluidics and lab-on-a-chip technologies could hold much future promise for the development of small-scale gas and liquid chromatographic systems combined with various designs of mass spectrometer; as do the rapid developments taking place with ion mobility mass spectrometers. Future systems would need to have reduced space and power needs and allow for a reduction in consumption of mobile phases. This could allow for the direct analysis of higher molecular mass compounds such as the range of marine toxins (Zielinski et al. 2009).

Radionuclides in the marine environment have been measured using a gamma-radiation probe device, which uses a NaI(Tl) crystal as a scintillation detector with a fully integrated spectrum analyzer. Under field conditions, this *in situ* method delivered comparable results to those based on chemical analysis for mixtures of nuclides originating from accidents (Wedekind et al. 1999). The approach has been further developed into the KATERINA system to measure marine radiation and has been deployed on ROVs and fixed platforms with data output transmitted via satellite to a base station (Tsabaris et al. 2008). More recently, the device has been used off the Peloponnesus Island, Greece, to monitor over a 25-day period the temporal variation of radionuclides (e.g., ^{40}K, ^{214}Bi representing ^{222}Rn, and ^{208}Tl representing ^{220}Rn) in a submarine groundwater discharge (Tsabaris et al. 2012). This allowed for the investigation of the mixing processes between groundwater discharges with seawater. The system has potential to be used in other such radon tracer studies in coastal regions to further our knowledge of the dynamics of seawater mixing processes. Another system called the gamma energy marine spectrometer (GEMS) has been described by Sartini et al. (2011) for measuring the decay of ^{40}K and other natural (e.g., ^{232}T, ^{238}U) and anthropogenic (^{137}Cs) radionuclides. The sensor is very stable and has deployed in the Ionian Sea for 6 months at a depth of 3200 m and was able to provide continuous time-series monitoring data of radioactivity in the deep sea.

72.3.3 Electrochemical Techniques

Electrochemical-based instruments are among the most widely used devices for *in situ* chemical analysis (Denuault 2009; Moore et al. 2009; Kimmel et al. 2012) and include conductometric, potentiometric, and amperometric/voltammetric electrode systems (Taillefert et al. 2000) and can be used for both single and multiple analyte measurements. Conductometric electrodes are used to measure salinity and are incorporated into a variety of commercially available CTD instruments (Table 72.1). Examples of potentiometric devices include those to measure pH, sulfide (e.g., the Ag_2S electrode), and pCO_2. Examples of amperometric devices include those for measuring concentrations of O_2 and N_2O. Electrodes to detect (free) trace metal concentrations so far lack the sensitivity required for use in the marine environment. Furthermore, the analyses are usually affected by changes in salinity.

It is now accepted (Tercier-Waeber et al. 2008) that the development of rugged, submersible none or little perturbing probes, for automatic, *in situ* measurements (i.e., measurements performed inside the water and sediment/soil matrix), is an important research need. Voltammetry is suited to this purpose since it can be used in low-cost, automated, and miniaturized equipment with low-energy requirements. However, significant developments are required to improve conventional voltammetric devices before they can be deployed truly *in situ* rather than being limited to use as onboard analyzers. One commercially available system is the voltammetric *in situ* profiling (VIP) system. The VIP (Figure 72.5) has the capacity for monitoring concentrations of trace metals down to 500 m with subnanomolar sensitivity. This system consists of several units: an online oxygen removal module, a multiparameter probe (Ocean Seven 316, Idronaut-Milan), and a calibration deck unit. Difficulties in calibration and validation are limitations to the future use of *in situ* electrochemical devices. Unlike in the laboratory, in field use, it is difficult to keep the electrodes in pristine condition, and this can affect reproducibility. As the systems measure in real time, it is difficult to calibrate the instrument *in situ* without undertaking a large number of spot water samples for verification. New developments, such as environmentally

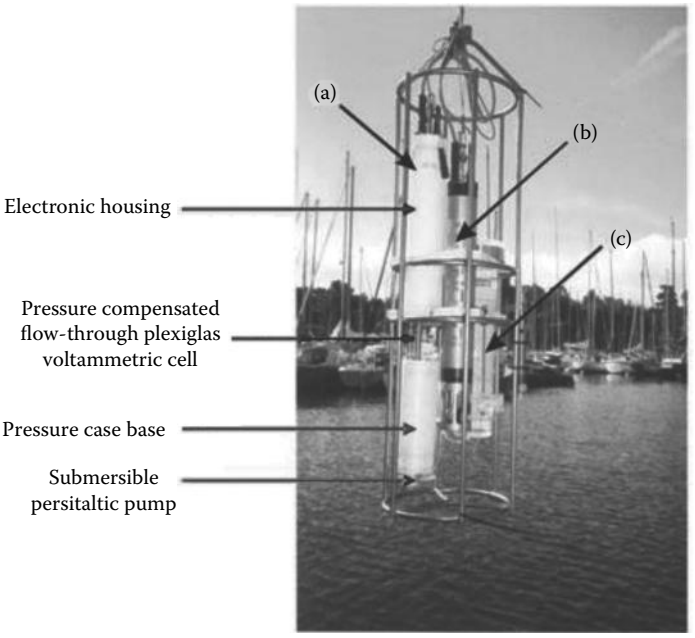


FIGURE 72.5 Standard version of the VIP system for *in situ* monitoring and profiling in freshwater and marine water columns: (a) voltammetric probe model 1, (b) multiparameter probe, and (c) online O_2 removal system. (From Tercier-Waeber and Taillefert, 2008: Reproduced by permission of The Royal Society of Chemistry, http://dx.doi.org/10.1039/B714439N.)

friendly disposable probes with microfabricated bismuth working electrodes for trace metal analysis (Zou et al. 2008), cobalt-based microelectrodes for phosphate (Lee et al. 2009), and screen-printed electrode sensors (Zaouak et al. 2010), are in the laboratory development phase; it will be then a challenge to adapt them for *in situ* operation. Gibbon-Walsh et al. (2011) used a vibrating, gold, microwire electrode (VGME) to detect low nanomolar levels of Mn and Zn in UK coastal waters. The VGME has advantages related to robustness, stability, and ease of use (no polishing, simple regeneration) facilitating on-site and potential *in situ* use. Using such approaches, it should be possible to simultaneously use arrays of sensors, and this will provide the necessary redundancy required for long-term deployments.

72.3.4 Biosensors

Biosensors are devices that integrate a biological recognition element with a transducer that converts changes associated with an interaction between the receptor and the target compound into an output signal that can be measured quantitatively. Several recent reviews (Farré et al. 2009; Roig et al. 2009; Kimmel et al. 2012) have provided useful definitions of biosensors and have differentiated them from biological systems (bioassays, biological early warning), where such an integral transduction mechanism is absent. Biosensors can be classified in a number of ways, usually on the basis of the type of recognition unit or the nature of the transducer. For instance, the biological component can be provided by DNA, enzymes, immunological systems, receptor proteins, and whole cells. The transduction component can be provided by acoustic, chemical, electrochemical (amperometric, conductometric, potentiometric) microbalance and optical (absorbance, bioluminescence, chemiluminescence, total internal reflection, surface plasmon resonance) and piezoelectric mechanisms (Kröger and Law 2005a). The advantages of biosensors, such as high specificity, sensitivity, in some cases biological relevance, and ability to work in a wide range of matrices, make them potentially very useful in environmental monitoring. The use of such devices that operate remotely to make repeated *in situ* measurements that could be transmitted to a laboratory is seen as the ideal way of meeting the legislative requirements for monitoring the quality of environmental waters. Many research, particularly in Europe, has been directed toward this target and continues to be so, for example, in European Union's Framework 7 Work Programme: The Ocean of Tomorrow 2013 call. However, so far there are few commercially available biosensors for use in the environment and even fewer that can be used *in situ* (Kröger and Law 2005b). There is currently a large gap between the research output and the commercial availability of appropriate instruments for environmental applications. In part, this is due to the potentially small market for such devices, the high costs of the development to manufacturing stages and of validation for *in situ* environmental use. This contrasts with the successful commercial development of instruments based on biosensors for use in medical applications (Kimmel et al. 2012).

An important environmental role for biosensors might be in the *in situ* measurement of general water quality (measured as levels of nutrients that can lead to eutrophication, BOD), sublethal effects (e.g., endocrine disruption), or general toxicity of the complex mixtures of toxicants that tend to be present in environmental waters (Rodriguez-Mozaz et al. 2005). Another area where biosensors might be favorably used is the detection of algal toxins that can cause damage to commercial seafood-rearing activities. Here, they could provide alarm systems that would justify the high development costs. The following gives some examples of where biosensors have been used successfully for environmental applications.

The biosensor RIver ANAlyzer (RIANA) and Automated Water Analyzer Computer Supported System (AWACSS) are two optically linked multichannel immunosensor devices. The former has been used in river water and in drinking water treatment plants for the detection of pesticides (Rodriguez-Mozaz et al. 2006). The latter has been used mainly in wastewater treatment plants and can measure up to 32 organic pollutants at low ng L^{-1} concentrations simultaneously, at a rate of 4 h^{-1} in real environmental samples (Tschmelak et al. 2005). One mature biosensor technology (Larsen et al. 2000) is based on the use of immobilized bacteria that convert NO_2^- or NO_3^- to N_2O. The bacteria and a reserve of an electron donor are separated from the sample by an ion-permeable membrane that

allows the substrates to diffuse into the sensor. The bacteria are separated from the electrochemical sensor for N_2O by a gas-permeable membrane. A commercial version of this sensor is available (www.unisense.dk) and can be modified by using different strains of denitrifying bacteria to measure either NO_2^- and NO_3^- or NO_2^-. This technology was originally developed for use in wastewater treatment plants but has been demonstrated to work in seawater. A fluorescence-based fiber-optic biosensor has been developed for determining *in situ* real-time concentrations of trace metals (copper and zinc) in seawater (Zeng et al. 2003; Zheng et al. 2004). The recognition unit is a protein molecule that is site-specifically labeled with a fluorophore that is attached to the distal end of an optical fiber. This binds free copper(II) with a high affinity and selectivity. The transducer reports the metal's concentration as a change in fluorescence intensity or lifetime, using a frequency domain approach. The sensor has a detection limit of 6 pg L^{-1} free copper(II) in seawater. However, these techniques have been mainly utilized in medical applications and are not available commercially. Spier et al. (2011) used a KinExA Inline Sensor (Sapidyne Instruments, United States), containing the monoclonal anti-PAH antibody, 7B2.3, that has specificity for 3–5-ringed PAHs. This device permitted near-real-time measurements in the field. A spatial study was conducted near a dredging site where contaminated sediments were being removed, and a temporal study was performed during a rainfall event. The method was able to determine aqueous PAH concentrations as low as 0.3 μg L^{-1}, within 10 min of sample acquisition, and to assess over 80 samples (not including standards and blanks) in less than 3 days. The field-based biosensor analyses gave a good correlation for the PAHs to conventional laboratory-based analytical methods. Here, the generation of high-resolution, near-real-time data guided management decisions in the field and determined proper sampling protocols for conventional analyses. There has been much concern of the presence of chemicals in the environment that can disrupt normal endocrine function, and work to develop biosensors for measuring endocrine disruptors (of sex and thyroid hormone systems) has been ongoing. Most work, however, is still in the laboratory stage. A biosensor based on a commercially available antibody was developed to detect sub-ng L^{-1} concentrations of testosterone in environmental water samples and was successfully incorporated in to the RIANA array of biosensors (Tschmelak et al. 2006).

With the exception of these few developments, most of the other systems are a long way from in field or *in situ* deployment for monitoring endocrine disruptors in environmental waters. Other recent developments have produced potentially useful materials such as molecularly imprinted polymers (MIPs) for use in biosensors. MIPs mimic the biological activity of antibodies and can be used as recognition units in sensors. They combine some of the desirable properties (specificity) of biological systems but are much more stable chemically, thermally, and mechanically (Rodriguez-Mozaz et al. 2005). MIPs have been incorporated in biomimetic sensors for a range of analytes including haloacetic acids in drinking water, monoamine naphthalenes in drinking water, explosives in natural waters, and soil and copper(II) ions in wastewater. These developments are exciting; however, the technology is still underdeveloped (Ávila et al. 2008), and it will be some time before these sensors become commercially available.

Currently, an array of laboratory-based bioassays is used to assess the overall toxicity of water samples. Biosensors offer the potential to replace the use of bioassays and to provide rapid, automated *in situ* monitoring of water quality. Several strategies have been used to develop biosensor systems that can provide rapid, sensitive measurements of the toxicological quality of environmental waters. One approach is to use an array of whole-cell sensors that respond to a range of toxic insults and signal damage to different cellular systems. One such system was developed by Lee et al. (2005) and comprised a set of 20 recombinant bioluminescent bacteria with different promoters fused with the bacterial *lux* genes. The bacteria were immobilized in a 384-well microplate and on chips, and luminescence was detected using a cooled charge-coupled device camera. The analysis time was 2 h. Different toxicants produced different patterns of response, depending on mode of action, among the strains of bacteria, and this has the potential for identifying the presence of particular types of pollutant in environmental samples. The use of biosensors for the measurement of general toxicity is a desirable goal and would enhance significantly the ability of regulators to monitor water quality in a way that would provide early warnings of problems

and timely information on the results of remediation actions. It would also allow a better regulation of waste discharges than is possible with infrequent spot water sampling and chemical analysis for a limited number of compounds of interest that provides no information on the interactions that can occur between toxicants. However, currently, there are no biosensor technologies available that could be fully applied *in situ* in the marine environment. Another priority area for research is the use of biosensors within the marine environment to detect harmful algae and their respective toxins. A DNA-based electrochemical sensor has been developed by Orozco and Medlin (2011), and this device shows some promise in being able to identify and quantify some toxic algal species in the laboratory-based experiments. Diercks et al. (2008) have also developed an *in situ* device for monitoring toxic algae. This is based on the use of multiprobe chip with an array of 16 gold electrodes and a semiautomated rRNA biosensor. The sandwich hybridization uses species-specific rRNA probes, and this allowed for the simultaneous detection of up to 14 target species of algae. Some devices are used for the screening seafood to detect phycotoxins at adequate sensitivities, but their limited availability, primarily as research tools, hinders their broader application as general environmental monitoring tools (Zielinski et al. 2009). Likewise in the future, biosensors could play an important role in the detection of other waterborne and fecal pathogens and play a key role on monitoring activities such as required by the European Union's Bathing Water Directive (2006/7/EC).

Scholin (2010) discussed the possibilities of developing *in situ* ecogenomic sensor devices that would detect, in real-time, molecular markers indicative of specific organisms, genes, or other biomarkers associated with an environmental context. For example, the instrument could sense in a time-series fashion molecular signatures that are already known, particularly defined sequences of DNA or RNA. The sampler would analyze intact cells and subcellular fractions. Typically, particles in a sample are concentrated and a series of reagents applied sequentially, followed by an extraction routine for further purification. The target molecules can then be detected on the basis of specific physical properties or revealed using a range of intermolecular reactions such antibody/antigen recognition, nucleic acid hybridization, and enzyme-mediated processes (Zehr et al. 2009). Signal transmission and quantification would be achieved by conventional optical, electrochemical, or spectrometric techniques. Researchers at Monterey Bay Aquarium Research Institute (MBARI) (http://www.mbari.org/ESP/default.htm) have developed the MBARI Environmental Sample Processor (ESP), which provides on-site (*in situ*) collection and analysis of water samples from the subsurface ocean. The instrument is an electromechanical/fluidic system designed to collect discrete water samples, concentrate microorganisms or particles, and automate application of molecular probes, which identify microorganisms and their gene products. The ESP also archives samples so that further analyses may be done after the instrument is recovered. The ESP has now been fully commercialized and is marketed as the Spyglass ESP (http://spyglassbio.com/content/products/autonomous-platforms); the system has been fully described by Paul et al. (2007) and Scholin (2010) and has been used on a number of field deployments (Greenfield et al. 2008; Ryan et al. 2011). Such *in situ* instruments offer much potential for the future (e.g., the ability of applying a number of molecular analytical techniques to a single sample); however, there are still a number of practical obstacles to overcome to fully link the upstream sample collection and processing requirements with the currently available molecular assays. Molecular analyses at present are also slow (typically 30 min or longer), and this limits response capability and can place an added drain on power requirements. Applications of such sensors in the future could be the early identification harmful algal bloom species by identifying and measuring specific biomarker toxins, for example, domoic acid and saxitoxin (Doucette et al. 2009; Scholin et al. 2009).

72.4 Conclusions and Future Developments

The importance of a healthy marine environment is widely recognized, and legislation is being extended to provide protection. The need for reliable and widespread monitoring data therefore becomes more urgent. The logistic problems associated with representative sampling over wide areas preclude the use

of strategies based on spot water sampling that were developed for monitoring freshwater. New strategies are needed, and although the task is daunting, some progress has been in developing methods that can collect information *in situ*. A number of sensors for substances such as nitrates and chlorophyll are available commercially from a number of companies, but progress has been slower for pollutants such as heavy metals, many organic compounds, and pathogenic agents. New compounds of interest will also evolve, such as the underwater detection of explosives and their signatures in relation to safety and security issues (Dock et al. 2010).

Many deployment devices (including Argo floats, AUVs, landers, and SmartBuoys) and associated communication systems are already available, but the range of routine measuring devices to use with them is relatively limited. Although outside of the scope of this chapter, significant progress is being made with remote sensing technology primarily through the use of satellites or in some cases airplanes, however, often these observations need to be validated from data ("ground truthing") obtained from *in situ* techniques. The main advantage of remote sensing methods is in spatial scale of measurements that can be achieved, and as such, there are clear benefits in the context of investigating concentrations of chlorophyll-a, climate change, or the extent of oil spill in the ocean, for example (Novoa et al. 2012). Furthermore, there is also some progress in attaching some simple design of sensors to marine fauna to track behavior and movement, and these data can be integrated with other information collected simultaneously by for instance satellites and linking this information to climatic changes. Such an approach also provides insights into physical conditions in the oceans at sites where access by conventional sampling platforms would prove difficult, for example, under Arctic or Antarctic sea ice. In addition, the use of electrochemical sensors that can be integrated into dry and wet suits worn by divers and recreational surfers/swimmers, thereby providing them with the ability to continuously assess their surroundings for environmental contaminants and security hazards, has recently been proposed (Malzahn et al. 2011).

Biofouling of optical and sensor surfaces remains to be solved and is often a neglected area of research by instrument manufacturers, as recognized in the recent EU FP7 call for the development of new antifouling materials. This is a particular challenge in nutrient-rich coastal waters where true operational lifetime can be only 1–2 weeks. Fouling can in some instances have a significant impact on the data obtained from *in situ* sensors, for example, some organisms can alter the local oxygen concentration near a sensing head, and some can exhibit natural fluorescence. The environmental situation in respect of biofouling is complex, and a number of different approaches have been suggested. These include mechanical scrapers or wipers, use of nontoxic (e.g., silicone greases) chemicals, the "uncontrolled" release of a biocide (e.g., leaching of tributyl tin and the use of copper coatings, meshes, and shutters), and "controlled" release methods based on localized seawater electrochlorination devices or automatic acid-dispensing systems (Delauney et al. 2010). Other solutions are being sought in the use of coatings and nanocoatings and those that peel away over a deployment period and biomimetic surfaces (e.g., peptide–peptoid conjugates and phosphorylcholine-based polymers) that were initially developed for medical applications. The use of natural occurring antifouling products from marine species has also been suggested.

Operational lifetimes of remotely deployed instruments are often limited by the available power supplies. Cabled observatories can provide the power to operate sensor networks for extended periods; however, the establishment of the infrastructure is expensive and therefore limited in scope. Many *in situ* sensors still rely on commercially available batteries. The use of renewable energy systems that can utilize naturally occurring chemicals and conditions in the water body is a potential option that holds some promise. These include the use of methane hydrate fuel cells, microbial fuel cells, sea-surface photovoltaic cells, and motion-to-electricity conversion techniques (Gong et al. 2011).

Recent developments in microfabrication, microfluidics, and integrated optics in many areas of analytical chemistry are being applied in the development of *in situ* monitoring devices. Lab-on-a-chip technologies have advantages of a small size and limited reagent and power requirements. The challenge is to ensure that these systems can attain the sensitivity needed for many marine applications where analytes are present at only trace concentrations. However, as the overall cost of the sensing system is

reduced, this would potentially enable the deployment of larger numbers of devices and thereby improve the spatial and temporal resolution and extent of offshore monitoring activities.

The proper calibration of sensors often still remains a critical issue. In addition, there is currently a lack of recognized verification and validation protocols and quality assurance and quality control procedures for many of the devices described previously. Their full validation in the field may in the future require, for example, the deployment of autosamplers alongside sensors so that reference samples over time could be collected. The future commercial development and wide acceptance of *in situ* technologies will depend on the availability of systems such as those available for laboratory-based analytical procedures. There is a need for innovations in the development of suitable reference materials and interlaboratory trials to fully validate any new sensor technologies. International collaborations and the development of specific scientific working parties and associated groups could help to address some of these accreditation, certification, and standardization issues (Waldmann et al. 2010). In addition, the need to continuously calibrate and recalibrate instruments over long-term deployments still remains an issue to be fully solved (Tan et al. 2010).

Most of the devices with the potential for use in the *in situ* monitoring of marine waters were developed for use in the laboratory where they could serve as less expensive, more rapid, and more sensitive replacements for conventional analytical methods. They could enable frequent screening of water quality to avoid the problems associated with infrequent regulatory sampling. The more expensive regulatory measurements could then be used only when required. However, much development and validation work is necessary to reach this goal. The long-term aim of using these technologies for *in situ* monitoring in the marine environment is even further away. Although many devices have been developed to proof of concept or prototype stage, the development work to convert these into robust, reliable, validated instruments that can be deployed in the marine environment for prolonged periods (weeks to months) has not been undertaken. This may be a consequence of the short-term nature of research funding and the failure to provide associated support for the development work necessary to take advantage of exciting research innovations. There have been many lost opportunities by failure of commercial manufacturers to capitalize on significant research developments.

There is little doubt, however, that in the future *in situ* sensors will play a significant role in ecological risk assessments, particularly those associated with larger spatial scales. If such activities are to be cost-effective, then they will require the continued application and development of sensors to cover (both spatially and temporally) an assessment of a wide range of stressors and must also include metadata storage and associated analytical capabilities. This will allow for improved data collection and permit more appropriate decision making on various marine hazards among the key environmental stake holders with the long-term aim of further protecting and restoring our natural habitats.

References

Aminot, A., Kérouel, R., and Birot, D. 2001. A flow injection-fluorometric method for the determination of ammonium in fresh and saline waters with a view to *in situ* analyses. *Water Res.* 35: 1777–1785.

Ávila, M., Zougagh, M., and Ríos, Á. 2008. Molecularly imprinted polymers for selective piezoelectric sensing of small molecules. *TrAC Trends Anal. Chem*. 27: 54–65.

Beaton, A.D., Sieben, V.J., Floquet, C.F.A., Waugh, E.M., Bey, S.A.K., Ogilvie, I.R.G. et al. 2011. An automated microfluidic colourimetric sensor applied *in situ* to determine nitrite concentration. *Sens. Actuat. B: Chem*. 156: 1009–1014.

Bellerby, R.G.J., Olsen, A., Johannessen, T., and Croot, P. 2002. A high precision spectrophotometric method for on-line ship board seawater pH measurements: The automated marine pH sensor (AMpS). *Talanta* 56: 61–69.

Boulart, C., Connelly, D.P., and Mowlem, M.C. 2010. Sensors and technologies for in situ dissolved methane measurements and their evaluation using technology readiness levels. *TrAC Trends Anal. Chem*. 29: 186–195.

Brewer, P.G., Malby, G., Pasteris, J.D., White, S.N., Peltzer, E.T., Wopenka, B., Freeman, J., and Brown, M.O. 2004. Development of a laser Raman spectrometer for deep-ocean science. *Deep-Sea Res. I* 51: 739–753.

van den Broeke, J., Langergraber, G., and Weingarter, A. 2006. Real-time detection of possible harmful events using UV/vis spectrometry. *Spect. Eur*. 18: 1–4.

Cai, W.J. and Reimers, C.E. 2000. Sensors for in situ pH and pCO_2 measurements in seawater and at the sediment-water interface. *In situ Monitoring of Aquatic Systems: Chemical Analysis and Speciation*, eds. J. Buffle and G. Horvai, Chichester, U.K.: Wiley, pp. 75–119.

Camilli, R. and Duryea, A.N. 2009. Characterizing spatial and temporal variability of dissolved gases in aquatic environments with in situ mass spectrometry. *Environ. Sci. Technol*. 43: 5014–5021.

Camilli, R. and Hemond, H.F. 2004. NEREUS/Kemonaut, a mobile autonomous underwater mass spectrometer. *TrAC Trends Anal. Chem*. 23: 307–313.

Campbell, L., Walpert, J.N., and Guinasso, N.L. 2008. A new buoy-based in situ optical early warning system for harmful algal blooms in the Gulf of Mexico. *Nova Hedwigia* 133: 161–170.

Clark, C.M., Hancke, K., Xydes, A., Hall, K., Schreiber, F., Klemme, J. et al. 2012. Estimation of volumetric oxygen concentration in a marine environment with an autonomous underwater vehicle. *J. Field Robotics* March 29, DOI: 10.1002/rob.21421.

Daniel, A., Birot, D., Lehaitre, M., and Poncin, J. 1995. Characterization and reduction of interferences in flow-injection analysis for the in situ determination of nitrate and nitrite in seawater. *Anal. Chim. Acta* 308: 413–424.

DeGrandpre, M.D., Hammar, T.R., Wallace, D.W.R., and Wirick, C.D. 1997. Simultaneous mooring-based measurements of seawater CO_2 and O_2 off Cape Hatteras, North Carolina. *Limnol. Oceanog*. 42: 21–28.

Delauney, L., Compère, C., and Lehaitre, M. 2010. Biofouling protection for marine environmental sensors. *Ocean Sci*. 6: 503–511.

Denuault, G. 2009. Electrochemical techniques and sensors for ocean research. *Ocean Sci*. 5: 697–710.

Dickson, A.G., Sabine, C.L., and Christian, J.R. (Eds.) 2007. *Guide to Best Practice for Ocean CO_2 Measurements*, Vol. 3. Sidney, British Columbia, Canada: PICES Special Publication, 191pp.

Diercks, S., Metfies, K., and Medlin, L.K. 2008. Development and adaptation of a multiprobe biosensor for the use in a semi-automated device for the detection of toxic algae. *Biosens. Bioelectron*. 23: 1527–1533.

Dock, M.L., Harper, R.J., and Knobbe, E. 2010. Combined pre-concentration and real-time in-situ chemical detection of explosives in the marine environment. In *Ocean Sensing and Monitoring II*, eds. W. Hou and R.A. Arnone, Vol. 7678. Proc. SPIE-The International Society for Optical Engineering.

Doucette, G.J., Mikulski, C.M., Jones, K.L., King, K.L., Greenfield, D.I., Marin III, R. et al. 2009. Remote, subsurface detection of the algal toxin domoic acid onboard the environmental sample processor: Assay development and initial field trials. *Harmful Algae* 8: 880–888.

Farré, M., Kantiani, L., Pérez, S., and Barceló, D. 2009. Sensors and biosensors in support of EU directives. *TrAC Trends Analyt. Chem*. 28: 170–185.

Fischer, M., Wahl, M., and Friedrichs, G. 2012. Design and field application of a UV-LED based optical fiber biofilm sensor. *Biosens. Bioelectron*. 33: 172–178.

Floch, J., Blain, S., Birot, D., and Treguer, P. 1998. In situ determination of silicic acid in seawater based on FIA and colorimetric dual-wavelength measurements. *Anal. Chim. Acta* 377: 157–166.

Gibbon-Walsh, K., Salaun, P., and van den Berg, C.M.G. 2011. Determination of manganese and zinc in coastal waters by anodic stripping voltammetry with a vibrating gold microwire electrode. *Environ. Chem*. 8: 475–484.

Glud, R.N., Gundersen, J.K., and Ramsing, N.B. 2000. Electrochemical and optical oxygen microsensors for in situ measurements. In: *In situ Monitoring of Aquatic Systems: Chemical Analysis and Speciation*, eds. J. Buffle and G. Horvai, pp. 19–73. Chichester, U.K.: Wiley.

Gong, Y.M., Radachowsky, S.E., Wolf, M., Nielsen, M.E., Girguis, P.R., and Reimers, C.E. 2011. Benthic microbial fuel cell as direct power source for an acoustic modem and seawater oxygen/temperature sensor system. *Environ. Sci. Technol*. 45: 5047–5053.

Greenfield, D.I., Marin, R., Doucette, G.J., Mikulski, C., Jones, K., Jensen, S. et al. 2008. Field applications of the second-generation environmental sample processor (ESP) for remote detection of harmful algae: 2006–2007. *Limnol. Oceanog. Methods* 6: 667–679.

Greenwood, N., Parker, E.R., Fernand, L., Sivyer, D.B., Weston, K., Painting, S.J. et al. 2010. Detection of low bottom water oxygen concentrations in the North Sea: Implications for monitoring and assessment of ecosystem health. *Biogeosciences* 7: 1357–1373.

Hardman-Mountford, N.J., Moore, G., Bakker, D.C.E., Watson, A.J., Schuster, U., Barciela, R. et al. 2008. An operational monitoring system to provide indicators of CO_2-related variables in the ocean. *ICES J. Marine Sci*. 65: 1498–1503.

Hydes, D.J., Hartman, M.C., Kaiser, J., and Campbell, J.M. 2009. Measurement of dissolved oxygen using optodes in a FerryBox system. *Estuarine Coastal Shelf Sci*. 83: 485–490.

Johnson, K.S., Coletti, L.J., and Chavez, F.P. 2006. Dielnitrate cycles observed with in situ sensors predict monthly and annual new production. *Deep-Sea Res. I* 53: 561–573.

Kimmel, D.W., LeBlanc, G., Meschievitz, M.E., and Cliffel, D.E. 2012. Electrochemical sensors and biosensors. *Anal. Chem*. 84: 685–707.

Körtzinger, A., Schimanski, J., Send, U., and Wallace, D. 2004. The ocean takes a deep breath. *Science* 306: 1337.

Kraft, M. and Mizaikoff, B. 2000. Amid-infrared sensor for monitoring of chlorinated hydrocarbons in the marine environment. *Int. J. Environ. Anal. Chem*. 78: 367–383.

Kröger, S. and Law, R.J. 2005a. Sensing the sea. *Trends Biotechnol*. 23: 250–256.

Kröger, S. and Law, R.J. 2005b. Biosensors for marine applications. We all need the sea, but does the sea need biosensors? *Biosens. Bioelectron*. 20: 1903–1913.

Kröger, S., Parker, E.R., Metcalfe, J.D., Greenwood, N., Forster, R.M., Sivyer, D.B., and Pearce, D.J. 2009. Sensors for observing ecosystem status. *Ocean Sci*. 5: 523–535.

Kromkamp, J.C., Dijkman, N.A., Peene, J., Simis, S.G.H., and Gons, H.J. 2008. Estimating phytoplankton primary production in Lake IJsselmeer (The Netherlands) using variable fluorescence (PAM-FRRF) and C-uptake techniques. *Eur. J. Phycol*. 43: 327–344.

Lacombe, M., Garçon, V., Thouron, D., Le Bris, N., and Comtat, M. 2008. Silicate electrochemical measurements in seawater: Chemical and analytical aspects towards a reagentless sensor. *Talanta* 77: 744–750.

Larsen, L.H., Damgaard, L.R., Kjær, T., Stenstrøm, T., Lynggaard-Jensen, A., and Revsbech, N.P. 2000. Fast responding biosensor for on-line determination of nitrate/nitrite inactivated sludge. *Water Res*. 34: 2463–2468.

Le Bris, N., Sarradin, P.-M., Birot, D., and Alayse-Danet, A.-M. 2000. A new chemical analyzer for in situ measurement of nitrate and total sulfide over hydrothermal vent biological communities. *Marine Chem*. 72: 1–15.

Lee, J.H., Mitchell, R.J., Kim, B.C., Cullen, D.C., and Gu, M.B. 2005. A cell array biosensor for environmental toxicity analysis. *Biosens. Bioelectron*. 21: 500–507.

Lee, W.H., Seo, Y., and Bishop, P.L. 2009. Characteristics of a cobalt-based phosphate microelectrode for *in situ* monitoring of phosphate and its biological application. *Sens. Actuat. B: Chem*. 137: 121–128.

Malzahn, K., Windmiller, J.R., Valdes-Ramirez, G., Schoning, M.J., and Wang, J. 2011. Wearable electrochemical sensors for in situ analysis in marine environments. *Analyst* 136: 2912.

Meyer, D., Prien, R.D., Dellwig, O., Connelly, D.P., and Schulz-Bull, D.E. 2012. In situ determination of iron(II) in the anoxic zone of the central Baltic Sea using ferene as spectrophotometric reagent. *Marine Chem*. 130: 21–27.

Michel, A.P., Lawrence-Snyder, M., Angel, S.M., and Chave, A.D. 2007. Laser-induced breakdown spectroscopy of bulk aqueous solutions at oceanic pressures: Evaluation of key measurement parameters. *Appl. Opt*. 46: 2507–2515.

Mizaikoff, B. 2003. Mid-IR fiber-optic sensors. *Anal. Chem*. 75: 258A–267A.

Moore, C., Barnard, A., Fietzek, P., Lewis, M.R., Sosik, H.M., White, S., and Zielinski, O. 2009a. Optical tools for ocean monitoring and research. *Ocean Sci*. 5: 661–684.

Moore, T.S., Mullaugh, K.M., Holyoke, R.R., Madison, A.S., Yücel, M., and Luther III, G.W. 2009b. Marine chemical technology and sensors for marine waters: Potentials and limits. *Annu. Rev. Marine Sci.* 1: 91–115.

Novoa, S., Chust, G., Sagarminaga, Y., Revilla, M., Borja, A., and Franco, J. 2012. Water quality assessment using satellite-derived chlorophyll-a within the European directives, in the southeastern Bay of Biscay. *Marine Pollut. Bull.* 64: 739–750.

Orozco, J. and Medline, L.K. 2011. Electrochemical performance of a DNA-based sensor device for detecting toxic algae. *Sens. Actuat. B: Chem.* 153: 71–77.

Owens, W.B. and Wong, A.P.S. 2009. An improved calibration method for the drift of the conductivity sensor on autonomous CTD profiling floats by θ–Sclimatology. *Deep-Sea Res. I* 56: 450–457.

Paul, J., Scholin, C., van den Engh, G., and Perry, M.J. 2007. In situ instrumentation. *Oceanography* 20: 70–78.

Pfannkuche, J., Lubecki, L., Schmidt, H., Kowalewska, G., and Kronfeldt, H.-D. 2012. The use of surface-enhanced Raman scattering (SERS) for detection of PAHs in the Gulf of Gdańsk (Baltic Sea). *Marine Pollut. Bull.* 64: 614–626.

Prien, R.D. 2007. The future of chemical in situ sensors. *Marine Chem.* 107: 422–432.

Queste, B.Y., Fernand, L., Jickells, T.D., and Heywood, K.J. 2012. Spatial extent and historical context of North Sea oxygen depletion in August 2010. *Biogeochemistry*, DOI: 10.1007/s10533-012-9729-9.

Reimers, C.E. 2007. Applications of microelectrodes to problems in chemical oceanography. *Chem. Rev.* 107: 590–600.

Rodriguez-Mozaz, S., López de Alda, M., and Barceló, D. 2006. Fast and simultaneous monitoring of organic pollutants in a drinking water treatment plant by a multi-analyte biosensor followed by LC–MS validation. *Talanta* 69: 377–384.

Rodriguez-Mozaz, S., López de Alda, M., Marco, M.-P., and Barceló, D. 2005. Biosensors for environmental monitoring: A global perspective. *Talanta* 65: 291–297.

Roig, B., Bazin, I., Bayle, S., Habauzit, D., and Chopineau, J. 2009. Biomolecular recognition systems for water monitoring. In: *Rapid Chemical and Biological Techniques for Water Monitoring*, eds. C. Gonzalez, R. Greenwood, and P. Quevauviller, Chichester, U.K.: Wiley, pp. 175–196.

Ryan, J., Greenfield, D., Marin, R., Preston, C., Roman, B., Jensen, S. et al. 2011. Harmful phytoplankton ecology studies using an autonomous molecular analytical and ocean observing network. *Limnol. Oceanog.* 56: 1255–1272.

Sakamoto-Arnold, C.M., Johnson, K.S., and Beecher, C.L. 1986. Determination of hydrogen sulfide in seawater using flow injection analysis and flow analysis. *Limnol. Oceanogr.* 31: 894–900.

Sartini, L., Simeone, F., Pani, P., Lo Bue, N., Marinaro, G., Grubich, A. et al. 2011. GEMS: Underwater spectrometer for long-term radioactivity measurements. *Nucl. Instrum. Methods Phys. Res. Sect. A: Accelerat., Spectrom., Detect. Associat. Equip.* 626–627: S145–S147.

Schlüter, M. and Gentz, T. 2008. Application of membrane inlet mass spectrometry for online and in situ analysis of methane in aquatic environments. *J. Am. Soc. Mass Spectrom.* 19: 1395–1402.

Schmidt, H., Ha, N.B., Pfannkuche, J., Amann, H., Kronfeldt, H.-D., and Kowalewska, G. 2004. Detection of PAHs in seawater using surface-enhanced Raman scattering(SERS). *Marine Pollut. Bull.* 29: 229–234.

Scholin, C. 2010. What are "ecogenomic sensors?" A review and thoughts for the future. *Ocean Sci.* 6: 51–60.

Scholin, C., Doucette, G., Jensen, S., Roman, B., Pargett, D., Marin III, R. et al. 2009. Remote detection of marine microbes, small invertebrates, harmful algae and biotoxins using the environmental sample processor (ESP). *Oceanography* 22: 158–167.

Short, R.T., Fries, D.P., Kerr, M.L., Lembke, C.E., Toler, S.K., Wenner, P.G., and Byrne, R.H. 2001. Underwater mass spectrometers for in situ chemical analysis of the hydrosphere. *J. Am. Soc. Spectrom.* 12: 676–682.

Short, R.T., Toler, S.K., Kibelka, G.P.G., Rueda Roa, D.T., Bell, R.J., and Byrne, R.H. 2006. Detection and quantification of chemical plumes using a portable underwater membrane introduction mass spectrometer. *TrAC Trends Anal. Chem.* 25: 637–646.

Sieben, V.J., Floquet, C.F.A., Ogilvie, I.R.G., Mowlem, M.C., and Morgan, H. 2010. Microfluidic colourimetric chemical analysis system: Application to nitrite detection. *Anal. Methods* 2: 484–491.

Spier, C.R., Vadas, G.G., Kaattari, S.L., and Unger, M.A. 2011. Near real-time, on-site, quantitative analysis of PAHs in the aqueous environment using an antibody-based biosensor. *Environ. Toxicol. Chem.* 30: 1557–1563.

Statham, P.J., Connelly, D.P., German, C.R., Brand, T., Overnell, J.O., Bulukin, E. et al. 2005. Spatially complex distribution of dissolved manganese in a fjord as revealed by high-resolution in situ sensing using the autonomous underwater vehicle autosub. *Environ. Sci. Technol.* 39: 9440–9445.

Statham, P.J., Connelly, D.P., German, C.R., Bulukin, E., Millard, N., McPhail, S. et al. 2003. Mapping the 3D spatial distribution of dissolved manganese in coastal waters using an in situ analyser and the autonomous underwater vehicle Autosub. *Underwater Technol.* 25: 129–134.

Taillefert, M., Luther, G.W., and Nuzzio, D.B. 2000. The application of electrochemical tools for in situ measurements in aquatic systems. *Electroanalysis* 12: 401–412.

Tan, C.Y., Jin, B., Ding, K., Seyfried, W.E., and Chen, Y. 2010. A long-term in situ calibration system for chemistry analysis of seawater. *J. Zhejiang Univ.-Sci. A* 11: 701–708.

Tedetti, M., Guigue, C., and Goutx, M. 2010. Utilization of a submersible UV fluorometer for monitoring anthropogenic inputs in the Mediterranean coastal waters. *Marine Pollut. Bull.* 60: 350–362.

Tengberg, A., Hovdenes, J., Andersson, H.J., Brocandel, O., Diaz, R., Hebert, D. et al. 2006. Evaluation of a lifetime-based optode to measure oxygen in aquatic systems. *Limnol. Oceanog. Methods* 4: 7–17.

Tercier-Waeber, M.L., Confalonieri, F., Koudelka-Hep, M., Dessureault-Rompre, J., Graziottin, F., and Buffle, J. 2008. Gel-integrated voltammetric microsensors and submersible probes as reliable tools for environmental trace metal analysis and speciation. *Electroanalysis* 20: 240–258.

Thouron, D., Vuillemin, R., Philippon, X., Lourenço, A., Provost, C., Cruzado, A., and Garçon, V. 2003. An autonomous nutrient analyzer for oceanic long-term in situ biogeochemical monitoring. *Anal. Chem*. 75: 2601–2609.

Tsabaris, C., Bagatelas, C., Dakladas, T., Papadopoulos, C.T., Vlastou, R., and Chronis, G.T. 2008. An autonomous in situ detection system for radioactivity measurements in the marine environment. *Appl. Radiat. Isotopes* 66: 1419–1426.

Tsabaris, C., Patiris, D.L., Karageorgis, A.P., Eleftheriou, G., Papadopoulos, V.P., Georgopoulos, D. et al. 2012. In-situ radionuclide characterization of a submarine groundwater discharge site at Kalogria Bay, Stoupa, Greece. *J. Environ. Radioactivity* 108: 50–59.

Tschmelak, J., Proll, G., Riedt, J., Kaiser, J., Kraemmer, P., Bárzaga, L., Wilkinson, J.S., Hua, P., Hole, J.P., Nudd, R., Jackson, M., Abuknesha, R., Barceló, D., Rodriguez-Mozaz, S., de Alda, M.J., Sacher, F., Stien, J., Slobodník, J., Oswald, P., Kozmenko, H., Korenková, E., Tóthová, L., Krascsenits, Z., Gauglitz, G. 2005. Automated water analyser computer supported system (AWACSS) Part I: Project objectives, basic technology, immunoassay development, software design and networking. *Biosens Bioelectron*. 20: 1499–1508.

Tschmelak, J., Kumpf, M., Käppel, N., Proll, G., and Gauglitz, G. 2006. Total internal reflectance fluorescence (TIRF) biosensor for environmental monitoring of testosterone with commercially available immunochemistry: Antibody characterization, assay development and real sample measurements. *Talanta* 69: 343–350.

Vuillemin, R., Le Roux, D., Dorval, P., Bucas, K., Sudreau, J.P., Hamon, M., Le Gall, C., and Sarradin, P.M. 2009. CHEMINI: A new *in situ* CHEmical MINIaturized analyzer. *Deep-Sea Res. I* 56: 1391–1399.

Waldmann, C., Tamburri, M., Prien, R.D., and Fietzek, P. 2010. Assessment of sensor performance. *Ocean Sci*. 6: 235–245.

Wangersky, P.J. 2005. Methods of sampling and analysis and our concepts of ocean dynamics. *Scie. Mar.* 69: 75–84.

Wedekind, C., Schilling, G., Grüttmüller, M., and Becker, K. 1999. Gamma-radiation monitoring network at sea. *Appl. Radiat. Isotopes* 50: 733–741.

Zaouak, O., Authier, L., Cugnet, C., Normandin, E., Champier, D., Rivaletto, M., and Potin-Gautier, M. 2010. Electroanalytical device for cadmium speciation in waters. Part 1: Development and characterization of a reliable screen-printed sensor. *Electroanalysis* 22: 1159–1165.

Zehr, J.P., Hewson, I., and Moisander, P. 2009. Molecular biology techniques and applications for ocean sensing. *Ocean Sci*. 5: 101–113.

Zeng, H.H., Thompson, R.B., Maliwal, B.P., Fones, G.R., Moffett, J.W., and Fierke, C.A. 2003. Real-time determination of picomolar free Cu(II) in seawater using a fluorescence based fiber optic biosensor. *Anal. Chem*. 75: 6807–6812.

Zheng, H.H., Thompson, R.B., and Fierke, C.A. 2004. Real time determination of free Zn(II) and Cu(II) in situ using fluorescence-based fiberoptic biosensors. *Biophys. J*. 86: 606A–607A.

Zielinski, O., Busch, J.A., Cembella, A.D., Daly, K.L., Engelbrektsson, J., Hannides, A.K., and Schmidt, H. 2009. Detecting marine hazardous substance and organisms: Sensors for pollutants, toxins and pathogens. *Ocean Sci*. 5: 329–349.

Zou, Z.W., Jang, A., MacKnight, E., Wu, P.M., Do, J., Bishop, P.L., and Ahn, C.H. 2008. Environmentally friendly disposable sensors with microfabricated on-chip planar bismuth electrode for *in situ* heavy metal ions measurement. *Sensors. Actuat. B: Chem*. 134: 18–24.

73

Mobile Environmental Sensing Using Smartphones

Siamak Aram
Politecnico di Torino

A. Toriono
Politecnico di Torino

F. Rugiano
Politecnico di Torino

E. Pasero
Politecnico di Torino

73.1 Introduction

In recent years, the tendency of people to share and record data via the Internet has grown increasingly. The same propensity has involved researchers to analyze distributed data. The explosion of information via blogs and social networks is an actuality that encourages lots of scientists and researchers to focus on this occurrence. New styles of life have appeared according to the advent of this phenomenon that affects business, commerce, government and politics, public discourse, health and medical, etc. This also made prominent changes on related technologies.

These decades, environmental information still remains one of the most interest areas for scientists because of its effect on people's real life. Environmental effects, changes, influences, and interactions in our life are important for the future. Consideration of these elements is valuable to have an applied definition of environmental sensing. However, understanding requires information that not only makes the complexities of studying the environment easier but also implies the limitations of existing related knowledge and uncertainties (Lein 2012).

The goal of this chapter is to have a general and practical definition of *mobile environmental sensing* that enables gathering and analyzing data to get environmental information. On the other hand, wireless sensor networks help to gather temperature, humidity, pressure, solar radiation, and so on from the environmental information. This chapter looks at the potency of *smartphones* that can be applied as sensors to remotely gather environmental data. Also, a new capacity of this technology is developing, in which persons or groups of people are involved in the data-gathering process. This characteristic called *participatory sensing* enables us to sense the environment's parameters using mobile phones. There are some sensors that are built into mobile phones that help to sense and interpret the condition of the environment.

Generally, using any ability such as sight, hearing, smell, taste, or touch from the outside of the body (Lein 2012) is an ordinary definition of the word *sensing*. In general, detecting or perceiving a

phenomenon by perceiving the human's external objects is identified as "sensing," and "remote sensing" technology is a simulated aptitude to detect and sense phenomena. Hussein (2006) defined it as follows:

> Collection of information about an object without coming into physical contact. Remote sensing is a technique to gather information. It is the quantitative analysis of digital information.

Also, Bureau for Africa's Office of Sustainable Development (2007) indicated that *remote sensing* is

> The science of obtaining and interpreting information from a distance, using sensors that are not in physical contact with the object being observed.

Remotely data gathering provides information that is not easily accessible. Time is also another important parameter that requires remote sensing systems to analyze data at the required intervals. Different systems are applied in different researches and projects to sense remotely. Unreachable information can become more attainable (but not completely) thanks to the aid of systems that make sensing the environments possible. The method or the way of sensing is different from its accuracy. On the other hand, more sensing means more information but not necessarily precise information. This leads to try to improve the process of sensing and also the accuracy.

For predicting, analyzing, and making decision on the environmental condition, *environmental remote sensing* can be defined. Since the environmental information itself and the process of data gathering are involved with uncertainty and complexity, remotely accessing this information is directed by different applied techniques to overcome subsequent problems. As an extension of these techniques, environmental remote sensing applies different approaches, methods, and technologies to obtain related information and also to reach the accuracy of gained data. Using mobile phones is a remote sensing approach to gather environmental information that allows the systems to remotely access data with one of the most popular and used devices (especially for small distances). For instance, Aram et al. (2012b) introduced a method by applying smartphone to acquire environmental information in small spaces. They showed that getting environmental information while moving can be easier by having warning.

73.2 Mobile Phone as a Sensor

The challenge of power consumption still remains from the first mobile phones, in the 1980s, which were not light, small, and smart. Mobile phones as a part of people's lives allow them to connect to each other via text and voice. More services such as the possibility of checking e-mail, playing games, and sending and receiving files were included in devices called personal digital assistants (PDAs). Smartphones appeared as a combination of primitive mobile phones and the PDAs' abilities and added more facilities (Zheng and Ni 2006). Table 73.1 shows the evolution of cell phones. This compares three different generations of cell phones. Size, shape, equipment, and power consumption have changed with

TABLE 73.1 Growth of Mobile Phones

Type	Analog	Digital	Smartphone
Weight	About 900 g	About 200 g	Less than 140 g
Processor	Simple tasks	Preliminary tasks	Advanced tasks such as multimedia, Internet communication, playback
Memory	Only for storing some setting and phone numbers	For storing more data in the range of megabytes	More than 32 MB with possibility to have extra memory
Bluetooth	—	A few of them	Yes
Battery	For talking short time and for standby time	For talking long time and for standby time	For talking in longer time and for standby time

each generation. There is technological evolution in different fields since the first generation of mobile phones. This fast growth was ongoing from the very low-end mobile phone category (Paananen 2011) to reach the smartphones that make it possible to perform a wide range of operations. Several attributes make these types of phones inseparable in daily life and also make them as one of most-used electronic devices (Firtman 2010):

- Multitasking operating systems
- Running a myriad of applications
- Having several features (such as Bluetooth)

Now, they are involved in various applications:

- Health and medical (Wei et al. 2005)
- Military
- Environmental
- Home and office (Kim and Paulos 2010)
- Commercial

In fact, this new generations' potential to compute, communicate, analyze, and monitor has many effects on the quality of life.

Now, the resultant combination is used in various areas such as business, health care, social networks, environmental monitoring, safety, and transport. These equipped devices enable different applications with the consideration of different domains according to their operating systems (Sugiharto 2010). This new mobile generation, especially smartphones and tablets, is being used to monitor quality of life parameters. They apply the following sensors:

- Accelerometers
- Compasses
- GPSs
- Gyroscope
- Microphones
- Cameras

These sensors are already embedded (Lane et al. 2010), but there are no sensors for air quality or the pollutants of the environment. These sensors are part of the mobile phones that can apply the internal application programming interfaces (APIs). Some of these embedded sensors are recognized by everybody such as the microphone, the camera, and the GPS module. There are also external sensors that are used besides the embedded sensors; two of them are also used as a prototype in this chapter. Generally, wireless sensor networks use sensors as preplanned sensors or equipped on carriers or vehicles, with immobile routing strategies; alternatively, nowadays, mobile devices that are working as data collectors, especially smartphones, can play a role as distributed movable sensors. Also, as smart devices, they can share information with groups or communities as an important parameter to aid for making decisions for governments, groups of people, and researchers in different communities (Staples 2011). Together, the computing, communication, and sensing, as potencies of the smartphones, make *participatory* or *opportunistic* operations possible (Das et al. 2010). For mobility, two types of sensors can be considered; the first type is *wearable sensors* that people can wear (Choudhury et al. 2008); the second is called *phone to web* (Mun et al. 2009), where it can also be called *phone sensors*. Some mobile devices such as PDAs and cell phones can connect to devices and transfer data via Bluetooth. In some projects, they apply storage such as memory cards to store data instead of transferring it online. On the other hand, they send the gathered data to other devices or a central station, either periodically or continuously. The second type applies mobiles to record information from the environment. Mobile phones can record required data and then send it in a short time to a certain station or upload them via the Internet to the web portal (Goldman et al. 2009). Users can gather the environmental information in different

situations such as walking, biking, driving, and running; also, the location and tagged (i.e., custom) data can be included. But it is not possible to apply camera, send text messages, or tag more information in these moving situations. Consequently, mobile phones and the web can communicate to aid to improve the quality of lifestyle and to help elders, their family, and doctors in the case of medical monitoring (Goldman et al. 2009).

There are different scenarios for monitoring the environment. These scenarios apply wireless sensor networks that can be followed by *personals*, *groups*, and *communities* (Lane et al. 2010) to have environmental information. Today, in doing so, the simplest way is to employ the mobile devices. Environmental sensors, such as temperature, humidity, solar radiation, and pressure, can monitor both outdoor and indoor sites. But usually the short distance data transferring is one of these types of networks' characteristics. The two important parameters that should be considered for designing any related networks' system are the low power and low cost. In general, four parameters have more effect on the sensor structures and consequently their platforms:

- Deployment (based on activities)
- Location (indoor or outdoor)
- The application
- Data

Sensing environmental information parameters such as temperature and humidity are useful parameters to control and monitor in different fields such as medical, social services, and agriculture. Mobile robot temperature sensing is a sample solution for temperature monitoring that is applied in airports and hospitals (Abdullah and Poh 2011). This robot can also apply to other applications such as heat temperature detection for firefighting. Also, there are wireless sensor networks' solutions with immobile distributed sensors (Kuchta et al. 2008; Rosi et al. 2010). Monitoring the environment of small places is possible via mobile phones instead of preplanned or robot sensor systems (Shan and Brown 2005).

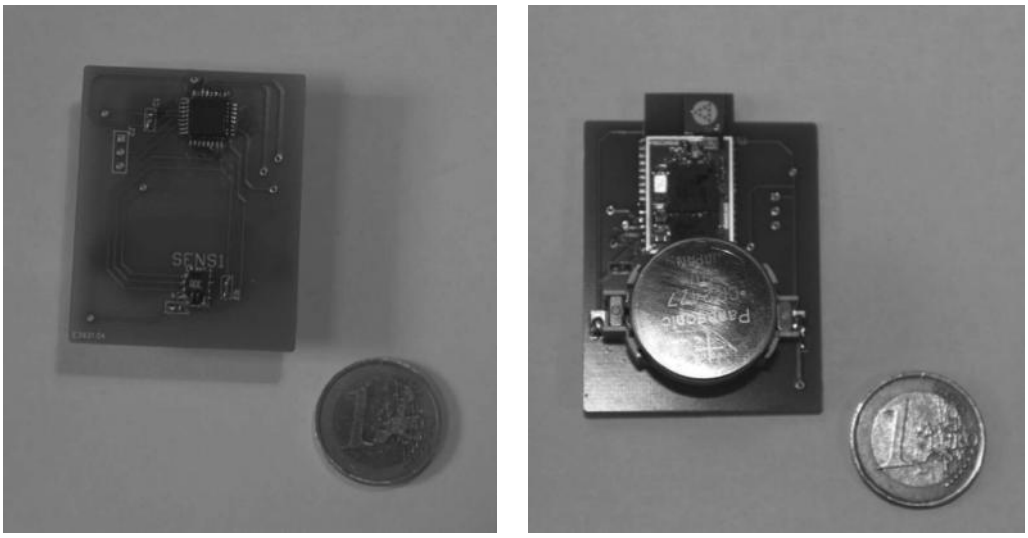
In this chapter, the possibility of applying smartphones for gathering data from other phones or sensors is presented. These days, temperature and humidity are two important parameters of monitoring climate condition to sense the changes of the environment in living or working places for the human. In doing so, distributed devices in different environments with high-resolution sensors and a wireless transmission apparatus for transferring data to smartphones can be a novel solution. Bluetooth (since it is embedded in almost all smartphones) is applied as a transmission tool that can work in the absence of a Wi-Fi connection. Smartphones easily communicate with other devices using the programmable tools that enable different kinds of applications. They can gather, analyze, and verify data. A prototype is proposed by Aram et al. (2012a) and improved by Aram et al. (2012b) by applying a Bluetooth-based sensor to sense *temperature* and *humidity* for the monitoring of environmental conditions using the android-based smartphone.

73.3 Prototype

73.3.1 Methods

73.3.1.1 First Version of the Bluetooth Sensor

The goal of the first version of the Bluetooth sensor was to demonstrate the feasibility of the architecture and to provide a device that is easy to interface to; this way showed how to have a simple data stream to develop the "client side" application more easily. The sensor is described by Aram et al. (2012a). It consists of a temperature and humidity sensor, a microcontroller, and a Bluetooth module. Figure 73.1 shows the first version of the sensor. The sensor used for this device is a Sensirion SHT11; it was chosen for its low-power consumption (about 80 μW) and its high accuracy (0.4°C for the temperature and 3% for the humidity, ranging from −40 °C to 125 °C and from 0% to 100%, respectively). It communicates with the microcontroller through a 2-wire serial interface. The microcontroller is a Silicon Labs C8051F314. It was chosen because its internal characteristics fit the requirements of the project; this microcontroller



(a) (b)

FIGURE 73.1 The electronic circuit related to the Bluetooth-based temperature and humidity acquisition system: (a) bottom view of the device including the temperature and humidity sensors and the microcontroller and (b) top view of the device including the battery holder and the Bluetooth module.

consumes about 1 mW at 1 MHz. As for the Bluetooth part, an integrated module was preferred to a custom circuit for two reasons: first of all, it is much simpler to use a module than to develop a Bluetooth interface and secondly because integrated solutions usually consume less power than discrete component solutions. An F2M03GLA Bluetooth module from Free2Move was chosen because of its simplicity and relatively low-power consumption (about 0.6 mW in sleep mode, with 200 mW peak consumption during transmission). Since the sensor has to be portable, it must be battery powered. A lithium battery (CR2247 from Motorola) was chosen because of its high capacity (1000 mAh).

This sensor was designed in order to be easy to interface to; consequently, the protocol that is implemented does not use low-power techniques to reduce consumption. The Bluetooth module is always on, giving the user the possibility to connect easily at any time. After opening a serial port, the user will start receiving an ASCII-encoded text string representing the measured temperature and humidity. Since no low-power techniques were adopted, this sensor consumes too much power, making it unsuitable for real measuring applications. The main issues to improve are

- Bluetooth module consumes too much power when turned on.
- Also the microcontroller can consume much less power.
- The protocol can be improved in order to transmit fewer bytes.

These were the guidelines that led to the development of the second version of the sensor.

73.3.1.2 Second Version of the Bluetooth Sensor

The first version consumed too much power. Consequently, the second version was designed to keep the power as low as possible. This is done in three ways: by changing the components, by changing the way they are handled, and by changing the protocol. As for the components, a new microcontroller was chosen. In doing so, a Texas Instruments MSP430F2132 was chosen because of its low-power features. Also, the sensor was changed by switching to a Sensirion SHT21. This sensor consumes even less power than the SHT11, and moreover, it has a common interface (I^2C) rather than a custom 2-wire connection. As a result, it is possible to use the integrated I^2C peripheral of the microcontroller so it can save some power. As for the Bluetooth module, the same feature is used as the first version, since it has the same power

consumption as the other modules available. As for the management of the devices, the microcontroller was given the possibility to turn off all other devices (the temperature sensor and the Bluetooth module). Additionally, since both the microcontroller and the sensor can exploit the battery until it decreases to about 2.3 V while the Bluetooth module requires at least 2.9 V, it was decided to use a dc–dc converter to provide 3 V from voltage sources between (slightly more than) 3 and 2.3 V. A charge pump converter from Maxim (MAX1759) was chosen because it could supply the current required by the module (75 mA at most) while requiring very few components (and no inductors). By using this converter, moreover, it was possible to switch off the module completely and also save the 0.6 mW of the power used for standby. As for the protocol, two issues were improved. First of all, transmitting ASCII messages is useful to develop since the information uses too many bytes. For instance, to transmit a number ranging from −40 to 120 (the temperature), in ASCII, it needs at least three characters (so three bytes), while sending it as a number uses only one byte. For this reason, pure numbers were used. The drawback is that it is easy to fine errors in ASCII strings (at the time of receiving nonprintable data, it is possible to recognize that the packet is corrupted), but with pure numbers, it is not possible. So a 1 byte CRC code is added at the end of every packet to enable the receiver to recognize bad packets. The second improvement of the protocol is to find periods when no listener is connected. In fact, it is useless to send data when no one listens for it. Therefore, instead of sending the information to one listener repeatedly, the sensor waits for its "clients" to send him requests. If there are no registered clients, the sensor doesn't even make a measurement or turn on the Bluetooth, which extends battery life. When a new client is registered, the sensor starts waking up at regular intervals to enable the client to make requests. If a client cannot be sensed, it will be removed from the list. These techniques extend the life of the battery. The intervals can be set directly by the user; since the power consumption is proportional to the measuring frequency, the user can select the best trade-off between high sampling rate and long battery life by modifying this parameter. The device is again powered by a 3 V lithium battery (CR2247) and is shown in Figure 73.2.

73.3.2 Smartphone Application for Sensing a Number of Sensors

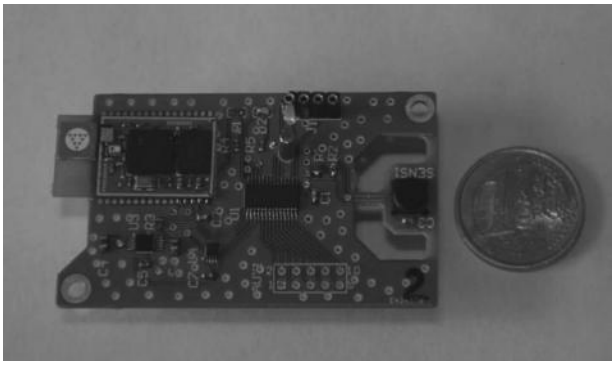
As a prototype, an application for sensing one sensor was developed by Aram et al. (2012b). It is called *aBluSen*. Figure 73.3 shows the data from a sensor. The application communicates, reads, and stores data from a paired Bluetooth-based sensor. This makes it possible to analyze data and also correct performance of the system. This application was developed to read data from several sensors at the same time. This new version makes it possible to communicate, read, and store data from distributed sensors within a short distance from the mobile phone.

The user can select the required sensors by himself or herself manually or automatically, or it is possible to organize system to communicate with all available sensors within a certain area automatically. The availability and connections of the sensors are checked continuously. The application can repeat the process for connecting, disconnecting, reading, and storing each sensor each minute. Figure 73.4 shows that it is possible to select several available sensors and run the system to work with selected ones.

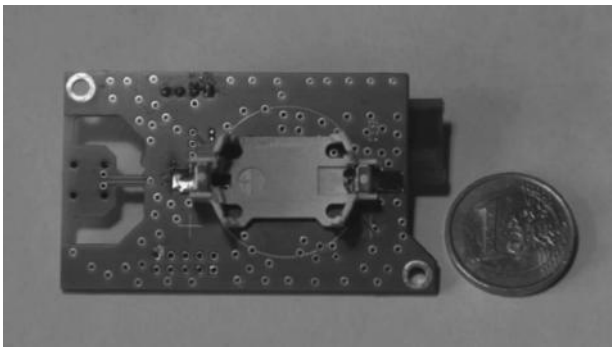
Unlike the first mobile phone application, the new version shows current and previous data of the current connected sensor and the last one (Figure 73.5). The system shows the temperature and humidity values of the previous connected sensor at the time the system tries to connect to the next sensor or reading the sensor. Thus, at the same time, there will be the values of last two read sensors on the screen.

Similar to the previous version of the application, the new one is also developed for the Android mobile phone with the updated version of the Android software development kit (SDK); Eclipse (together with its ADT Plugin) was chosen as the IDE. *aBluSen* has four main parts:

- Configuration
- Discovering
- Selecting sensors
- Display



(a)



(b)

FIGURE 73.2 The electronic circuit of the new Bluetooth-based temperature and humidity acquisition system. (a) Top view of the device including the temperature and humidity sensor, the microcontroller, and the Bluetooth module with the clear change in size. (b) Bottom view of the device including the battery holder.

Discovering and selecting parts both have a menu; Discovering is for scanning the available sensors, while Selecting makes it possible to select the desired Bluetooth-based sensors that gather the environmental information. *aBluSen* can distinguish between non-sensor Bluetooth-based devices and sensors. The application checks the name of the devices as a method of detecting sensors among all available Bluetooth devices. In the Configuration part, the user can set the number of sensors that can be simultaneously connected, sampling frequency, and also the way the system works (e.g., manually or automatically). As shown in Figure 73.5, the Display part shows the temperature and humidity values as environmental information related to the current and previous sensor, number of available and selected sensors, and sensor information. After the first communication, sensors are set as paired sensors. Immediately after doing the select procedure, the application tries to connect to the first sensor. Some communications and sensors may need two attempts for connecting to the sensor. There are two methods of connections:

- Standard method
- Reflection method

If the first method of connecting fails, the application tries to start the second method. The reflection method avoids the complete failure of connection with the standard method for connecting. After successful connection, the *aBluSen* starts reading data from the opened port. At the time of connecting to a sensor, the mobile phone keeps listening to the opened port to read input streams. In the Configuration part, the size of the buffer can change to permit handling the amount of data during the time of reading. The default value for the buffer size is 64. The read data from the sensors are not the data that are shown

FIGURE 73.3 The "Display" part of the aBluSen previous version shows the temperature and humidity values obtained from the Bluetooth-based acquisition system.

FIGURE 73.4 "Select Sensor" part of the new android application (aBluSen) for selecting several available sensors.

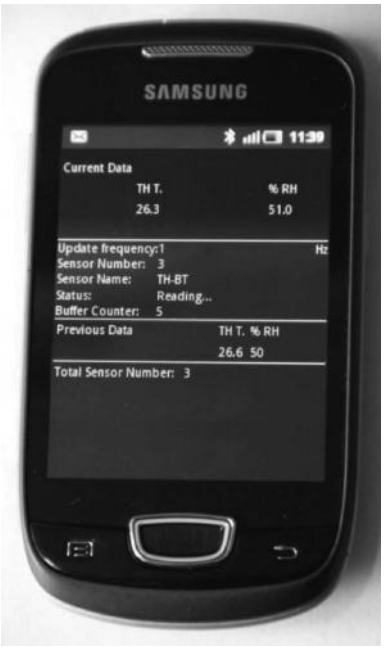


FIGURE 73.5 "Display" part of the new aBluSen application for the android's mobile phone that shows the current and last data of the temperature and humidity values of the two sensors at the same time.

in the Display. The data streams need to parse. The *tokenization* process breaks the streams of data to the desired values of temperature and humidity from the several lines of data that are reached from the connected sensor. A sample line of data is like

$$T = + ., +25.6{,}46$$

The temperature and the humidity are two values that are required to break from the lines of data. The temperature and humidity values are +25.6 and 46, respectively, in this sample. At the time of data reading, the data of the sensors are stored in a text file in the android-based mobile phone. Then the text file is used for analyzing. The interval for the process of reading from the connected sensor is 1 min as a default. In this experiment, the interval is the default. This frequency sample can be modified to communicate and read to/from the sensors via the Configuration part of the application. After each interval, the current communication of the sensor and the mobile phone disconnects, and immediately the application tries to connect the mobile phone to the other selected sensor via Bluetooth. Again, all the processes repeat for the connected sensor. Connecting, reading, storing, and disconnecting are all steps that are repeated continuously according to the configured interval.

73.4 Evaluation

Two experiments were carried out; the first to test the reliability of the sensors and the second to measure their power consumption. Results of the experiments are presented in this section. In order to test the reliability of this system, three temperature and humidity Bluetooth sensors were placed in a climate chamber (Angelantoni Challenge 250); this instrument is able to control environmental conditions such as temperature (from −40 °C to 180 °C) and humidity. The sensors were connected to the smartphone application in order to log the environmental data. They were set with a sampling frequency of one sample per minute. The climate chamber was set to follow a precise temperature profile. For the first 20 min, the temperature was kept constant at 25 °C; then the chamber was set to reach −20 °C with a temperature gradient

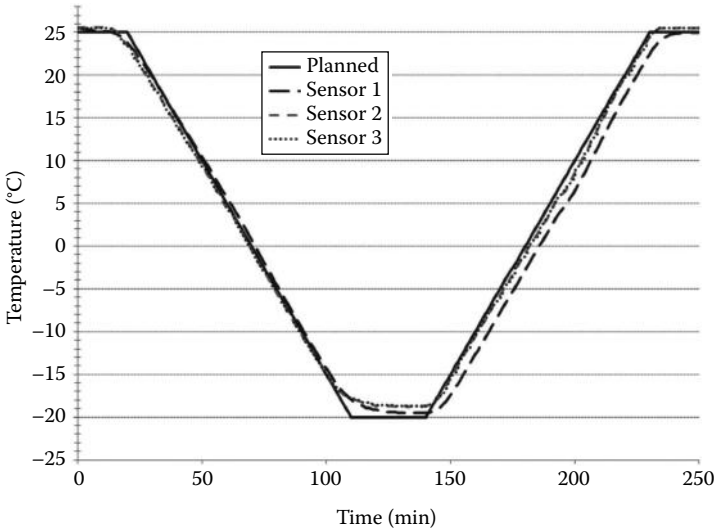


FIGURE 73.6 Temperature values that are obtained by the three Bluetooth-based temperature and humidity acquisition systems during the first experiment.

of −0.5 °C/min (90 min). This temperature was kept for 30 min. Then the chamber was set to reach 25 °C with a gradient of 0.5 °C/min (another 90 min). At the end of the experiment, this temperature was kept constant for 20 min. The overall experiment time was 250 min. The second experiment aimed at measuring the power consumption of the sensor and, consequently, the estimated battery lifetime. The sensor was powered by a power supply, with a constant voltage of 3 V, and the current was measured using a digital multimeter (Agilent 34401A), set in order to acquire current values with a sampling frequency of four samples per second. The second version of the sensor was tested, because it was the only one designed to save power. Figure 73.6 shows the results of the first experiment. The graph represents the temperature values collected by the sensors. The blue line follows the temperature profile imposed by the climate chamber.

As for the second experiment, current values collected by the multimeter are shown in Figure 73.7. The plot shows a 1 min measure; at $T = 0$, the sensor woke up and waited for the client to connect (first peak). At the connection (second peak), the sensor transmitted the values to the client; the operation lasted for about 5 s (due to the pairing required by Bluetooth) and then, since there were no other listeners, put itself into sleep mode for the remaining 50 s.

In low-power mode, the current absorbed by the device is about 1 μA, while during the 10 s of the "active phase," the average current is 27 mA. These values lead to some considerations. First of all, it is required

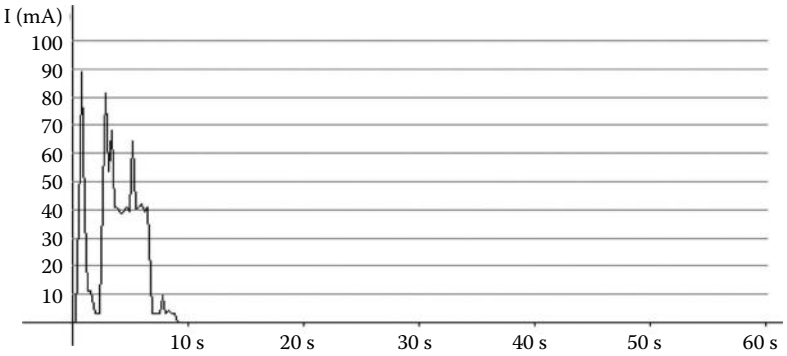


FIGURE 73.7 Current values for the Bluetooth-based temperature and humidity acquisition systems during the second experiment.

that the active phase lasts the shortest time possible. This is achieved by using "control messages," which means that the client sends to the sensor information such as "start measuring" or "disconnect from me." The second consideration is that the sleep time does not influence the current consumption. In fact, in a 1 min measurement, the sleep phase affects current consumption about 0.02%. Neglecting this term permits calculating the battery duration in terms of number of measurements (since the consumption is directly proportional to the frequency of measurement). Collected data, averaged over a 1 h period, indicate that the chosen battery (Motorola CR2477, 1000 mAh) should last for about 14,000 measurements; setting the system to acquire one data every 5 min, the battery should last for about 48 days.

This work studies a low-power wireless sensor network for acquiring temperature and humidity values from a number of sensors. Employing low-cost and low-power components, the Bluetooth as a transceiver (for transferring acquired environmental information) and a smartphone using an application for communicating, reading, and storing data automatically are the major issues of these two recent works.

A laboratory test using a climate chamber was applied to verify the performance of the system. The climate chamber changed the environmental conditions. The outcomes indicated that all involved elements together correctly reported the environmental condition in order to follow the climate chamber changes, and also, the smartphone's android-based application acquired and stored the data properly, which were transferred by the Bluetooth connections. In the first experiment, the smartphone application gathered the environmental data continuously with no absence of signal from any device. Also, the results from the second experiment indicated that the current consumption of the Bluetooth-based temperature and humidity acquisition system is 1 μA while waiting for a transmission, and it is 27 mA (as the average) during the time of transmitting to a smartphone. The results show that the device can transfer data to a smartphone for more than 4 months by using a time interval of 15 min between each two measures using the CR2247 battery. Therefore, the proposed wireless sensor network is working properly.

This prototype shows an approach that is useful for monitoring climatic conditions for small environments using smartphones. The environmental information of small places such as laboratories, home rooms, and medical spaces can be gathered by smartphones for analyzing and also alarming at the time the conditions change or go over fixed thresholds. Also, it is possible to have an application for detecting fire in small environments. Nowadays, the role of smartphones has changed by applying scientific research in projects using their sensors. Aram et al. (2012a,b) applied the smartphone to communicate with external sensors to gather and make possible analyzing obtained environmental data. They considered the power consumption because of the possibility of mobility for long working times that make the sensors—and consequently required information—available. Miniaturization of the proposed device can be considered for improvement of this work. It can be followed by using a smaller microcontroller with the radio part of the Bluetooth module, which needs only one external antenna. Also, reducing the power consumption of the device by considering the following issues can be focused for the future:

- Modifying the use of the Bluetooth protocol
- Having a Bluetooth module with a lower-power consumption (since most consumption of the power is because of the Bluetooth as a transceiver)
- Improving the operating system and software layer of the sensor and mobile phone

References

Abdullah, MFL and Poh, LM, 2011, Mobile robot temperature sensing application via Bluetooth, *Int. J. Smart Home*, 5(3), 39–48.

Aram, S, Troiano, A, and Pasero, E, 2012a, Environment sensing using smartphone, *Proc. Seventh IEEE Sensors Appl. Symposium (SAS)*, Brescia, Italy, pp. 110–113.

Aram, S, Troiano, A, Rugiano, F, and Pasero, E, 2012b, Low power and Bluetooth-based wireless sensor network for environmental sensing using smartphones, *3rd Intelligent Syst. Quality Life Information Services Workshop*, Halkidiki, Greece.

Bureau for Africa's Office of Sustainable Development, 2007, *Environmental Guidelines for Small-Scale Activities in Africa: Environmentally Sound Design for Planning and Implementing Development Activities*, 2nd edn., Bureau for Africa's Office of Sustainable Development Economic Growth, Environment and Agriculture Division, Washington, DC.

Choudhury, T, Consolvo, S, Harrison, B, Hightower, J, LaMarca, A, LeGrand, L, Rahimi, A et al., 2008, The mobile sensing platform: An embedded activity recognition system, *IEEE Pervasive Comput*., 7(2), 32–41.

Das, T, Mohan, P, Padmanabhan, VN, Ramjee, R, and Sharma, A, 2010, Prism: Platform for remote sensing using smartphones, *Proc. 8th ACM MobiSys*, New York, pp. 63–76.

Firtman, M, 2010, *Programming the Mobile Web*, O'Reilly Media, Sebastopol, CA.

Goldman, J, Shilton, K, Burke, J, Estrin, D, Hansen, M, Ramanathan, N, Reddy, S, Samanta, V, Srivastava, M, and West, R, 2009, Participatory sensing: A citizen-powered approach to illuminating the patter that shape our world, White Paper, Foresight and Governance Project, Washington, DC.

Hussein, K, 2006, *Remote Sensing of the Environment*, PDF file, Cooperative Institute for Research in Environmental Science, University of Colorado at Boulder, Boulder, CO.

Kim, S and Paulos, E, 2010, In air: Sharing indoor air quality measurements and visualizations, *Proc. 28th Int. Conf. Human Factors Comput. Syst*., New York, pp. 1861–1870.

Kuchta, R, Vrba, R, and Sulc, V, 2008, Smart platform for wireless communication—Case study, *Proc. 17th ICN*, Cuncum, Mexico, pp. 117–120.

Lane, ND, Miluzzo, E, Lu, H, Peebles, D, Choudhury, T, and Campbell, AT, 2010, A survey of mobile phone sensing, *IEEE Commun. Mag*., 48(9), 140–150.

Lein, JK, 2012, *Environmental Sensing: Analytical Techniques for Earth Observation*, Springer Science, New York.

Mun, M, Reddy, S, Shilton, K, Yau, N, Burke, J, Estrin, D, Hansen, M, Howard, E, West, R, and Boda, P, 2009, Peir, the personal environmental impact report, as a platform for participatory sensing systems research, *Proc. 7th ACM MobiSys*, New York, pp. 55–68.

Paananen, T, 2011, Smartphone cross-platform framework, Bachelor's thesis, JAMK University of Applied Sciences, Jyväskylä, Finland.

Rosi, A, Bicocchi, N, Castelli, G, Corsini, A, Mamei, M, Zambonelli, F, Berti, M et al., 2010, Environmental monitoring and task-driven computing, *IEEE Pervasive Comput*., 9(4), 48–50.

Shan, Q and Brown, D, 2005, Wireless temperature sensor using Bluetooth, *Proc. IWWAN*, London, U.K.

Staples, E, 2011, EnviSensor: Environmental data collection via participatory sensor networks utilizing mobile devices, Master's thesis, University of Edinburgh, Edinburgh, U.K.

Sugiharto, H, 2010, *Current and Future Mobile Platforms*, PDF file, viewed July 31, 2012, SNET Master Project, Berlin Institute of Technology, Germany.

Wei, D, Zhu, X, Uchida, M, Ding, S, and Cohen, M, 2005, A mobile phone based wearable vital signs monitoring system, *Fifth Int. Conf. Comput. Information Technol*., Shanghai, China, pp. 950–955.

Zheng, P and Ni, LM, 2006, *Smart Phone & Next Generation Mobile Computing*, Elsevier Oxford, U.K.

74

Space Dosimetry

Jarvis A. Caffrey
Oregon State University

David M. Hamby
Oregon State University

74.1 Introduction

The advance of space exploration has been fraught with inherent risk throughout its history. Health effects from space radiation for off-world astronauts remain one such risk that has historically received little attention but in fact may be the most constraining factor for long-duration missions that depart from the relative safety of the Earth's magnetosphere. With interplanetary and lunar-based missions rapidly approaching on the horizon, it is now critically important to understand the impacts of the space radiation environment and to develop instrumentation capable of properly expressing that hazard. A summary is provided here for the status of those instruments and their recent developments and to briefly introduce the basis for space radiation hazards. Readers are encouraged to explore the cited literature for further details.

74.1.1 Space Radiation Environment

The space radiation environment is generally characterized by three contributing sources. Galactic cosmic rays (GCRs) represent the permanent baseline of radiation outside of low-Earth orbit (LEO) and the protective magnetosphere. The primary concern for GCR is the prevalence of high-Z (charge), high-energy particles (HZE) likely originating from distant ultra-energetic events such as supernova explosions. The flux profile of GCR peaks for particles with energies of around 1 GeV [1] and extends with decreasing probability to energies of up to 10^{20} eV [2]. HZE particles penetrate shielding and tissue while producing cascades of secondary particles that cause widely varied dose profiles throughout the spacecraft.

The sun periodically releases large masses of solar energetic particles (SEPs) in what are referred to as solar proton events (SPEs). Large solar storms are comparatively rare but can potentially deliver extreme doses sufficient to cause deterministic effects including radiation sickness. Events can often be forecast to some extent, but the high risks call for detectors that can rapidly alert crew to take shelter when needed.

Within LEO the main dose contributors are protons and electrons trapped in belts within the Earth's magnetosphere. Orbiting spacecraft typically maintain altitude below the most intense portion of the belt so that the main dose contribution occurs as it passes through a dip in the magnetic field known as the South Atlantic Anomaly (SAA) [1].

74.1.2 Space Dosimetry Units

Radiation dose is traditionally measured and quantified as average energy deposition within a given volume of known mass, deemed "absorbed dose" (D). Values of absorbed dose caused by photons have been studied and correlated to known health effects. Other radiations that cause dense regions of ionization are known to cause greater biological harm, or equivalent dose (H_T), for identical quantities of absorbed dose in a given tissue, T. In typical dosimetry applications, this difference in relative biological effectiveness (RBE) is corrected by simple radiation weighting factors (w_R) applied to the value of absorbed dose ($D_{T,R}$) for each radiation type, R. These are applied to the value of absorbed dose ($D_{T,R}$) averaged over an organ with w_R equal to 1 for photons and electrons, 2 for protons, 20 for heavy ions, etc., so that

$$H_{T,R} = w_R \cdot D_{T,R} \tag{74.1}$$

For mixed radiation fields, the total equivalent dose (H_T) is calculated by summing the various contributions:

$$H_T = \sum_R H_{T,R} = \sum_R w_R \cdot D_{T,R} \tag{74.2}$$

This can then be corrected with a set of tissue weighting factors (w_T) applied for each organ to determine the effective dose (E) that serves as a risk estimate for long-term effects [3]:

$$E = \sum_T w_T \cdot H_T \tag{74.3}$$

This method is not suited to space environments dominated by a varied flux of charged particles, where applying the basic radiation weighting factor of 20 for all HZE interactions would be overly conservative in most cases [3]. A method is instead recommended by NCRP 132 [4] for space dosimetry in LEO that employs the factor for radiation quality (Q) as a function of the linear energy transfer (LET) as provided by ICRP-60 [5]. Dose equivalent is traditionally expressed simply as

$$H = D \cdot Q(L) \tag{74.4}$$

where

(L) represents LET = dE/dx

$Q(L)$ is unity for (L) less than 10 keV/μm, which increases gradually to 30 for (L) up to 100 keV/μm and decreases thereafter due to the increased probability of cell death that does not contribute to cancer effects

The use of quality factor as a correction for absorbed dose results in point-location values of "dose equivalent" (H), rather than the "equivalent dose" as obtained by the weighting factor method. To take advantage of existing tissue weighting factors for data that include radiation quality, dose is determined as a tissue-averaged value as recommended by ICRU [6]:

$$H_T = \frac{1}{m} \int_m dm \int Q(L) \cdot \phi_T(L) \cdot L \, dL \tag{74.5}$$

where $\phi_T(L)$ represents the fluence of particles through organ T, in the differential increment of LET $(L + dL)$ [7,8]. Investigations continue for development of dosimetry methods in deep space [9] and research has shown significant evidence toward the importance of track structure and charge rather than simple LET to approximate RBE [7,10].

74.1.3 Challenges in Space Dosimetry

Radiation detectors can provide essential information regarding the environment experienced by astronauts, but such measurements serve to clarify only one aspect of a complex process. The stochastic nature of long-term radiation effects such as carcinogenesis has resulted in large uncertainties to risk estimates. This is true both for low-dose exposures where existing population cancer rates are already high and for astronaut exposures in mixed HZE fields with a small sample population and a number of other contributing factors [7]. Effects in deep tissue and organs are not easily measured, though biodosimetric methods can observe dose-dependent cytogenetic effects with methods that continue to improve rapidly [11,12].

Phantom experiments have been used to form relationships between shallow skin dose (the measurable quantity for dosimetry) and depth dose measured throughout a volume of water or tissue-equivalent resin. Several phantom experiments have been performed including the NASA [13] and MATROSHKA [14] human torso phantoms, as well as the early Mir [15] and later MATROSHKA-R [16] spherical phantoms. Computational transport models have also performed important roles in depth dose analysis for organs based upon measurable quantities [17,18] and are likely to become increasingly important as computing power and accessibility steadily improve. Ground-based studies for high-energy interactions and detector performance are also improving the scientific basis for space dosimetry [19–22]. Information gleaned from these studies can be used to extrapolate effective dose based upon flux profile data gathered by personnel dosimeters, area monitors, or in exploratory surveys.

Instrumentation for space dosimetry is then left with a series of challenges for development. The importance of particle discrimination and LET measurement is among the primary technological hurdles, generally requiring advanced processing techniques and extensive calibration studies in accelerator beams [22,23]. Beyond these performance requirements, limitations for weight, size, power consumption, reliability, and cost must be considered. Several instruments have been developed and expanded upon by a global variety of researchers. These include the operational devices from NASA for monitoring and dose records on the International Space Station (ISS) such as TLD/CR-39 (Figure 74.1),

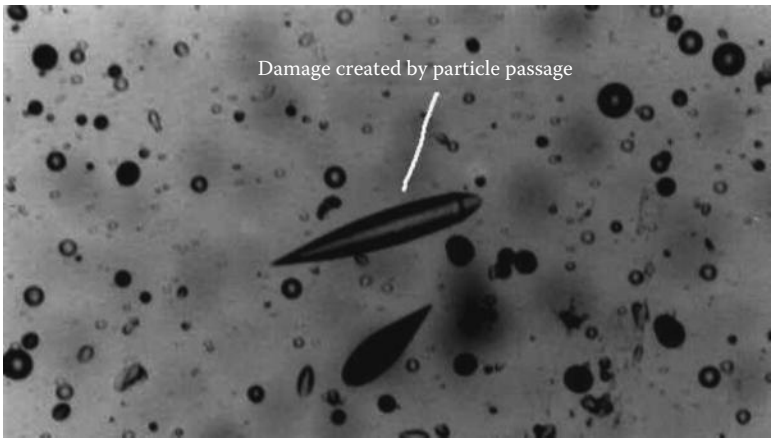


FIGURE 74.1 A CR-39 PNTD from the DOSMAP experiment reveals an etched particle track under microscope. (Courtesy of NASA/JPL-Caltech/SwRI. NASA Dosimetric Mapping (DOSMAP), *International Space Station—Research and Technology*, NASA, April 26, 2013, http://www.nasa.gov/mission_pages/station/research/experiments/1003.html, accessed August 8, 2013.)

tissue-equivalent proportional counter (TEPC), and CPDS (soon replaced by the radiation assessment detector [RAD]), in addition to the R-16, DB8, and Pille devices operated within the Russian module [23]. Beyond these existing operational instruments, a host of science-driven experiments have operated to improve the knowledge base of the space radiation environment while many others proceed in development to improve or replace the fleet of operational dosimeters.

74.2 Passive Dosimetry

Passive dosimeters accumulate a record of incident radiation that can later be processed to estimate the dose integrated over the time of exposure. They are typically small and economical and operate with no demands on the limited power onboard spacecraft. They have been an effective and practical choice for space dosimetry, particularly in measuring the widely varied radiation burden throughout spacecraft [24]. Passive dosimeters have also been central to phantom experiments conducted by a diverse group of researchers, and ground-based studies such as HAMLET continue to improve the understanding of results obtained therein [14,21,25,26].

74.2.1 Thermoluminescent Dosimeters

Thermoluminescent dosimeters (TLDs) have played a key role in space dosimetry throughout the history of manned spaceflight [27]. They are lightweight, affordable, and provide a time-integrated dose that translates easily to an administrative dose record. Space dosimetry introduces several problems not encountered in terrestrial TLD dosimetry due to the complex radiation field with significant high-LET contributions. TLDs respond poorly to high-LET radiation, as the dense ionization tracks tend to saturate electron/hole trapping centers and the deposited energy is then dissipated by competing mechanisms [28,29]. Efforts from many researchers have attempted to extend the function of TLDs by estimating LET and particle type through various methods and materials with limited success [28,30–32]. Their poor sensitivity to high-LET radiation limits TLDs to dose estimates from particles of less than 10 keV/μm [21,33].

TLD dosimeters are typically returned to Earth for laboratory processing, thus accumulating some dose throughout a mission and limiting its capability for long-term use. The notable exception is a portable readout TLD system known as "Pille" developed by the Hungarian KFKI Atomic Energy Research Institute. Pille dosimeters are constructed of $CaSO_4$:Dy crystal laminated to a resistive plate within a glass vacuum bulb. Ten such dosimeters can be worn or permanently stationed within a compact reader device that manages heating, light measurement, and dose records for each dosimeter. "Semipassive" systems such as the Pille offer a reasonable compromise for portability and power consumption for long-duration missions in which passive dosimeters must be read out in transit, though like other TLDs the Pille is not sensitive to high-LET particles [24,34,35].

74.2.2 Optically Stimulated Luminescent Dosimeters

Optically stimulated luminescent dosimeters (OSLDs) share the same fundamental mechanisms as TLDs for dose accumulation and readout but are instead stimulated by optical laser. Stimulation sensitivity to natural UV light adds to the risk of dose fading, so extra care must be taken to maintain total darkness for the crystal until readout is complete. OSLDs for space applications are constructed with Al_2O_3:C and have shown to be comparable to TLD in most cases [36,37]. Light stimulation methods require less power consumption than heating, provide greater flexibility for portable readout systems [38], and show improved capabilities for monitoring various effects such as light emission decay time. A possible advantage for OSL over TL is its capability to release a fraction of trapped electrons/holes by briefly pulsed stimulation to estimate dose without resetting (bleaching) the crystal. Dose record could therefore be monitored at periodic intervals while maintaining the dose record for long-term monitoring. A summary of OSLD with its history of development and applications in space is provided in reference [39].

74.2.3 Nuclear Track Detectors

Plastic nuclear track detectors (PNTDs) are the primary means of passive dosimetry for high-LET radiation. Among these, CR-39 (Figure 74.1) resin is the most widely used and studied. Researchers continue to investigate improvements in processing methods [40–44] and particle identification [45–48]. High-LET particles produce latent tracks as they traverse the resin, where ionization weakens molecular bonds. Following irradiation, the resin is etched in a caustic solution to dissolve the material. Latent tracks dissolve at a faster rate than the remaining plastic to form distinct conical pits that are clearly visible under microscope, either processed manually through observation or through automated scanning [44]. The profile of etched cones can be correlated with particle LET and charge, then coupled to the track density to determine fluence and dose [49].

74.2.4 Combined Passive Dosimeters

Estimating the equivalent dose received by individuals in mixed radiation fields generally requires a combination of dosimeters with varying response characteristics. CR-39 (Figure 74.1) is sensitive primarily to charged particles with LET between 5 and 1000 keV/μm and is typically coupled with TLD or OSLD sensitive to radiation below 10 keV/μm. This combined passive dosimetry method is widely used in space dosimetry and other mixed radiation fields to produce reliable measurements for dose of record [26,27,37,50,51].

74.2.5 Superheated Drop Emulsions

Superheated drop (bubble) detectors offer a novel approach to fast neutron detection in space. A fluid of polymer liquid droplets is encapsulated within a clear vial and sealed at a low-pressure (superheated) state. The polymer remains liquid at ambient temperature but readily vaporizes to form visible bubbles when ionized by recoil protons of fast neutron collisions [52,53]. A portable readout machine is used to measure the quantity of bubbles formed during irradiation, a value that is calibrated directly to dose equivalent. The device can then recompress the fluid to reset the dosimeter for a limited number of repeated irradiations. A single dosimeter has a limited capacity for bubble formation and requires resetting at regular intervals; thus, it behaves as a semipassive system requiring a powered device. Bubble detectors have been operated within the MATROSHKA-R spherical phantom for neutron depth dose measurements [25], as neutron-sensitive area dosimeters, and as a neutron spectrometer by exposing multiple detectors of varying threshold energies. Investigations on the ISS have pointed toward a greater sensitivity to fast neutrons than to charged particles [54,55].

74.3 Active Dosimetry

Active radiation detectors require power and are typically larger than passive measurement systems, but their ability to immediately resolve fluctuations in the radiation environment could be a lifesaving feature in the event of an unforeseen SPE. Recent progress in microelectronics has enabled the development of personal electronic dosimeters capable of LET measurement, though at this time no such devices are used as operational space dosimeters.

74.3.1 Tissue-Equivalent Proportional Counters

A variety of TEPC dosimeters have flown on missions including the International Space Station (ISS) and Mir Orbital Station beginning in 1994 [56]. The TEPC microdosimeter serves as the primary operational area radiation monitor on the ISS and is integrated into an alarm system to warn astronauts of radiation storms that may require refuge in the station's radiation shelter. The system has undergone

several design revisions over the past two decades, beginning as a cylindrical detector to simplify charge collection characteristics and recently changing to a two-detector spherical shell system. Propane is used as a fill gas at very low pressure to emulate the radiation effects inside of a micron-scale volume to approximate a single tissue cell.

In the case of the TEPC, lineal energy (y) is calculated for each event by dividing deposited energy over the mean chord length. Detector response calibrations are used to correlate the measured lineal energy with particle LET and, through previous research [57], determine the radiation quality (Q) or RBE for final calculation of dose equivalent (H). The occurrence of interactions in the detector wall and uncertainty in the use of mean chord length can result in deviations from true values of ionization density, especially when considering the nonuniform angular distribution of chords through cylindrical detectors [58,59].

The latest TEPC revision uses two independent spherical detection volumes that improve the uncertainty associated with mean chord length. Irregularities in the electric field caused by the varying anode/cathode distance within the chamber are mitigated in the new TEPC by constructing the spherical shell as a stack of several annular rings each with adjustable applied voltage [60].

74.3.2 Compact Diode Dosimeters

Semiconductors provide several advantages for use as compact active dosimeters. Solid diodes perform well in very small form factors and can operate with far lower voltage requirements than gas detectors. Power consumption is rather minimal by comparison, and battery-operated devices have a distinct advantage in noise isolation from the station power supply. Semiconductor materials are not tissue equivalent, however, and they require correction factors to apply Si-based measurements into tissue dose [61].

The Liulin-4 single-diode detector has been used extensively for detection in European space missions with remarkable reliability. Its measurements are generally limited to absorbed dose and dose rate in the absence of a known deposited track length, though it does attempt some estimate of dose equivalent by separating low-energy interactions from high-energy depositions assumed to originate from neutrons [62,63]. A separate single-diode instrument was flown recently onboard the Lunar Reconnaissance Orbiter as part of the CRaTER experiment to demonstrate miniaturization capabilities of a simple dosimeter [64].

Advances in microfabrication have led to the development of various instruments that facilitate LET measurements within silicon microdosimeters. Silicon-on-insulator (SOI) microdosimeters such as the MIDN are constructed as an array of microscopic sensitive volumes with thickness of 10 μm or less to permit reasonable estimates of lineal energy and LET [65–69].

An alternative approach involves monitoring energy deposition in a pixelated silicon wafer, essentially serving as an active nuclear track detector. The Timepix/Medipix device is one such pixelated detector approaching applications for space dosimetry based upon work from the CERN Medipix collaboration [70]. Dose equivalent is measured by calculating the primary particle's tangential track length and total energy deposition in all pixels [71–73]. Some measure of particle identification can be performed based upon track structure [74], and neutron dosimetry may be viable with proton recoil measurement under a polyethylene converter [75].

Other pixelated, or rather voxelated, devices based upon scintillator technology have been proposed that take advantage of the recent progress in CMOS technology. The advancement of solid-state photomultiplier (SSPM) technology has allowed for the replacement of large power-demanding photomultiplier tubes with small and efficient silicon sensors. SSPMs have been developed from the miniaturization of technology based upon the avalanche photodiode, where each photodiode micropixel operates in a Geiger mode that activates when struck by light photons. When coupled to a scintillating crystal, the quantity of radiation imparted can be measured as a function of the number of activated pixels. Improvements in microfabrication allow for a high density of pixels and therefore good dynamic range and high resolution for spectroscopy. A 3D voxel structure of small scintillators, each coupled to an independent SSPM, may provide a capable neutron-sensitive particle tracking dosimeter with minimal power requirements [76,77].

74.3.3 MOSFET Dosimeters

Metal oxide semiconductor field effect transistor (MOSFET) systems have become a very popular method of dosimetry for space applications. Standard MOSFETs are constructed by application of a silicon oxide layer upon a *p*- or *n*-doped silicon substrate between two electrical contacts, deemed the source and drain electrodes. In the case of a *p*-channel MOSFET (*p*MOSFET), the silicon substrate is doped with electron-donor (*n*-type) impurities, while nMOSFET systems utilize a *p*-type substrate. During irradiation, a positive bias voltage is applied to the gate electrode placed in the SiO_2 region, so that electron/hole pairs are separated after ionization. The electrons produced in the sensitive SiO_2 are drawn away by the positive gate bias, while holes migrate and trap within the Si/SiO_2 interface. During readout, an applied gate voltage is gradually increased while the source and drain electrodes are monitored for continuity. As the amount of ionization causes the number of trapped holes to increase, so too does the negative gate voltage required to induce a channel of conductivity in the silicon substrate [8].

MOSFET has been used extensively for dosimetry experiments and radiation monitoring in space, as they are generally very compact and require minimal power. In practice, they can operate similarly to the portable readout "semipassive" dosimeters such as the Pille TLD or bubble-detector systems, where the integrated dose is measured periodically such as in EVA dosimetry [78]. They are often operated as "semiactive" detectors, however, such as within the MATROSHKA-R phantom where the small detector region was placed at precise locations and wired to an automated measurement system for continuous cycling of readout and measurement [79]. Some performance features of MOSFET dosimeters include permanent dose storage, extremely small size, and adjustable sensitivity based upon applied gate voltage during irradiation. They show very good sensitivity to photons and charged particles across a wide range of energy and can be made sensitive to neutrons particularly with the addition of converter material. Various improvements are underway for space dosimetry applications, particularly for stability and for preventing source–drain leakage current after heavy irradiation [80].

74.3.4 Particle Telescopes

Systems of multiple detectors operated with coincident signal processing can reveal far more information about an incident particle than any individual detector. As mentioned previously, a single large-area semiconductor may accurately measure absorbed dose as a function of deposited energy but cannot reliably estimate LET due to the ambiguity of the particle trajectory and path length inside the sensitive volume.

Charged particle telescopes are capable of measuring LET for charged particles incident within a restricted solid angle as they pass through a series of detectors activated in coincidence. Isolating measurements to known trajectories allows for a reasonable estimate of the particle track length through each detector, where narrower fields of view serve to improve the uncertainty of this estimate. The charge collected within an individual detector is then divided over the thickness of the active region to calculate the particle's rate of energy loss, or dE/dx. Individual detectors within a telescope may still operate between coincidence events to independently accumulate measurements for absorbed dose or flux [81].

The early Voyager missions included cosmic ray telescope experiments with multiple stacked sets of silicon detectors that have operated reliably since their launch in 1977 [82,83]. Similar telescopes of large-area planar silicon detectors have been operated with various improvements, including the German DOSTEL 2D instrument flown in both Mir and ISS [27,84] to map the dose profile within the spacecraft. The Bulgarian Liulin instrument was expanded from previous work to build the Liulin-5 telescope within the MATROSHKA-R spherical phantom experiment [16,81,85], and the Liulin-Phobos telescope onboard the ill-fated Fobos-Grunt sample-return mission intended for Mars' moon in 2011. The CRaTER device currently onboard the Lunar Reconnaissance Orbiter uses a similar configuration as Liulin-5 but fills the volumes between detectors with tissue-equivalent plastic [86–88].

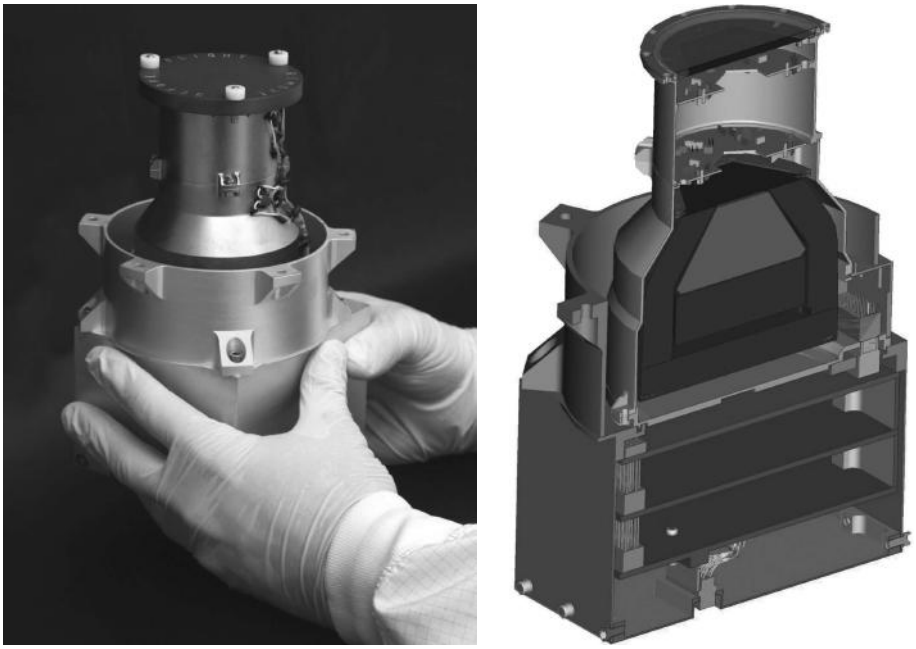
A telescope stack can further improve the accuracy of measured LET by including position-sensitive detectors (PSDs) to establish a more precise particle trajectory. Double-sided strip detectors (DSSDs) are commonly used for this purpose, in which perpendicular arrays of strip detectors create a detection grid for penetrating charged particles. DSSDs were used to construct the Japanese RRMD-III telescope spectrometer to measure real LET and dose equivalent [59,89,90].

The Italian SilEye devices use strip detector telescopes for particles traversing the eyes and brain to investigate the anomalous light flashes first noted by Apollo astronauts [91]. The SilEye3/Alteino device and the later ALTEA device were operated onboard the ISS as astronaut experiment facilities, then continued operation as area monitors under the Altcriss and ALTEA-DOSI projects. A set of six telescope units were mounted on a helmet that also contained an EEG cap for electrophysiological measurements. Astronauts operated the experiment by wearing the device and activating pushbutton sensors to indicate the observation of light-flash events [92]. The large telescope device is capable of identifying energetic particles with charges (Z) from 4 to 28 and energy greater than 25 MeV per nucleon [93,94].

Three DSSD detectors were also incorporated into NASA's charged particle directional spectrometer (CPDS) put into service on the ISS beginning in 2002 [95,96]. The intricate device housed three other silicon detectors for coincidence triggering, six lithium-drifted silicon (Si(Li)) detectors for spectroscopy and LET measurement below which sat a sapphire crystal with a small photomultiplier tube for Cherenkov detection. The instrument could perform spectroscopy and LET measurement in the Si(Li) detector stack across a wide field of view for particles that only activated the top two coincidence and DSSD detectors. Particles entering within a narrower field of view could penetrate all 12 silicon detectors and deposit into the Cherenkov detector allowing the CPDS to identify particles with charge of $Z < 12$. The CPDS was implemented in both the interior (IV-CPDS) and exterior (EV-CPDS) of the ISS as part of NASA's radiation monitoring program but was troubled by problems that resulted in the premature failure of one EV-CPDS as well as the IV-CPDS.

A similar spectrometer, MARIE, was included in the Mars Odyssey mission launched in 2001 [97–99]. The MARIE detector suffered from a few problems including excessive preamplifier gain that saturated with high-energy events, slow readout for pulse height that caused excessive system dead time, and a design flaw in the Cherenkov detector [83]. Regardless, the MARIE device successfully operated for over 2 years until a large SPE ended its operation in 2003 [7]. An advanced telescope design concept, the combined ion and neutron spectrometer (CINS), includes a charged particle stack based upon the MARIE design but features improved electronics [83]. It proposes to use two advanced neutron spectrometers, including a thick Si(Li) semiconductor fast neutron spectrometer as well as an Eljen boron-loaded scintillating spectrometer for mid-energy neutrons based upon the gamma ray and neutron spectrometer within the MESSENGER Mercury spacecraft [100].

The RAD (Figure 74.2), built by Southwest Research Institute, represents the latest implementation of telescope spectrometers. It is currently operating onboard the Mars Science Laboratory (MSL) Curiosity rover to measure the flux profile at the Mars surface in support of future human exploration and Martian life investigations. The RAD is also slated to replace the IV-CPDS as an operational area monitor inside the ISS by 2015. The MSL RAD houses three silicon diodes as a charged particle telescope, each segmented into several complex regions that provide position dependence sufficient to minimize path length uncertainty. Three scintillators housed beneath the diode stack measure energy deposition from charged particles arriving through the telescope and also perform spectroscopy for neutrons and photons arriving from all angles. The thallium-doped cesium iodide (CsI(Tl)) scintillator sits directly below the telescope and is shaped to match its field of view for charged particle collection. It also responds well to photons and, to some degree, neutrons from all directions. An organic plastic scintillator constructed of Bicron BC-432 m sits below the CsI(Tl) crystal, primarily intended for neutron detection by recoil proton production, but also sensitive in some degree to photons. A BC-432 m anticoincidence shield surrounds the inner scintillators to eliminate charged particle background, except for those traversing the telescope counted for spectroscopy or particle identification. Photodiodes, nearly identical to the diodes used in the telescope, are bonded to the scintillators for light measurement, thus requiring only a single power supply for the entire



(a) (b)

FIGURE 74.2 (a) RAD—before installation on MSL Curiosity rover. (Courtesy of NASA/JPL-Caltech/SwRI. Mars Science Laboratory: Radiation Assessment Detector (RAD), *Mars Science Laboratory: Curiosity Rover*, NASA Jet Propulsion Laboratory: California Institute of Technology, n.d., http://mars.jpl.nasa.gov/msl/mission/instruments/radiationdetectors/rad/, Accessed August 8, 2013.) (b) RAD—cutaway diagram. (Courtesy of NASA/JPL-Caltech/SwRI. Mars Science Laboratory: Radiation Assessment Detector (RAD), *Mars Science Laboratory: Curiosity Rover*, NASA Jet Propulsion Laboratory: California Institute of Technology, n.d., http://mars.jpl.nasa.gov/msl/mission/instruments/radiationdetectors/rad/, Accessed August 8, 2013.)

system [101]. Neutron and gamma contributions are separated in the central scintillators by evaluating the unique response of each scintillator with a statistical maximum-likelihood inversion technique developed for that purpose [102]. Other independent neutron detectors are described in the following section.

74.3.5 Neutron Detector

Neutrons are notoriously difficult to detect efficiently, as their lack of charge prevents the occurrence of direct ionization to produce measurable electrical signals. Neutron detection must then rely upon the production of secondary recoil ions from scatter interactions, through photon emission after activating capture interactions or through the products of fission interaction. High-energy neutrons often produced as secondaries within spacecraft may proceed to scatter and maintain relatively high energy until entering a low-Z (especially hydrogenous) material where its kinetic energy can be transferred away as energetic recoil ions. Unfortunately, humans often serve as that hydrogenous material, and neutrons can therefore introduce a significant contribution to dose equivalent, approximated to be around 30% of total dose equivalent with significant uncertainty [7,103,104].

74.3.5.1 Bonner Sphere

Bonner sphere detectors provide a method for determining neutron energy spectra by surrounding a ^{3}He proportional counter with polyethylene spheres of varying thickness. ^{3}He proportional counters respond well to neutrons across a broad range of energy by way of the ^{3}He (n,p) ^{3}H reaction. Each moderating sphere yields a unique spectral response as energetic neutrons are moderated to varying degrees.

The Bonner Ball Neutron Detector (BBND) experiment measured the neutron energy spectrum onboard the ISS in 2001. Six ^{3}He proportional counters were installed within a large assembly, four of which included polyethylene spheres of varying thickness while two counters remained unmoderated. Gadolinium shields were mounted on one unmoderated detector as well as upon the smallest moderator serving as controls to reject the entry of prethermalized neutrons. An unfolding technique was then applied to the combined spectra to develop a single differential neutron energy spectrum [89].

74.3.5.2 Neutron Telescope

Several neutron telescope detectors have been developed within the Russian space program. The Board Telescope of Neutrons (BTN) system is mounted onboard the Russian segment of the ISS and is identical to the high-energy neutron detector (HEND) instrument operated on the Mars Odyssey orbiter. The detectors operate with three long-tube ^{3}He proportional counters oriented in each of the three axes sensitive to epithermal neutrons outside of the space station. Each counter is encased in polyethylene of varying thickness, where the directionality of the system provides information regarding the source of incident neutrons. A $C_{14}H_{12}$ stilbene scintillating crystal measures high-energy neutrons of 400 keV–10 MeV based upon a calibrated response for recoil protons, while a CsI:Tl^{3+} anticoincidence shield blocks/reduces background from charged particles [105].

A separate neutron telescope design, the Lunar Exploration Neutron Detector (LEND), is included on the Lunar Reconnaissance Orbiter to measure the neutron field around the moon and the flux of neutrons emitting from the lunar surface. LEND is constructed with eight ^{3}He proportional counters, four of which are located within parallel holes of a collimation module and covered with cadmium foil to restrict measurements to epithermal neutrons only. A fifth collimation tube houses a stilbene scintillator, again used for high-energy neutron measurements inside an anticoincidence shield. The collimation module is constructed of polyethylene and ^{10}B to moderate then capture any off-angle incident neutrons, restricting the collimated detectors' field of view to targeted locations on the lunar surface. The four remaining ^{3}He proportional counters are mounted outside of the collimator to measure the surrounding neutron flux, two of which are cadmium-wrapped to eliminate thermal neutron contribution [106].

74.4 Summary

Space dosimetry is a complex and developing science that will play a critical role in the forthcoming exploration of space. Astrophysics, health physics, materials science, electrical engineering, particle physics, and a host of other disciplines play vital roles in developing an understanding for the radiological risks of space travel. A broad array of instruments has been developed in this pursuit, either as operational monitors for crew dosimetry or as experiments to profile various aspects of the radiation environment. The importance of determining quality factor for incident radiation has added to the challenge of producing instruments capable of fulfilling the rigorous demands of spaceflight. Rapid developments in microelectronics and crossover from research in related fields have expedited the development of capable devices, but as with most aspects of radiation dosimetry, there is a pressing need for further research.

References

1. E. V. Benton and E. R. Benton, Space radiation dosimetry in low-Earth orbit and beyond, *Nucl. Instrum. Methods Phys. Res. B., Beam Interact. Mater. Atoms*, 184(1–2), 255–294, September 2001.
2. J. Cronin, Cosmic rays: The most energetic particles in the universe, *Rev. Modern Phys*., 71(2), S165–S172, March 1999.
3. ICRP, The 2007 recommendations of the International Commission on Radiological Protection, *Ann ICRP*, 37(2–4):1–332, January 2007.
4. NCRP, Report No. 132: Radiation protection guidance for activities in low-Earth orbit, Bethesda, MD, 2000.

5. ICRP, *1990 Recommendations of the International Commission on Radiological Protection*. Pergamon Press, Oxford, U.K., 1991.
6. ICRU, Quantities and units in radiation protection dosimetry (Report 51), Bethesda, MD, 1993.
7. M. Durante and F. A. Cucinotta, Physical basis of radiation protection in space travel, *Rev. Modern Phys*., 83(4), 1245–1281, November 2011.
8. G. F. Knoll, *Radiation Detection and Measurement*. John Wiley & Sons, New York, 2010.
9. NCRP, Report No. 153: Information needed to make radiation protection recommendations for space missions beyond low-Earth orbit, Bethesda, MD, 2006.
10. R. Katz, S. C. Scharma, and M. Homayoonfar, The structure of particle tracks, in *Topics in Radiation Dosimetry*, F. H. Attix (ed.), Academic Press, New York, pp. 317–383, 1972.
11. F. A. Cucinotta, M. Y. Kim, V. Willingham, and K. A. George, Physical and biological organ dosimetry analysis for international space station astronauts, *Radiat. Res*., 170(1), 127–138, 2008.
12. K. A. George, J. Rhone, L. J. Chappell, and F. A. Cucinotta, Cytogenetic biodosimetry using the blood lymphocytes of astronauts, *Acta Astronaut*., 92(1), 97–102, November 2013.
13. G. D. Badhwar, W. Atwell, F. F. Badavi, T. C. Yang, and T. F. Cleghorn, Space radiation absorbed dose distribution in a human phantom, *Radiat. Res*., 157(1), 76–91, January 2002.
14. G. Reitz, T. Berger, P. Bilski, R. Facius, M. Hajek, V. Petrov, M. Puchalska et al., Astronaut's organ doses inferred from measurements in a human phantom outside the international space station, *Radiat. Res*., 171(2), 225–235, February 2009.
15. T. Berger, M. Hajek, L. Summerer, N. Vana, Y. Akatov, V. Shurshakov, and V. Arkhangelsky, Austrian dose measurements onboard space station MIR and the International Space Station—Overview and comparison, *Adv. Space Res*., 34(6), 1414–1419, January 2004.
16. J. Semkova, R. Koleva, S. Maltchev, N. Bankov, V. Benghin, I. Chernykh, V. Shurshakov, V. Petrov, S. Drobyshev, and I. Nikolaev, Depth dose measurements with the Liulin-5 experiment inside the spherical phantom of the MATROSHKA-R project onboard the International Space Station, *Adv. Space Res*., 49(3), 471–478, February 2012.
17. W. Atwell, Anatomical models for space radiation applications: An overview, *Adv. Space Res*., 14(10), 415–422, October 1994.
18. T. Sato, A. Endo, L. Sihver, and K. Niita, Dose estimation for astronauts using dose conversion coefficients calculated with the PHITS code and the ICRP/ICRU adult reference computational phantoms, *Radiat. Environ. Biophys*., 50(1), 115–123, March 2011.
19. J. Norbury, J. Miller, and A. Adamczyk, Review of nuclear physics experiments for space radiation, NASA Technical Paper, 217179, 2011.
20. J. W. Norbury, J. Miller, A. M. Adamczyk, L. H. Heilbronn, L. W. Townsend, S. R. Blattnig, R. B. Norman, S. B. Guetersloh, and C. J. Zeitlin, Nuclear data for space radiation, *Radiat. Meas*., 47(5), 315–363, May 2012.
21. P. Bilski, T. Berger, M. Hajek, and G. Reitz, Comparison of the response of various TLDs to cosmic radiation and ion beams: Current results of the HAMLET project, *Radiat. Meas*., 46(12), 1680–1685, December 2011.
22. N. Yasuda, Y. Uchihori, E. R. Benton, H. Kitamura, and K. Fujitaka, The intercomparison of cosmic rays with heavy ion beams at NIRS (ICCHIBAN) project, *Radiat. Protect. Dosim*., 120(1–4), 414–420, January 2006.
23. T. Berger, Radiation dosimetry onboard the International Space Station ISS, *Zeitschrift für Medizinische Physik*, 18(4), 265–275, December 2008.
24. G. Reitz, Space radiation measurements on-board ISS—The DOSMAP experiment, *Radiat. Protect. Dosim*., 116(1–4), 374–379, December 2005.
25. R. Machrafi, K. Garrow, H. Ing, M. B. Smith, H. R. Andrews, Y. Akatov, V. Arkhangelsky et al., Neutron dose study with bubble detectors aboard the International Space Station as part of the Matroshka-R experiment, *Radiat. Protect. Dosim*., 133(4), 200–207, February 2009.

26. D. Zhou, E. Semones, D. O'Sullivan, N. Zapp, M. Weyland, G. Reitz, T. Berger, and E. R. Benton, Radiation measured for MATROSHKA-1 experiment with passive dosimeters, *Acta Astronaut.*, 66(1–2), 301–308, January 2010.
27. G. Reitz, Past and future application of solid-state detectors in manned spaceflight, *Radiat. Protect. Dosim.*, 120(1–4), 387–396, January 2006.
28. Y. S. Horowitz, D. Satinger, E. Fuks, L. Oster, and L. Podpalov, On the use of LiF:Mg,Ti thermoluminescence dosemeters in space—A critical review, *Radiat. Protect. Dosim.*, 106(1), 7–24, January 2003.
29. T. Berger and M. Hajek, TL-efficiency—Overview and experimental results over the years, *Radiat. Meas.*, 43(2–6), 146–156, February 2008.
30. N. Vana, W. Schöner, M. Fugger, and Y. Akatov, Absorbed dose measurement and LET determination with TLDs in space, *Radiat. Protect. Dosim.*, 66(1–4), 145–152, July 1996.
31. Y. S. Horowitz, A. Horowitz, L. Oster, S. Marino, H. Datz, and M. Margaliot, Investigation of the ionisation density dependence of the glow curve characteristics of LIF:MG,TI (TLD-100), *Radiat. Protect. Dosim.*, 131(4), 406–413, January 2008.
32. P. Bilski, On the correctness of the thermoluminescent high-temperature ratio (HTR) method for estimating ionization density effects in mixed radiation fields, *Radiat. Meas.*, 45(1), 42–50, January 2010.
33. M. Puchalska and P. Bilski, An improved method of estimating ionisation density using TLDs, *Radiat. Meas.*, 43(2–6), 679–682, February 2008.
34. I. Apáthy, Y. A. Akatov, V. V. Arkhangelsky, L. Bodnár, S. Deme, I. Fehér, A. Kaleri et al., TL dose measurements on board the Russian segment of the ISS by the 'Pille' system during Expedition-8, -9 and -10, *Acta Astronaut.*, 60(4–7), 322–328, February 2007.
35. I. Feher, S. Deme, B. Szabo, J. Vagvolgyi, P. P. Szabo, A. Csoke, M. Ranky, and Akatov YuA, A new thermoluminescent dosimeter system for space research, *Adv. Space Res.*, 1(14), 61–66, January 1981.
36. E. Yukihara, G. Sawakuchi, S. Guduru, S. W. S. McKeever, R. Gaza, E. R. Benton, N. Yasuda, Y. Uchihori, and H. Kitamura, Application of the optically stimulated luminescence (OSL) technique in space dosimetry, *Radiat. Meas.*, 41(9–10), 1126–1135, October 2006.
37. F. Vanhavere, J. Genicot, D. O'Sullivan, D. Zhou, F. Spurný, I. Jadrníčková, G. O. Sawakuchi, and E. G. Yukihara, DOsimetry of BIological EXperiments in SPace (DOBIES) with luminescence (OSL and TL) and track etch detectors, *Radiat. Meas.*, 43(2–6), 694–697, February 2008.
38. J. L. Genicot, M. Moyaerts, and W. Schroeyers, Description of a portable devices developed at SCK•CEN for OSL and TL dosimetry, *Radiat. Meas.*, 46(12), 1578–1581, December 2011.
39. S. W. S. McKeever, Optically stimulated luminescence: A brief overview, *Radiat. Meas.*, 46(12), 1336–1341, December 2011.
40. C. E. Johnson, E. R. E. V. Benton, and N. Yasuda, Analysis of short-range tracks and large track fluences in CR-39 PNTD using atomic force microscopy, *Radiat. Meas.*, 44, 742–745, 2009.
41. A. Kulkarni, C. K. Vyas, H. Kim, P. C. Kalsi, T. Kim, and V. Manchanda, Online optical monitor of alpha radiations using a polymeric solid state nuclear track detector CR-39, *Sens. Actuators B: Chem.*, 161(1), 697–701, January 2012.
42. P. A. Mosier-Boss, L. P. G. Forsley, P. Carbonnelle, M. S. Morey, J. R. Tinsley, J. P. Hurley, and F. E. Gordon, Comparison of SEM and optical analyses of DT neutron tracks in CR-39 detectors, *Radiat. Meas.*, 47(1), 57–66, January 2012.
43. S. Kodaira, D. Nanjo, H. Kawashima, N. Yasuda, T. Konishi, M. Kurano, H. Kitamura et al., Mass spectrometry analysis of etch products from CR-39 plastic irradiated by heavy ions, *Nucl. Instrum. Methods Phys. Res. Sect. B: Beam Interact. Mater. Atoms*, 286, 229–232, November 2011.
44. F. Coppola, M. Durante, G. Gialanella, G. Grossi, L. Manti, M. Pugliese, and P. Scampoli, Development of an automated scanning system for the analysis of heavy ions' fragmentation reaction by nuclear track detectors, *Radiat. Meas.*, 44(9–10), 802–805, October 2009.
45. P. A. Mosier-Boss, S. Szpak, F. E. Gordon, and L. P. G. Forsley, Triple tracks in CR-39 as the result of Pd-D Co-deposition: Evidence of energetic neutrons, *Die Naturwissenschaften*, 96(1), 135–142, January 2009.

46. S. Ota, N. Yasuda, L. Sihver, S. Kodaira, M. Kurano, S. Naka, Y. Ideguchi, E. R. Benton, and N. Hasebe, Charge resolution of CR-39 plastic nuclear track detectors for intermediate energy heavy ions, *Nucl. Instrum. Methods Phys. Res. Sect. B: Beam Interact. Mater. Atoms*, 269(12), 1382–1388, June 2011.
47. S. Kodaira, N. Yasuda, H. Kawashima, M. Kurano, S. Naka, S. Ota, Y. Ideguchi, N. Hasebe, and K. Ogura, Detection threshold control of CR-39 plastic nuclear track detectors for the selective measurement of high LET secondary charged particles, *Radiat. Meas.*, 46(12), 1782–1785, December 2011.
48. E. R. Benton, C. E. Johnson, J. DeWitt, N. Yasuda, E. V. Benton, M. H. Moyers, and A. L. Frank, Observations of short-range, high-LET recoil tracks in CR-39 plastic nuclear track detector by visible light microscopy, *Radiat. Meas.*, 46(5), 527–532, May 2011.
49. D. Zhou, D. O'Sullivan, E. Semones, N. Zapp, S. Johnson, and M. Weyland, Radiation dosimetry for high LET particles in low Earth orbit, *Acta Astronaut.*, 63(7–10), 855–864, October 2008.
50. M. Hajek, T. Berger, N. Vana, M. Fugger, J. K. Pálfalvi, J. Szabó, I. Eördögh, Y. A. Akatov, V. V. Arkhangelsky, and V. A. Shurshakov, Convolution of TLD and SSNTD measurements during the BRADOS-1 experiment onboard ISS (2001), *Radiat. Meas.*, 43(7), 1231–1236, August 2008.
51. J. K. Pálfalvi, Y. Akatov, J. Szabó, L. Sajó-Bohus, and I. Eördögh, Detection of primary and secondary cosmic ray particles aboard the ISS using SSNTD stacks, *Radiat. Protect. Dosim.*, 120(1–4), 427–432, January 2006.
52. H. Ing and H. C. Birnboim, A bubble-damage polymer detector for neutrons, *Nucl. Tracks Radiat. Meas.* (*1982*), 8(1–4), 285–288, January 1984.
53. H. Ing, Neutron measurements using bubble detectors-terrestrial and space, *Radiat. Meas.*, 33, 275–286, 2001.
54. M. B. Smith, Y. Akatov, H. R. Andrews, V. Arkhangelsky, I. V. Chernykh, H. Ing, N. Khoshooniy et al., Measurements of the neutron dose and energy spectrum on the International Space Station during expeditions ISS-16 TO ISS-21, *Radiat. Protect. Dosim.*, July 2012.
55. B. J. J. Lewis, M. B. Smith, H. Ing, H. R. Andrews, R. Machrafi, L. Tomi, T. J. Matthews et al., Review of bubble detector response characteristics and results from space, *Radiat. Protect. Dosim.*, 150(1), 1–21, June 2012.
56. G. Badhwar, W. Atwell, G. Reitz, R. Beaujean, and W. Heinrich, Radiation measurements on the Mir Orbital Station, *Radiat. Meas.*, 35(5), 393–422, October 2002.
57. NCRP, Report No. 137: Fluence-based and microdosimetric event-based methods for radiation protection in space, Bethesda, MD, 2001.
58. B. B. Gersey, T. B. Borak, S. B. Guetersloh, C. Zeitlin, J. Miller, L. Heilbronn, T. Murakami, and Y. Iwata, The response of a spherical tissue-equivalent proportional counter to iron particles from 200–1000 MeV/nucleon, *Radiat. Res.*, 157(3), 350–360, March 2002.
59. T. Doke, T. Hayashi, J. Kikuchi, T. Sakaguchi, K. Terasawa, E. Yoshihira, S. Nagaoka, T. Nakano, and S. Takahashi, Measurements of LET-distribution, dose equivalent and quality factor with the RRMD-III on the Space Shuttle Missions STS-84, -89 and -91, *Radiat. Meas.*, 33(3), 373–387, June 2001.
60. D. Perez-Nunez and L. A. Braby, Replacement tissue-equivalent proportional counter for the International Space Station, *Radiat. Protect. Dosim.*, 143(2–4), 394–397, February 2011.
61. E. R. Benton, E. V. Benton, and A. L. Frank, Conversion between different forms of LET, *Radiat. Meas.*, 45(8), 957–959, September 2010.
62. T. Dachev, P. Dimitrov, B. Tomov, Y. Matviichuk, F. Spurny, O. Ploc, K. Brabcova, and I. Jadrnickova, Liulin-type spectrometry-dosimetry instruments, *Radiat. Protect. Dosim.*, 144(1–4), 675–679, March 2011.
63. T. Dachev and B. Tomov, Monitoring Lunar radiation environment: RADOM instrument on Chandrayaan-1, *Curr. Sci.*, 96(4), 25–28, 2009.
64. J. E. Mazur, W. R. Crain, M. D. Looper, D. J. Mabry, J. B. Blake, A. W. Case, M. J. Golightly, J. C. Kasper, and H. E. Spence, New measurements of total ionizing dose in the lunar environment, *Space Weather*, 9(7), July 2011.

65. D. Roth, P. McNulty, W. J. Beauvais, R. A. Reed, and E. G. Stassinopoulos, Solid state microdosimeter for radiation monitoring in spacecraft and avionics, *IEEE Trans. Nucl. Sci*., 41(6), 2118–2124, December 1994.
66. P. D. Bradley, A. B. Rosenfeld, and M. Zaider, Solid state microdosimetry, *Nucl. Instrum. Methods Phys. Res. Sect. B, Beam Interact. Mater. Atoms*, 184(1–2), 135–157, September 2001.
67. V. L. Pisacane, Q. E. Dolecek, H. Malak, F. A. Cucinotta, M. Zaider, A. B. Rosenfeld, A. Rusek, M. Sivertz, and J. F. Dicello, Microdosemeter instrument (MIDN) for assessing risk in space, *Radiat. Protect. Dosim*., 143(2–4), 398–401, February 2011.
68. A. Wroe, A. Rosenfeld, M. Reinhard, V. Pisacane, J. Ziegler, M. Nelson, F. Cucinotta, M. Zaider, and J. Dicello, Solid state microdosimetry with heavy ions for space applications, *IEEE Trans. Nucl. Sci*., 54(6), 2264–2271, December 2007.
69. A. L. Ziebell, W. H. Lim, M. I. Reinhard, I. Cornelius, D. A. Prokopovich, R. Siegele, A. S. Dzurak, and A. B. Rosenfeld, A cylindrical silicon-on-insulator microdosimeter: Charge collection characteristics, *IEEE Trans. Nucl. Sci*., 55(6), 3414–3420, December 2008.
70. *The Medipix Home Page*, CERN, n.d., http://medipix.web.cern.ch, Accessed August 8, 2013.
71. D. Turecek, L. Pinsky, J. Jakubek, Z. Vykydal, N. Stoffle, and S. Pospisil, Small dosimeter based on Timepix device for International Space Station, *J. Instrum*., 6(12), C12037, December 2011.
72. L. S. Pinsky, N. Stoffle, J. Jakubek, S. Pospisil, C. Leroy, A. Gutierrez, H. Kitamura et al., Penetrating heavy ion charge and velocity discrimination with a TimePix-based Si detector (for space radiation applications), *Nucl. Instrum. Methods Phys. Res. Sect. A: Accelerat., Spectrom., Detect. Associat. Equip*., 633(1), 226–229, May 2011.
73. L. S. Pinsky, A. Empl, S. Hoang, N. Stoffle, J. Jakubek, Z. Vykydal, D. Turecek et al., Preparing for the first Medipix detectors in space, in *2012 IEEE Aerospace Conf*., Big Sky, MT, pp. 1–6, 2012.
74. T. Holy, E. Heijne, J. Jakubek, S. Pospisil, J. Uher, and Z. Vykydal, Pattern recognition of tracks induced by individual quanta of ionizing radiation in Medipix2 silicon detector, *Nucl. Instrum. Methods Phys. Res. Sect. A: Accelerat., Spectrom., Detect. Associat. Equip*., 591(1), 287–290, June 2008.
75. M. A. R. Othman, M. Petasecca, S. Guatelli, J. Uher, D. G. Marinaro, D. A. Prok, M. I. Reinhard et al., Neutron dosimeter development based on medipix2, *IEEE Trans. Nucl. Sci*., 57(6), 3456–3462, December 2010.
76. E. B. Johnson, E. Chapman, P. Linsay, S. Mukhopadhyay, C. J. Stapels, J. F. Christian, and E. Benton, Tissue-equivalent solar particle dosimeter using CMOS SSPMs, *2009 IEEE Aerospace Conf*., Big Sky, MT, pp. 1–7, March 2009.
77. C. J. Stapels, E. B. Johnson, X. J. Chen, T. H. Prettyman, E. R. Benton, and J. F. Christian, Space neutron spectrometer design with SSPM-based instrumentation, *Nucl. Instrum. Methods Phys. Res. Sect. A: Accelerat., Spectrom., Detect. Associat. Equip*., 652(1), 342–346, October 2011.
78. I. Thomson, EVA dosimetry in manned spacecraft, *Mutation Res*., 430(2), 203–209, December 1999.
79. A. Hallil, M. Brown, Y. Akatov, V. Arkhangelsky, I. Chernykh, V. Mitrikas, V. Petrov et al., MOSFET dosimetry mission inside the ISS as part of the Matroshka-R experiment, *Radiat. Protect. Dosim*., 138(4), 295–309, March 2010.
80. M. G. Faruk, R. Wilkins, R. C. Dwivedi, D. Kalaria, M. Patel, S. Binzaid, and J. O. Attia, Proton and neutron radiation effects studies of MOSFET transistors for potential deep-space mission applications, in *2012 IEEE Aerospace Conf*., pp. 1–13, 2012.
81. J. Semkova, R. Koleva, G. Todorova, N. Kanchev, V. Petrov, V. Shurshakov, I. Tchhernykh, and S. Kireeva, Instrumentation for investigation of the depth-dose distribution by the Liulin-5 instrument of a human phantom on the Russian segment of ISS for estimation of the radiation risk during long term space flights, *Adv. Space Res*., 34(6), 1297–1301, January 2004.
82. D. E. Stilwell, W. D. Davis, R. M. Joyce, F. B. McDonald, J. H. Trainor, W. E. Althouse, A. C. Cummings, T. L. Garrard, E. C. Stone, and R. E. Vogt, The voyager cosmic ray experiment, *IEEE Trans. Nucl. Sci*., 26(1), 513–520, 1979.

83. C. Zeitlin, R. Maurer, D. Roth, J. Goldsten, and M. Grey, Development and evaluation of the combined ion and neutron spectrometer (CINS), *Nucl. Instrum. Methods Phys. Res. Sect. B: Beam Interact. Mater. Atoms*, 267(1), 125–138, January 2009.
84. R. Beaujean, J. Kappa, S. Burmeister, F. Petersen, and G. Reitz, Dosimetry inside MIR station using a silicon detector telescope (DOSTEL), *Radiat. Meas.*, 35, 433–438, 2002.
85. J. Semkova, R. Koleva, V. Shurshakov, V. Benghin, S. Maltchev, N. Kanchev, V. Petrov, E. Yarmanova, and I. Chernykh, Status and calibration results of Liulin-5 charged particle telescope designed for radiation measurements in a human phantom onboard the International Space Station, *Adv. Space Res.*, 40(11), 1586–1592, January 2007.
86. H. E. Spence, A. W. Case, M. J. Golightly, T. Heine, B. A. Larsen, J. B. Blake, P. Caranza et al., CRaTER: The cosmic ray telescope for the effects of radiation experiment on the Lunar reconnaissance orbiter mission, *Space Sci. Rev.*, 150(1–4), 243–284, January 2010.
87. Y. Charara, L. Townsend, H. Moussa, R. Hatcher, C. Dudney, S. McKee, P. Mckinnis, and K. Ottinger, Calculated energy loss spectra in the CRaTER detector for selected cosmic ray ions, *2007 IEEE Aerospace Conf.*, Big Sky, MT, pp. 1–12, 2007.
88. N. A. Schwadron, T. Baker, B. Blake, A. W. Case, J. F. Cooper, M. Golightly, A. Jordan et al., Lunar radiation environment and space weathering from the Cosmic ray telescope for the effects of radiation (CRaTER), *J. Geophys. Res.*, 117, 1–13, March 2012.
89. T. Sakaguchi, T. Doke, N. Hasebe, T. Hayashi, T. Kashiwagi, J. Kikuchi, S. Kono et al., LET distribution measurement with a new real-time radiation monitoring device-III onboard the Space Shuttle STS-84, *Nucl. Instrum. Methods Phys. Res. Sect. A: Accelerat., Spectrom., Detect. Associat. Equip.*, 437(1), 75–87, November 1999.
90. T. Doke, T. Fuse, K. Hara, T. Hayashi, J. Kikuchi, S. Suzuki, and K. Terasawa, Measurement of linear energy transfer distribution at CERN-EU high-energy reference field facility with real-time radiation monitoring device III and its comparison with dosimetric telescope, *Jpn. J. Appl. Phys.*, 43(6A), 3576–3581, June 2004.
91. V. Bidoli, M. Casolino, E. D. Grandis, M. P. D. Pascale, G. Furano, A. Morselli, L. Narici et al., In-flight performance of SilEye-2 experiment and cosmic ray abundances inside the Mir space station, *J. Phys. G: Nucl. Particle Phys.*, 27(10), 2051–2064, October 2001.
92. L. Narici, Heavy ions light flashes and brain functions: Recent observations at accelerators and in spaceflight, *New J. Phys.*, 10(7), July 2008.
93. V. Zaconte, F. Belli, V. Bidoli, M. Casolino, L. D. Fino, L. Narici, P. Picozza et al., ALTEA: The instrument calibration, *Nucl. Instrum. Methods Phys. Res. Sect. B: Beam Interact. Mater. Atoms*, 266(9), 2070–2078, May 2008.
94. L. Di Fino, M. Casolino, C. De Santis, M. Larosa, C. La Tessa, L. Narici, P. Picozza, and V. Zaconte, Heavy-ion anisotropy measured by ALTEA in the International Space Station, *Radiat. Res.*, 176(3), 397–406, 2011.
95. K. Lee, J. Flanders, E. Semones, T. Shelfer, and F. Riman, Simultaneous observation of the radiation environment inside and outside the ISS, *Adv. Space Res.*, 40(11), 1558–1561, January 2007.
96. G. D. Badhwar, Radiation measurements on the International Space Station, *Phys. Med.* 17(Suppl 1), 287–291, January 2001.
97. G. D. Badhwar, Martian radiation environment experiment (MARIE), *Space Sci. Rev.*, 110(1/2), 131–142, 2004.
98. C. Zeitlin, T. Cleghorn, F. Cucinotta, and P. Saganti, Overview of the Martian radiation environment experiment, *Adv. Space*, 33(12), 2204–2210, 2004.
99. R. S. Saunders, R. E. Arvidson, G. D. Badhwar, W. V. Boynton, P. R. Christensen, F. A. Cucinotta, W. C. Feldman et al., 2001 Mars odyssey mission summary, *Space Sci. Rev.*, 110(1/2), 1–36, 2004.
100. J. O. Goldsten, E. A. Rhodes, W. V. Boynton, W. C. Feldman, D. J. Lawrence, J. I. Trombka, D. M. Smith et al., The MESSENGER gamma-ray and neutron spectrometer, *Space Sci. Rev.*, 131(1–4), 339–391, November 2007.

101. D. M. Hassler, C. Zeitlin, R. F. Wimmer-Schweingruber, S. Böttcher, C. Martin, J. Andrews, E. Böhm et al., The radiation assessment detector (RAD) investigation, *Space Sci. Rev.*, 170(1–4), 503–558, July 2012.
102. J. Köhler, B. Ehresmann, C. Martin, E. Böhm, A. Kharytonov, O. Kortmann, C. Zeitlin, D. M. Hassler, and R. F. Wimmer-Schweingruber, Inversion of neutron/gamma spectra from scintillator measurements, *Nucl. Instrum. Methods Phys. Res. Sect. B: Beam Interact. Mater. Atoms*, 269(22), 2641–2648, November 2011.
103. E. Benton, E. Benton, and A. Frank, Neutron dosimetry in low-earth orbit using passive detectors, *Radiat. Meas.*, 33(3), 255–263, June 2001.
104. G. D. Badhwar, J. E. Keith, and T. F. Cleghorn, Neutron measurements onboard the space shuttle, *Radiat. Meas.*, 33(3), 235–241, June 2001.
105. V. I. Tret'yakov, I. G. Mitrofanov, Y. I. Bobronitskii, A. V. Vostrukhin, N. A. Gunko, A. S. Kozyrev, A. V. Krylov et al., The first stage of the 'BTN-Neutron' space experiment onboard the Russian segment of the International Space Station, *Cosmic Res.*, 48(4), 285–299, August 2010.
106. I. G. Mitrofanov, A. Bartels, Y. I. Bobrovnitsky, W. Boynton, G. Chin, H. Enos, L. Evans et al., Lunar exploration neutron detector for the NASA lunar reconnaissance orbiter, *Space Sci. Rev.*, 150(1–4), 183–207, December 2009.

75

Capacitance Sensors for Nondestructive Moisture Determination in Agricultural Products and Biofuel Materials

Chari V. Kandala
United States Department of Agriculture

Naveen Puppala
New Mexico State University

75.1 Introduction

Previous research showed that the dielectric properties of hygroscopic materials can be used to estimate the MC present in these materials (Nelson, 1977; Klein, 1981). The variation in dielectric constant with MC for shelled yellow field corn (Figure 75.1) was found to be more pronounced between 1 and 5 MHz (Nelson, 1978). Thus, $(\varepsilon_{r1} - \varepsilon_{r2})$, the difference in the dielectric constants at 1 and 5 MHz or any other higher frequency, should be a good indicator of the moisture present in the material. The difference in capacitance of a parallel-plate system of plate area A and separation d at two frequencies can be written as

$$C_1 - C_2 = (\varepsilon_{r1} - \varepsilon_{r2}) \left(\frac{\varepsilon_0 A}{d} \right) \tag{75.1}$$

where

ε_{r1} and ε_{r2} are the dielectric constants of the material between the plates at the two frequencies
ε_0 is the permittivity of free space (8.854×10^{-12} F/m)

It was found earlier that though $(C_1 - C_2)$ was a good estimate of the MC, it was highly influenced by the sample size of the peanuts held between the plates (Kandala and Nelson, 1990). The volume of space that a sample of odd-shaped material, such as in-shell peanuts, occupies between two parallel plates

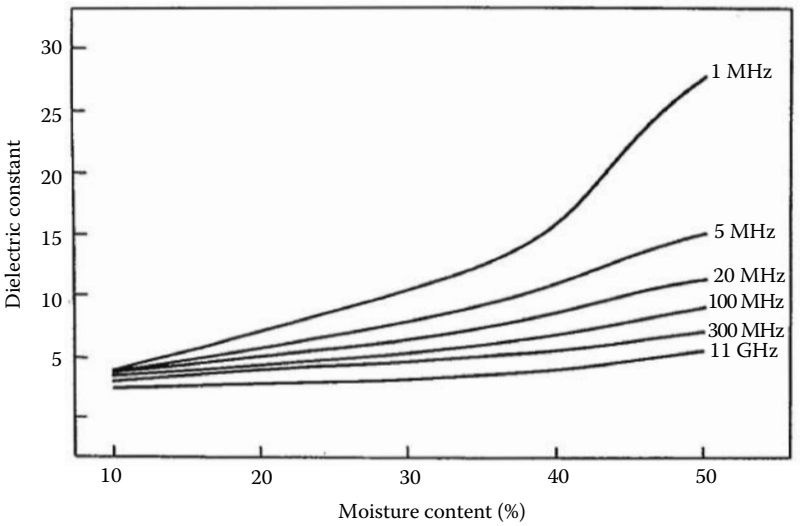


FIGURE 75.1 Moisture dependence of the dielectric constant of shelled yellow-dent field corn. (From Nelson, S.O., *J. Microwave Power*, 12(1), 67, 1977.)

varies each time they are placed between the plates. Air gaps between the peanut pods, and between the pods and the capacitor walls, occur differently. Variations in two other related electrical parameters, phase angle (θ) and dissipation factor (D), were also measured at these two frequencies. The capacitance and dissipation factor increased with moisture but were lower at higher frequencies. The phase angle decreased with moisture but was higher at the higher frequencies. Empirical equations were developed to estimate MC of corn, wheat, shelled peanut, and in-shell peanut samples from the differences in capacitance ($C_1 - C_2$), phase angle ($\theta_1 - \theta_2$), and dissipation factor ($D_1 - D_2$) at 1, 5, and/or 9 MHz without the need to weigh the samples (Kandala and Nelson, 2007). This minimized the errors due to the void space within the sample.

Dissipation factor is a measure of energy loss that results from subjecting a dielectric to an alternating current electric field. It is related to the Q factor of the peanut material and is a measure of the energy stored in the electric field relative to energy dissipated in any one period. The power dissipated depends on the equivalent resistance of the complex circuit, and thus the variation in the impedance values ($Z_1 - Z_2$) of the system at the two frequencies adequately represents the term ($D_1 - D_2$). Thus, from the values of ($C_1 - C_2$), ($\theta_1 - \theta_2$), and ($Z_1 - Z_2$) measured with a two parallel-plate system fitted inside a cylinder, the MC of the grain and nut samples could be estimated to an acceptable accuracy (Kandala et al., 2008). For a single parallel-plate system, the following semiempirical equation worked well:

$$\text{MC} = A_0 + A_1(C_1 - C_2) + A_2(\theta_1 - \theta_2) + A_3(Z_1 - Z_2) + A_4(C_1 - C_2)^2 + A_5(\theta_1 - \theta_2)^2 + A_6\,(Z_1 - Z_2)^2 \quad (75.2)$$

where

C_1, θ_1, Z_1 and C_2, θ_2, Z_2 are the capacitance, phase angle, and impedance at 1 and 5 MHz (in most of the cases), respectively

$A_0 \ldots A_5$ are calibration constants

The calibration constants were evaluated by making measurements on several samples of known MC values within the required moisture range, and applying a least squares computation. Statistical significance of the terms used in Equation 75.2 was verified by the t statistic, requiring that the probability of a greater absolute value of t under the null hypothesis had to be less than or equal to 0.05 for each variable. The numerical values of the constants were found to be different for different types of grain and nuts.

75.2 Moisture Meters

An impedance meter called the CI meter (Chari's impedance meter) designed and developed by the first author was used to measure the two parameters, impedance (Z) and phase angle (θ). Three frequencies 1, 5, and 9 MHz are generated by crystal oscillators as shown in the block diagram (Figure 75.2) drawn for 1 MHz. The electronic circuits for the three frequencies are similar except for slight variations in the values of certain capacitors and resistors. These signals are applied to the parallel-plate electrode system (z) alternately, by switching through a multiplexer. Initially, at 1.0 MHz, the current, flowing through this system with an impedance Z, is fed into an op amp. The same current flows through the range resistor R_r. The output voltage of the op amp and the original 1 MHz signal from the oscillator are rectified and measured as e_{m1} and e_{r1}, respectively. The current through Z is calculated as e_{m1}/R_r and the magnitude of the impedance of the parallel-plate system with the peanut kernel between them is obtained as

$$|Z_1| = R_r\left(\frac{e_{r1}}{e_{m1}}\right) \tag{75.3}$$

The phase angle at 1 MHz is determined by comparing the signal emerging out of the op amp with that of the original signal, using a phase detector. However, the phase detector can compare signals of the same amplitude only. To keep the amplitude of the output signal from the op amp constant and equal to the amplitude of the original signal, a comparator is used. The comparator would output a square wave, and a filter is used to convert it to a sine wave. The original signal from the oscillator is attenuated to the same amplitude as this signal and the two signals are fed into the phase detector. The phase detector compares the two signals and gives an output voltage e_{p1} proportional to the phase angle θ_1 between the two. The computer then switches the multiplexer to allow the 5 MHz signal to pass through the parallel-plate system. The signals are processed through a circuit similar to the 1 MHz circuit but with a range resistor of a different value. The impedance magnitude $|Z_2|$ and the phase angle θ_2 are determined at this frequency as was done for 1 MHz. Similar measurements were made with 9 MHz after switching to this frequency through the multiplexer and the values of impedance $|Z_3|$ and phase angle θ_3 are determined at this frequency. From the values of Z and θ, the real and imaginary parts of the impedance R and X at

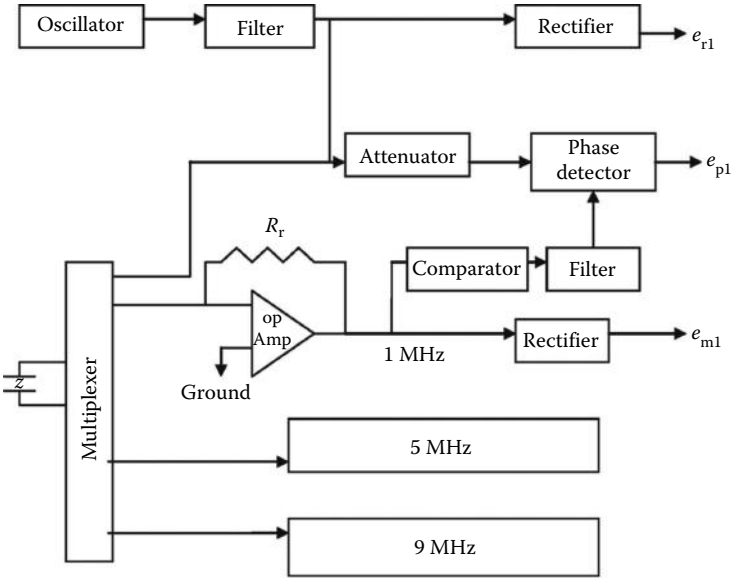


FIGURE 75.2 Block diagram of the CI meter electronic circuit to measure impedance and phase angle.

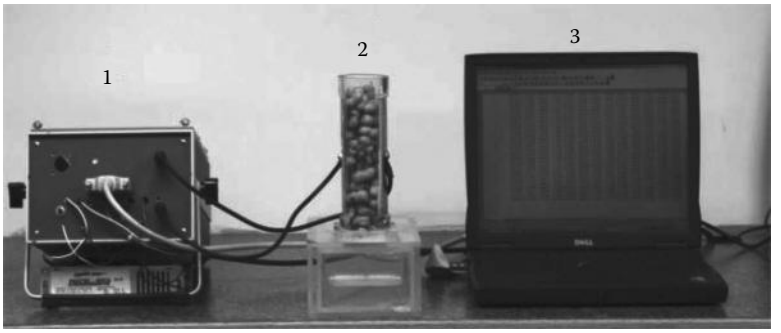


FIGURE 75.3 CI meter measuring system: (1) CI meter, (2) cylinder with electrodes filled with in-shell peanuts, and (3) computer.

each frequency are calculated as $R = |Z|\cos\theta$ and $X = |Z|\sin\theta$. The value of capacitance C of the parallel-plate system with the peanut sample between them is given as

$$C = -\frac{1}{2\pi fX} \tag{75.4}$$

The layout of the CI meter measuring system is shown in Figure 75.3. The power supply consists of two 12 V rechargeable lead–acid batteries from which the voltages, required to operate the circuits, were derived.

A laptop computer was used to register data from the system, compute the calibration constants, and calculate the moisture content.

A cylindrical acrylic tube, fitted with two sets of parallel-plate electrodes, serves as the test cell for moisture sensing. The test cell is 190 mm long, with an internal diameter of 50 mm and has a wall thickness of 7 mm. The electrode assembly for the test cell consists of two rectangular aluminum plates, 140 mm long and 50 mm wide. The electrodes are cemented to the inner walls of the test cell, as shown in Figure 75.1, about 25 mm from the ends. The distance of separation between the parallel plates for the test cell was 42 mm. The test cell would rest on top of a rectangular acrylic box and in a circular hole, centered about 35 mm from the front side of the box. This box is provided with an acrylic drawer that slides in and out of the box manually. The top of this drawer was covered 75 mm from the front end, with an acrylic plate. When the drawer is pushed all the way into the box, this plate covers the hole and prevents the sample placed in the cylinder from dropping into the drawer. Except for the electrodes, no metal parts were used in the assembly of the electrode system or the sample collecting system to prevent any interaction with the RF signal used in the measurements. With the drawer pushed all the way in, the cylinder was filled with the grain samples and the impedance measurements were taken. After the completion of the measurements, the drawer is pulled out slowly, allowing the sample to fall into the drawer. The drawer was emptied before another sample was placed in the cylinder for measurement. With the grain sample occupying the space between the electrodes, the analyzer measured the impedance and phase angle of this electrode system at 1, 5, and 9 MHz, and a computer controlled and collected the data from the analyzer. A Mettler* AE 163 electronic balance was also interfaced to this computer to record the wet and dry weights of the peanut pod samples to obtain their air-oven moisture values. Using these measurements in an empirical equation, the computer was programmed to calculate the moisture content of the sample.

* Mention of company or trade names is for purpose of description only and does not imply endorsement by USDA.

75.3 Moisture Measurement of Agricultural Products

75.3.1 Moisture Measurements of In-Shell Peanuts

Peanuts of the Georgia Green cultivar, harvested in 2007, dried, cleaned, and stored at 4 °C at the National Peanut Research Laboratory, were used for these studies (Kandala and Sundaram, 2010). The initial moisture content* of these peanuts was about 7%. These peanuts were divided into 12 sublots and were placed in glass jars. Appropriate quantities of water were added to each jar, to raise the moisture levels in steps of about 2% moisture, to obtain a total of 11 moisture levels ranging between 8% and 24%. The jars were sealed and allowed to equilibrate at 4 °C for 1 week. The jars were periodically rotated during this period to aid uniform distribution of moisture. Thus, 12 moisture levels with nominal MC values of 8, 9, 11, 12, 13, 15, 16, 17, 18, 19, 20, and 23 were available for calibration and validation. The jars were removed from cold storage and the pods were allowed to reach room temperature in the jars before the measurements were made.

The peanut pod samples, after conditioning, were separated into two groups. The calibration group consisted of samples taken from the nominal MC levels of 8%, 11%, 13%, 16%, 18%, 19%, and 23% and the validation group was from nominal MC levels of 9%, 12%, 15%, 17%, and 20%. The MC values of the samples in each of these 12 moisture levels (called the sublots) were obtained by the standard air-oven method (ASAE Standards, 2003). Three samples, each weighing about 100 g, were taken from each sublot and placed in small aluminum containers and their wet weights were taken. The containers were placed in a hot-air oven at 130 °C for 4–6 h depending on their initial MC value. At the end of the heating period, the containers were removed and weighed again to determine the dry weight of the samples. Moisture content was determined for each sample as the percentage ratio of the weight loss to the original wet weight of the sample. The average of the three moisture values, for each sublot thus obtained, were labeled as the oven MC value of that sublot. Impedance measurements were made on 30 samples from each sublot. Peanut samples were transferred from the jars into the cylindrical electrode system, until the space between the two plates of the cylinder was completely filled. The cylinder accommodated about 100–150 g of peanut pods. The room temperature during the measurements varied from 21 °C to 23 °C. Peanut pods from one of the jars were transferred into the cylinder with the drawer sitting fully inside the box, until the space between the two plates of the cylinder was filled. In this position, measurements of impedance (Z) and phase angle (θ) were made at 1 and 5 MHz, using the CI meter. The sample was then collected in the drawer by gently pulling it out and tapping on the cylinder for the peanuts to drop down. The drawer was emptied and reset in the box. The procedure was repeated on all peanut samples (sublot) in the rest of the jars.

From the measured values of impedance and phase angle the capacitance value for each sample was obtained using Equation 75.4. From the capacitance value, measured values of impedance and phase angle, and the oven determined MC value of the samples in the calibration lots, the semiempirical equation developed by multilinear regression (MLR) model using statistical software (*The Unscrambler*) was

$$MC_{peanut} = -9.71 + 0.56(\theta_2 - \theta_1) - 3.71(C_2 - C_1) - 25.13(Z_2 - Z_1) \tag{75.5}$$

where

C_1, θ_1, and Z_1 are the capacitance, phase angle, and impedance values at 1 MHz
C_2, θ_2, and Z_2 are the capacitance, phase angle, and impedance values at 5 MHz

Using Equation 75.5, the MC of each pod sample in the seven calibration lots was calculated, averaged over the 30 samples in each MC level, and the results are shown in Table 75.1 along with the MC values determined by the air-oven method. The standard deviations and the difference between the oven

* Moisture contents are expressed in percent wet basis throughout this chapter.

TABLE 75.1 Comparison of Mean MC Values Determined by the Air-Oven and CI Meter Methods for Calibration Groups of Peanuts (Average of 30 Samples in Each Group)

Oven, % MC	CI Meter, % MC	Standard Deviation	Difference	Predictability, %
8.01	8.13	0.08	−0.12	100
11.05	11.17	0.10	−0.14	93
13.55	13.96	0.29	−0.41	93
16.12	15.40	0.51	0.72	88
17.81	17.79	0.01	0.02	93
19.15	19.80	0.46	−0.65	92
22.94	22.37	0.40	0.57	90

TABLE 75.2 Fitness Measures for the Calibration Set of Peanuts

R^2	RMSEC	SEC	Bias ($\times 10^{-6}$)
0.97	0.82	0.82	−2.59

and calculated MC values (difference) are also shown. The predictability shown in the last column of Tables 75.1 and 75.2 is defined as the percentage of samples for which moisture content was calculated within 1% of the air-oven MC values. An R^2 value of 0.97, a low value of SEC (shown in Table 75.2), relatively small differences in the calculated and the oven MC values, and acceptable standard deviation values suggest the suitability of Equation 75.5 for MC predictions. The standard error of calibration (SEC)* was 0.82 as shown in Table 75.2 which contains fitness measures for the calibration set. The mean calculated values agree very closely with the oven MC values.

The MC values of the pod samples in the five validation lots calculated by using Equation 75.5 are shown in Table 75.3. The calculated mean moisture values were compared with the air-oven values and the results are summarized in Table 75.3, along with standard deviations, difference, and predictability. The predictability was better than 95% for all the levels tested. The errors in individual samples in a moisture level could be due to nonhomogeneity of moisture distribution in the sample. The standard error of prediction (SEP)† was 0.70 as shown in Table 75.4. The bias value of −0.35, close to zero, indicates the closeness of the mean calculated values to the standard values confirming the fitness of the prediction Equation 75.3 (Moore, 2000).

Comparison of the MC values determined by the air-oven and CI meter methods, using the parallel-plate cylindrical sample holder, is shown in Table 75.3. Note that the average values predicted by the

* $$\text{SEC} = \left(\frac{1}{n-p-1} \sum_{i=1}^{n} e_i^2 \right)^{\frac{1}{2}}$$

where
n is the number of observations
p is the number of variables in the regression equation with which the calibration is performed
e_i is the difference between the observed and reference value for the ith observation

† $$\text{SEP} = \left(\frac{1}{n-1} \sum_{i=1}^{n} (e_i - \bar{e})^2 \right)^{\frac{1}{2}}$$

where
n is the number of observations
e_i is the difference in the moisture content predicted and that determined by the reference method for the ith sample
$\bar{e}$ is the mean of e_i for all of the samples

TABLE 75.3 Comparison of Mean MC Values Determined by the Air-Oven and the CI Meter Methods for Validation Groups of Peanuts

Oven, % MC	CI Meter, % MC	Standard Deviation	Difference	Predictability, %
9.2	8.42	0.55	0.78	96
12.35	11.71	0.46	0.64	96
14.84	14.40	0.31	0.44	96
16.76	16.52	0.17	0.24	100
19.98	20.39	0.29	−0.41	95

TABLE 75.4 Fitness Measures for the Validation Set of Peanuts

R^2	RMSEC	SEP	Bias
0.96	0.78	0.70	−0.35

CI meter agreed well with the air-oven values. The time taken for one sample measurement was ~1 min as compared to several hours by the gravimetric methods. There is no need to shell, weigh, or measure the volume of the peanuts in this method.

75.3.2 Moisture Measurements of Corn

Yellow-dent field corn samples harvested in 2007 from Nebraska were used in this study (Kandala and Sundaram, 2010). The initial moisture content of the corn lot, as determined with an electronic moisture meter, was about 6%. This corn lot was divided into 12 sublots and was placed in glass jars. Similarly, as done for peanuts, appropriate quantities of water were added to each jar, to raise the moisture levels in steps of about 1% moisture, to obtain a total of 12 moisture levels ranging between 6% and 20%. Thus, 12 moisture levels with nominal MC values of 6, 7, 8, 9, 10, 11, 13, 15, 16, 17, 18, and 19 were available for calibration and validation.

From the CI meter measured values of impedance and phase angles, the capacitance value for each sample was obtained using Equation 75.4. Using the capacitance value, measured values of impedance and phase angle, at the frequencies 1 and 5 MHz, and the oven-determined MC value of the samples in the calibration lots, a MLR model was developed as shown below:

$$MC_{corn} = -8.395 + 1.17(C_2 - C_1) - 0.89(\theta_2 - \theta_1) + 15.90(Z_2 - Z_1) \tag{75.6}$$

where

C_1, θ_1, and Z_1 are the capacitance, phase angle, and impedance values at 1 MHz

C_2, θ_2, and Z_2 are the capacitance, phase angle, and impedance values at 5 MHz

Using the MLR Equation 75.6, MC values of the samples in both calibration and validation sets were determined. The fitness measures for the calibration set are shown in Table 75.5. The calibration set had an R^2 value of 0.98 and the SEC was 0.61.

TABLE 75.5 Fitness Measures for the Calibration Set of Corn

R^2	RMSEC	SEC	Bias ($\times 10^{-6}$)
0.98	0.61	0.61	+0.19

TABLE 75.6 Comparison of Average MC of Corn Determined by the CI Meter and the Air-Oven Methods (Calibration Set)

Oven, % MC	CI Meter, % MC	Standard Deviation	Difference	Predictability, %
6.85	7.25	0.28	−0.40	100
8.90	8.68	0.16	0.22	100
11.00	11.01	0.003	−0.01	100
13.22	13.48	0.18	−0.26	97
15.67	15.18	0.35	0.49	100
17.86	17.91	0.04	−0.05	90
18.21	18.21	0.002	−0.002	97

The average MC values (of 30 samples) at each moisture level for the calibration set, as calculated by the MLR Equation 75.6 from the CI meter measurements, along with the corresponding air-oven moisture values, their standard deviations, and the difference between the two moisture values, are shown in Table 75.6.

The predictability shown in Tables 75.6 and 75.8 are the percentage of samples in each moisture group whose values are predicted within 1% of their air-oven values. For the validation set, the fitness measures are shown in Table 75.7. The validation set had an R^2 value of 0.96 and the SEP was 0.70. Here too, a low bias value of 0.15 indicates the closeness of the mean calculated values to the standard values confirming the fitness of the prediction (Equation 75.6).

The average MC values (30 samples) of each of the six moisture levels of the validation group (which are not used in the calibration) as predicted by Equation 75.6 from the CI meter measurements are shown in Table 75.8 along with their standard air-oven values. Also shown are the standard deviations, the difference between the two moisture values, and the predictabilities. The SEP is very close to the SEC value and both the calibration and the validation lots have a good R^2 value. The predictability for the validation sets are above 93% for any moisture level and averaged 98.6% over all the levels.

The overall predictability for peanuts (Table 75.3) was 96.6% while for corn it was 98.6% (Table 75.8). Variations in temperature and relative humidity, while measurements were made, may have caused some variations in the predicted values of the MC. The oil content and protein in the peanuts might have some influence on the predicted MC values. Including temperature as a variable in the prediction equations for both the crops could improve the predictability. In the case of peanuts, measuring the total oil content (TOC) of peanuts and including the TOC also, as a variable in their prediction equation, could further improve the MC predictions.

TABLE 75.7 Fitness Measures for the Validation Set of Corn

R^2	RMSEC	SEP	Bias
0.96	0.71	0.70	+0.15

TABLE 75.8 Comparison of Average MC of Corn Determined by the CI Meter and the Air-Oven Methods (Validation Set)

Oven, % MC	CI Meter, % MC	Standard Deviation	Difference	Predictability, %
7.34	7.65	0.22	0.31	100
8.45	8.15	0.22	−0.30	93
9.44	9.55	0.08	0.11	100
15.5	15.49	0.004	−0.006	100
17.32	16.99	0.23	−0.33	100

75.3.3 Moisture Measurements of Wheat

A total of six varieties of wheat, planted and harvested around the Texas panhandle and at the New Mexico State University Station near Clovis, were used in this study (Kandala et al., 2012). The wheat varieties are Tam111, Duster, Scoutt 66, Endurance, Jagger, and Hatcher (Brent, 2011), planted during October 2010 and harvested during July 2011. All the sample lots were stored at 4 °C and 40% relative humidity. The MC of the samples initially was about 9%. These lots were divided into 22 sublots and were placed in glass jars. Appropriate amounts of water were added to the samples to develop 22 moisture levels between 9% and 25%. Thus, from the available quantities, six MC levels of Tam111, five of Duster, four of Hatcher, five of Jagger, and one each of Scoutt 66 and Endurance were developed for calibration and validation. Moisture content of each subsample was determined by the standard air-oven method by drying triplicate samples of 10 g each for 19 h at 130 °C (ASAE Standards, 1994). Samples were weighed before and after drying and the MC of each sample was obtained as the percentage ratio of the weight loss to the original wet weight of the sample.

Impedance measurements were made with the CI meter on 30 samples from each sublot. From the measured values of C, θ, and Z on the calibration sublots with the CI meter, and using unscrambler procedures for regression analysis, the values obtained for the constants $A_0 \ldots A_6$ in Equation 75.2 were

$$A_0 = -27.235, \quad A_1 = 0.0032, \quad A_2 = 4.734, \quad A_3 = 20.476, \quad A_4 = -0.000088, \quad A_5 = -0.405, \quad A_6 = -2.694$$

These constants along with the values of impedance, phase angle, and capacitance (obtained from Equation 75.4) were used in Equation 75.2 to calculate the MC of each of the 30 samples from the calibration sublots. The calculated values were averaged over the 30 samples in each moisture group and were compared with their air-oven values, and are shown in Table 75.9 along with their standard deviations, differences, and predictability.

Shown in Table 75.10 are the fitness measures for the calibration lots. The SEC was 0.49% MC (with $p = 6$ and $n = 210$). The predictability was good at all the MC levels. The calculated values averaged over 30 samples from each moisture group agreed well with their air-oven values, and the predictability ranged from 97% to 100%. The standard deviation for 30 samples at any moisture level was under 0.5%. An R^2 value of 0.99, an SEC of 0.49, a low bias value, and a high rate of predictability suggest that this model is suitable for MC predictions. This was further verified from the fitness measures obtained for the validation sets.

TABLE 75.9 Calibration Sublots: Comparison of CI Meter and Air-Oven MC Measurements

	Moisture Content (%)				
		Predicted by CI Meter and Equation 75.2			
Wheat Variety[a]	Oven Value	Mean[b]	Standard Deviation	Difference (Oven—Predicted)	Predictability[c]
Tam111	9.11	9.29	0.45	−0.18	100
Duster	11.33	11.65	0.23	−0.32	100
Duster	13.56	13.04	0.28	0.52	100
Tam111	16.67	16.38	0.23	0.29	100
Scoutt 66	19.78	20.43	0.26	−0.65	97
Endurance	22.67	22.38	0.30	0.29	100
Jagger	24.67	24.63	0.16	0.04	100

[a] All from 2010 harvest.

[b] Mean of 30 sample measurements.

[c] Predictability is the % of samples in each moisture group for which the predicted MC is within 1% of the air-oven value.

TABLE 75.10 Fitness Measures for the Calibration Set of Wheat Samples

R^2	RMSEC	SEC	Bias
0.99	0.49	0.49	0.001

TABLE 75.11 Validation Sublots: Comparison of CI Meter and Air-Oven MC Measurements

	Moisture Content (%)				
		Predicted by CI Meter and Equation 75.2			
Wheat Variety	Oven Value	Mean[a]	Standard Deviation	Difference (Oven-Predicted)	Predictability[b] (%)
Hatcher	9.33	9.68	0.18	−0.35	100
Hatcher	9.67	9.63	0.23	0.04	100
Hatcher	9.78	9.62	0.25	0.16	100
Duster	10.00	9.81	0.23	0.19	100
Tam111	10.67	10.30	0.23	0.37	100
Hatcher	12.14	11.33	0.19	0.81	97
Tam111	12.89	12.11	0.28	0.78	87
Tam111	14.89	14.14	0.29	0.75	90
Jagger	16.67	16.89	0.20	−0.22	100
Duster	17.78	17.89	0.32	−0.11	100
Jagger	18.22	18.30	0.35	−0.08	100
Tam111	18.44	19.16	0.26	−0.72	93
Jagger	19.11	19.15	0.30	−0.04	100
Duster	20.00	20.56	0.24	−0.56	100
Jagger	23.56	23.69	0.25	−0.13	100

[a] Mean of 30 samples in each moisture group.

[b] Predictability is the % of samples in each moisture group for which the predicted MC is within 1% of the air-oven value.

Shown in Table 75.11 are the MC values for 15 validation sets consisting of three varieties used in the calibration, and an additional variety, Hatcher, not used in the calibration. The moisture levels of the three varieties used in the validation are different from the levels used in the calibration. The MC values were predicted using the measured values of impedance and phase angle, and the calibration constants in Equation 75.2, averaged over 30 samples in each group, and compared with their air-oven values. Also shown are the standard deviations and the predictability of each moisture group. The predictability was 87% or better for any moisture level and averaged about 98% over the 15 validation levels.

Shown in Table 75.12 are the fitness measures for the validation group of samples. The validation set also came up with a good R^2 value of 0.98 and the SEP* was 0.60. The SEP value is close to the SEC value, and the R^2 value compares well with that of the calibration group. A low bias indicates the closeness of the mean calculated values to the air-oven values confirming that the prediction model is dependable.

* $$\text{SEP} = \left(\frac{1}{n-1} \sum_{i=1}^{n} (e_i - \bar{e})^2 \right)^{\frac{1}{2}}$$

where

n is the number of observations

e_i is the difference in the moisture content predicted and that determined by the reference method for the ith sample

$\bar{e}$ is the mean of e_i for all of the samples

TABLE 75.12 Fitness Measures for the Validation Sublots

R^2	RMSEP	SEP	Bias
0.98	0.62	0.60	−0.146

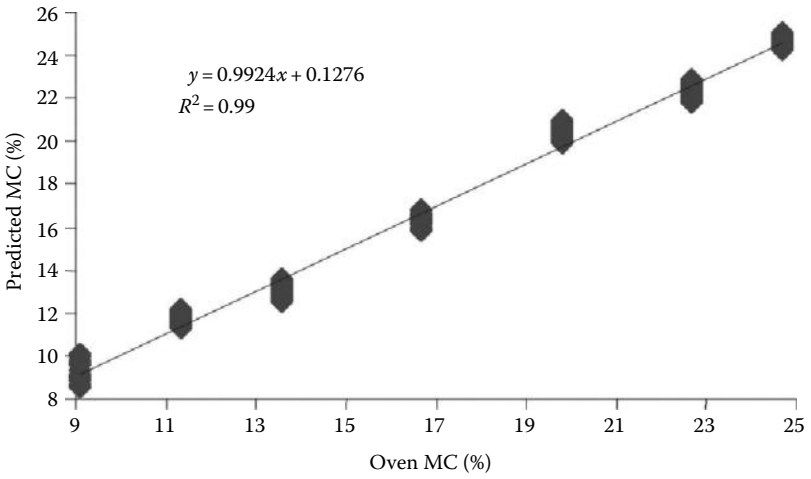


FIGURE 75.4 Comparison of oven and CI meter MC values for the wheat calibration lots.

The SEC value of 0.49 and the SEP value of 0.60 were better than the 0.88 for SEC and 0.91 for SEP obtained earlier with small wheat samples (Kandala et al., 1996). Shown in Figure 75.4 are the comparison of the predicted and the air-oven values of the seven calibration wheat lots. An R^2 value of 0.99 further indicates the fitness of the calibration equation. Similarly in Figure 75.5, the MC values of the validation samples of wheat are compared with their respective air-oven values and plotted. The predicted MC values compare well with their corresponding air-oven values with an R^2 value of 0.99. The CI meter predicted the MC values of different varieties quite well with a single calibration equation, developed with randomly picked MC levels from the different available varieties. Moreover, the measurement is rapid and nondestructive and there was no need to measure the volume or weight of the wheat samples.

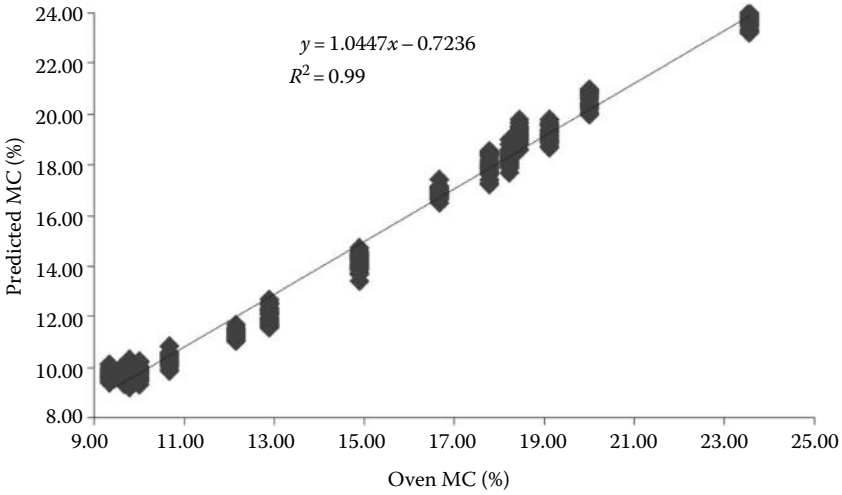


FIGURE 75.5 Comparison of oven and CI meter MC values for the wheat validation lots.

75.4 Moisture Measurement of Wood Chips and Powders

Measurement of MC of wood chips and related powders is important in the pulp and biofuel industries (Biermann, 1993; Nystrom and Dahlquist, 2004). Measurements of phase angle and impedance were made on pine and hard wood chips placing them between the parallel plates of the cylindrical sample holder, used above for grain and peanuts. With the phase angle and impedance values of the calibration samples, measured with the CI meter, and their known MC values, regression analysis was done and calibration equations were developed, separately, for pine and hard wood chips. Using these equations, MC of pine and hard wood samples in the moisture range between 5% and 50% was determined and compared with their air-oven values. The results are shown in Figures 75.6 and 75.7. For the seven moisture groups (30 samples in each group) of pine wood validation samples, the CI meter predicted the MC values close to their oven values with an R^2 value of 0.99 and an SEP of 1.37. Similarly, for seven moisture groups of hard wood samples used for validation the CI meter predicted the MC values close to the oven values with an R^2 of 0.99 and an SEP of 1.56.

Measurement of MC for biofuel materials in the form of powders was also done using the CI meter using a parallel-plate system. The plates were held horizontally with the lower plate held rigidly while the upper plate could slide up and down by about 20 mm. The sample powder was packed into a Petri dish about 20 mm diameter and 10 mm thick, and the lid was closed. The Petri dish was placed between the plates with both plates in contact with it, and measurements of phase angle and impedance were made with the CI meter. From these measured values of the calibration samples, and their MC values obtained by the standard air-oven method, regression equations were developed. Using the regression equations, MC predictions were made for the validation samples. Measurements were done for pine wood and switchgrass powders, in the MC range of 6%–25%, with the CI meter, and the results are shown in Figures 75.8 and 75.9.

The CI meter predicted well the MC of five moisture groups of pine wood powder, used for validation, with an R^2 value of 0.96 and an SEP of 1.45. Similarly, MC of five moisture groups of switchgrass powder predicted by the CI meter gave an R^2 value of 0.98 and an SEP value of 1.38, indicating that the CI meter is useful in rapidly and nondestructively predicting the MC of the biofuel materials either in the form of chips or in the powder form.

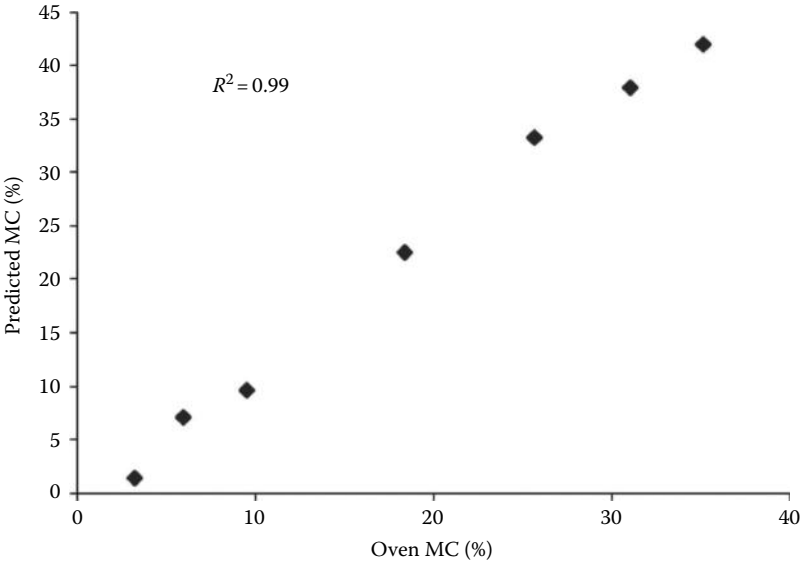


FIGURE 75.6 Comparison of oven and predicted MC values: pine chips.

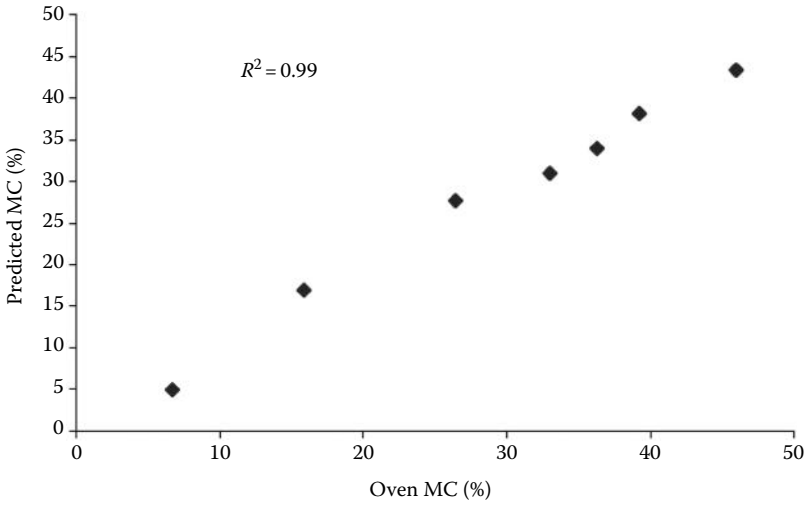


FIGURE 75.7 Comparison of oven and predicted MC values: hard wood chips.

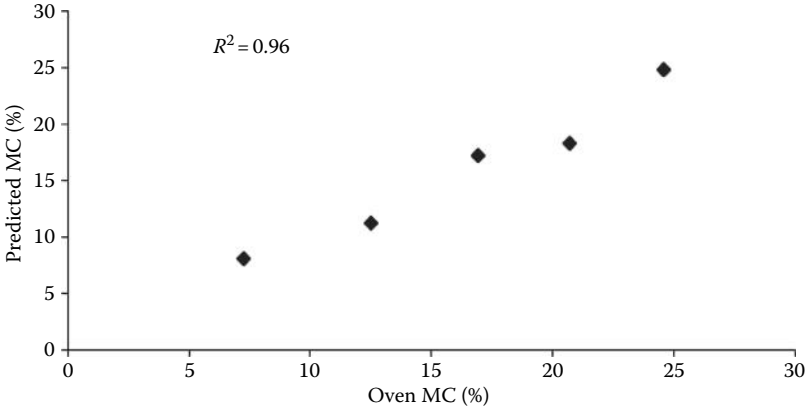


FIGURE 75.8 Comparison of oven and predicted MC values: pine wood powder.

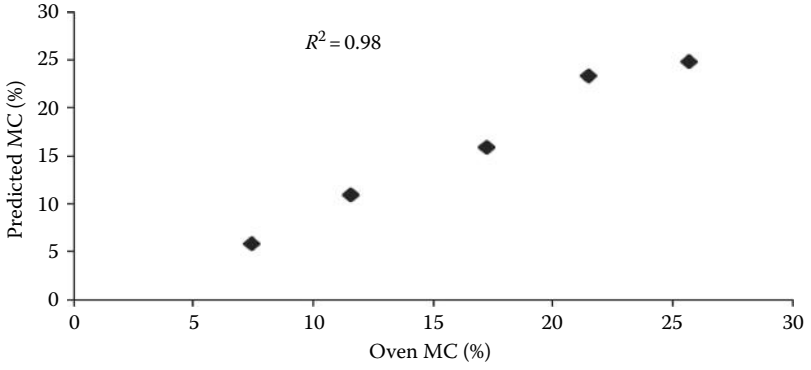


FIGURE 75.9 Comparison of oven and predicted MC values: switchgrass powder.

75.5 Conclusions

From the measurements of impedance and phase angle using a low-cost impedance meter and a capacitance sensor, the average moisture content of 150–200 g of in-shell peanuts, corn, and wheat samples could be determined. The weight or volume of the samples need not be measured. Ability to measure the MC of peanuts while they are in their shells would eliminate the need to shell them and clean them saving considerable time and labor. The MC of the corn samples tested was between 6% and 20% and the MC of the peanut samples was between 8% and 24% (most useful range for drying purposes in each case). The predicted MC values were within 1% of the air-oven values for over 98% of the samples tested for corn and 96% for in-shell peanuts. In the case of wheat, the moisture range of the samples used was between 9% and 25%. Calibration was done with five different wheat varieties, and validation was done on all these varieties and one additional variety. For 98% of the samples tested from the six varieties, the predicted MC values were within 1% of their air-oven values. This meter performed well for all the six varieties of wheat in predicting the MC with a single calibration equation. Using impedance measurements taken at two frequencies has helped in eliminating the errors due to air gaps that would occur randomly when the sample is filled between the plates. The CI meter would also be useful to rapidly and nondestructively determine the MC in materials used as biofuels such as pine and hard wood chips, and powders such as pine and switchgrass. This low-cost instrument would be useful for the grain, peanut, as well as pulp and biofuels industry for nondestructive and rapid measurement of moisture content using the principles of capacitance sensing.

References

ASAE Standards. 1994. *Moisture Measurement—Unground Grain and Seeds*, 41st edn., ASAE S352.2. American Society of Agricultural Engineering, St. Joseph, MI.

ASAE Standards. February 2003. *Moisture Measurements: Peanuts*, ASAE S410.1. American Society of Agricultural Engineering, St. Joseph, MI, pp. 604–605.

Biermann, C. J. 1993. *Essentials of Pulping and Paper Making*. Academic Press, San Diego, CA.

Brent, B. 2011. Wheat variety trials conducted in the Texas and New Mexico High Plains. Available at http://amarillo.tamu.edu/files/2011/11/wheat-variety-Discriptions

Kandala, C. V. K., C. L. Butts, and M. C. Lamb. 2008. Moisture content determination for in-shell peanuts with a low-cost impedance analyzer and capacitor sensor. *Trans. ASAE*, 51(4): 1377–1381.

Kandala, C. V. K. and S. O. Nelson. 1990. Measurement of moisture content in single kernels of peanuts: A nondestructive electrical method. *Trans. ASAE*, 33(2): 567–572.

Kandala, C. V. K. and S. O. Nelson. 2007. RF Impedance method for nondestructive moisture content determination for in-shell peanuts. *Meas. Sci. Technol.*, 18: 991–996.

Kandala, C. V. K., S. O. Nelson, and K. C. Lawrence. 1996. Determining moisture content in small wheat samples by dual-frequency RF impedance sensing. *Trans. ASABE*, 39(3): 1185–1188.

Kandala, C. V. K. and J. Sundaram. 2010. Nondestructive measurement of moisture content using a parallel-plate capacitance sensor for grain and nuts. *IEEE Sensors J.*, 10(7): 1282–1287.

Kandala, C. V. K., J. Sundaram, N. Puppala, and V. Settaluri. 2012. Nondestructive measurement of moisture content of different varieties of wheat using a single calibration with a parallel-plate capacitance sensor. *Trans. ASAE*, 55(4): 1–5.

Klein, A. 1981. Microwave determination of moisture in coal—A comparison of attenuation and phase measurement. *J. Microwave Power*, 16(3–4): 289–304.

Moore, D. S. 2000. *The Basic Practice of Statistics*, 2nd edn. W.H. Freeman & Co., New York.

Nelson, S. O. 1977. Use of electrical properties for grain-moisture measurement. *J. Microwave Power*, 12(1): 67–72.

Nelson, S. O. 1978. Frequency and moisture dependence of the dielectric properties of high-moisture corn. *J. Microwave Power*, 13(2): 213–218.

Nystrom, J. and E. Dahlquist. 2004. Methods for determination of moisture content in wood-chips for power plants—A review. *Fuel*, 83: 773–779.

The Unscrambler. Version 9.7. Camo Software Inc., Woodbridge, NJ.

76

Radio Location, Radio Navigation, and GPS Systems

Halit Eren
Curtin University

76.1 Introduction

Radio location is concerned with the use of radio waves for position sensing and locating objects. For this purpose, the radio signals can be used in an active or passive mode. In the passive mode, the characteristics of the radio signals received by a particular object situated in the field are used to determine the location. In the active case, the signals generated by an object yield the information about its location.

There are numerous radio location methods employed, being

- Backscattered properties waves, for example, radio-frequency identification (RFID), or radar
- Received signal strength (RSSI)
- Time of arrival (TOA)
- Differential TOA (DTOA) by receiving additional signals
- Angle of arrival (AOA)
- Phase differences and changes in frequencies, such as the Doppler shift
- Multilateration of multiple receivers that are receiving information from the same source
- Global Positioning Systems (GPS)
- Fingerprint and trilateration signals from the cellular telephony, such as TDMA, CSMA, and TSM
- Combinations of the earlier methods

Both active or passive radio location techniques are used in real-time locating systems (RTLS) for tracking and navigation purposes. There are many commercially available RTLS based on GPS, RFID, TOA, DTOA, and other techniques. Some of these will be explained in the rest of this chapter.

76.2 Radio Location Fundamentals

Radio location technology is used extensively in military applications, industrial and transportation systems, and general consumer applications. One of the most extensively employed technologies is radio navigation, which can be classified by its range, scope, error budget, and cost. The range classifications can be identified to be short, medium, and long range, within which exact limits are rather indefinite. The scope classifications can be either self-contained or externally supported, or active (transmitting) or passive (not transmitting) modes of operations.

Utilization of electromagnetic radio waves is the common denominator in all navigation systems discussed here. Particularly in the long-range applications, understanding of their behavior in the Earth's atmosphere is very important in the design, construction, and use of all kinds of navigation equipment—from the complex satellites and radar navigation and location systems to simple handheld devices that many consumers use in their cell phones. Therefore, fundamentals behind the electromagnetic wave propagation will be explained briefly.

When an FM radio wave is generated within the Earth's atmosphere, the wave travels outward. The waves may be absorbed or reflected from surfaces of materials they encounter. The absorption and scattering of electromagnetic waves take place for many reasons, one of which is caused by excitation of electrons within the molecules in the propagation media. The behavior of an electromagnetic wave is dependent on its frequency and corresponding wavelength. Figure 76.1 shows the frequency spectrum of electromagnetic waves. They are classified as *audible waves* at the lower end of the spectrum, *radio waves* from 5 kHz to 300 GHz, and *visible light* and various other types of rays at the upper end of the spectrum.

For practical purposes, the radio wave spectrum is broken into eight bands of frequencies; these are *very low frequency* (VFL) <30 kHz, *low frequency* (LF) 30–300 kHz, *medium frequency* (MF) 300 kHz–3 MHz, *high frequency* (HF) 3–30 MHz, *very high frequency* (VHF) 30–300 MHz, *ultrahigh frequency* (UHF) 300 MHz to 3 GHz, *super high frequency* (SHF) 3 to 30 GHz, and *extremely high frequency* (EHF) 30–300 GHz.

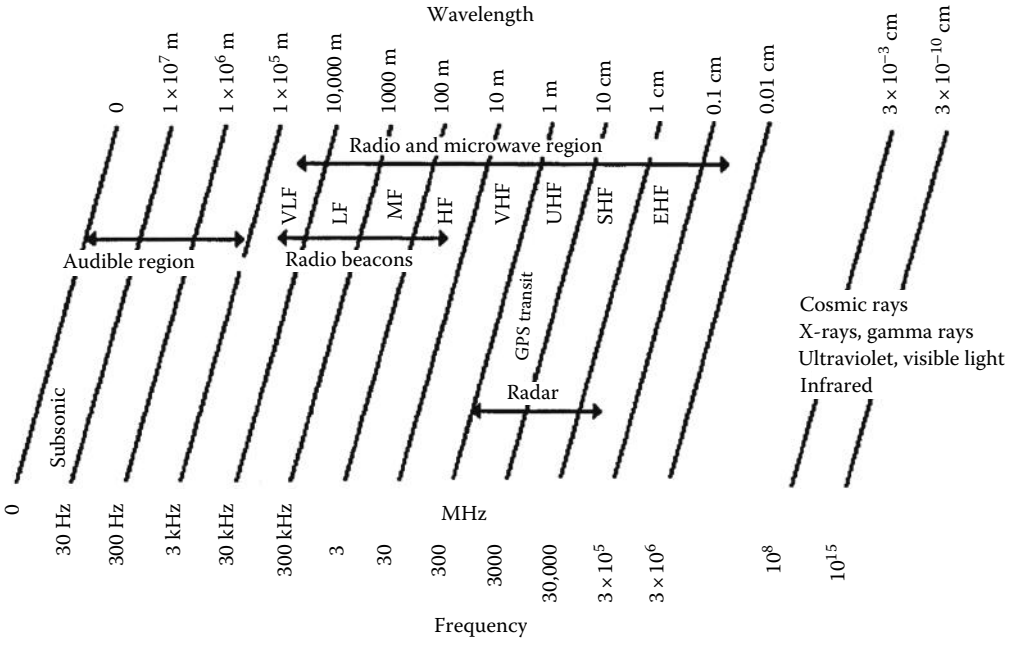


FIGURE 76.1 Electromagnetic wave frequency spectrum. Radio beacons operated in the VLF, LF, and MF ranges. GPS, Transit, and Glonass use UHF frequencies. Wavelengths <10 cm are not suitable for satellite systems, but they are used in radars.

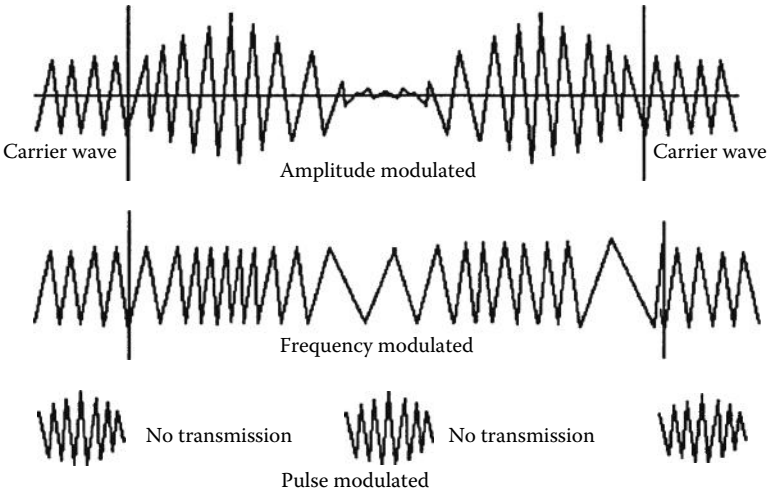


FIGURE 76.2 Amplitude, frequency, and pulse modulation of RF carrier waves. Amplitude modulation is suitable for broadcasting radio stations. Frequency modulation is used in commercial radio broadcasts. The pulse modulation is used in satellite systems, radars, and long-range navigation aids.

For easy identification, the frequencies above 1 GHz are further broken down by letter designators as L-band (1–2 GHz), S-band (2–4 GHz), C-band (4–8 GHz), X-band (8–12.5 GHz), and K-band (12.5–40 GHz). Since severe absorption of radar waves occurs near the resonant frequency of water vapor at 22.2 GHz, the K-band is subdivided into lower K-band (12.5–18 GHz) and upper K-band (26.5–40 GHz). Most navigation radars operate in the X- and S-bands, and many weapons fire control radars operate in the K-band range.

The radio waves are transmitted as continuous or modulated waves. A carrier wave (CW) is *modulated* to convey information in three basic forms: amplitude, frequency, and pulse modulation, as shown in Figure 76.2. The amplitude modulation (AM) modifies the amplitude of the CW with a modulating signal. In frequency modulation (FM), the frequency of the CW is altered in accordance with the frequency of the modulating wave. FM is used in commercial radio broadcasts and the sound portion of the television broadcast. Pulse modulation is different from AM and FM in that there is usually no impressed modulation wave employed. In this modulation, the continuous wave is broken up into very short bursts or "pulses," separated by periods of silence during which no wave is transmitted. This is used in satellite navigation systems, surface search radars, and long-range radio navigation aids.

When an electromagnetic wave encounters an obstruction, *diffraction* takes place marked by a weak reception of the signal within the "shadow" zone. Two waves acting on the same point will also result in *interference*. The degree of interference depends on the phase and frequency relationship. For example, two waves of the same frequency with a 180° phase difference will result in a null at that point. Also, under certain conditions, a portion of the electromagnetic energy in radio waves may reflect back toward the Earth's surface to form the *ionosphere*. The ionosphere is a layer of charged particles located about 90–400 km high from the Earth's surface; such reflected waves are called *sky waves*.

In the study of radio wave propagation, there are four *ionosphere layers* of importance, as shown in Figure 76.3. The D-layer is located about 60–90 km and is formed during daylight. The E-layer is about 110 km. It persists through the night with decreased intensity. The F_1-layer is between 175 and 200 km; it occurs only during daylight. The F_2-layer is between 250 and 400 km; its strength is greatest in the day but it combines with the F_1-layer later to form a weak F-layer after dark. The layers in the ionosphere are variable, with the pattern seeming to have diurnal, seasonal, and sunspot periods. The layers may be highly conductive or may entirely hinder transmissions, depending on the frequency of the wave, its angle of incidence, height, and intensity on various layers at the time of transmission. In general, frequencies in the MF and HF bands are most suitable for ionosphere reflections during both day and night.

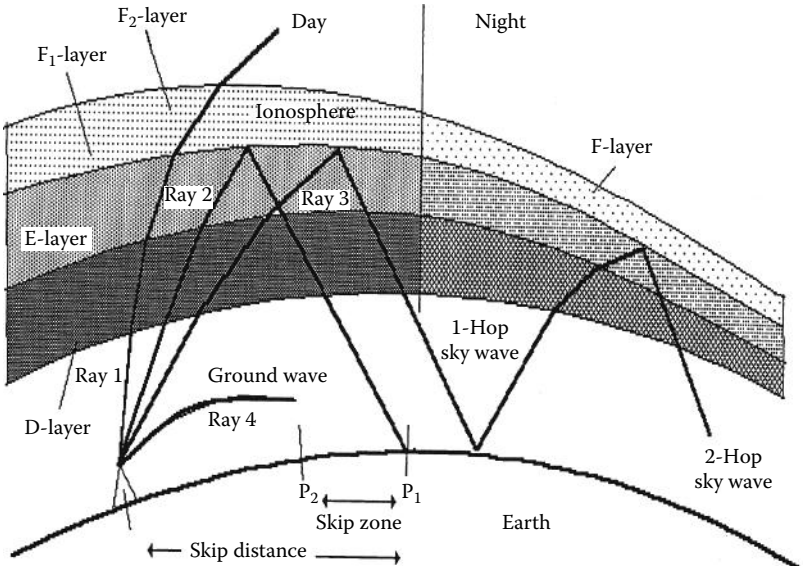


FIGURE 76.3 The four layers of the ionosphere and its effect on radio propagation. The four layers are produced by the ionization of molecules of particles in the atmosphere by ultraviolet rays of the sun. The effect of ionosphere on the radio waves is shown by reflections, also termed hops. The frequency of the electromagnetic wave is important in its behavior through the ionosphere.

Because of the higher resistance of the Earth's crust as compared to the atmosphere, the lower portions of radio waves parallel to the Earth's surface are slowed down, causing the waves to bend toward Earth. A wave of this type is termed a *ground wave*. The ground waves exist because they use the Earth's surface as a conductor. They occur at LF since LF causes more bending in conformity to Earth's shape. The ultimate range of such ground waves depends on the absorption effects. Sometimes, in the lower atmosphere, *surface ducting* occurs by multiple hopping, thus extending the range of a ground wave well beyond its normal limits. It is associated with higher radio and radar frequencies. This phenomenon is common in tropical latitudes. Behavior patterns of waves transmitted at various angles are illustrated in Figure 76.3.

76.2.1 Accuracy of Radio Location

There are a number of random effects that influence the accuracy of radio locations: atmospheric disturbances along the transmission path, errors in transmitters and receivers, clocks, inaccuracy in electronic circuitry, multipath, reflections, fading, etc. As a result, a series of positions determined at a given time and location usually results in a cluster of points near the true position. There are two measures commonly used to describe the accuracy: the first is the *circular error probable* (CEP)—a circle drawn on the true position whose circumference encompasses 50% of all indicated positions, and the second technique, more common, is the *root mean square* (rms), where

$$\text{rms} = \sqrt{\sum_{n=1}^{N} \frac{(E_n)^2}{N}} \tag{76.1}$$

where

E is the distance between actual and predicted positions
N is the number of predicted positions

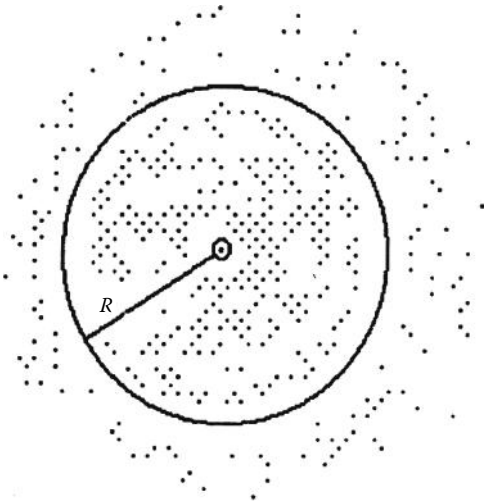


FIGURE 76.4 The rms radius circle that encompasses 68% of all measured positions. The variations in the measurements are due to a number of factors, including ionosphere conditions, precise location of satellites, and inefficiencies in electronic circuits (2 rms encompasses 95% of all indicated positions).

A circle, shown in Figure 76.4, with one rms value is expected to contain 68% of all the indicated positions. Another circle of radius equal to 2 rms should contain 95% of all the indicated positions, for isotropic scattering, or errors.

In electronic navigation systems, three types of accuracy are important: (1) *predictable* or *absolute accuracy*—the accuracy of a position with respect to the geographic coordinates of the Earth; (2) *repeatable accuracy*—the accuracy with which the user can return to a position whose coordinates have been determined at a previous time with the same navigation system; and (3) *relative accuracy*—the accuracy with which a user can measure position relative to that of another user of the same system at the same time.

76.3 Radio Navigation

In the 1930s, improved radio direction-finding techniques and better equipment led to the establishment of systems called *radio beacons*. These systems consisted of small radio transmitters located in fixed places to provide radio bearings that could be used in all weather conditions. Position findings by these beacons became known as *radio navigation*. Continued refinements in the state-of-the-art electronics technology and a better understanding of electromagnetic wave propagation led to the subsequent development of radar and longer-range radio navigation systems.

Essentially, radio beacons are single transmitters, transmitting a continuous wave at low power, usually modulated by audible Morse code characters for identification. The transmitted signal is received by an on-board receiver incorporating a radio direction finder (RDF) to be processed further.

Many radio navigation systems were developed and perfected by various countries over many decades since 1930s. Some well-known systems are Loran-C, Decca, Consol, Omega, GEE, Alpha, Lorenz-VOR, and a few others. However, most of these radio navigation systems are outdated and shut down as they are no longer commercially or militarily can be used in an economic way. Therefore, these systems will not be explained any further. However, some still find applications in a limited way such as the VOR. As the conventional radio location systems are replaced by the technologies such as the GPS or other satellite navigation systems, communication satellites, cell phones, and the like, more attention will be on these modern techniques.

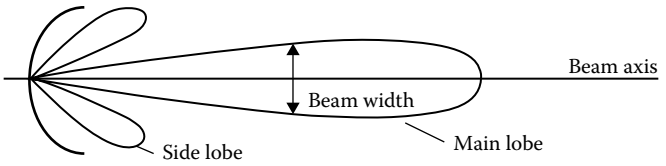


FIGURE 76.5 A surface search radar beam. High-frequency electromagnetic waves are formed by a parabolic antenna. The receiving antenna is rotated 360° to scan the entire surrounding area. The location of the target is determined by the reflected back signals and the orientation of the antenna.

76.3.1 Radar Navigation

The word is derived from *radio detection* and *ranging*. It works on the basic principle of reflection, which determines the time required for an RF pulse to travel from a reference source to a target and return as an echo. Most surface search and navigation radar high-frequency electromagnetic waves are formed by a parabolic antenna into a beam form, as shown in Figure 76.5. The receiving antenna is rotated to scan the entire surrounding area, and the bearings to the target are determined by the orientation of the antenna at the moment the echo returns. The standard radar is made up of five components: transmitter, modulator, antenna, receiver, and indicator. They operate on pulse modulation.

Radars are extremely important devices for air-control applications. Nowadays, airborne beacon radar systems are well developed in traffic alert and collision avoidance systems (TCAS). In this system, each plane constantly emits an interrogation signal, which is received by all nearby aircrafts that are equipped appropriately. The signal triggers a transponder in the target aircraft, which then transmits some information concerning three-dimensional (3D) location and identification.

There are a number of different radar-based radio navigation systems. These systems are primarily used in the avionics industry to provide information for pilots on the position of the aircraft based on the signals transmitted from the ground stations located worldwide. This information is used as the primary or the secondary navigation aid for course determination and correction, direction finding, and distance measuring.

An example of such radio location systems is the VHF Omni directional Range (VOR). In this system, a ground station transmits two signals; one is constant in all directions while the other varies the phase relative to the first. The receiver senses the phase difference between two signals, thus yielding the information on the relative position of the aircraft with respect to the ground station. The information is suitably displayed on the cockpit panel.

An application of radar is the *ground penetrating radar* (GPR) used for locating objects at soil subsurface, ice, and water. In GPR, microwave band pulses are generated and the reflected signals are processed to locate objects, materials, cracks, and structures. When the generated pulses encounter a buried object or a boundary, the reflected signals vary characteristics, which can be processed suitably to reveal information about the subsurface structures. Usually, image processing techniques are used for interpretation and display. The GPR finds applications in military, geologic explorations, soils and bedrocks formation studies, archeological discoveries, underground mining, oceanography, and ice structure determinations. The penetration of signals is dependent on power level and dielectric properties of the soil. It can penetrate soil up to 15–20 m.

76.3.2 Satellite Navigation Systems

The use of satellites is a highly developed technology utilized extensively throughout the world. In the past three decades, it has progressed from quasi-experimental in nature to one with routine provisions of new services for both military and civilian use. The satellite relay systems take advantage of the unique characteristics of *geostationary satellite orbits* (GSOs). The design of satellite systems is well understood, but the technology is still dynamic. The satellites are useful for long-distance communication services, for services across oceans or difficult terrain, and point-to-multipoint services such as television distribution.

Satellite navigation systems use electronic receivers to determine the location of an object within a few meters using time signals transmitted from the line-of-sight satellites. The receivers calculate the time as well as position (longitude, latitude, and altitude) of the object.

Today, there are two fully operational Global Navigation Satellite Systems (GNSS): GPS of United States and GLONASS of Russia. There are a number of projects run by different countries such as the Deidou and COMPAS of China, Galileo of European Union, Doris of France, QZSS of Japan, and IRNSS of India.

Frequency allocation for satellites is controlled by the International Telecommunication Union (ITU). In the United States, the Federal Communications Commission (FCC) makes the frequency allocations and assignments for nongovernment satellite usage. The FCC imposes a number of conditions regarding construction and maintenance of in-orbit satellites.

There are many satellite systems, known as ComSats, operated by different organizations and different countries mainly developed for communications and data transmissions. These include Iridium of Motorola, Globalstar of Loral Corporation, CS series of Japan, Turksat of Turkey, Aussat of Australia, Galaxy and Satcom of the United States, Anik of Canada, TDF of France, etc. Some of the communication satellite systems are suitable for navigation purposes. However, satellite systems specifically designed for navigation are limited in numbers as indicated earlier.

The first generation of the satellite system was the *Navy satellite system* (Navsat), which became operational in January 1964, following the successful launch of the first transit satellite into polar orbit.

The system was declared open for private and commercial use in 1967. Civil designation of the name of the system is *Transit Navigation Satellite System*, or simply *Transit*. Later, this system evolved to become the modern Navsat GPS system, which will be discussed in detail. Most of the operational principles discussed here are inherited by the GPS system.

The Transit system consists of operational satellites, plus several orbiting spares, a network of ground tracking stations, a computing center, an injection station, naval observatory time signals, and receiver–computer combinations. The transit satellites are in circular polar orbits about 1075 km above ground with periods of revolution of about 107 min. Because of the rotation of the Earth beneath the satellites, every position on Earth comes within range of each satellite at least twice a day, at 12 h intervals. As originally intended, if at least five satellites are operational at any given time, the average time between fix opportunities would vary from about 95 min near the equator to about 35 min or less above 70° North and South.

The Transit system is based on the Doppler shift of two frequencies, 150 and 400 MHz, transmitted simultaneously by each satellite moving its orbit at a tangential velocity of about 7.5 km/s. Two frequencies are used so that the effects of the ionosphere and atmospheric refraction on the incoming satellite transmission can be compensated for by the receivers. Each frequency is modulated by a repeating data signal lasting 2 min, conveying the current satellite time and its orbital parameters and other information. Within the receiver, a single Doppler signal is created by processing the two signals transmitted by the satellite. By plotting the frequency of this signal versus time, a characteristic curve of the type shown in Figure 76.6 is obtained. Since the frequency of the transmitted signal is compressed as the satellite approaches, according to what proportion of the velocity vector is seen by the user receiver, the curve begins at time T_1 at a frequency several cycles higher than the transmitted frequency.

Tracking stations record Doppler observations and memory readout received during each satellite pass to relay them to a computer center. Updated orbital position and time data communications are relayed to an "injection" station from the computer center for transmission to satellite in a burst once each 12 h. Enough data are supplied in this 15 s injection message to last for 16 h of consecutive 2 min broadcasts describing the current orbital positions of the satellite.

The system accuracy depends on the accuracy of the satellite orbit computation, the effect of ionosphere refraction, the precision of the receiver speed, and heading determination. Under optimal conditions, the system is capable of producing fixes with a maximum rms error of about 2–3 m for the stationary receivers anywhere on Earth. Errors can be reduced to give accuracy within 3 cm yielding an

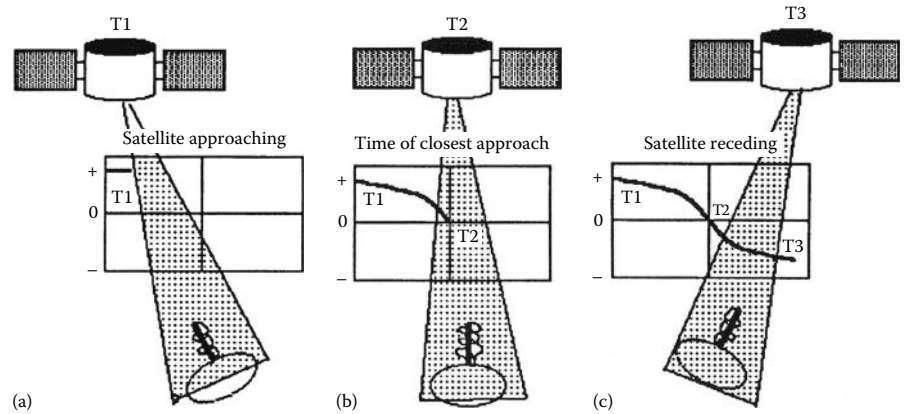


FIGURE 76.6 Transit satellite Doppler curve. (a) As the satellite approaches toward the receiver, the frequency of the received signal increases due to the Doppler shift. (b) At the time of closest approach, the transmitted and received frequencies are the same. (c) The frequencies received from a receding satellite result in lower values. This is also applicable in other position-sensing satellites such as GPS, Glonass, Starfix, etc.

absolute accuracy of 20–30 cm. Also, the time signal transmitted as a "beep" at the end of each 2 min transmission cycle coincides with even minutes of Coordinated Universal Time, which can be used as a chronometer check.

76.4 Global Satellite Navigation Systems

76.4.1 GPS

Global Satellite Navigation Systems are second-generation satellites evolved primarily from the Naval GPS. They provide a continuous 3D position-finding capability (i.e., latitude, longitude, and altitude), in contrast to the periodic two-dimensional information of the Transit system. Twenty-four operational satellites, as shown in Figure 76.7, constitute the system. Each satellite orbit is circular, about 2200 km high and inclined at angles of 55° with respect to the Earth's axis. Some of the older satellites are taken out of service to be replaced by their advanced counterparts.

The position determination using the GPS system is based on the ability of the receivers to accurately determine the distance to the GPS satellites above the user's horizon at the time of fix. If accurate distances of two such satellites and the heights are known, then the position can be determined. In order to do this, the receiver would need to know the exact time at which the signal was broadcast and the exact time that it was received. If the propagation speeds through the atmosphere are known, the resulting range can be calculated. The measured ranges are called *pseudoranges*. Nowadays, normally, information is received from at least four satellites, leading to accurate calculations of the fix. The time errors plus propagation speed errors result in range errors, common to all GPS receivers. Time is the fourth parameter evaluated by the receiver if at least four satellites can be received at a given time. If a fifth satellite is received, an error matrix can be evaluated additionally.

Each GPS satellite broadcasts simultaneously on two frequencies for the determination and elimination of ionosphere and other atmospheric effects. The Navstar frequencies are at 1575.42 and 1227.6 MHz, designated as L1 and L2 in the L-band of the UHF range. Both signals are modulated by 30 s navigation messages transmitted at 50 bits/s. The first 18 s of each 30 s frame contain *ephemeris* data for that particular satellite, which defines the position of the satellite as a function of time. The remaining 12 s is the *almanac* data, which define orbits and operational status of all satellites in the system. The GPS receivers store and use the ephemeris data to determine the pseudorange, and the almanac data to help determine

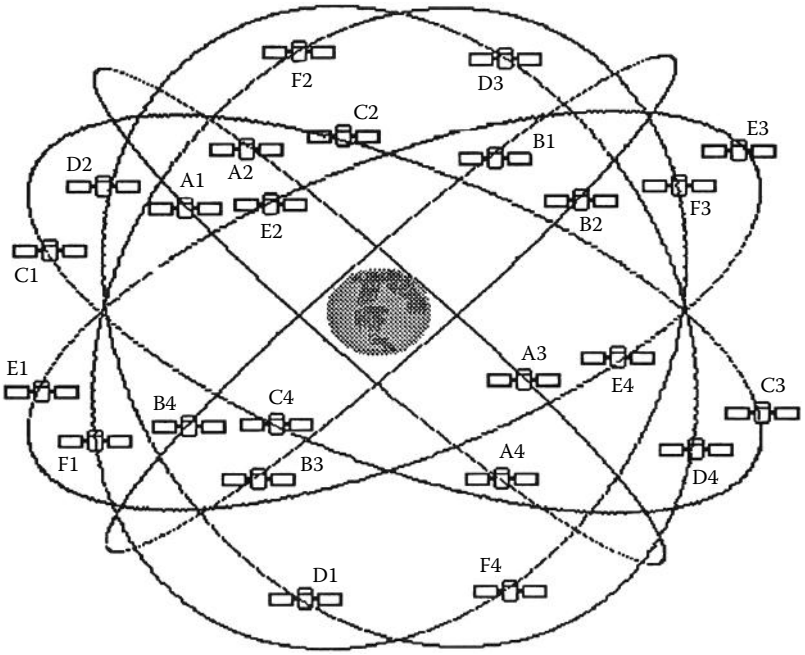


FIGURE 76.7 Operational GPS satellite coverage for navigation. Four satellites orbit in circular form. There are six such orbits inclined at angles of 60° from each other. In this arrangement, any point on Earth can see at least four satellites at any given time. This yields great accuracy in the position determination of the target, even only with C/A codes received.

the four best satellites to use for positional data at any given time. However, the "best four" philosophy has been overtaken slowly by an all-in-view philosophy.

The L1 and L2 satellite navigation signals are also modulated by two additional binary sequences called *CIA code* for acquisition of coarse navigation and the other *P-code* for precision ranging. The L1 signal is modulated both by the C/A and P-codes, and the L2 only by the P-code. The P-code, however, is not available for civilian users. The P-code is redesignated to be a Y-code, decipherable only by high-precision receivers having access to encrypted information in the satellite message. Civilian users have access to the so-called *Standard Positioning Services* (SPS), while U.S. and NATO military can use *Precise Positioning Service* (PPS) with better accuracy.

In enhancing SPS accuracy, differential techniques may be applied, as shown in Figure 76.8, to the encrypted GPS signals. Since the reference receiver is at a known location, it can calculate the correct ranges of pseudoranges at any time. The differences in the measured and calculated pseudoranges give the correction factors. Accuracy <1 m can be obtained in the stationary and moving measurements. Recently, differential receivers became commonly available, giving higher accuracy in subcentimeter ranges. They are getting cheaper day by day and finding applications in many areas such as airplanes, common transport vehicles, cars, geologic surveying, orienteering, farming, etc.

76.4.2 Glonass

There are a number of other satellite navigation systems similar to GPS of the United States, such as Russian *Glonass*. The Glonass consists of 31 satellites but 24 are in continuous use orbiting in circular form 1500 km above the ground. The accuracy of the system is about 10 m rms. Glonass satellites transmit details of their own position and a time reference. The carrier frequencies are in L-band, around 1250 MHz (L2) and 1600 MHz (L1). Only the L1 frequency carries the Civil C/A code.

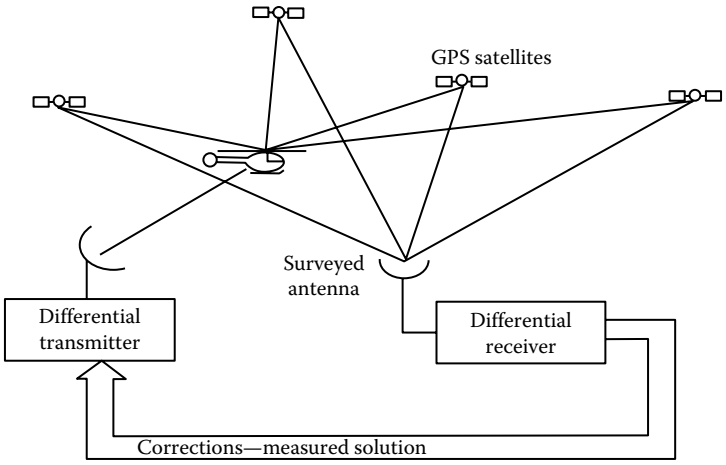


FIGURE 76.8 Differential GPS operation. Satellite and target positions are sensed by ground-fixed receivers or mobile receivers with exact known positions. Errors of the target position due to signals received from the satellites are corrected using the information from the fixed receivers. Using this method, very high accuracy within a few centimeter ranges can be obtained.

The radio-frequency carriers used by Glonass are channelized within bands 1240–1260 MHz and 1597–1617 MHz, the channel spacing being 0.4375 MHz at the lower frequencies and 0.5625 MHz at the higher frequencies. The number of channels is 24. The data message is formatted in frames of 3000 bits, with a duration of 30 s. The ephemeris data are transmitted as a set of position, velocity, and acceleration coordinates in a Cartesian Earth-centered, Earth-fixed (ECEF) coordinate system. The new ephemeris data are available every half hour, valid for the following quarter hour. The data are sent at a 50 baud rate and superimposed on a pseudorandom noise (PRN) code. The low-precision code has length 511 bits as compared to 1023 bits for Navstar. Glonass accuracy is as good as that for the GPS system. Glonass and GPS have different coordinate frames and different time frames that are being coordinated together.

76.4.3 Starfix

Another interesting satellite-based system—privately funded, developed, launched, and maintained—is the Starfix positioning system. This system is designed primarily for oil surveying. The system consists of a master site, which generates satellite ephemeris data, and four satellites in geosynchronous orbits. The system is said to have a precision of 2.5 m rms.

76.5 GPS Receivers and Their Uses

GPS receivers are devices constructed from electronic circuits, display units, and an antenna. They are supported by appropriate software. The devices are tuned to the frequencies transmitted by the satellites, and use a highly stable clock for timing purposes. Currently, there are three basic types of GPS receivers designed and built to address various user communities. These are called *slow sequencing*, *fast sequencing*, and *continuous tracking* receivers. The least complicated and lowest cost receiver for most applications is the slow sequencing type, wherein only one measurement channel is used to receive sequential L1 C/A code from each satellite every 1.2 s, with occasional interrupts to collect ephemeris and almanac data. Once the data are received, computation is carried out within 5 s, making this system suitable for stationary or near-stationary fixes.

Fast sequencing receivers have two channels: one for making continuous pseudorange measurements and the other for collection of the ephemeris and almanac data. This type is used in medium dynamic applications such as ground vehicles.

A number of companies produce highly sophisticated GPS receivers. EURONAV GPS receivers operate on two and five channels for military applications. They provide features such as precise time, interfacing with digital flight instruments, RS-422 interface, altimeter input, self-initialization, etc.

Software implementation satellite management functions, having different features, are offered by many manufacturers. In the case of DoD NAVSTAR GPS receivers, for example, three functional requirements are implemented: (1) database management of satellite almanac, ephemeris, and deterministic correction data; (2) computation of precise satellite position and velocity for use by navigation software; and (3) using satellite and receiver position data to periodically calculate the constellation of four satellites with optimum geometry for navigation. The DoD receivers are divided for three functions as Satellite Manager (SM), Satellite-Data-Base-Manager (SDBM) SV-Position Velocity Acceleration (SVPVA), and Select-Satellites (SS).

Differential navigation is also applied where one user set is navigating relative to another user set via a data link. In some cases, one user has been at a destination at some prior time and is navigating relative to coordinates measured at that point. The true values of this receiver's navigation fix are compared against the measured values, and the differences become the differential corrections. These corrections are transmitted to area user sets in real time, or they may be recorded for postmission use so that position fixes are free of GPS-related biases. The differential navigation and GPS systems find applications in en-route navigations for commercial and civil aviation, and military application.

GPS find wide applications in civilian uses. Some of these applications are

- Cellular telephony for position detections
- Time setting and clock synchronizations
- Emergency services and disaster relief
- Person, wildlife, animal tracking
- Vehicle detection, tracking and navigation systems
- Geotagging, cartographies, and map making
- Navigation of ships especially in shallow waters
- Aircraft position tracking
- Orienteering, tours, recreational activities, display of information about objects such as in museums
- Surveying, land management, tracking terrain changes, and terrain and property boundaries
- Tectonics, seismic geophysical explorations, landslides, and earthquakes
- Navigation, velocity determination, and orientation measurements
- Robotics, self-navigation, and autonomous mobile systems

All these applications are supported by software in a number of ways. For example, in some applications, the vector-based maps are used for route calculations and direction finding. In others, the GPS unit sends details of the locations based on time interval and distance traveled for tracking purposes.

Most commercial software run on Windows, Linux, Mac OS X, or Android platforms. There is much commercially available software to support GPS-based location determination and navigation, some of which are Destinator, Navigon, NDrive, ROUTE 66, Google Earth, Google Maps, Navit, VZ Navigator, and Tom Tom Mobile. There is also open source software such as Navit and Waze.

76.6 Transponders

Transponders are transducers that respond to incoming signals by generating appropriate reply messages. Recent developments in technology have made the configuration of transponders possible using elaborate and powerful on-board signal processing. This enhanced the capacity by improving

the link budgets, by adjusting the antenna patterns, and by making the satellite resources available on a demand basis—called the "switch board in the sky concept."

Increased interest in deep sea exploration has brought acoustic transponders to the forefront as an important navigation tool. They provide 3D position information for subsurface vehicles and devices.

Some transponders are based on radar signals that respond to radar illumination. Transponders are programmed to identify friend or foe or, in some cases, simply inform ground stations about the position of aircraft.

Transponders are used for emergency warning. The U.S. and Canadian satellites carry Sarsat transponders, and Russian satellites carry Cospas transponders. They are also used as warning devices in collision avoidance systems in aircraft and land vehicles.

76.7 RFID Radio Location Systems

RFID uses electromagnetic propagation to transmit information from a tag attached to an object. The tag contains nonvolatile memory to digitally store information that can be read from a distance. The system requires a transmitter and a receiver. The reader transmits an encoded signal to read the tag. The tag responds with its identification codes and with the information stored in it.

RFID tags can be active or passive or battery-assisted passive. An active tag contains an on-board battery to power the system. Passive tags use the radio energy transmitted by the reader as its energy source. A battery-assisted passive tag uses a small battery on-board to activate the system in the presence of a request from the reader. Tags may be read-only or may be read/write. In read-only tags, the information is factory set and cannot be changed. Read/write tags may be written by users.

Today's technology allows the production of small RFIDs. Hitachi produces RFID chips as small as 0.05 mm × 0.05 mm. These small-sized chips are capable of typically storing 38-digit numbers using 128 bit read-only memory (ROM). They can store information up to 2 kB of data.

RFID systems are used in

- Asset management
- Inventory systems
- Product tracking
- Transportation and logistics
- Passports
- Prevention of counterfeits such as chips in casinos
- Hospitals and healthcare
- Antitheft devices such as wallets, luggage, bags, keys, etc.
- Pharmaceutical industry
- Food industry and food safety
- Livestock or wildlife identification
- Libraries, books, CDs
- Telemetry
- Electronic toll collections
- Human or animal implantations
- Personal safety and alarm systems
- Tracking personal or commercial properties, etc.

A typical application of RFID is in healthcare. They are used in multiple ways, from keeping historical information about patients and collecting the status of vital life signs to equipment tracking in hospitals, clinics, and medical centers. For example, the locations of beds, wheelchairs, surgical and other medical equipment can be realized and tracked. Active RFID tags together with other sensors are used to detect any movement of equipment as well as any tampering and environmental conditions where they are kept.

Bibliography

Antalovic, D., Numerical investigation of algorithms for multi-antenna radiolocation, *Proceedings of 2009 IEEE 70th Vehicular Technology Conference* (*VTC 2009 Fall*), Anchorage, AK, 5 pp., 2009.

Finkenzeller, K., *RFID Handbook*, Wiley, Chichester, U.K., 2002.

Gao, S., Zhao, L., and Gao, L., Research progress on satellite navigation with inertial information-aided, *Lecture Notes in Electrical Engineering*, v 139 LNEE, Springer-Verlag, New York, pp. 153–158, 2012.

Grewal, M. S., Weill, L. R., and Andrews, A. P., *Global Positioning Systems, Inertial Navigation and Integration*, 2nd Edn., Wiley-Interscience, Hoboken, NJ, 2007.

Hegarty, C. J. and Chatre, E., Evolution of the Global Navigation Satellite System (GNSS), *Proc. IEEE*, 96(12), 1902–1917, 2008.

Hofman-Wellenhof, B., Legat, K., and Wieser, M., *Navigation: Principles of Positioning and Guidance*, Springer-Verlag, Wien, Austria, 2003.

Kolawole, M. O., *Radar Systems, Peak Detection and Tracking*, Newness, Oxford, U.K., 2002.

Last, D., Radionavigation—The future direction, *19th International Conference Applied Electromagnetics Communications (ICECom'07)*, IEEE, Dubrovnik, Croatia, pp. 59–62, 2007.

Melbourne, W. G., *Radio Occultations Using Earth Satellites*, Wiley-Interscience, Hoboken, NJ, 2005.

Pearson, C., McGuire, M., and Coady, Y., The application of Wi-Fi radiolocation research to mobile devices, *2009 IEEE Pacific Rim Conference on Communications, Computers, and Signal Processing (PacRim)*, Victoria, British Columbia, Canada, pp. 768–773, 2009.

Partial List of Manufacturers/Suppliers

B&D Industrial, P.O. Box 4325, Macon, GA 31208-4325, Phone: 866-462-3133 (toll free), Fax: 478-742-7836, http://bdindustrial.rtrk.com/

Beacon Industries, Inc./World Class Products, 12300 Old Tesson Rd., Saint Louis, MO 63128, Phone: 314-487-7600, 800-454-7159 (toll free), Fax: 314-487-0100, www.beaconteчnology.com

Boeing Co., 2201 Seal Beach Blvd., P.O. Box 4250, Seal Beach, CA 90740, Phone: 562-797-2020, www.boeing.com

Cetacea Sound, Inc., 2950 Airways Blvd., Memphis, TN 38116, Phone: 901-363-3856, 800-489-1759 (toll free), Fax: 901-367-1420, www.cetacea.com

Garmin Corporation, 1200 E. 151st St., Olathe, KS 66062-3426, Phone: 913-397-8200, 800-525-6726 (toll free), Fax: 913-397-8282, www.garmin.com/garmin/cms/site/us

GPS Integrated Systems, Inc., 1414 W Belt N Ste 110, Houston, TX 77043, Phone: 713-973-8889, 866-736-8477 (toll free), Fax: 713-973-9065, gps-is.com/

Grainger Industrial Supply, 100 Grainger Pkwy., Lake Forest, IL 60045, Phone: 847-535-1000, 888-361-8649 (toll free), Fax: 847-535-9123, www.grainger.com/thomasnet

Mobile Epiphany, 3131 S. Vaughn Way, Suite 650, Aurora, CO 80014, Phone: 720-282-2044, www.mobileepiphany.com/

Navhouse Co., 10 Loring Drive, Bolton, ON L7E1J9, Phone: 905-857-8102, Fax: 905-857-8104, www.navhouse.com

PSI Company, 2612 Pacific Highway, Unit #A1, Tacoma, Washington 98424, Tel: 1-800-826-2907, http://www.psicompany.com/

Raveon Technologies Corp., 990 Park Center Drive, Suite C, Vista, CA 92081, Phone: 760-727-8004, Fax: 760-598-8004, www.raveontech.com

Reltima GPS Fleet Management Systems, 800 W. Cummings Park, Suite 6550, Woburn, MA 01801, 800 W. Cummings Park, Suite 6550, Woburn, MA 01801, Phone: 781-569-2051, Fax: 781-569-2185, http://www.reltima.com/

RFind Systems, Inc., 2076 Enterprise Way #102, Kelowna, BC V1Y6H7, Phone: 250-862-3412, Fax: 250-869-4800, www.rfind.com

Sealite USALLC/Avlite, 172 Lilly Pond Rd., Gilford, NH 03249, Phone: 603-737-1311, 888-628-2869 (toll free), Fax: 603-737-1320, www.navbuoy.com

Select Controls, Inc., 45 Knickerbocker Ave. Bohemia, NY 11716-3110, Phone: 866-929-5312 (toll free), Fax: 631-567-9013, www.select-controls.com

Smiths Industries Aerospace, Information Management Systems Div., 3290-T Patterson Ave., Grand Rapids, MI 49518, Phone: 616-241-7000, Fax: 616-241-7533, www.smiths-industries.com/

Space Systems/Loral, 3825 Fabian Way, Palo Alto, CA 94303, Phone: 650–852–4000, Fax: 650-852-6417, www.ssloral.com

Target Systems Technologies Corp., 12–1 Technology Dr., Setauket, NY 11733, Phone: 631-689-7183, Fax: 631-689-2127, http://target-sys-tech.com/

Trimble Navigation, Ltd., 936 Stewart Dr., Sunnyvale, CA 94088, Phone: 408-481-8000, www.trimble.com

Unimar, Inc., 4944 Verplank Rd., P.O. Box 220, Clay, NY 13041, Phone: 315-699-4400, 800-739-9169 (toll free), Fax: 315-699-3700, www.unimar.com

Universal Avionics Systems Corp., 11351 Willows Rd. N.E., Redmond, WA 98052, Phone: 425-602-5000, Fax: 425-556-7808, http://www.uasc.com/

Warren-Knight Instrument Co., 2045 Bennett Dr., Philadelphia, PA 19116-3019, Phone: 215-464-9300, Fax: 215-464-9303, http://www.warrenknight.com/

77

Air Pollution Monitoring

Michael Bennett
Manchester Metropolitan University

77.1 Introduction

There is a huge range of atmospheric pollutants and for a given pollutant there may be several commercially available detection systems. In a document of this size, it is clearly not feasible to discuss in detail the theory of operation of every possible system. After some general remarks about air pollution monitoring, a broad introduction to the physics of molecular spectroscopy (the most widely used detection principle) will be presented. The detection of individual pollutants will then be discussed. Most of the practical examples given will refer to British experience, since this reflects the author's background. The principles involved, however, are universally applicable. The EPA's website (Table 77.3) provides an useful gateway to American legislation and practice.

Before spending money on a detection system, it is essential that the user understands the purpose of the measurement. This will affect both the choice of instruments and the location of monitoring sites. In general, active monitors with short response times will be more expensive than passive, slow response instruments. A survey intended to determine spatial variations of an ambient pollutant might therefore be best to employ very many cheap, slow response instruments. A classic example would be national surveys carried out in England and Wales in the early 1990s to measure indoor radon pollution [1]. The National Radiological Protection Board (NRPB—now part of the Health Protection Agency) supplied householders with two passive detectors, each consisting of a strip of plastic inside a protective

TABLE 77.1 List of a Selection of Manufacturers of Point Samplers with Their Appropriate Model Numbers

Manufacturer	SO_2	NO_x	O_3	CO	THC/CH_4
Beijing Zhong Sheng (formerly Dasibi)	4208	2208	1208	3208	
Casella	CM2050	CM2041	CM2010	CM2030	
Ecotech	EC9850[a]	EC9841[a]	EC9810[a]	EC9830[a]	
Environnement SA	AF22M	AC32M	O342M	CO12M	HC51M
Horiba	APSA-370	APNA-370	APOA-370	APMA-370	
Monitor Labs (Teledyne)	TML50 T100	TML41 T200		TML30 T300	
Signal-Ambitech	Ambirak-477	Ambirak-447	Ambirak-427	Ambirak-416	300HM
Thermo-Scientific	43i	42i-D	49i	48i	55i

Note: This list is not intended to be exhaustive and no particular recommendation is implied.
[a] Also Serinus series.

container. These are left in a living room and a bedroom for 3 months, at the end of which they are mailed back to the NRPB for analysis. (The number of α-particle trails in the plastic is counted.) The householder is then reassured (or not) as to the safety of the dwelling; the authorities, meanwhile, can build up a national map of Rn concentration in relation to geology, building type, etc. The essence of the system is that the individual detectors are so cheap—the Agency still (2012) provides a pair on request to individual householders at <£50—that the authorities can afford many thousands, such a number being necessary to provide realistic spatial coverage.

Identification of a particular source in the presence of background emissions, however, requires a sensor with a response time of minutes, or 1 h at most, supplemented by meteorological measurements. This need arises since at temperate latitudes the wind direction changes by order of 15° in 1 h and 90° in 24 h. Unless a source dominates local pollution, its impact is unlikely to be demonstrable through daily sampling.

Some applications may require response times of seconds or less. Odors are a very common cause of complaint. In this case the complainant perceives fluctuations in concentration over periods of a few seconds, while he may become inured to steady concentrations over periods of minutes [2]. It is also the case that many toxic gases (e.g., H_2S) are more dangerous in high doses over short periods than at a steady concentration over some hours [3]. Flammability, of course, also depends upon peak rather than mean concentrations.

Central to obtaining reliable information from such measurements is the protocol for the siting of instruments [4,5]. This naturally depends upon the objectives of the survey. The UK national survey for smoke and SO_2 [6], for example, was highly successful in demonstrating the effects on background pollutant concentrations of the 1956 Clean Air Act and (from the late 1960s) the availability of natural gas (cf. Figure 77.1). Sites were chosen so as not to be dominated by individual local sources, being typically in backyards or on the roofs of municipal buildings. By the early 1980s, however, urban air pollution had become increasingly dominated by road traffic emissions and the earlier network of sites no longer seemed appropriate [7,8]. Many roadside measurements have now also been made in addition to more general urban or rural surveys [9]. Note that a *typical* site is not necessarily a *representative* one [4]. The former is what we experience about our daily lives. The latter is chosen through a protocol so that measurements from one site may be compared with those from another. Depending upon the application, representative sites may be more or less typical.

A protocol is important not only for comparing one place with another but also for comparisons between different times. One should not be dismissive of the old technology of the national SO_2 survey: the conservatism of the design allows us to rely on the time series of concentrations over >60 years. Clearly, there are advantages in keeping up with the latest technology. But, if we are interested in long-term trends, we must be very sure of the relative responses of "improved" instruments and methods.

The deployment of fast response instruments around individual sources gives rise to somewhat different problems. Given the high cost of such instruments, we must be sure that statistically useful

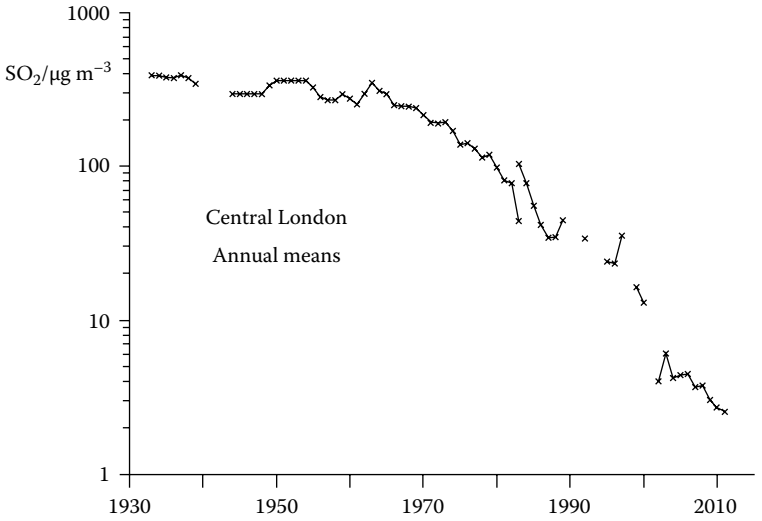


FIGURE 77.1 Annual average SO_2 concentrations in Central London, 1933–2011. Data for 1933–1983 are from a composite of sites in Lambeth and Westminster; for 1983–1999 from the Islington9 site. Data for 2002–2011 are from a modern instrument in Horseferry Road, Westminster.

measurements of environmental impact will be made over the course of the survey. Concentrations should ideally be measured simultaneously upwind and downwind of the source. For an elevated buoyant source, it is desirable to measure downwind concentrations over a range of distances: atmospheric dispersion models are not yet so foolproof that the calculated distance from the stack to the peak ground-level concentration may be relied upon. At a given point on the ground, pollution from such a source may typically be detectable during only a few percent of hours. Mobile laboratories are therefore commonly used to enhance the capture of data, but it is then essential that some statistical and climatological analysis be applied to extrapolate the measurements obtained to long-term means at fixed target points on the ground. The protocol and analysis should be planned before the instruments are purchased.

77.2 Spectroscopic Principles

The most common principle employed for the detection of contaminants in air involves the interaction of the trace species with light. The advantages are obvious: with careful choice of system, the interaction can be made specific to the chosen molecule, and the photon will then transport itself from the point of interaction to the detector. Reference [10] provides a general undergraduate-level introduction to molecular spectroscopy, while Ref. [11] describes specialized applications to atmospheric measurement.

The specificity of the interaction of light with a molecule arises from quantum theory. Light is absorbed or emitted as photons, of energy hc/λ, while individual atoms or molecules may only exist at discrete levels of energy or angular momentum. Light is thus only absorbed or emitted by gaseous species at discrete values of wave number, $1/\lambda$ (usually quoted in cm^{-1}). Generally, different molecular energy levels arise through three processes:

- *Rotational*: A molecule may spin with an angular momentum equal to a discrete value of $h/2\pi$. Transitions between angular momentum states are typically equivalent to photon energies of 1–100 cm^{-1}, i.e., microwave wavelengths.
- *Vibrational*: The atoms in a molecule may vibrate relative to one another. Transition energies between modes are typically of order 10^3 to 10^4 cm^{-1}, i.e., infrared (IR) wavelengths.
- *Electronic*: The electrons may occupy different orbitals within an atom. Transitions between these would typically take place at UV wavelengths.

Photons may initiate transitions between combinations of the above levels; observed molecular spectra thus tend to be extremely complicated, having structure over all scales of wave number. Atomic spectra, e.g., from a monatomic gas like Hg, are simpler, since there are no vibrational or rotational modes. Nevertheless, even for molecular spectra, "selection rules" permit only a restricted number of transitions to occur: a single photon can only effect a transition if the initial and final dipole moments of the molecule are different. (The direction of the change of dipole depends upon the polarization of the associated photon.) Conservation of angular momentum also requires that a molecule's spin may only change by one unit of $h/2\pi$ at a time.

Photon energies are broadened by temperature (which causes a Doppler shift in frequency) and by pressure (which shortens the lifetime of the excited state). In practice, of course, any practicable measurement system also has a finite resolution. Parts of the spectrum may thus contain so many possible transitions that it becomes impractical to distinguish individual lines. The specificity of molecule–photon interactions promised by quantum theory is thus seen to be somewhat limited in practice. Spectroscopic identification of a particular species relies upon the existence of resolvable structure in an accessible part of the spectrum where there is no serious interference from other common gases. In general, this search must be solved individually for each analyte, with no guarantee of the existence of a satisfactory solution.

77.3 Absorption Techniques

Detection of a specific molecule may be through either its absorption or emission of light at a particular wavelength. Consider monochromatic light of flux I passing through a gas of concentration χ. In so far as each interaction between a photon and a molecule is an independent event, then the rate of interaction is separately proportional to the number of photons and the number of gas molecules. Thus,

$$\frac{\partial I(x,\lambda)}{\partial x} = -\sigma(\lambda) I(x,\lambda)\, \chi(x) \tag{77.1}$$

where $\sigma(\lambda)$ is the absorption cross section. For constant gas concentration, this may be integrated as a function of path length, x, to give the Beer–Lambert law:

$$I(x,\lambda) = I(0,\lambda) e^{-\sigma(\lambda)\chi x} \tag{77.2}$$

The argument of the negative exponential is known as the optical density at this wavelength. The gas burden, χx, along the optical path is thus given by

$$\chi x = \frac{\ln(I(0,\lambda)/I(x,\lambda))}{\sigma(\lambda)} \tag{77.3}$$

and since $\sigma(\lambda)$ may be measured in the laboratory, we now have a measurement of the gas concentration.

The advantage of absorption techniques is that they require minimal disruption to the observed system. This is of benefit in ambient monitoring where, for example, a long-path measurement may be set up over many hundreds of meters without problems of safety or power. (This may be of particular value when monitoring, for example, at an airport.) Equally, in emissions monitoring it is advantageous to be able to measure gas concentrations in situ in the flue. Long-path techniques may also be applied in point samplers through the use of a multipass cell, e.g., a White cell [11]. The response time would then be limited by the volume of the cell in relation to the sample throughput.

Because of instrumental offsets and interference from other species, measurement of transmission at a single frequency is unlikely to give an accurate estimate of the gas concentration. More commonly, absorption is measured at several frequencies. Standard techniques are

- *Differential optical absorption spectroscopy* (*DOAS*): A broadband source is used in conjunction with a high-resolution spectroscope tuned around distinctive absorption features of the target gas. These days, the spectrum would typically be registered on a linear array of charge-coupled devices. The difference between online and off-line absorption (or between the absorption with and without the analyte gas being present) then gives the gas concentration. The measured absorption spectrum may then be fitted numerically to the target spectrum, allowing some correction for interferent species.

  A limitation with the method is that it is, as the name implies, *differential*. Molecular spectra are not usually so sharp that an absolute measurement can be made. Instead, one is usually looking for differences from a reference spectrum, obtained when it is believed that the analyte gas was at a very low concentration. This may be difficult to achieve in a real situation—and it may then be difficult to freeze the optical response of the system until the next reference can be obtained. In practice also, it is not a trivial task to fit the observed spectrum to that of the target gas. In the presence of noise and several interferent gases, the calculated concentration of the target gas will depend on which other gas spectra are included in the fit.
- *Fourier transform infrared* (*FTIR*): The output from the gas cell is put into an interferometer. The Fourier transform of the output signal as a function of phase lag is then the absorption spectrum. As with DOAS, fitting programs may be applied to estimate the mix of pollutants responsible for this spectrum. With an appropriate optical design, the full range of phase delays may be obtained simultaneously. There are two practical advantages to working in phase space rather than frequency space:
  - Rather than the image of a slit on the detector surface, we have the collimated focus of the whole aperture—in effect the real image of the radiation source. The photon throughput is thus much larger (Jacquinot advantage).
- All wavelengths are measured simultaneously and the device can thus obtain a spectrum more quickly for a given resolution (Fellgett advantage). With modern CCD arrays, this is less of an issue.
- Nondispersive infrared (NDIR). Optical filters are applied to limit the input spectrum to a window where interference effects from other gases are small. No further spectral separation ("dispersion") is then applied. Gross absorption by the target gas is measured and may be calibrated by switching gas cells into the light path. Clearly, this requires a careful choice of wavelength window. For CO detection, for example, a wavelength around 4.7 μm is frequently used, since this lies between the absorption bands of CO_2 and water vapor—both of which would also be present in high concentrations in combustion gases.
- *Tunable laser diode absorption spectroscopy* (*TLDAS*): This uses a very narrowband source (viz. a tunable diode laser) and a broadband receptor. The frequency of the source will typically be oscillated close to an absorption line of the target gas: the first harmonic of the signal is then proportional to the integrated concentration along the light path.

77.4 Emission Techniques

The converse of looking for the absorption of light of a given frequency by the target gas is to excite molecules of the gas and then examine the light emitted as they return to their ground state. The signal is passed through a narrowband filter and measured with a photomultiplier tube. There are several standard techniques:

- *Flame photometry*: A flame (typically of H_2) is burned in the sample gas. The heat breaks up and ionizes the target molecules which then relax to their ground state. Since we are now looking at fragments of the molecule rather than the molecule itself, the technique is not completely specific. Its advantage is that it tends to be very fast, being ultimately limited by the time scale for the ions to pass through the flame. A related technique, flame ionization detection (FID), though not strictly spectroscopic, is to measure the conductivity of the flame which arises from such ionization.

- *Chemiluminescence*: A reactive gas is added to the sample gas. Light from the excited products of the reaction is detected.
- *UV fluorescence*: The sample gas is excited with UV light and the subsequent fluorescence measured. A related technique, photoionization detection (PID), is to measure the ionization current arising from such UV irradiation. This then detects all species whose first ionization potential is less than the photon energy of the lamp. The technique is extremely fast [12].

77.5 Particulate Sampling

Sampling of particulate matter in air gives rise to a different set of considerations from the sampling of trace gases. Fundamentally, we are dealing with a much more diverse set of analytes. Clearly, one SO_2 molecule (say) is much the same as any other SO_2 molecule. The same cannot be said of an aerosol particle: particles may vary by mass, by morphology, or by chemical composition. All of these factors will determine their toxicity and how they will disperse in the environment.

At its simplest, we can classify aerosol particles by their sedimentation velocity. This is sensible because it gives an indication of how quickly they will be lost to deposition, and how easily they may be inhaled to the deep lung: since gravitational and centrifugal accelerations are equivalent, a massive particle cannot follow the tortuous streamlines from mouth to alveoli. In practice, it is conceptually more convenient to classify a particle by its equivalent "aerodynamic" diameter, this being the diameter that a spherical particle of unit density would need to have for it to have the same sedimentation velocity as the reference particle. Thus, aerodynamic and physical diameters will be similar for (say) a small droplet of ammonium nitric acid aerosol; for soot aggregate, however, the aerodynamic diameter will be much smaller than the envelope diameter. Particles of aerodynamic diameter <10 μm (PM_{10}) tend to travel with the air flow. Since, moreover, their Brownian diffusivity is very small, they exhibit very slow deposition [13]. At <2.5 μm ($PM_{2.5}$), such particles may be inhaled deeply into the lungs; as such, they are known as "respirable aerosol." Epidemiological studies [14] have demonstrated that urban aerosol in this size fraction induces significant mortality, even at modest concentrations.

Note that just because $PM_{2.5}$ has been shown to increase mortality in an urban environment does not mean that all fine aerosols are equally toxic. It is physiologically implausible, for example, that aerosols of soluble ions, e.g., sea salt or sulfate, should cause problems for public health at concentrations of 10 $\mu g/m^3$.

If we add an electron to each particle and then apply an electric field, the resultant electrostatic force takes the place of gravity, measuring the drift ("mobility") of the particle then provides a direct measure of the aerodynamic diameter, even of ultrafine particles. Summing up the rate of arrival of the electrons, i.e., the electric current, in a Faraday cup electrometer can then give us a number density of fine or ultrafine particulate (cf. Dekati or Grimm Aerosol Technik). Scanning the applied voltage and looking at the current as a function of arrival time gives us a number spectrum as a function of aerodynamic diameter ("differential mobility analyzer") down to diameters of a few nanometers.

Particles of diameters >10 μm exhibit significant inertial and gravitational effects. They may thus deposit through impaction as the air flows around small obstacles. For particles >50 μm, sedimentation becomes dominant. The sampling method must thus be tuned to the range of particle sizes which it is wished to monitor. For a general review, see the book by Vincent [15].

Spectroscopic methods are unlikely to be effective. The interaction of light with small particles (i.e., of diameter much less than the wavelength of the scattered radiation) is described by the Rayleigh scattering formula, while for larger spherical particles it is described by the more complicated Mie scattering analysis [16], with the maximum backscatter per unit aerosol mass occurring when the wavelength of the radiation is comparable to the circumference of the particles. For nonspherical particles, scattering can be modeled numerically. In any case, the response tends not to show very specific behavior as a function of particle composition.

77.6 Slow Ambient Monitoring

We will discuss first systems with sampling times of 1 day and upward. These would typically be used for background monitoring.

77.6.1 Diffusion Tubes

A diffusion tube is the classic inexpensive, slow response instrument which may be deployed in large numbers to quantify spatial variations of ambient pollutants. Typically, it consists of a sealed perspex tube with a removable cap and an active substrate at the closed end. It is deployed in the field with the open end downward and after the prescribed exposure time, typically 7 days, it is returned to the laboratory for analysis. The system was originally developed as a personal monitor of NO_2 exposure [17] but has since been widely used for ambient urban surveys (e.g., [18]).

Assuming that the substrate is a perfect trap for the target gas, the diffusive flux of gas into the tube is given by $DA\chi/L$, where D is the diffusivity of the target species, A the internal cross-sectional area of the tube, and L the length of the tube. So long as the diffusivity remains constant over the period of measurement, the total deposition to the substrate is therefore a measure of the mean concentration over the period of exposure. Theoretically, D should vary with temperature and there may also be some circulation within the tube arising from wind across its mouth. A typical commercially available model has $L = 71$ mm and an internal diameter of 12 mm, so such circulation is suppressed. Studies have shown acceptable correlation between samples from diffusion tubes and long-term means from adjacent point samplers [19].

The system employed for NO_2 is a substrate of triethanolamine deposited on a stainless steel mesh. The sample is dissolved in orthophosphoric acid and the nitrite detected colorimetrically using the Greiss/Saltzman technique at a cost of less than £10 per tube. More costly systems are also available (e.g., from Gradko) for SO_2, benzene, xylene, toluene, and a wide range of other gases.

77.6.2 Drechsel Bottles

The classic bubbler (a "Drechsel" bottle) consisted of a Pyrex® bottle containing a solution through which filtered sample gas was bubbled. For the detection of SO_2, a solution of hydrogen peroxide was used so that

$$H_2O_2 + SO_2 \rightarrow 2H^+ + SO_4^{2-}$$

In polluted areas, the H^+ ions were then detected by titration, it being assumed that all the acidity came from atmospheric SO_2. In cleaner, rural areas, this became insufficiently sensitive (NH_3 emissions from livestock could give apparently negative SO_2 concentrations!) and the sulfate ions were then measured using ion chromatography. Sampling times were typically 24 h for a sensitivity of a few ppb.

As implemented in the U.K. national surveys of smoke and SO_2, eight such bottles were mounted in a case with a separate pump and flow meter. At some preset time each day, the airflow was switched from one bottle to the next; at the end of the week, the operator then visited, replaced all the bottles, and returned the old ones to the laboratory for analysis. In this system, the air was first drawn through a separate filter for each day of measurement. This captured particulate in the range 5–25 μm diameter; the blackness of the stain was measured photometrically to provide a 24 h sample of "black smoke." The system is now obsolete, being replaced by UV fluorescence systems for SO_2 and an Aethalometer® for black carbon (see later).

The bubbler system was simple, inexpensive, and required minimal operator training. It was therefore possible to install many hundreds of such sites and to run them for a considerable period. In central London, for example, such systems were first installed in 1933. Figure 77.1 shows how urban SO_2

levels have dropped by two orders of magnitude over the subsequent 80 years. Concentrations in central London were still well in excess of 300 μg/m^3 in the early 1950s. (The reader may wonder how the author survived his early childhood.) With the arrival of natural gas for domestic heating in the late 1960s, concentrations had dropped to 40 μg/m^3 by the early 1980s. By that stage, smoke and SO_2 emissions from domestic coal burning were no longer a serious issue and the network was scaled down to around 250 active sites [9]. Concentrations continued to fall rapidly with power generation from natural gas being permitted from 1991, and ultralow sulfur fuel for motor vehicles being introduced over the following decade. They are now below 3 μg/m^3. Figure 77.1 also betrays how data capture fell off in the 1990s. Black smoke and SO_2 were no longer a priority of central government and the network was gradually wound up. The total number of sites sponsored by central government has remained at a similar level, however, but now split between various networks (urban background, rural background, black carbon, etc.) under the supervision of AEA technology and using modern technology. The Westminster site whose data are used in Figure 77.1 is one of the urban background sites. Other networks, sponsored by local authorities or commercial interests (e.g., airports), also operate. Reference [20] supplies helpful advice on the choice of modern instruments for such networks.

In practice, the real cost of any monitoring system is the manpower and the site. These are small if the system is sited in a municipal building and operated by existing staff. At a remote site, however, a secure cabin with power must be provided and a nonautomatic system must be visited weekly. The practical difficulties of identifying, securing, and maintaining a site should never be underestimated when designing a monitoring network.

77.6.3 Deposition Gauges [21]

Traditionally in the United Kingdom, measurements of atmospheric dust were made using the passive British Standard dust deposition gauge, while measurements of "smoke" were made using the active filter system described in the previous section. The standard dust deposition gauge consisted essentially of a bowl of diameter 300 mm and depth 225 mm mounted 1.2 m above the ground; the sample was washed into a bottle beneath the gauge, from where it was filtered, dried, and weighed. After having been in use for the better part of a century, the collection efficiency of the standard gauge was tested and found to be in fact rather poor. In very light winds, 50% of particulate at 100 μm and 80% at 200 μm might be captured, but these efficiencies fell off very rapidly with increasing wind speed. Such a gauge would thus give no more than an indication of nuisance from coarse aerosol—respirable aerosol could not be captured.

Later work showed that a gauge shaped like an upside-down Frisbee® has a far superior capture performance. With the addition of a foam substrate, collection efficiencies in excess of 80% in light winds, even for particle diameters as small as 50 μm, can be achieved and these efficiencies remain good at moderate wind speeds. Such gauges are now commercially available (www.hanby.co.uk).

77.6.4 Active Aerosol Sampling

Measurement of particulate matter of aerodynamic diameter <50 μm in air requires an active sampling system.

A conventional (e.g., Andersen) high-volume sampler for total suspended particulate (TSP) consists of a filter and a powerful blower. The system is mounted in a housing with a pitched roof, the air being drawn in under the eaves, i.e., at a height of about 1.1 m. Air sampling rates are in excess of 1 m^3/min. The filter may be changed and weighed, or otherwise analyzed, after 24 h.

Size discrimination can be introduced into such samplers by requiring the air to follow a tortuous path. Small particles then follow the streamlines while large particles are deposited. As a first stage, the high-volume sampler may have an inlet impaction chamber, which only allows the PM_{10} or $PM_{2.5}$ fraction to be subsequently captured in the filter.

In a more sophisticated system, the filter may be replaced with a cascade impactor. This consists of a stack of plates with successively smaller perforations. The holes are staggered so that air passing through one hole impacts on the following plate. As the perforation size diminishes, the air speed increases and the size of particle which can escape diminish. By covering each plate with a removable membrane, the particulate captured in each size fraction can be analyzed and, in principle, weighed. The classic Andersen systems are now available from Thermo-Scientific.

Low-flow systems (1 m^3/h) are also available which permit measurement of PM_{10} or $PM_{2.5}$. Inlet heads, certified by the EPA, are available for various cutoff diameters. Reference [22] quotes capture efficiencies for such devices as a function of particle size.

77.7 Fast Ambient Monitoring

The alternative to a mechanically based, slow response sampler which must be visited regularly is an automatic, fast response system, self-calibrating if possible, which is logged on site and may be interrogated remotely. Routine visits are then minimized, while visits can be made urgently if there is an alarm. We list later instruments currently available for various common gases. Most of these systems are designed to be bench or rack-mountable, weighing typically in the range 15–30 kg. Autocalibration versions are available for most of them. Most manufacturers also supply appropriate logging systems with user-friendly software. The lower detectable limit (LDL) is conventionally two times the standard deviation of the noise, while the response time is typically the rise time to 95% of final value; with some instruments, there may also be a significant lag time. Obviously, by averaging over a longer sampling period, the LDL can be reduced. The price of most gas monitors falls in the range of £5000–£10000 plus sales tax (U.K., August 2012), depending on model and manufacturer. Hydrocarbon monitors are slightly more; aerosol monitors are twice as expensive.

77.7.1 SO_2

For the last 20 years, most ambient monitors for SO_2 have used UV fluorescence; when irradiated, SO_2 molecules reemit light in the range 220–240 nm. The measured light intensity is then proportional to the SO_2 concentration. Currently available systems have response times of order 1–2 min and LDL of order 0.3–1 ppb. Most manufacturers also supply a converter which allows H_2S to be measured as a separate channel.

77.7.2 NO_x

Most ambient monitors for NO_x employ chemiluminescence. Ozone generated within the instrument is mixed with the sample air. NO in the sample then reacts very rapidly to form NO_2, with the emission of IR radiation (peaking at 1200 nm). This is a first-order reaction; IR emission is therefore proportional to the NO concentration. Total NO_x may be measured by first passing the sample air over a catalytic converter to reduce any NO_2 present to NO. By alternating this conversion, or through the use of a dual channel system, NO_2 concentrations may be found by difference. Inevitably, the measurement of NO_2 will be noisier than that of NO or NO_x separately. Currently available systems have response times of order 30–90 s and LDL of order 0.4–1 ppb. Most manufacturers also supply a converter which allows NH_3 to be measured as a separate channel.

77.7.3 O_3

Most ambient monitors for O_3 now use modulated UV absorption. O_3 has a broad absorption spectrum in the UV with a peak around 254 nm. Practical instruments alternate the gas in the absorption cell between sample air and a deozonized reference gas; from the Beer–Lambert law, the log ratio of the

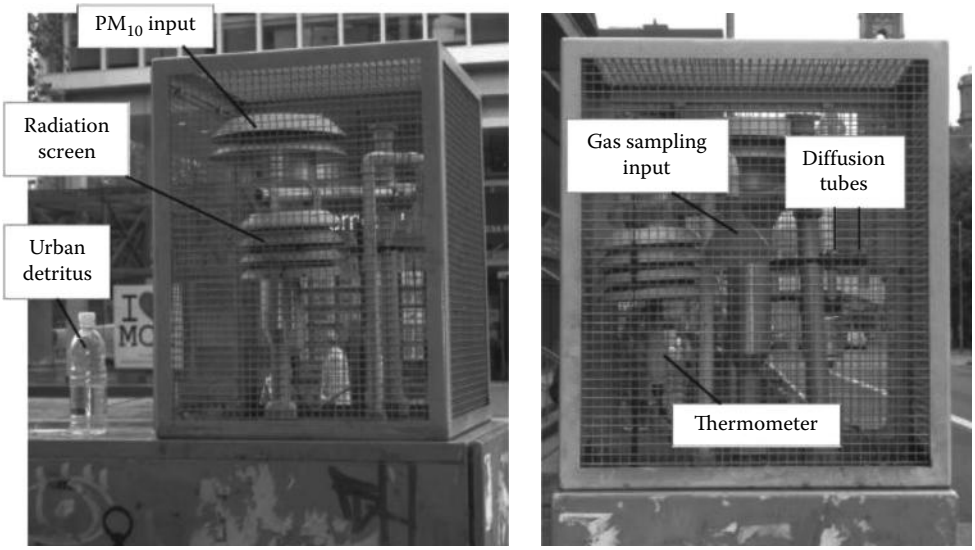


FIGURE 77.2 Sampling inputs for an air quality monitoring site in Central Manchester. The site monitors NO, NO_2 (chemiluminescence), and PM_{10} (β-absorption). The instruments are secured in the steel street-side box beneath: vandalism is a chronic problem at unmanned sites. This location is heavily trafficked, with the NO_2 concentrations being well in excess of the long-term limit of 40 μg/m^3 required under European law; peak diurnal concentrations of PM_{10} appear to be marginally within the limit.

signal must then be proportional to the O_3 concentration. Currently available systems have response times of order 20–50 s and LDL of order 0.4–1 ppb.

77.7.4 CO and CO_2

The standard method of measuring ambient CO or CO_2 is to use modulated IR absorption (NDIR); modulation of the signal allows detector offsets to be removed and may be achieved by alternating between sample and reference gas in the absorption chamber. A conventional way of doing this is through the deployment of a rotating wheel with gas cells containing N_2 and CO (or CO_2). Currently available systems have response times of order 20–60 s and LDL of order 40–100 ppb CO. An interesting variant of IR absorption is "photoacoustic detection," where the warming of the analyte is measured rather than the loss in IR flux. The irradiation is modulated—the varying heating of the sample gas is then registered as an oscillating pressure in the measurement chamber.

77.7.5 Hydrocarbons

The instruments available for the detection of ambient hydrocarbons tend to be somewhat more diverse than for the gases listed earlier. The most widely used technique is FID. A hydrogen flame is burned in the sample gas between two electrodes and the flame conductivity measured. This tends to be proportional to the concentration of ·CH ions in the flame, though the calibration varies between hydrocarbon species. Most commercial instruments allow CH_4 to be measured separately by first passing the sample gas across an oxidation catalyst. Nonmethane hydrocarbons (NMHC) are thereby removed from the sample; they may be estimated by difference from the total hydrocarbon (THC) channel. In principle, a FID is very fast, with some commercial instruments having response times as short as a few seconds. There is, however, a trade-off between sensitivity and response time, a sampling time of 1 min or more being required to give a LDL of 20 ppb.

TABLE 77.2 List of a Small Selection of Manufacturers of Cross-Duct Emission Systems and the Technique Employed

Supplier	Technique	Gases
Codel	NDIR	CO, CO_2, SO_2, NO, NO_2, HCl, H_2O, CH_4
Servomex	TDLAS	HCl, HF, NH_3, CO, CO_2, H_2S, HCN, CH_4
Sick AG	DOAS (UV)	SO_2, NO, NO_2, NH_3
	TDLAS	HF, HCl, NH_3
Teledyne Monitor Labs	Modulated UV	NO, SO_2
OPSIS	DOAS (UV or IR)	Many
	TDLAS	HF, HCl, NH_3, CO, CO_2

Note: The list is not intended to be exhaustive and no particular recommendation is implied.

Many of the manufacturers listed in Table 77.2 supply direct-throughput FIDs for ambient. Thermo-Scientific offers an instrument (55i) which combines a FID with a cycled gas chromatograph, thereby better guaranteeing separation between methane and NMHC, though at the cost of a longer cycle time. The Photovac range of instruments, now supplied by Inficon, includes a handheld FID (MicroFID), which is designed to be intrinsically safe and intended for, e.g., workplace monitoring or leak detection. As such it is extremely fast: <3 s to 90% response.

An alternative technique for hydrocarbons is PID. This is nonspecific, with the detector also responding to inorganic gases of low ionization potential (e.g., NH_3). It is, however, very fast (with response times from 0.2 s for an LDL of 10 ppb) and the instruments can be made to be economical, lightweight, and intrinsically safe. A typical application might be as a handheld leak detector in a chemical plant, or as a personal monitor of toxic solvents. Instruments are marketed by Photovac and RAE Systems.

77.8 Lightweight Chemical Sensors

Where weight and power consumption are critical issues, techniques based on surface chemistry may be used. Unisearch Associates markets a NO_x sensor ("Luminox") based on wet chemiluminesence: sample air is drawn over a wick containing a proprietary solution which reacts with the target gas. Light emitted in the reaction is monitored. The system is fast (response time $\leq$0.5 s), sensitive (LDL $\leq$0.1 ppb), and lightweight.

Small, lightweight sensors based on miniaturized fuel cells have also been developed [23] for the detection of NO, NO_2, CO, and other gases (cf. Alphasense in Table 77.3). These are prone to drift, but with suitable analysis this can be allowed for. The point is that such sensors are cheap enough to be distributed in large numbers around a site, permitting real-time spatial mapping of pollutant distribution.

77.8.1 Aerosols

Measuring atmospheric aerosol concentrations at ambient levels in real time is more demanding than measuring trace gases because the pollutant is concentrated into a relatively small number of tiny particles, inaccessible to spectroscopic examination. Optical methods are nevertheless useful. The most straightforward method is to look at the absorption of radiation across a test cell, preferably at several wavelengths. This is the principle of operation of the Aethalometer marketed by Magee Scientific and currently used in the United Kingdom black carbon monitoring network. The measured signal arises from the combined absorption from many particles within the test cell; some discrimination between elemental carbon and organic carbon compounds is provided by comparing the absorption at different wavelengths.

A more sophisticated technique is that of the optical particle counter (OPC, e.g., from Grimm or Turnkey Instruments). In this case, red or IR light from a laser diode is measured as it is scattered by individual particles as they pass through the system. The intensity and distribution of the scattered

TABLE 77.3 List of Useful Contact Details

AEA Technology Plc., Harwell, Didcot, Oxon., OX11 0QJ, United Kingdom
+44 (0) 796 870 7784, www.aeat.co.uk

Aerodyne Research Inc., 45 Manning Road, Billerica, MA 01821-3976
+1 978-663-9500, www.aerodyne.com

Alphasense Ltd., Sensor Technology House, 300 Avenue West, Skyline 120, Great Notley, Essex, CM77 7AA, United Kingdom
+44 (0) 1376 556 700, www.alphasense.com

Beijing Zhong Sheng Tai Ke, Building #2, Yuquanhuigu, 3, Minzhuang Road, Haidian District, China
+86-10-88850188, www.bjzstk.com/h/h/English

Casella Measurement, Regent House, Wolseley Road, Kempston, Beds, MK42 7JY, United Kingdom
+44 (0) 1234 844100, www.casellameasurement.com

Codel, Station Building, Station Road, Bakewell, Derbyshire, DE45 1GE, United Kingdom
+44 (0) 1629 814351, www.codel.co.uk

Dekati Ltd., Osuusmyllynkatu 13, FIN-33700 Tampere, Finland
+358 3 3578, www.dekati.fi

Ecotech Pty Ltd, 1492 Ferntree Gully Road, Knoxfield, VIC, 3180, Australia
+61 (3) 9730 7800, www.ecotech.com

Environment Agency, Horizon House, Deanery Road, Bristol, BS1 5AH, United Kingdom
+44 (0) 370 850 6506, www.environment-agency.gov.uk

Environmental Protection Agency, Ariel Rios Building, 1200 Pennsylvania Ave. NW, Washington, DC 20460
+1 (202) 272-0167, www.epa.gov

Environnement SA, 111 bd Robespierre, BP 4513, 78304 Poissy, France
+33 1 39 22 38 00, www.environnement-sa.com

Enviro Technology Services plc, Kingfisher Business Park, London Road, Stroud, Gloucestershire, GL5 2BY, United Kingdom
+44 (0) 1453 733200, www.et.co.uk

Department of Food, Environment and Rural Affairs, Whitehall Place, London, SW1A 2HH, United Kingdom
+44 (0) 845 933 5577, www.defra.gov.uk

Gasmet Technologies Oy Pulttitie 8 A, FI-00880 Helsinki, Finland
+358 9 7590 0400, http://www.gasmet.fi

Gradko International Ltd., St. Martins House, 77 Wales Street, Winchester, Hants., SO23 0RH, United Kingdom
+44 (0) 1962 860331, www.gradko.com

Grimm Aerosol Technik, Dorfstrasse 9, 83404 Ainring, Germany
+49 (0) 8654/578-0, www.grimm-aerosol.com

Health Protection Agency, Centre for Radiation, Chemical & Environmental Hazards, Chilton, Didcot, Oxon, OX11 0RQ, United Kingdom
+44 (0) 1235 831600, http://www.hpa.org.uk

Horiba, 17671 Armstrong Ave., Irvine, CA 92614
+1 949 250 4811, www.horiba.com

Infinicon AG, Hintergasse 15B, CH-7310 Bad Ragaz, Switzerland
www.inficon.com

Kipp & Zonen, Delftechpark 36, 2628 XH Delft, The Netherlands
+31 (0) 15 2755 210, www.kippzonen.com

Land Instruments Int., Stubley Lane, Dronfield, Derbyshire S18 1DJ, United Kingdom
+44 (0) 1246 417691, www.landinst.com

Magee Scientific Corp., 1916A M.L. King Jr. Way, Berkeley, CA 94704
+1 510 845 2801, www.mageesci.com

National Physical Laboratory, Hampton Road, Teddington, Middx, TW11 0LW, United Kingdom
+44 (0) 20 8977 3222, www.npl.co.uk

OPSIS AB, Box 244, SE-244 02 Furulund, Sweden
+46 46 72 25 00, www.opsis.se

PCME Ltd., Clearview Building, 60 Edison Road, St. Ives, Cambs., PE27 3GH, United Kingdom
+44 (0) 1480 468200, www.pcme.com

RAE Systems Inc., 3775 North First Street, San Jose, CA 95134
+1 877-723-2878, www.raesystems.com

Servomex Limited, Jarvis Brook, Crowborough, East Sussex, TN6 3FB, United Kingdom
+44 (0) 1892 652181, www.servomex.com

Sick AG, Erwin-Sick-Str. 1, 79183 Waldkirch, Germany
+49 7681 202-0, www.sick.com

Signal Group Ltd., Ambitech Division, Regal Way, Faringdon, Oxon. SN7 7BX, United Kingdom
+44 (0) 1367 242660, www.signal-group.com

Spectrasyne Ltd., 11 Shackleton Close, Bowerhill, Melksham, Wilts., SN12 6EY, United Kingdom
+44(0) 1225 702303, www.spectrasyne.ltd.uk

TABLE 77.3 (continued) List of Useful Contact Details

Teledyne—Monitor Labs, 35 Inverness Drive East, Englewood, CO 80112 +1(303) 792 3300, www.monitorlabs.com	Turnkey Instruments Ltd., 1-2 Dalby Court, Gadbrook Business Centre, Northwich, Cheshire, CW9 7TN, United Kingdom +44 (0) 1606 44520, http://www.turnkey-instruments.com
Thermo Fisher Scientific, 81 Wyman Street, Waltham, MA 02454 +1 781-622-1000, www.thermoscientific.com/air	Unisearch Associates, 96 Bradwick Drive, Concord, Ontario, L4K 1K8, Canada +1 (905) 669-3547, www.unisearch-associates.com
TSI Incorporated, 500 Cardigan Road, Shoreview, MN 55126 +1 855-520-3738, http://www.tsi.com	United Sciences Testing, 201 Commonwealth Drive, Warrendale, PA 15086 +1 724-778-6934, www.ustirata.com

Note: For reasons of space, only a primary address has been given for each contact. Most of the manufacturers will have subsidiaries or agents in your own country; these should normally be listed on their websites.

light can be used to estimate the size of each individual particle. An OPC can thus provide both a particle size spectrum and an estimate of the total aerosol volume per unit volume of sampled air. The technique, however, has inherent limitations both for large and small particles. For particles smaller than the illuminating wavelength, the scattered intensity per unit mass is proportional to the particle volume. Particles smaller than ~0.1 μm are undetectable. Geometric scattering from large particles on the other hand gives a scattered intensity per unit mass inversely proportional to the particle diameter and thus also falls off with particle size. More seriously, the counting statistics can become very poor. An ambient aerosol concentration of 10 $\mu g/m^3$ made up of spheres of 10 μm in diameter and unit density corresponds to a number concentration of only ~20 particles/cm^3. For coarse particulate, the volume of input air actually passing through the optical field over the desired averaging time must thus be of order cm^3. This cuts both ways, of course: if the sampled volume is too large, then the detection system will not be able to keep up with the signals from the vast numbers of tiny particles.

It is also difficult to deliver coarse aerosol from the sample inlet to the detector. Sample lines should be short, straight, and preferably vertical.

Direct gravimetric methods are also employed, based either on β-absorption or on an oscillating filter. The sensitivity required for such a measurement may be appreciated if we consider a sampler drawing 3 L/min which is intended to measure aerosol concentrations over a sampling period of 6 min. A detection limit of 5 $\mu g/m^3$ then requires that the sample mass be measured to a precision of 90 ng.

In the β-absorption technique (e.g., Horiba, model APDA-371), the sample air is drawn through a filter tape which is exposed to β-radiation. As atmospheric aerosol is deposited on the filter, its transparency to the radiation gradually diminishes. Differentiating the transmitted signal with respect to time gives an aerosol concentration at close to real time. The tape is advanced every hour, in effect rezeroing the measurement and capturing a timed sample for subsequent analysis if so desired.

In the tapered element oscillating microbalance system (TEOM—Thermo-Scientific, model series 1400ab) system, the filter is balanced on a hollow tube through which the sample air is drawn. The tube is clamped at its base and induced to vibrate at a frequency which depends on the mass of the filter. Again, the differential of the filter mass with respect to time provides the aerosol concentration.

Whichever system is employed, care must be taken in the choice of inlet port so that the desired fraction of ambient particulate is sampled. Typically, a PM_{10} sampling head would be employed.

It is not always the mass of fine aerosol that is of regulatory or scientific interest. Often the number density is of more interest. As noted in an earlier section, ultrafine aerosol particles may be detected and counted electrostatically. An alternative technique is the condensation particle counter (CPC, e.g., from Grimm or TSI). In this device, the sampled air is passed through a warm chamber saturated with butanol or water and then into a cool chamber. The supersaturated vapor then condenses onto

the particles, which grow into droplets large enough to be detected optically. As with the electrostatic technique, this can count particles of aerodynamic diameter down to a few nm. If size discrimination is required, then the system can be used as the detector in a differential mobility analyzer.

Fine particles can also be differentiated by accelerating them through a critical orifice, i.e., in a supersonic jet, into a low pressure chamber. Smaller particles will then travel faster than larger ones. This is the principle underlying Aerodyne's aerosol mass spectrometer. Having separated the incoming fine aerosol into a spectrum of different sizes (40–1000 nm) using a aerodynamic lens and a chopper, the individual particles can then be volatilized and analyzed by one of a range of different detector modules. It is thus possible almost in real time to obtain both a size spectrum of ambient aerosol and a chemical profile of the particles in each size range.

77.9 Remote Monitoring

While point samples of air pollution can be very valuable, they are limited both in their position, which is normally near ground level, and in representing a single point rather than a broader spatial sample. A variety of optical techniques are available which permit more general measurement of pollution in the atmosphere.

77.9.1 Long-Path Measurements

With the use of a broadband source and a retroreflector, spectroscopic analysis of the returned light can provide spatially averaged measurements of a wide range of pollutants. OPSIS is the market leader.

For ambient monitoring, the OPSIS system applies DOAS to an absorption spectrum ranging between the UV and the near IR. Standard systems are available for monitoring three (NO_2, SO_2, O_3) or five (+ toluene, benzene) species. Modules may be added for a range of trace gases. The system may also be used for emissions monitoring. An absorption band in the IR may then be used and the calibration must be optimized to the optical density of the particular flue.

77.9.1.1 Overhead Burden

If we are interested in concentrations well away from the ground, a retroreflector may be impractical. In this case, spectral measurements of sunlight scattered from the sky may provide us with an estimate of the total overhead burden ("column") of a pollutant. "Cospecs" (analog correlation spectrometers) were available from the early 1970s for the measurement of SO_2 or NO_2 burden. These are most valuable in conjunction with mobile surveys of elevated emissions (power stations or volcanoes), since they allow the investigator to know when he is beneath the plume. With very careful calibration, and knowledge of the wind speed, they may also be used to estimate the pollutant flux. The instrument has no absolute zero level and so has difficulty distinguishing very broad plumes. While the Cospec is no longer manufactured, much more sophisticated digital systems are now deployed, for example, on satellite platforms to monitor regional SO_2 emissions [24].

The classic measurement of overhead pollutant burden is that of stratospheric O_3 using the Dobson spectrophotometer. This measures the difference in irradiance at two wavelengths 20 nm apart close to the peak of UV absorption by O_3: other things being equal, this difference is proportional to the overhead burden of O_3. The system was used to detect the ozone hole over Antarctica in 1984 [25]. The detection method had then remained unchanged since the original instruments of 60 years previously, though the optics had been refined and the valve electronics replaced with solid state. This continuity of method was important in giving confidence in the surprising observation of the ozone hole.

The Dobson instrument is now obsolete; an improved instrument, the Brewer spectrophotometer, which also measures column SO_2, is available from Kipp & Zonen.

77.9.2 Lidar

Light detection and ranging (lidar) is the optical equivalent of Radar [11,26]. A pulse of laser light is directed into the atmosphere. Backscattered light is collected by a telescope and directed to a photomultiplier tube. The strength of the return signal corresponding to a given range $R = ct/2$ is given by the lidar equation:

$$V(R) = \frac{CW}{R^2}\sigma_b\, n(R) \exp\left\{-2\int_0^R (\sigma(\lambda)\chi(r) - \sigma_e\, n(r))dr\right\} \tag{77.4}$$

where

C is a system constant
W is the pulse energy
$n(R)$ is the density of scatterers
σ_b and σ_e are the backscatter and extinction cross sections for scattering, respectively
$\sigma(\lambda)$ is the absorption cross section from a tracer gas of concentration χ

If the extinction is small, scanning the laser beam through an elevated plume allows a cross section of the scatterer density to be built up. The range resolution of this cross section depends on the frequency at which the signal is digitized: a 60 MHz digitizer gives a range resolution of 2.5 m. The time resolution of the measurement depends on the pulse repetition rate of the laser. At 30 Hz, a 2 s scan would normally give adequate lateral spatial resolution. *A priori* calibration for particulate density requires knowledge of σ_b, which in turn depends upon the particle size and refractive index. In practice, an independent gravimetric measurement must be made at some point in the scanned area.

Alternatively, shots may be alternated between two nearby wavelengths (λ_1 and λ_2) having very different values of $\sigma(\lambda)$. Neglecting variations in the other terms, the lidar equation may be rearranged to give

$$\chi(R) = \frac{1}{2(\sigma(\lambda_1) - \sigma(\lambda_2))}\frac{\partial}{\partial R}\ln\left\{\frac{V_2(R)}{V_1(R)}\right\} \tag{77.5}$$

Since the differential absorption $\sigma(\lambda_1) - \sigma(\lambda_2)$ may be measured in the laboratory, we now have a remote, range-resolved measurement of the tracer gas. This technique is known as differential absorption lidar (DIAL) [11] and has been applied to a wide range of atmospheric pollutants (SO_2, NO_2, O_3, CH_4, hydrocarbons, Hg, etc.). Its particular strength is in identifying and quantifying fugitive emissions, e.g., CH_4 from cattle or leaks in refinery plant [27]. Because of the greater subtlety of the measurement, it does not in general have the spatial or temporal resolution of a simple backscatter lidar. For a given DIAL system at a given range, the product of spatial, temporal, and concentration resolution is approximately fixed; optimal performance in one is at the cost of worse performance in the others. If the signal is rather weak, there is a temptation to average measurements over a period greater than the timescale for a significant change in the target gas concentration. Great care should be taken in this case since the nonlinearity of Equation 77.5 can then lead to very serious distortions [28].

DIAL services are available from (for example) National Physical Laboratory (NPL) or Spectrasyne in the United Kingdom.

77.10 Emissions Monitoring

Industrial plants are in general authorized to release effluent to the atmosphere on the condition that such emissions are monitored reliably enough to satisfy the regulator that authorized limits have not been exceeded. This has increasingly led to the installation of continuous emissions monitoring systems (CEMS) for prescribed substances, with measurements being automatically displayed and archived. Years ago, we used to hear stories of power stations turning off their electrostatic precipitators at night in order to save a little bit of power—this is no longer feasible.

Aside from satisfying regulatory conditions, such monitoring can also be of value to the operator in controlling his plant (e.g., optimizing the excess O_2 in a combustion plant). Note that any measured value of pollutant concentration must be converted to standard conditions of temperature, pressure, humidity, and excess O_2 before comparison with a regulatory standard.

To be credible, any such monitoring must be made to a defensible standard. In the United Kingdom, a measurement certification scheme (MCERTS) is run by the Environment Agency. Beyond merely regulating emissions, the Agency does its best to help industry to make reliable measurements and to work within the regulatory system successfully. Among other activities, it has published a series of technical guidance notes (up to M21 [29] at the time of writing). The system certifies the instruments to be used, the monitoring protocols, staff training, etc.

No respectable business wishes to pollute its environment. By training and advising industrial emitters, the Agency makes its own task easier. Similar arrangements are in place in other industrial countries. In the United States, certified methods are promulgated by the Emissions Monitoring Center of the EPA. In Germany, certification is the responsibility of the four Technical Inspection Associations (*Technischer Überwachungs-Verein*—TÜV), which in fact operate as technical consultancies worldwide and in contributing to the development of International Standards.

77.10.1 Isokinetic Sampling of Aerosol

The gold standard for the measurement of aerosol emissions is to take a sample of the effluent and weigh the particulate therein. This is done by introducing a sample line into the flue, pumping the gas through a filter, and measuring the mass gained per metered volume of flue gas. It is essential that the gas speed at the sample inlet is the same as the flow velocity in the flue: if it is too fast (or slow), large particulate will be under (or over) sampled. To ensure a representative sample, such measurements must be made at several diameters downstream of any bends or confluences in the flue. It is also necessary to take several samples (4 or 8) over the cross section of the flue. Even taking every precaution, the precision of the method may be as poor as 20% of the mean value. In practice, the nominal calibration against continuous monitors may be found to be *very much* worse than this from trial to trial.

Since each sample may take 30 min, this procedure clearly does not give a real-time measurement of particulate load. In poorly designed plants, there may also be severe practical difficulties in gaining access to the flue with heavy, powered sampling equipment at a sufficient height for the flow to be uniform. Nevertheless, this is the only direct gravimetric method of measuring particulate concentration in a flue. For this reason, regular (at least annual) such measurements would be required for the calibration of any continuous monitors.

77.10.2 Aerosol Sampling: *In Situ* Methods

In situ methods are particularly attractive for emissions measurements because the pollutant is measured directly without any possibility of loss or transformation in the path from flue to instrument.

The most obvious way of continuously measuring the dust load in a flue is to pass a beam of light across the flue and measure the obscuration. It may be assumed that the transmission obeys the Beer–Lambert law and that the optical density is proportional to the particulate load. Practical instruments,

of course, require a degree of sophistication to allow for instrumental offsets. It will not generally be possible to turn off the plant at regular intervals to check the instrument zero! Typically, the light source will be modulated and a modulated signal measured. Purge air is normally blown over transmitter and receptor windows (and the retroreflector if used) in order to keep the light path clear. The sensitivity of the system will nevertheless drift with time (and will change with the composition of particulate) so regular gravimetric calibration is essential.

Such cross-duct monitors are simple and robust, and they provide a measurement across the full width of the flue. They are limited to optical densities in the approximate range 0.01–2.5. Systems are available from Codel, Land Combustion, Grimm, United Sciences, and others.

For very narrow ducts, or for very low particle densities, a more sensitive technique must be chosen. As with ambient particulate monitoring, it is possible to use sideways scattering from a light beam. The transmission and reception optics are now set up so that there is no direct path between the two. Light scattered from particles in the flue may, however, be detected. This gives a true zero level in the absence of particulate, and a signal proportional to particle density. Care must be taken to ensure that the light path is representative of the particle distribution in the flue. By switching the light path, it may be possible to route it through a filter of known optical density, thereby permitting online calibration. Such systems are available from Erwin-Sick and from United Sciences.

An alternative technique is the triboelectric monitor. In this device, a metal rod is inserted across the flow. Particles striking the rod transfer static charge to it. The leakage current (or, better, the rms component thereof) is thus proportional to the particle flux along the duct. This method is fast, simple, and robust: it is little affected by the accumulation of dirt on the probe and is insensitive to the presence of steam (which would, of course, preclude optical measurements of particle density). In very wet flows, there are possible problems with leakage across the insulated base of the probe: these may be overcome with the use of purge air or extended insulation. It should be noted that, with this device, calibration is required not only for gravimetric particle density but also for the flow speed in the duct. If this speed varies significantly, the implicit particle density will be in error. Systems are available from PCME or Codel.

77.10.3 Gas Sampling: Spectroscopic Methods

The cross-beam technique can be applied to *in situ* sampling for pollutant gases. Practical devices are usually based on nondispersive absorption of IR or UV. DOAS may also be used. Sophisticated instrumental design is necessary to ensure

- Minimal interference from other species present in the effluent.
- Minimal interference from IR emissions from the hot flue gases. Typically, a modulated system would be used.
- Insensitivity to instrumental drifts (dirty windows, misaligned optics, aging of source, and detector). This may be achieved with dual light-path systems, where only one path is sensitive to the target gas. The differential signal then gives a robust measurement of concentration in the flue. This may be calibrated online by switching in cells containing a known concentration of the target gas.

Clearly, the operator will not wish to turn off his plant so that zero and span on the monitor can be measured directly. An alternative is to encase the beam path in a filtered chamber within the flue. The chamber can then be filled with clean or span air as necessary. This has the further advantage that dust can then be excluded from the beam.

77.10.4 Gas Sampling: Extractive Techniques

If flue gases are to be sampled by an extractive technique, great care must be taken in delivering the sample from the flue to the monitor. The simplest technique is in effect to take the probe sample chamber

outside the flue and deliver the flue gas to it using a heated, filtered, sample line. The same cross-duct technique can then be used as if the sample chamber were physically in the flue itself. In many systems, however, the same techniques are used for extractive emissions monitoring as are used for ambient gas monitoring; in many cases, even the same instruments might be used. Since process gases tend to be hotter, wetter, dustier, and more acidic than ambient air, there will often be a need to change the condition of the sample so that it is acceptable to the monitor. The aim of sampling technology is to deliver an acceptable sample to the instrument at a target gas concentration which bears a known relation to the gas concentration in the flue [30].

Generally, the sample may be conditioned by dilution, filtering, cooling, or drying. The techniques chosen will depend on the gas being measured and on the instrument being used. At its simplest, conditioning may amount to the sample gas being diluted with several hundred times its volume of clean air. Normal ambient monitors could then be used, but separate measurements would have to be made of O_2 and H_2O in order to express the measured concentration in standard form.

To avoid clogging of the sample line, it is usual to filter the sample as soon as practically possible. To avoid condensation, the normal practice is to keep the line heated. If, however, the instrument cannot tolerate flue gas temperatures, cooling will be necessary at some stage. If this is done without dilution, water will condense out and subsequent measurements will be on a dry gas basis. In itself, this is advantageous but any acid gases will remain with the water and be lost. Alternatively, the sample may be dried by passing it through a permeation drier. Such driers cannot tolerate acid gases. SO_2 (for example) must therefore be measured on a wet gas basis, with temperature and condensation being controlled by appropriate dilution.

Considerable skill is involved in designing and implementing continuous extractive emissions monitoring systems. Instruments are available for hydrocarbons and for the conventional range of stack gases from Codel, Horiba, Land Combustion, Teledyne Monitor Labs, Servomex, Sick AG, Signal-Ambitech, and Thermo-Scientific, among others.

Defining Terms

Aerosol: Particulate matter so finely divided in a gas that it remains in suspension.

Burden: The integrated concentration of an analyte along a defined path.

Column: The burden along a vertical path between the surface and the edge of the atmosphere.

Concentration: The mass or volume of an analyte in a unit sample volume. The volume may need to be defined in terms of temperature, pressure, humidity, etc.

Deposition: The transfer of material from the atmosphere to a surface.

Monitoring: A routine series of measurements made for control purposes.

Protocol: A written list of instructions for obtaining a reliable measurement.

References

1. B. M. R. Green, J. C. H. Miles, E. J. Bradley, and D. M. Rees, Radon atlas of England and Wales, National Radiological Protection Board, Report no. W26, 2002.
2. M. Bennett and S. Christie, An application of backscatter lidar to model the odour nuisance arising from aircraft tyre smoke, *Int. J. Environ. Pollut*., 44: 316–328, 2011.
3. R. F. Griffiths, The effect of uncertainties in human toxic response on hazard range estimation for ammonia and chlorine, *Atmos. Environ*., 18: 1195–1206, 1984.
4. R. E. Munn, *The Design of Air Quality Monitoring Networks*, London, U.K.: Macmillan, 1981.
5. B. B. Kebbekus and S. Mitra, *Environmental Chemical Analysis*, 1st edn., Newark, NJ: Blackie Academic & Professional, 1998.

6. Warren Spring Laboratory, *National Survey of Air Pollution 1961–1971*, Vol. 1, London, U.K.: HMSO, 1972.
7. D. J. Ball and R. Hume, The relative importance of vehicular and domestic emissions of dark smoke in Greater London in the mid-1970s, the significance of smoke shade measurements and an explanation of the relationship of smoke shade to gravimetric measurements of particulate, *Atmos. Environ*., 11: 1065–1073, 1977.
8. M. Bennett, C. Rogers, and S. Sutton, Mobile measurements of winter SO_2 levels in London 1983–84, *Atmos. Environ*., 20: 461–470, 1986.
9. V. Bertorelli and R. G. Derwent, Air quality A to Z: A directory of air quality data for the United Kingdom in the 1990s, Meteorological Office.
10. C. N. Banwell and E. M. McCash, *Fundamentals of Molecular Spectroscopy*, 4th edn., London, U.K.: McGraw-Hill, 1994.
11. M. W. Sigrist (ed.), *Air Monitoring by Spectroscopic Techniques*, New York: Wiley-Interscience, 1994.
12. R. F. Griffiths, Emissions and environmental monitoring using energetic UV radiation: A new development in portable ambient monitoring, *Monitor '93*, Manchester, Spring Innovations Ltd., Bramhall, U.K., pp. 57–62, 1993.
13. G. A. Sehmel, Particle and gas dry deposition: A review, *Atmos. Environ*., 14: 983–1011, 1980.
14. Environmental Protection Agency, Provisional assessment of recent studies on health effects of particulate matter exposure, Report no. EPA/600/R-06/063, National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, 2006.
15. J. H. Vincent, *Aerosol Sampling. Science and Practice*, New York: John Wiley & Sons, 1989.
16. H. C. van de Hulst, *Light Scattering by Small Particles*, New York: Wiley, 1957.
17. E. D. Palmes, A. F. Gunniston, J. DiMattio, and C. Tomczyk, Personal sampler for nitrogen dioxide, *J. Amer. Ind. Hyg. Assoc*., 37: 570–577, 1976.
18. Airparif, Campagne de mesure autour de l'aéroport de Paris Charles de Gaulle, Airparif, Pôle Etudes 7, rue Crillon, 75004 Paris, 2009.
19. G. W. Campbell, J. R. Stedman, and K. Stevenson, A survey of nitrogen dioxide concentrations in the United Kingdom using diffusion tubes, July–December 1991, *Atmos. Environ*., 28: 477–486, 1994.
20. D. Mooney, A guide for local authorities purchasing air quality monitoring equipment, Report no. AEAT/ENV/R/2088 Issue 2, UK Department for Environment Food and Rural Affairs, 2006.
21. D. J. Hall, S. L. Upton, and G. W. Marsland, Designs for a deposition gauge and a flux gauge for monitoring ambient dust, *Atmos. Environ*., 28: 2963–2979, 1994.
22. D. Mark and D. J. Hall, Recent developments in airborne dust monitoring, *Clean Air*, 23: 193–217.
23. M. I. Mead, O. A. M. Popoola, G. B. Stewart, P. Landshoff, M. Calleja, M. Hayes, J. J. Baldovi, M. W. McLeod, T. F. Hodgson, J. Dicks, A. Lewis, J. Cohen, R. Baron, J. R. Saffell, and R. L. Jones, The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks, *Atmos. Environ*., 70: 186–203, 2013.
24. K.-P. Heue, C. A. M. Brenninkmeijer, A. K. Baker, A. Rauthe-Schőch, D. Walter, T. Wagner, C. Hőrmann et al., SO_2 and BrO observation in the plume of the Eyjafjallajőkull volcano 2010: CARIBIC and GOME-2 retrievals, *Atmos. Chem. Phys*., 11: 2973–2989, 2011.
25. J. C. Farman, B. G. Gardiner, and J. D. Shanklin, Large losses of total ozone reveal seasonal ClO_x/NO_x interactions, *Nature*, 315: 207–210, 1985.
26. D. J. Carruthers, H. Edmunds, M. Bennett, P. T. Woods, M. J. T. Milton, R. Robinson, B. Y. Underwood, and C. J. Franklin, Validation of the UK-ADMS dispersion model and assessment of its performance relative to R-91 and ISC using archived lidar data, Department of the Environment, Research Report no. DOE/HMIP/RR/95/022, March 1996.
27. R. H. Partridge, P. T. Woods, M. J. T. Milton, and A. J. Davenport, Gas standards and monitoring techniques for measurements of vehicle and industrial emissions, occupational exposure and air quality, *Monitor '93*, Manchester, Spring Innovations Ltd., Bramhall, U.K., pp. 23–35, 1993.

28. M. Bennett, The effect of plume intermittency upon differential absorption lidar measurements, *Atmos. Environ*., 32: 2423–2427, 1998.
29. Environment Agency, Stationary source emissions—A procedure to use an alternative method for measuring emissions of sulphur dioxide, using instrumental techniques, Technical Guidance Note (Monitoring) M21, 2010.
30. K. Honner, Continuous extractive gaseous emissions monitoring, *Monitor '93*, Manchester, Spring Innovations Ltd., Bramhall, U.K., pp. 53–56, 1993.

78

Water Quality Measurement

Kathleen M. Leonard
The University of Alabama in Huntsville

78.1 Introduction

Maintaining and verifying water quality is important in many environmental applications. The most obvious is in drinking water applications, but industrial and municipal wastewater, natural surface/groundwater, industrial process waters, and closed-loop control systems all require a certain range or maximum concentration of species to operate properly. Therefore, the development of instrumentation for monitoring and detecting these contaminants is a highly competitive area. Of course, the chemical species of interest depend on the ultimate use of the water, but this section deals with some of the water quality sensors and instrumentation being currently used.

Table 78.1 provides a list of the most commonly monitored chemical species in the area of water quality. The third column categorizes the parameters in the general categories of drinking water, wastewater (municipal), industrial, storm water, and ambient water quality. The "ultimate reference" source for accepted methods of determining chemical concentrations in water is known as the *Standard Method for the Examination of Water and Wastewater* [1]. This book has been updated regularly since 1905 and contains the current techniques recommended by Water Environment Federation (WEF) and the U.S. Environmental Protection Agency (EPA).

The next section introduces the theory behind some of the major types of water quality sensors, including electrical, optical, and chemical separation. Since each technique has its own particular water quality applications, a brief discussion of advantages and disadvantages is included. This section is limited to commercially available instruments commonly used both inside and outside the laboratory, although emphasis is placed on field (i.e., portable) instruments. Future trends will be addressed in the conclusion section.

TABLE 78.1 Common Chemical Species Monitored in Water Quality

Chemical Parameter	Chemical Symbol	Primary Use
Ammonium	NH_4	DW, WW
Arsenic	As	DW
Barium	Ba	DW
Bicarbonate	HCO_3^-	DW, WW
Cadmium	Cd	DW, WW
Calcium	Ca^{2+}	DW
Carbonate	CO_3^{2-}	DW
Chlorinated hydrocarbons	Various herbicides and pesticides	DW, SW
Chloride	Cl^-	DW, SW
Fluoride	F^-	DW
Chromium	Cr	DW
Dissolved oxygen	O_2	WQ
Hydrogen (pH)	H^+	DW, WW, WQ
Iron	Fe	DW
Lead	Pb	DW
Magnesium	Mg^{2+}	DW
Mercury	Hg	DW
Nitrate	NO_3^-	DW, WW, SW
Nitrogen	N	DW, WW
Petrochemicals	Various	
Phosphates	PO_4^{3-}	WW, SW, WQ
Potassium	K^+	WW
Selenium	Se	DW
Silver	Ag	DW
Sodium	Na^+	DW
Sulfate	SO_4^{2-}	DW, WW
Gross measures		
Alkalinity	As $CaCO_3$	DW
Biochemical oxygen demand	BOD	WW, WQ
Conductivity		
Chemical oxygen demand	COD	WW, WQ
Hardness	As $CaCO_3$	DW
Particle counts		WW, WQ
pH	H^+	DW, WW, WQ
Total organic carbon	TOC	SW, WW, WQ
Other parameters		
Color		
Dissolved solids	In ppm	SW, DW
Turbidity	In NTU	DW, WQ
Microbiological contaminants	Counts/100 mL	DW, WW, WQ
Radiological contaminants	In curies	DW

Note: DW, drinking water; WW, waste water; SW, storm water; WQ, water quality.

78.2 Theory

78.2.1 Electrical Methods of Analysis

There are various electric mechanisms employed for determining water quality concentrations. For example, in the area of electrochemical sensors, there are potentiometric, amperometric, and conductometric devices [2]. The pH sensor for determining hydrogen ion concentration is probably the most widely used instrument in this category.

78.2.1.1 Potentiometric Sensors

This type of sensor is based on the relationship of the electrochemical cell and the chemical activity of a sample, based on the general form of the Nernst equation:

$$E_{\text{cell}} = E^0 - \left(\frac{RT}{nF}\right)\ln\left\{\frac{(\text{red})}{(\text{ox})}\right\} \tag{78.1}$$

for the general reaction: ox + ne^- ↔ red. ox and red indicate the oxidized and reduced species [3]. The term RT/F has a value of 0.059. E^0 is the standard cell potential and E_{cell} is the adjusted cell potential. In short, the activities of the oxidized and reduced species determine the potential of the electrode. If the relationship of the reacting species is known, it is used to measure concentration of that species in another solution. In most electrochemical instruments, it is important to maintain a reference electrode whose potential remains constant for all cell conditions or one that can be easily calibrated for other species. These sensors are commonly known as ion-selective electrodes (ISE) in environmental applications (also known as ion-sensitive field-effect transistors or ISFET). They are all based on the activity level of a specific ion within a solution. The key to the ISE is the use of an ion-specific membrane with channel size proportional to the concentration of the ion. It is necessary to maintain a reference potential to ensure that the "ion concentration will be directly related to the substrate potential" [4]. Specific electrodes can be based on a gas electrode, metal electrode, oxidation–reduction electrode, membrane electrode, glass, liquid membrane, crystalline membrane, or electrode with metal contacting a slightly soluble salt [5]. Examples of this type of sensor and instrumentation for gross measurements include pH sensors, dissolved oxygen (DO), hardness, and dissolved solids.

Commercially available ISE include sensors for ammonia, chlorides, cyanide, iodide, fluoride, nitrates, potassium, and sodium. The instruments labeled as multiparameter water quality monitors, such as that available from Solomat, are usually composed of 1 m with separate probes for each measurement parameter. However, some of these multiparameter monitors have been upgraded with a multichannel probe that contains the separate probes within a single housing.

A membrane probe in popular usage is the DO probe based on the principle of polarography. It consists of a gas-permeable membrane over a silver anode and a platinum cathode within a cavity full of electrolyte (usually KCl). Figure 78.1 shows the placement of the anode and electrode in a typical polarographic membrane probe. Since the oxygen will be reduced at the cathode, a current is induced. The amount of oxygen can then be correlated to the current when a potential is applied to the anode and cathode.

The pH instrument is one of the most commonly used potentiometric sensors in water quality applications. It is based on a glass electrode that develops a potential related to hydrogen ion activity of the solution. Since this probe was initially developed in the early part of the 1900s, there are numerous probes and meters on the market. Small, battery-operated, handheld units are excellent choices for fieldwork. However, for more precise applications, advanced (i.e., expensive) laboratory-grade instruments are also available. Most pH probes require standardization and may have other upkeep requirements after prolonged usage.

78.2.1.2 Amperometric Sensors

The amperometric classification of electrical sensors is based on the measurement of current through a working electrode. An empirical relationship can be used to enhance the performance of these electrodes.

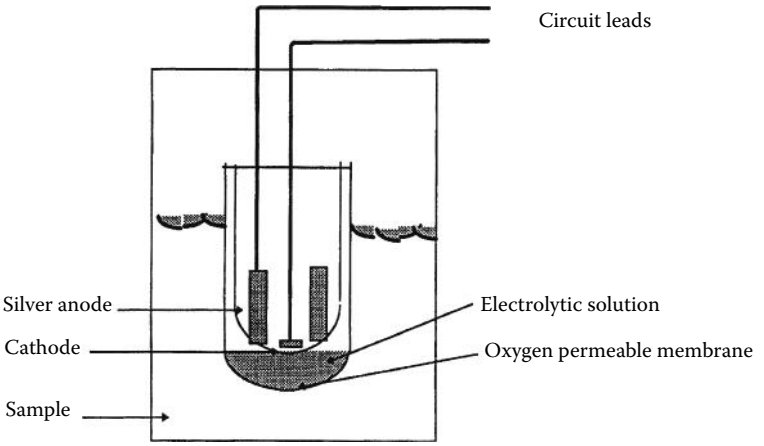


FIGURE 78.1 Simplified diagram of dissolved oxygen sensor based on polarography. The oxygen-permeable membrane surrounds the anode and cathode to deter other chemical species from entering the electrolytic solution.

An example of such a probe is an electrode for chlorine, consisting of a silver anode and platinum cathode within an electrolyte reservoir. A chemically specific membrane allows only the ions of interest within the probe cavity near the cathode. Reduction occurs at the cathode and then the silver anode reacts with the electrolyte and, through oxidation, positively charged ions produce a current.

78.2.1.3 Conductometric Sensors

The conductometric sensors are based on measurements of concentration of a sample due to modulation in conductivity. For example, conductivity is usually measured with a direct reading of electrical conductivity in water, which is then correlated with the number of ions present. The probe usually consists of two voltage electrodes (+ and –) and two current electrodes (also + and –) supported by some type of housing. The newer instruments have a temperature compensation option and data storage. Most will also provide a reading of total dissolved solids. Conductometric sensors are also available for resistivity, salinity, and temperature measurements.

The most commonly available DO sensors are galvanic electrodes with replaceable membranes. The current is proportional to the DO concentration under steady-state conditions [1]. One of the drawbacks of this type of probe has been the need for periodic calibration, either in air or in a saturated medium. Some of the newer products are self-calibrating, thus eliminating this error source. There are many companies that sell these products and the prices range from modest ($300) to high, depending on extras. If the application will be messy (e.g., wastewater), the added expense for the deluxe model, such as a self-cleaning probe, may be realized very quickly. A company that has been producing DO instruments for many years is YSI. It has a "deluxe" model with microprocessor control and RS232 interface for continuous reading.

78.2.2 Optical Sensors

This category involves any type of measurement system that relies on an optical property to measure a chemical concentration, including reflection, colorimetry, fluorescence, or absorption of light. The general equation relating wavelength of light (λ) to the energy of a photon (E) is given by the Planck's equation:

$$E = \frac{hc}{\lambda} \tag{78.2}$$

where

h is Planck's constant (6.626×10^{-34} J s)

c is the velocity of light (3.0×10^{8} m/s)

Some of these mechanisms are included in the category of "gross optical measurements." This category includes any method that uses reflection/refraction of light beams, colorimetry, or absorption of light. For example, in order to determine the efficiency of a water treatment system due to deteriorating filter capacity or to optimize coagulant dosages, a particle counting method is used. This type of measurement can give not only quantity, but also size ranges. This method is helpful in guarding against the intrusion of cryptosporidium cysts and guardia in a drinking water facility.

78.2.2.1 Nephelometry

In this method, light is directed to the sample and the light reflected at right angles to the beam is measured (Tyndall effect). This type of analysis is the accepted method for quantifying turbidity (cloudiness) since the intensity of light scattered by a water sample can be compared to the scattering of a standard (usually Formazin polymer) under similar conditions. The units measured are known as nephelometric turbidity units (NTUs), which are approximately comparable to the older "candle" units. A turbidimeter consists of a light source and a photoelectric sensor that measures the light scattered at 90° from the source.

78.2.2.2 Absorbance Methods

Certain chemical species exhibit natural light absorbance patterns over a range of wavelengths. These include nitrates/nitrites, heavy metals, unsaturated organics, and aromatics. This approach is valid even in multicontaminant solutions, since no two substances have exactly the same absorbance pattern. Several commercially available instruments take this technology to water quality applications with online capabilities. For example, in the water/wastewater area, there are flow-through monitors that measure a broad range of wavelengths of ultraviolet light simultaneously. This pattern is checked against a "signature" for the contaminant of interest and both the presence and concentration are verified.

Some contaminants do not have a natural absorbance, so a "conditioning" step is required in order to use this method. This usually involves the addition of a secondary chemical, or indicator, that will absorb in the presence of the contaminant. This is commonly known as induced light absorbance. Ammonia, phosphates, and chlorine absorbance probes are available with this conditioning step.

There are several advantages to using absorbance, including low maintenance, especially when compared to ISEs, high reliability, and automatic compensation for turbidity (which can cause havoc in other systems). One of the major drawbacks of the available models is the need to buy a separate unit for each contaminant, at a comparatively higher cost than ISE.

78.2.2.3 Colorimetry

This simple method works by correlating the color of a solution with the concentration of a specific chemical. The theory is that light absorption (A), which is the amount of light intensity (I/I_0) absorbed, is related to concentration based on Beer's law [5]:

$$A = \log\left(\frac{I}{I_0}\right) = k^n C \tag{78.3}$$

where

k^n is the constant for a particular solution
C is the concentration of the solution

Beer's law is valid for most water quality ranges of concentration, but it can be verified using dilutions of a specific contaminant. The first and still simplest colorimetric methods involve the use of color comparator tubes (or Nessler tubes) for comparing a range of standards to water samples. However, this method is susceptible to large discrepancies due to human error in selecting variations of tints of a color. An instrument that improves upon the accuracy of colorimetry is a

photoelectric colorimeter. A light source is directed through a filter (making it monochromatic), then passed through the sample cell and on to the photoelectric detector. The obvious advantage over the comparator is the elimination of the human factor for identifying variations in color. Another adaptation is the spectrophotometer, which employs a diffraction grating to select the wavelength for a wide variety of uses. It is a very versatile instrument since both the incoming and outgoing wavelengths are selectable. Newer models incorporate computer chips for automatic programming and data storage.

Colorimetric tests include single analyte meters for chlorine and DO, and direct readout multianalyte photometers for many parameters. For example, the DPD colorimetric method is commonly used to measure total chlorine concentrations in water and wastewater. However, it must be used with caution, since turbidity, other organic contaminants, monochloromine, etc. will cause interference.

78.2.2.4 Fluorimetry

This method utilizes the fluorescence, either natural or induced, of a compound. Fluorescent chemicals absorb radiation of a specific wavelength and emit at another. Fluorescent tracers are commonly used in water quality studies to determine direction of flow for both surface and groundwater systems. Fluorescent tracers include Rhodamine B, Fluorescein, and Pontacyl Pink B. The detection instrument is known as a fluorimeter, which range from moderate cost for a stationery (single chemical) wavelength model (<$4000) to high cost for a unit with variable wavelength and detection capabilities.

78.2.2.5 Remote Fiber-Optic Spectroscopy

In the area of in situ chemical analysis, remote fiber spectroscopy (RFS) shows much promise for a variety of monitoring applications. RFS is a fiber-optic application of existing spectroscopic techniques that can be applied to the monitoring of chemical species in remote locations. In this method, light of an appropriate frequency is launched into a single optical fiber or fiber bundle. The light is guided to the region of interest through the fiber by the mechanism of total internal reflection. An optical sensor (optrode) designed for a specific chemical or chemical group is positioned at either the distal end of the fiber or a length along the fiber. The altered signal is returned through the same fiber to the spectrometer for signal decoupling and analysis. The major advantages to fiber-optic chemical sensors (FOCS) include their small diameter, in situ nature, resistance to harsh environments, and imperviousness to electric interference. These properties make FOCS ideal for many environmental applications.

The distinguishing feature of most RFS systems is the type of sensor employed. There are basically two approaches to the sensor, either intrinsic or extrinsic. The intrinsic method utilizes the fiber itself as the active element, for example, using change in index of refractions from the fiber to the media as the parameter being measured. The extrinsic method uses the fiber optic only as light pipe, and usually involves incorporating a chemical onto the fiber as the active element. Most of the related literature for FOCS deals with design and advantages of various types of optrodes that exploit either absorbance, surface enhanced Raman scattering, or fluorescence. Although FOCS have been mentioned in literature for the past decade, there has been little success in developing a multicontaminant sensor for practical applications.

The direct fluorescence approach (measuring the natural fluorescence of the media) is an attractive technique for applications where aromatic chemicals are to be monitored. However, since many chemicals that act as contaminants are nonfluorophores, a fluorescent intermediate can be incorporated in the FOCS. Ideally, the intermediate should be sensitive to either a group of chemicals or a specific compound, and exhibit a spectral response that could be exploited to identify the compound.

As mentioned previously, the standard pH instrumentation is based on a glass pH electrode; however, optical methods have obvious advantages in corrosive conditions and fluctuating temperatures [6]. Optical-based pH sensors are generally based on the changes in the absorbance or fluorescence of an

indicator dye that has been immobilized on a fiber optic. Recently, pH sensors based on the immobilization of fluorescein dye in a sol–gel matrix have been developed. Sol–gel chemistry is a method of producing a porous glass matrix at a low temperature using a hydrolysis of an alkoxide (usually TEO). An example of such a probe is a pH probe based on fluorescein dye in a sol–gel matrix [7].

Optical sensors for O_2 and DO, which are based on fluorescence or phosphorescence quenching, have the advantage of size, nonconsumption of oxygen, and resistance to interference due to flow rates or sample stirring over traditional DO electrodes. The major limitations are finding an indicator dye with a strong sensitivity to oxygen quenching. There are many such studies being performed; however, most are still in the research stage. For example, polycyclic aromatic hydrocarbons such as pyrene and long-wave absorbing dyes have been investigated due to their ability to be quenched by oxygen [6]. A fiber optics–based sensor that is on the market detects VOCs, such as carbon tetrachloride, TCE, and BTX in water systems (FiberChem). A similar application is a fluorescence-based system for measuring petroleum products in the groundwater.

78.2.3 Other Sensors

78.2.3.1 Infrared Spectrophotometry

Most organic chemical compounds exhibit absorption in the infrared. This attribute can be used to determine atomic groupings based on quantum mechanics [5]. An example is a unit that measures total organic content (TOC) using UV-promoted persulfate oxidation and nondispersive infrared detection. This methodology allows for online TOC monitoring. Correlations can be made to convert values of TOC to biochemical oxygen demand (BOD) and chemical oxygen demand (COD) for particular samples. Drawbacks to this method include pretreatment to remove inorganic carbon and the high cost of instrumentation.

78.2.3.2 Respirometry

The standard method to calculate BOD_5 is to measure oxygen consumption (based on DO concentration) over a 5 day period. However, there are other respirometers based on extremely small differences in DO or CO_2 concentrations using external gas sensors. This method is great for respirometry rates of animals, biodegradable contaminants, and sediment oxygen demand studies.

78.2.4 Chemical Separation Techniques

Separation techniques are based on the phase partitioning of molecules within a mixture. Gas chromatography refers to a mobile phase of a vapor, while liquid chromatography refers, of course, to a liquid mobile phase. There are several commonly used types such as gas–liquid chromatography, high-performance liquid chromatography (HPLC), and ion chromatographs. An excellent reference book for these methods is *Chemistry for Environmental Engineering* [5]. The advantages of these instruments are the ability to detect components in complex mixtures and high sensitivity. Chromatography instruments have long been considered standard in environmental laboratories, but the size of the instruments has been a drawback for field applications. However, size constraints are being challenged with the availability of equipment with size reductions of over 90%.

78.2.4.1 Mass Spectrometry

In environmental applications, these are normally used in conjunction with gas chromatography and are referred to as GC/MS. The attraction of these instruments is the ability to detect a variety of compounds and also determine the mass fragmentation patterns of a complex organic structure [1]. The MS essentially ionizes the substances with an electron beam. The ions are then accelerated through a series of lenses according to their mass-to-charge ratios. The charged fragments are detected by a electron

multiplier, which results in a mass spectra for the particular compound. In addition to the old-fashioned laboratory-scale GC/MS, there are portable units now commercially available that can be taken into the field for on-site environmental characterizations. These are very useful for ambient air toxics analysis, process diagnostics, and "real-time" groundwater plume movement. However, a newer technique of capillary electrophoresis (CE) has been gaining in popularity over capillary GC since it requires less solvent and can be extremely sensitive.

78.2.4.2 Atomic Absorption Spectrometry

Since metals have their own characteristic absorption wavelengths, they are usually measured by the atomic absorption method where a light beam is directed into a sample being aspirated onto a flame. The resulting light is then sent to a monochromator with a detector that measures the amount of light absorbed by the flame. This method exhibits high sensitivity and is applicable to all of the commonly found metals in water. The drawback of this type of instrument is that it is not portable and is more of a laboratory method.

78.3 Instrumentation and Applications

Table 78.2 lists some of the instruments that are commercially available for specific water applications. This is not an exhaustive list of all types and manufacturers, but is indicative of the chemical parameters and types of instruments that are useful for field use. For example, the chemical separations category is very sparse in the table; although there are numerous laboratory instruments, very few are useful in the field.

Many of the manufacturers provide helpful technical information on theory and applications for their products. For example, Hach Company has many excellent publications for a variety of water/wastewater applications. They sell and develop instruments that vary from inexpensive kits (e.g., for colorimetry) to spectrophotometers with computer compatibility. Additionally, due to the broad range of diversity in the water quality area, there are many small companies that produce instruments for specific applications. Table 78.3 provides contact information for each of the companies listed in Table 78.2. Another excellent source of information is the annual listing of industrial manufacturers by WEF.

78.3.1 Data Evaluation

In the past, the better instruments may have had a self-storing data system, usually in the form of paper graph or a retrievable magnetic tape. Although this allowed for continuous sampling, it was not without problems. The ease of obtaining data, in situ data gathering, and digital data acquisition are all attributes that are attractive in environmental applications. New technology has made these wishes a reality. For example, radio telemetry is a wireless technology to link remote sensors to dedicated computers or Supervisory Control and Data Acquisition (SCADA) systems. This capability allows for truly automated data acquisition and real-time uplinks. Other systems can be hardwired to computers (usually via RS 232 boards) or through modems to a remote acquisition site.

78.4 Trends in Water Quality Measurements

Analytical Chemistry publishes a biennial review dealing with recent literature in the area of water analysis [8]. These review articles, while not in-depth on a single topic, are quite extensive in topic matter and really illuminate trends in environmental sensing. For example, in a recent review, the area of in situ analysis contains a significant amount of coverage. Since many water quality applications would benefit from in situ, real-time data, it will be an area of research and development growth for the next 5 years.

TABLE 78.2 Instruments for Remote Water Quality Analysis

Product Name	Water Quality Parameters	Company
Biological analyzers: Portable		
Potatlite	Extreme Portable kit WE 1000	RB Instruments
Ion-selective electrodes: single parameter and combination types		
AdvantEDGE	Dissolved Oxygen Electrode Sensor	HF Scientific, Inc.
HQ30d	Portable Dissolved Oxygen Meter Package	Hach
D63/5440D	pH/ISE Meter	Thermo Scientific Orion
SCAMP	DO, temperature	Precision Measurement Engineering
9110/9120	Dissolved oxygen	Royce Instrument Corporation
Professional Plus with ISE probes D	Ammonia, conductivity	YSI
Nephelometers: Turbidity		
1110-TUX	Temperature, pH, Conductivity, Dissolved oxygen, Turbidity, Selective ions, Redox and water flow	Solomat
Model 2100Q	Turbidity	Hach Inc.
Conductivity sensors		
DataSonde 4A	Multiprobe Conductivity, TDS, salinity, resistivity	Rickly Hydrological Co
2100Q	Turbidity	Hach Inc.
Smart Sensor	Conductivity, pressure	Greenspan Technology Pty Ltd.
YSI 5000 series	DO	YSI Inc.
Optical		
AdvantEDGE sensor	DO	HF Scientific, Inc.
AccUView	UV transmission	HF Scientific, Inc.
711	Turbidity and suspended solids	Royce Instrument Corporation
PC3400	Online particle counter	Chemtrac, Inc.
pH probes: with data storage and computer compatibility		
Orion VERSA STAR Meter multiparameter	Thermo scientific	
Respirometry		
OXITOP IS RESPIROMETER SYSTEMS	Global Water	
Micro-Oxymax	Oxygen, CO_2, NH_3, CO	Columbus Instruments International Corp.
Spectrometer: in-line types only		
DR 5000™	UV-Vis	ALToptronic
	Spectrophotometer	Hach Inc.
FPA 1000 Series	Chlorine, oils in water, H_2S	Tytronics
Tytronics Sentinel	NH_3, Fl, Cy, NO_3, Na	Galvanic Applied Sciences, Inc.
UV absorbance instruments		
ChemScan	Nitrate, ammonia, iron, turbidity	Applied Spectrometry
Multiparameter systems		
AQUALAB	Phosphate, ammonia, nitrate, pH, conductivity	Greenspan Technology
Chemscan 6101	Over 30 parameters	ASA Analytics
HI 9828	Multiparameter Meter	Hanna Instruments
DataSonde	4 Temperature, conductance, TDS, resistivity	Hydrolab
YSI 4 series	Over 10 parameters	YSI Inc.

TABLE 78.3 Companies that Make Water Quality Sensors

Applied Spectrometry Associates, Inc.
W226 N55G Eastmound Drive
Waukesha, WI 53186
Tel: (414) 650-2280
http://www.chemscan.com/

Campbell Scientific, Inc.
815 W. 1800 N., Logan, UT 84321-1784
Phone: 435-753-2342, Fax: 435-750-9540
http://www.campbellsci.com/

Campbell Scientific
411 Bayswater Road
Garbutt, QLD 4814, Australia
Tel: +61 (0)7-441-7700
http://www.campbellsci.com.au/

Davis/Inotek
11212 Indian Trail, Dallas, TX 75229
Phone: 972-243-7000, 800-492-6767 (toll free), Fax: 972-243-7814, 888-818-3981 (toll free)
http://www.davis.com/

FiberChem Incorporated
Tel: (509) 376-5074
http://www.fiber-chem.com

Great Lakes Instruments
9020 West Dean Road
P.O. Box 23056
Milwaukee, WI 53224
http://www.hach.com/gli

Greenspan Technology Pty Ltd.
24 Palnerin Street
Warwick, Queensland, 4370 Australia
Tel: 61-76-61-7699
http://www.greenspan.com.au/

HF Scientific
3170 Metro Parkway
Fort Meyers, FL 33916-7597
Tel: (941) 337-2116
http://www.hfscientific.com/

H & S Enterprises
148 South Dowlen #120
Beaumont, TX 77707

Hydrolab Corporation
P.O. Box 50116
Austin, TX 78763
http://www.hachhydromet.com/

Ocean Sensors
9883 Pacific Heights Blvd., Suite E
San Diego, CA 92121
Tel: (619) 450-4640
http://www.oceansensors.com/

Precision Measurement Engineering
1827 Hawk View Drive
Encinitas, CA 92024
Tel: (619) 942-5860
http://www.pme.com/

Solomat
26 Pearl Street
Norwalk, CT 06850
Tel: (203) 849-3111
http://www.lms-water.com/

Royce Instrument Corporation
13555 Gentilly Road
New Orleans, LA 70129
Tel: (800) 347-3505
http://www.roycetechnologies.com/

Ryan Herco Flow Solutions
3010 N. San Fernando Blvd. Burbank, CA 91504
Phone: 818-841-1141, 800-848-1141 (toll free)
Fax: 818-973-2650
http://www.rhfs.com/

Tytronics Inc.
25 Wiggins Avenue
Bedford, MA 01739-2323
Tel: (617) 275-9660

Water On-Line Analysis
675 Third Ave., New York, NY 10017
Phone: 646-839-5528
http://www.biosensores.com/

YSI Inc.
1725 Brannum Lane
Yellow Springs, OH 45387
Tel: (800) 765-4974
http://www.ysi.com/

Defining Terms

A light absorption
c velocity of light (3.0×10^8 m/s)
C concentration of the solution
E energy of a photon
E_{cell} cell potential adjusted

E^0 standard cell potential
F Faraday's constant (96,500°C/equivalent)
I light intensity
I_0 initial light intensity
k'' constant for a particular solution
h Planck's constant (6.626×10^{-34} J s)
n number of moles
ox oxidized species
R ideal gas law constant (8.31 J/mol/K)
red reduced species
T absolute temperature
λ wavelength of light

References

1. *Standard Methods for the Examination of Water and Wastewater*, 22nd edn., American Public Health Association, American Public Health Association, American Water Works Association and Water Environment Federation, Washington, DC, 2012.
2. J. Janata, M. Josowicz, and D. DeVaney, Chemical sensors, *Anal. Chem*., 66, 207R–228R, 1994.
3. D. Snoeyik, V. Jenkins, J. Ferguson, and J. Leckie, *Water Chemistry*, 3rd edn., Wiley, New York, 1980.
4. D. Banks, *Microsystems, Microsensors & Microactuators: An Introduction*, Banks, Surrey, England 1996.
5. C. Sawyer, P. McCarty, and G. Parkin, *Chemistry for Environmental Engineering*, McGraw-Hill, New York, 1994.
6. O. Wolfbeis, Ed., *Microengineering, MEMS, and Interfacing: A Practical Guide*, CRC Press, Boca Raton, FL, 2006.
7. P. Wallace, Y. Yang, and M. Campbell, Towards a distributed optical fiber chemical sensor, *Chemical, Biochemical and Environmental Fiber Sensors VII, SPIE*, Vol. 2508, 1996.
8. P. McCarthy, R. Klusman, S. Cowling, and J. Rice, Water analysis, *Anal. Chem*., 67-12, 525R–562R, 1995.

Further Information

HACH, *Water Analysis Handbook*, 7th edn., Hach Company, Loveland, CO, 2012.
USA BlueBook, Utility Supply of America, Northbrook, IL.
Water Environment Federation, Alexandria, VA, Annual Review.

79

Satellite Imaging and Sensing

Robert F. Cromp
National Aeronautics and Space Administration

79.1 What Can Be Seen from Satellite Imagery?

Satellite imaging and sensing is the process by which the electromagnetic energy reflected or emitted from the Earth (or any other planetary) surface is captured by a sensor located on a spaceborne platform. The Sun as well as all terrestrial objects can be sources of energy. Visible light, radio waves, heat, ultraviolet, and x-rays are all examples of electromagnetic energy. Since electromagnetic energy travels in a sinusoidal fashion, it follows the principles of wave theory, and electromagnetic waves are categorized by their wavelength within the electromagnetic spectrum. Although it is continuous, different portions of the electromagnetic spectrum are usually identified and referred to as (from shorter to longer wavelengths) cosmic rays, γ rays, x-rays, ultraviolet, visible (0.4 μm, 0.7 μm), near-infrared (near-IR), mid-infrared (mid-IR), thermal infrared (thermal-IR, above 3 μm), microwave (1 mm, 1 m), and television/radio wavelengths (above 1 m). Figure 79.1 shows the electromagnetic spectrum and these subdivisions.

79.2 General Sensor Principles

Sensors are often categorized as "passive" or "active." All energy observed by "passive" satellite sensors originates either from the Sun or from planetary surface features, while "active" sensors, such as radar systems, utilize their own source of energy to capture or image specific targets.

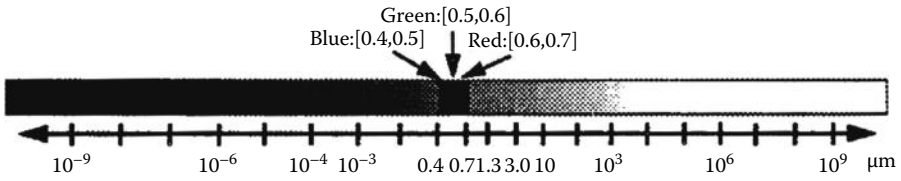


FIGURE 79.1 Electromagnetic spectrum.

79.2.1 Passive and Active Sensors

All objects give off radiation at all wavelengths, but the emitted energy varies with the wavelength and with the temperature of the object. A "blackbody" is an ideal object that absorbs and reemits all incident energy, without reflecting any. If one assumes that the Sun and the Earth behave like blackbodies, then according to the Stefan–Boltzmann law, their total radiant exitance is proportional to the fourth power of their temperature. The maximum of this exitance, called dominant wavelength, can be computed by Wien's displacement law (see Refs. [1–4] for more details on these two laws). These dominant wavelengths are 9.7 μm for the Earth (in the IR portion of the spectrum) and 0.5 μm for the Sun (in the green visible portion of the spectrum). It implies that the energy emitted by the Earth is best observed by sensors that operate in the thermal-IR and microwave portions of the electromagnetic spectrum, while Sun energy that has been reflected by the Earth predominates in the visible, near-IR, and mid-IR portions of the spectrum. Most passive satellite sensing systems operate in the visible, IR, or microwave portions of the spectrum. Since electromagnetic energy follows the rules of particle theory, it can be shown that the longer the wavelength, the lower the energy content of the radiation. Thus, if a given sensing system is trying to capture long wavelength energy (such as microwave), it must view large areas of the Earth to obtain detectable signals. This obviously is easier to achieve at very high altitudes, thus the utility of spaceborne remote sensing systems.

The most common active satellite sensor is radar (acronym for "radio detection and ranging"), which operates in the microwave portion of the electromagnetic spectrum. The radar system transmits pulses of microwave energy in given directions, and then records the reflected signal received by its antenna. Radar systems were initially employed by the military as a reconnaissance system because their main advantage was to operate day or night and in almost any weather condition. They are very important in satellite remote sensing because microwave radiations are hardly affected by atmospheric "screens" such as light rain, clouds, and smoke. The time it takes for the radar signal to return to the satellite is also measured by instruments such as altimeters which are very useful in determining surface height measurements.

79.2.2 Polar Orbiting and Geostationary Earth Sensing Satellites

Satellite remote sensing systems are also characterized by the different Earth orbiting trajectories of a given spacecraft. These two modes are usually referred as "polar orbiting" and "geostationary" (or "geosynchronous") satellites. A polar orbit passes near the Earth's North and South poles. Landsat, SPOT, and NOAA are near-polar satellites; their orbits are almost polar, passing above the two poles and crossing the equator at a small angle from normal (e.g., 8.2° for Landsat-4 and -5). If the orbital period of a polar orbiting satellite keeps pace with the Sun's westward progression compared to the Earth rotation, these satellites are also called "sun synchronous." This implies that a sun-synchronous satellite always crosses the equator at the same local sun time. This time is usually very carefully chosen, depending on the application of the sensing system and the type of features that will be observed with such a system. It is often a trade-off between several Earth science disciplines such as atmospheric and land science. Atmospheric scientists prefer observations later in the morning to allow for cloud formation, whereas the researchers performing land studies prefer earlier morning observations to minimize cloud cover.

A geostationary satellite has the same angular velocity as the Earth so its relative position is fixed with respect to the Earth. Examples of geostationary satellites are the "Geostationary Operational Environmental Satellite" (GOES) series of satellites that orbit at a constant relative position above the equator.

79.2.3 Sensor Characteristics

79.2.3.1 Spectral Response Patterns

The design of new satellite instruments is based on the principle that targets of interest can be identified based on their spectral characteristics. For example, different Earth surface features, such as vegetation or water, present very distinctive reflectance or emittance curves that are a function of the energy wavelength. These curves are often called the "spectral signatures" of the objects being observed. Although these curves are very representative of each feature and can help identify them, they do not correspond to unique and absolute responses. Because of different reasons, such as atmospheric interactions, temporal or location variations, the response curves of a given object observed under different conditions might vary. For this reason, these curves are often called "spectral response patterns" instead of "spectral signatures." Figure 79.2 shows an example of such reflectance patterns for several features: fir tree, clear lake water, barley, and granite.

79.2.3.2 Atmospheric Interactions

Earth satellite sensors are designed to take into consideration the fact that all observed radiation must pass at least once through the atmosphere; therefore, the energy interactions of the atmosphere must be considered during the design phase. The distance through which the radiation passes through the atmosphere is called "path length." The effect of the atmosphere depends on the extent of the path length and on the magnitude of the energy signal. The two main atmospheric effects are known as "scattering" and "absorption."

Scattering is the unpredictable redirection of radiation by particles suspended in the atmosphere. The type and the amount of scattering mainly depend on the size of the particles but also on the wavelength of the radiation and the atmospheric path length. If these particles are smaller than the radiation wavelength, this effect is known as "Rayleigh scatter." This scattering especially affects the shorter visible

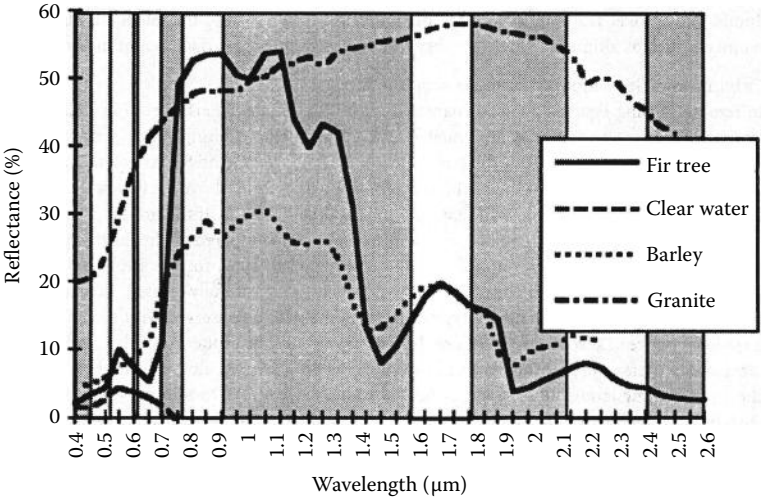


FIGURE 79.2 Examples of spectral response patterns for four different types of features: fir tree, clear water, barley (example of crop), and granite (example of rock). White areas show the portions of the spectrum corresponding to the 7 channels of Landsat-Thematic Mapper (TM-4&5).

wavelengths of the sunlight (i.e., blue visible wavelength) and it explains why the sky appears blue to the human eye. In the evening, when the path length is longer, the effect of the Rayleigh scatter is only visible on the longer red wavelengths of the sunlight and the sky appears red or orange. If the particles are about the size of the radiation wavelength, the scatter is known as "Mie scatter"; this scattering effect is often due to water vapor and dust. When the atmospheric particles are larger than the radiation wavelengths, a "nonselective scatter" occurs; all visible wavelengths radiations are scattered equally and this type of scattering explains why clouds appear white.

Atmospheric absorption occurs in specific wavelengths at which gases such as water vapor, carbon dioxide, and ozone absorb the energy of solar radiations instead of transmitting it. "Atmospheric windows" are defined as the intervals of the electromagnetic spectrum outside these wavelengths, and Earth remote sensors usually concentrate their observations within the atmospheric windows. As an example, white areas of Figure 79.2 show the portions of the spectrum (i.e., the "channels" or "bands") from visible to mid-IR used by the Landsat-Thematic Mapper (TM).

79.2.3.3 Spectral, Radiometric, Spatial, and Temporal Resolutions

Although the spectral response patterns are not absolute, they play an important role in the design of new sensors. When a new sensor is being designed, the type of features to observe and the accuracy with which they will be mapped define which wavelengths are of interest, the widths of the wavelength intervals to be used, what is the accuracy to be achieved in these bandwidths, and what is the "smallest" or "faintest" feature that might be detected by the sensor. Following the examples of Figure 79.2, the best wavelength interval to distinguish between vegetation and granite will be the (1.55 μm, 1.75 μm) wavelength interval. The earlier sensor requirements correspond to the "resolutions" of the sensor by which it is usually identified—spectral, radiometric, spatial, and temporal resolutions. The term "resolution" is usually employed to define the smallest unit of measurement or granularity that can be recorded in the observed data. The spectral resolution of a sensor is defined by the bandwidths utilized in the electromagnetic spectrum. The radiometric resolution defines the number of "bits" that are used to record a given energy corresponding to a given wavelength. The spatial resolution corresponds to the area covered on the Earth's surface to compute one measurement (or one picture element, "pixel") of the sensor. The temporal resolution (or frequency of observation), defined by the orbit of the satellite and the scanning of the sensor, describes how often a given Earth location is covered by the sensor.

79.2.3.4 Signal-to-Noise Ratio

Sensors are also characterized by their signal-to-noise ratio (SNR) (i.e., the noise level relative to the strength of the signal). In this case, the "noise" usually refers to variations of intensity that are detected by the sensor and that are not caused by actual variations in feature brightness. If the noise level is very high compared to the signal level, the data will not provide an accurate representation of the observed features. At a given wavelength λ, SNR is a function of the detector quality, the spatial resolution of the sensor, as well as its spectral resolution (see Ref. [1] for a detailed formula). To maintain or improve the SNR and therefore improve the radiometric resolution of the sensor, a trade-off must be made between spatial and spectral resolutions; in particular, improving spatial resolution will decrease the spectral resolution. Of course, other factors such as atmospheric interactions will also affect the SNR.

79.2.3.5 Multispectral and Hyperspectral Sensors

The remote sensing industry is experiencing a rapid increase in the number of spectral bands of each sensor. The first Landsat sensors (Landsat-1 and -2) were designed with four bands in the visible and near-IR portions of the spectrum. Landsat-4 and -5 were refined with seven bands from visible to thermal-IR. Then, Landsat-6 and -7 were planned with an additional panchromatic band, which is highly sensitive over the visible part of the spectrum. In general, most Earth remote sensors are *multispectral*; that is, they utilize several bands to capture the energy emitted or reflected from Earth features. The addition of panchromatic imagery, which usually has a much better spatial resolution than multispectral

imagery in the visible part of the spectrum, provides higher quality detail information. Multispectral and panchromatic data, usually acquired simultaneously, are coregistered and can be easily merged to obtain high spatial and spectral resolution. Coregistered multispectral–panchromatic imagery is available from sensors such as the Indian satellite sensor, IRS-1, and the French sensor, SPOT.

Ideally, if a sensor had an infinite number of spectral channels (or bands), each observed area on the ground (or pixel) could be represented by a continuous spectrum and then identified from a database of known spectral response patterns. Adding more bands and making each of them narrower is the first step toward this ideal sensor. But, as previously explained in the previous section, due to technology limitations, it was very difficult until recently to increase the number of bands without decreasing the SNR. Due to recent advances in solid-state detector technology, it is now possible to increase significantly the number of bands without decreasing the SNR, thus seeing the rise of new types of sensors, called *hyperspectral*. Although the boundary between multispectral and hyperspectral sensors is sometimes defined as low as 10 bands, hyperspectral imaging usually refers to the simultaneous detection in hundreds to thousands of spectral channels. The aim of hyperspectral sensors is to provide unique identification (or "spectral fingerprints") capabilities for resolvable spectral objects. Potential applications include agricultural yield monitoring, urban planning, land use mapping, mining and mineral deposits, disaster relief/assessment, tactical military operations, and forest fire protection management. The NASA Airborne Visible InfraRed Imaging Spectrometer (AVIRIS) simultaneously collects spectral information from visible to IR ranges (from 0.4 to 2.5 μm) in 224 contiguous spectral bands. Each band has an approximate bandwidth of 10 nm (or 0.01 μm), with a spatial resolution of about 20 m. The instrument flies aboard a NASA ER-2 airplane at ~20 km above sea level. The science objectives of the AVIRIS project are mainly directed toward understanding processes related to the global environment and climate change.

79.2.3.6 Super Resolution

When data are collected from a satellite sensor, each sample of information represents an area on the ground that might correspond to several features with different spectral responses and the final information represents a "mixture" of several disparate information. When a sensor does periodic imaging over the same area, the direction of observation slightly changes, and even when successive data are correctly registered (i.e., in a perfect correspondence), two respective samples might represent two slightly different areas on the ground. Super resolution is an area of research that aims at combining such information from different directions but taken under similar lighting conditions in the goal of improving spatial and spectral resolutions. As an example, some recent work [5] utilizes a Bayesian method for generating subpixel resolution composites from multiple images with different alignments. Software products have also been proposed by some commercial systems, such as ERDAS-Imagine and ENVI, to help in the unmixing of spectral bands (see Ref. [6] for more detail).

79.2.4 Direct Readout Data

Examples of different spectral and spatial resolutions are given in Figure 79.3 with a summary of current Earth remote sensing systems that operate from visible to thermal-IR wavelengths. Figure 79.3 provides information about their bandwidths and their spatial resolutions. The table also indicates if the data from these sensors can be acquired by direct readout. Sensors can transmit data in two modes. The first mode, called "direct readout," transmits data as soon as it "sees" them, and any receiving station within that satellite footprint can receive these data. In the second mode, the sensor records whatever it sees for playback at a later time, and specialized receiving stations (or "ground stations") are required to receive these data since they are transmitted at higher rates than direct readout. The data received by these ground stations can cover a larger extent. Historically, the cost of acquiring and processing Earth remotely sensed data has limited satellite data collection to a small number of expensive ground stations around the world, operated by the owners of the satellites. Direct readout

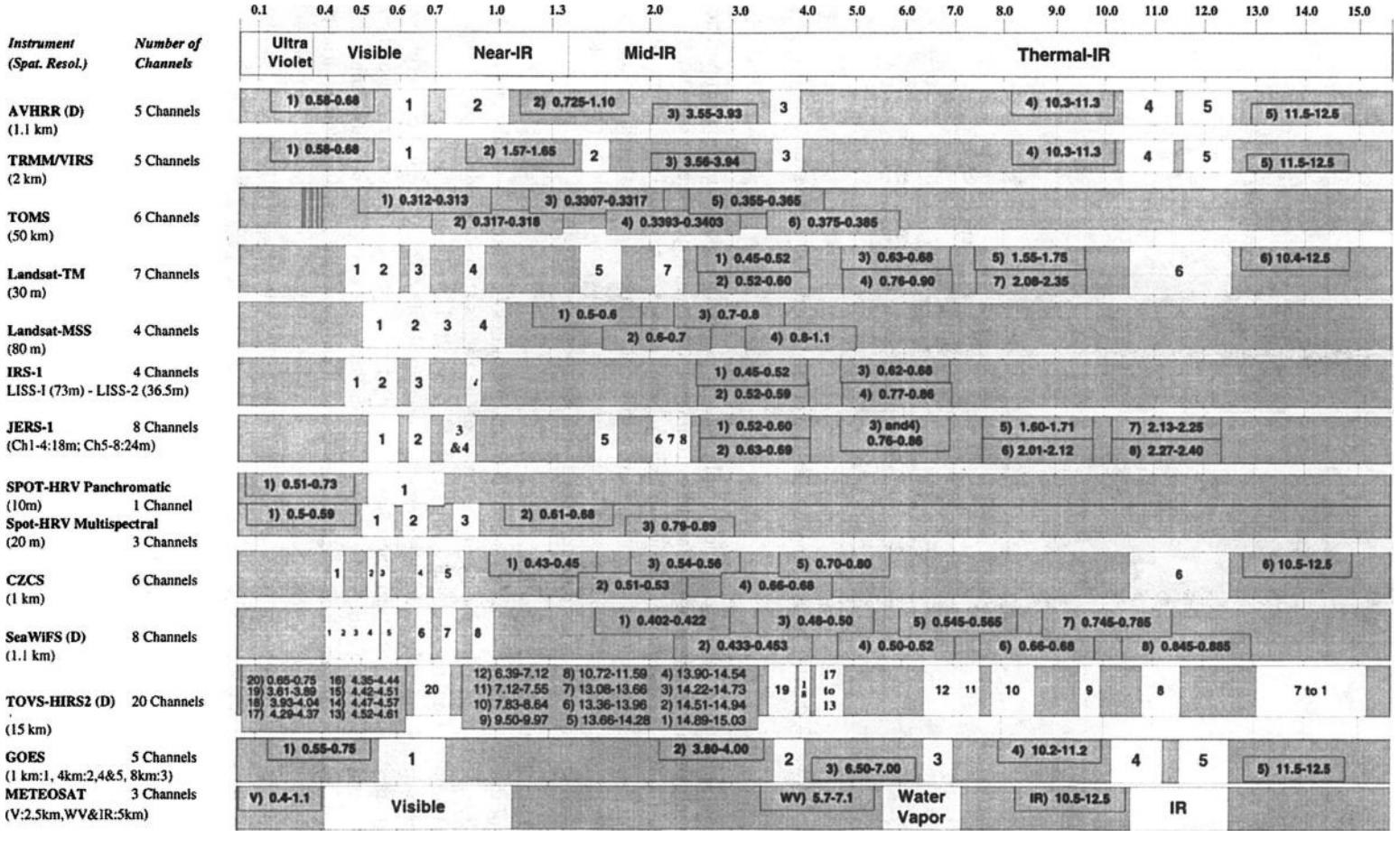


FIGURE 79.3 Summary of the main current earth science satellite data operating from UV to thermal-IR "D": Direct Broadcast.

sensors were mostly confined to meteorological applications that required timely data such as weather forecasting, severe weather identification and tracking, and disaster prediction and assessment. But this situation is changing; due to new technology, costs have been greatly reduced and direct readout data that were initially expensive and beyond the reach of the nontraditional user are now generating a growing interest. A small industry has evolved to design, install, and upgrade ground stations around the world that acquire direct readout data.

79.2.5 Typical Attributes Measured from Space

This section does not intend to be an exhaustive up-to-date description of all Earth and space applications of satellite imaging and sensing, but rather it presents a few representative applications and their associated satellite sensors. The References section as well as the World Wide Web (WWW) offer more extensive references to other applications and sensors.

79.2.5.1 Earth Science Applications

Over the past few decades, a number of international global change research programs have been initiated whose goals are to understand the relation between human activities and the global Earth systems processes and trends. Mission To Planet Earth (MTPE [7]) is a multiagency program, whose goal is to achieve this kind of understanding, especially through improved satellite observations. As part of NASA's Mission to Planet Earth program, the Earth Observing System (EOS [8]) will launch, over the next 2 decades, several platforms of sensors aimed at ecology, oceanography, geology, snow, ice, hydrology, cloud, and atmospheric studies. Each platform will carry one or several instruments, thus globally covering a wide range of spectral, spatial, and temporal resolutions. Europe and Japan have similar programs, such as the ADvanced Earth Observation Satellite (ADEOS [9]), developed by the Japanese space agency, NASDA, in collaboration with France and the United States. ADEOS, launched in 1996, flew for 9 months before it failed, and included remote sensing instruments for observing the Earth's atmosphere, land surfaces, and oceans. Other examples of such programs are the ERS and ENVISAT satellites from European Space Agency (ESA) and the Indian satellites, IRS. ERS-1 and -2 were launched in 1991 and 1995, respectively, and ENVISAT was launched in 1999. All these satellites carry on the same platform with different instruments making simultaneous observations. The first satellite in the IRS series was launched in 1988, and in 9 years, India has designed and launched six remote sensing satellites. For more information on the IRS series, see Ref. [10].

In all the earlier programs, studies concentrate on global processes occurring in the atmosphere (especially lower parts of the atmosphere), on the Earth surface (terrestrial studies), and in the oceans (hydrospheric studies). Results of all these studies will contribute to international programs such as the World Climate Research Programme, the International Geosphere–Biosphere Programme, and the International Human Dimensions of Global Environmental Change Programme.

Several of these instruments show application promise for regional and local community interests, in helping farmers to monitor and control their agricultural productivity (weather, disease control), in early warning and in rescue efforts in case of severe storms (e.g., hurricanes), and in predicting the spread of diseases based on vegetation data combined with socioeconomic data.

79.2.5.2 Examples of Atmospheric Studies

One of the key issues in climate research is to understand global atmospheric changes and how human activity affects the composition and chemistry of the Earth's atmosphere. To create accurate models, a large number of multiyear global studies must be conducted.

The atmosphere is divided into several layers. From the Earth's surface up to interplanetary space, which starts at about 1000 km, these layers are called troposphere, stratosphere, mesosphere, thermosphere, and exosphere. Figure 79.4 shows a simplified diagram of the different atmospheric layers. Each of these layers is characterized by differences in chemical composition that produce variations in temperature. The two

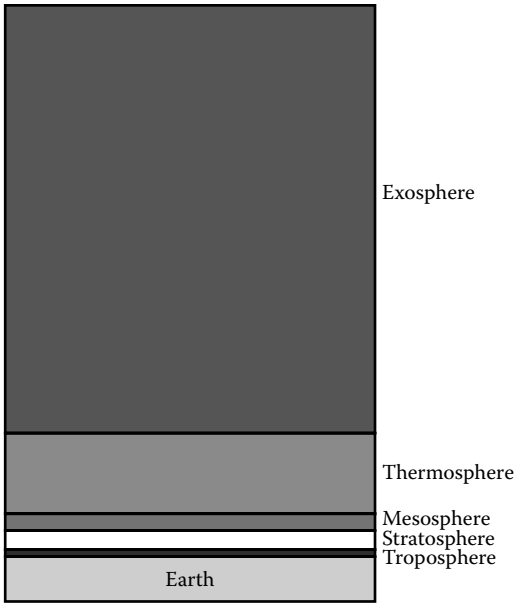


FIGURE 79.4 Simplified diagram of the different layers of the atmosphere.

lower layers, troposphere (up to 10 km) and stratosphere (10–50 km above the Earth), are particularly important since 99% of the water vapor in the atmosphere is included in the troposphere, and 90% of the ozone of the atmosphere is included in the stratosphere. All weather phenomena occur within the troposphere, with some turbulence sometimes extending to the lower stratosphere. The concentration of ozone, which should stay mainly concentrated in the stratosphere, is being studied in both the troposphere and stratosphere.

Ozone is a relatively unstable molecule made up of three oxygen atoms. Depending on the altitude where it is found, ozone is referred to as "good" or "bad" ozone. The largest concentration of ozone is located in the stratosphere, at an altitude between 20 and 30 km, and plays a major role in the evolution and the protection of life on Earth. Since it absorbs most of the harmful ultraviolet radiation from the Sun, stratospheric ozone protects life on Earth. When found closer to the Earth's surface, ozone may be harmful to lung tissue and plants. Recent studies have shown that the proportions of ozone in the air are increasing compared to decreasing amounts of protective ozone. However, studies still need to determine if these changes are due to human activity or if they are part of regular natural cycles.

The mission of the total ozone mapping spectrometer (TOMS) is to provide global measurements of total column ozone as well as of sulfur dioxide on a daily basis. The TOMS instrument measures the reflectivity of the atmosphere in six near-UV wavelengths (see Figure 79.3) and provides differential UV absorption and surface reflectivity data. From these measurements, total ozone is computed by searching precomputed albedo tables, which depend on solar zenith angle, view angle, latitude, surface reflectance, and surface pressure; a lower amount of radiation measured by TOMS corresponds to higher concentrations of ozone. Maps of volcanic eruptions are a by-product of TOMS sulfur dioxide measurements. The first TOMS instrument was flown on Nimbus 7 in 1978; successive ones were launched on a Russian Meteor spacecraft in 1991, on an Earth Probe satellite in 1994, and on the Japanese ADEOS satellite in 1996. See Refs. [11–16] for more information on TOMS and ozone measurements.

The TOMS measurements are also being compared to the ozone measurements provided by the National Oceanic and Atmospheric Administration (NOAA) series of the Television Infrared Observing Satellite (TIROS) Operational Vertical Sounder (TOVS) data. The TOVS sounding unit consists of three

instruments, including the High-Resolution Infrared Sounder-2 (HIRS-2) whose channels are shown in Figure 79.3. TOVS-type instruments have been flying since 1978. These instruments provide information about the structure of the atmosphere, vertical temperature and moisture profiles, as well as cloud amounts and heights. Through analysis of this data, the TOVS Pathfinder data set is created and contains 74 layers of measurements on attributes such as temperature, water vapor, ozone level, precipitation, cloud coverage, etc. taken at various atmospheric pressure levels (e.g., 1000 mb, 850 mb). A full global coverage of TOVS data is produced twice daily, and a 16 year global data set for climate studies is being gathered [17].

Another satellite that obtains atmospheric data is the Upper Atmosphere Research Satellite (UARS). UARS was launched in 1991 and performs a comprehensive study of the stratosphere and furnishes important new data on the mesosphere and the thermosphere. UARS operates 585 km above the Earth in a near circular orbit inclined 57° to the equator. This orbit permits UARS sensors to provide global coverage of the stratosphere and mesosphere and measurements are made approximately every 36 days. The 10 UARS chemistry and dynamics sensors are making measurements of temperature, pressure, wind velocity, and gas species concentrations. All these simultaneous measurements will help define the role of the upper atmosphere in our climate and its variability.

The Tropical Rainfall Measuring Mission (TRMM [18]) is a joint project between the United States and Japan. The goal of this project is to measure precipitation at tropical latitudes and to provide accurate mapping of tropical rainfall. The mission consists of three instruments: precipitation radar, a multichannel microwave radiometer, and a visible–IR scanner. The data provided by TRMM will be very important to verify and develop climate models.

The French space agency, CNES, has also developed the POLDER (POLarization and Directionality of the Earth's Reflectances) instrument, which flew on ADEOS. This is the first French/Japanese cooperative project in the area of Earth observation. A second, identical instrument is to be flown on ADEOS-2, successor to ADEOS, in 1999. POLDER is a wide field-of-view imaging radiometer that will provide global, systematic measurements of spectral, directional, and polarized characteristics of the solar radiation reflected by the Earth/atmosphere system, as well as aerosols, land and sea surfaces, and water vapor measurements.

NOAA's Advanced Very High Resolution Radiometer (AVHRR [19]) is very useful to study biomass burning in the tropics, and the interactions of smoke particles with clouds. More generally, information from the five AVHRR channels (see Figure 79.3) is integrated into clouds and climate models.

Weather images are an everyday occurrence televised all over the world. Several weather satellites are operated by several countries. In the United States, NASA and NOAA are operating the GOES series of geostationary satellites, which provide global weather data every 30 min since 1974, positioned at 36,000 km above the Earth. GOES image and sounder data are also used for climate studies. In Europe, the Meteosat weather satellites are developed and launched by ESA, and financed and owned by Eumetsat, an international organization of 17 European weather services. Meteosat-1 was launched in 1977, followed by five others in 1981, 1989, 1991, and 1993. Three of them are currently in service, each equipped with an imaging radiometer. Figure 79.3 shows the spectral ranges of operations of these two series of geostationary satellites. Several channels in the visible, water vapor, and thermal-IR spectral bands provide important information about cloud coverage, storm formation and evolution, as well as Earth radiation.

79.3 Examples of Terrestrial Studies

79.3.1 Land-Cover Applications

There are two basic types of data considered most important for global change research [20]; the data for documenting and monitoring global change, and the data for discovering the dynamical interplay among the various elements that define our environment. Previous studies show that global studies of land transformations require extrapolation among several scales (spatial, radiometric, and temporal).

This extrapolation is especially important to control the minimum detectable change, whether spatial, spectral, or temporal. This accuracy in change detection, which is based on the properties of the sensing systems [21], can be especially essential in distinguishing between nature- and human-induced changes.

Getting accurate quantitative information about the distribution and the areal extent of the Earth's vegetation formations is a basic requirement in understanding the dynamics of the major ecosystems. Among all land transformations most critical to study for global change research, the assessment of tropical forests is one of the most important [22–25]. The tropical forest biome forms 7% of the Earth land surface, and its extensive loss could have a major impact on the future of the Earth (habitat fragmentation, species extinction, soil degradation, global climatic modifications, etc.). Previous studies have shown that in the last 2 decades, 50% of the areal extent of tropical forests might have been lost to deforestation [23]. At present, there is a wide range of estimates of the areal extent of tropical forests and of their rates of deforestation. Therefore, there is a great need to produce accurate and up-to-date measurements concerning the Tropical Forest worldwide. A range of different sensors must be utilized for such applications.

Other examples of land-cover applications include agriculture and crop forecasting, water urban planning, rangeland monitoring, mineral and oil exploration, cartography, flood monitoring, disease control, real estate tax monitoring, detection of illegal crops, etc. In many of these applications, the combination of remote sensing data and Geographic Information Systems (GISs; see Refs. [26,27]) shows great promise in helping the decision-making process.

Most instruments utilized to observe land features are on-board low Earth orbit satellites and are multispectral sensors with two or three bands in the visible part of the spectrum and at least one band in the IR [6].

The Landsat series of satellites is the oldest land monitoring satellite system. Initiated in 1967, the Earth Resource Technology Satellites (ERTS) program was planning a series of six satellites to perform a broad-scale, repetitive survey of the Earth's land areas. After the launch in 1972 of the first ERTS-1, the ERTS program was renamed "Landsat." As of 1997, five Landsat satellites have been launched, each one carrying two instruments. The payload of Landsat-1 and -2 included a Return Beam Vidicon (RBV) camera and a Multispectral Scanner (MSS), while Landsat-4 and -5 still use the MSS and the TM. The first RBV system consisted of three television-like cameras with a ground resolution of 80 m, each looking respectively, at the green, red, and near-IR portions of the spectrum. On Landsat-3, the RBV was 30 m panchromatic. MSS quickly became of primary interest due mainly to its capability of producing multispectral data in a digital format. The four MSS spectral bands are shown in Figure 79.3, and the spatial resolution of MSS data is about 80 m. Very early on, the utility of MSS data was recognized for such applications as agriculture, mapping, forest monitoring, geology, as well as water resource analysis. The same MSS system was kept on Landsat-4 and -5, but the RBV system was replaced by the TM system. Like MSS, TM is a MSS, but includes spatial, spectral, and radiometric improvements over MSS. With seven bands instead of four (see Figure 79.3), TM covers a larger portion of the visible wavelengths, and includes two mid-IR and one thermal-IR bands. Data are quantized over 256 levels (8 bits) instead of the 64 levels for MSS, and the spatial resolution of a TM pixel is about 30 m. TM data are usually chosen to perform classification of land-cover features, manmade or natural. In vegetation and change detection applications, leaf segmentation is studied with TM visible channel data, while cell structure can be seen in near-IR, and leaf water content is found in the mid-IR channel data. The two mid-IR bands (5 and 7) are also useful for geologic applications. All the Landsat satellites are placed in low Earth orbit (at an altitude of about 900 km for Landsat-1 to -3 and 705 km for Landsat-4 and -5) and in a near-polar, sun-synchronous orbit. Landsat-4 and -5 cross the equator at 9:45 a.m. to hopefully take advantage of cloud-free imagery. Landsat-4 and -5 have a 16 day repeat cycle and their orbits are 8 days out of phase. Landsat-6 failed to achieve orbit in 1993; Landsat-7 is planned to be launched in 1998 and includes an improved TM instrument, the Enhanced Thematic Mapper (ETM), which will also include a panchromatic band at a spatial resolution of 15 m. For a more in-depth description of Landsat systems, see Refs. [1,29,30]; for more applications and analysis of Landsat data, see Ref. [30].

As previously mentioned, NOAA's AVHRR is primarily used for atmospheric applications but is also utilized for land surface applications. Having a near-polar, sun-synchronous orbit (at 833 km above the Earth's surface), the AVHRR instrument provides global data with a 1.1 km spatial resolution at nadir, and includes five bands, with daily or twice-daily (for thermal-IR) coverage. Since 1978, AVHRR data are available at full resolution (called Local Area Coverage [LAC]) or subsampled to 4 km (known as Global Area Coverage [GAC]). Because of its high temporal resolution, the AVHRR instrument is very useful in applications such as flood, storm, or fire monitoring, as well as volcanic eruption. Because of its GAC, AVHRR is also often utilized for studying geologic or physiographic features, vegetation conditions, and trends at a global, continental, or regional level, snow cover mapping, soil moisture analysis, and sand storms and volcanic eruptions worldwide. A popular parameter extracted from AVHRR data is the Normalized Difference Vegetation Index (NDVI), computed from GAC data as NDVI = (Channel 2 – Channel l)/(Channel 1 + Channel 2). GAC data are processed daily and composited on a weekly basis to produce a global map showing vegetation vigor. An example of NDVI applications is the monitoring of the Sahara desert extent. AVHRR data are also used for sea surface temperature.

The first Système Pour l'Observation de la Terre (SPOT), designed by CNES, was launched in 1986. SPOT-2 and SPOT-3 were launched, respectively, in 1990 and 1993. The SPOT satellites fly in a near-polar, sun-synchronous low Earth orbit at an altitude of 832 km and cross the equator at 10:30 a.m. SPOT's repeat cycle is 26 days, but due to its off-nadir viewing capability (viewing angle up to 27°), SPOT has a "revisiting" capability with which the same area can be seen up to five times every 26 days. This off-nadir viewing capability also enables some stereo imaging possibilities. The SPOT payload includes two identical high-resolution visible (HRV) imaging instruments that can be employed in panchromatic or multispectral modes, with pointing capabilities. Spectral coverage of these two modes is given in Figure 79.3. The panchromatic mode is 10 m spatial resolution, while multispectral data have a 20 m spatial resolution. Whereas Landsat is a scanning mirror-type instrument, SPOT employs a push-broom system with a linear array of detectors simultaneously acquiring all data pixels in one image line, which minimizes geometric errors. SPOT data are very useful for applications involving small features. Due to its increased spatial resolution, revisit and pointing capabilities, simultaneous panchromatic and multispectral data, and stereo data capabilities, SPOT opens a new range of applications, such as topographic mapping, studies of earthflows (e.g., land, rock, and mudslides), urban management, and military applications. SPOT-4 is planned for launch in 1998, and SPOT-5 in 2002. Among the planned improvements, a mid-IR channel will be added to SPOT-4, which will also carry a new AVHRR-type instrument, the European Vegetation instrument.

Since 1988, India has launched a series of five satellites, the IRS series. These satellites were designed in support of India's agriculture and exploration businesses, and they seem to be successful in this challenge of bringing remote sensing to the users (see Ref. [10]). For land applications, IRS-1A, -IB, and -1C all carry the LISS instrument, which is a MSS very similar to Landsat-TM. LISS-2 acquires imagery in four bands similar to bands 1–4 of Landsat-TM (from visible to near-IR) at the spatial resolution of 36.5 m (see Figure 79.3 for wavelengths description). LISS-3, carried on IRS-1C, also acquires imagery in four bands, but the visible blue band has been suppressed and replaced by a mid-IR band similar to TM/band 5. Due to their similarity to Landsat data, IRS/LISS-2 data could be used as complements or replacements to Landsat data if needed until Landsat-7 is launched. IRS-1C also carries a 5 m panchromatic instrument whose data are coregistered with LISS-2 data. For more details on IRS data, see Refs. [10,27].

Other instruments are also available. JERS-1, designed by Japan, was launched in 1992, and its payload includes both an SAR instrument and an optical imaging system; see Figure 79.3 for its spectral channels from visible to mid-IR wavelengths, with spatial resolutions of 18 and 24 m. MOMS, the German Modular Optoelectronic Multispectral Scanner, has been flying as a research instrument on U.S. Space Shuttle missions, and has a spatial resolution ranging from 4.5 to 13.5 m; see Ref. [27] for more details on these different instruments.

Among the first EOS instruments to be launched is the Moderate Resolution Imaging Spectrometer (MODIS). MODIS is being developed to provide global monitoring of the atmosphere, terrestrial

ecosystem, and oceans, and to detect climate change. MODIS will cover the visible to IR portions of the spectrum with 36 channels at spatial resolutions of 250 m to 1 km. Many interesting land studies will be performed by fusing together AVHRR, Landsat, and MODIS data.

The fusion of several of these types of data is becoming a very important issue [26]. Already, sensors such as SPOT or LISS-3 present the advantage of acquiring coregistered panchromatic and multispectral data. It would be of great interest to combine data from sensors with different spectral and spatial resolutions, as well as different viewpoints. The combination of coarse resolution viewing satellites for large-area surveys and finer resolution sensors for more detailed studies would offer the multilevel information necessary to assess accurately the areal extent of features of interest (e.g., tropical forests). The fusion of multispectral data with SAR data would provide information on ground cover reflectance with the shape, roughness, and moisture content information from SAR. Of course, multidata fusion requires very accurate registration of the data, as will be described in the next section.

79.3.2 Geologic Studies

Other examples of terrestrial studies are the mapping of geologic features, such as geologic faults and earthquake sites, or volcanic eruptions. Although many geologic features lie beneath the surface of the Earth, remote sensing (aerial or satellite) provides a valuable tool to perform geologic mapping, landforms and structures analysis, as well as mineral exploration. This is due to the fact that topography and soil properties provide clues to underlying rocks and structural deformations. Landsat and SPOT gather data about the effects of subsurface geologic phenomena on the surface. These data are especially useful to recognize some specific landforms (such as volcanoes), to depict topographic features, to discriminate some geologic facies and rock unit distribution patterns, and more generally to provide regional overviews of surface geology. In mineral exploration, rock or soil alteration can be detected by spaceborne sensors and may indicate the presence of mineral deposits or oil reservoirs. Other types of sensors that are very useful for geologic applications are radar sensors, such as the two radar systems, SIR-C and X-SAR, carried on the Space Shuttle Endeavour in 1994. These sensors captured in real time the eruption of a volcano in Russia and an earthquake in Japan [31]. For more information on geologic applications, see Refs. [2,32].

79.3.3 Geophysics Studies

Other satellites, such as the LAGEOS-1 and -2, have proved very useful in geophysics for the study of the Earth's gravity field, tectonic plate motion, polar motion, and tides. LAGEOS sensors are reflector orbs covered with laser beams. For more information on these studies, see Refs. [33,34].

79.3.4 Ocean Studies

Oceans cover 75% of the Earth's surface and contain most of the energy of the planet. Although their role in climate evolution is very important, it is still poorly understood. By understanding chemical, physical, and biological processes in oceans, scientists will be able to model the interactions between oceans and the atmosphere and determine how these interactions affect Earth temperature, weather, and climate.

An example of interaction between oceans and the atmosphere is illustrated by the phenomenon known as El Niño/Southern Oscillation, which occurs in the tropical Pacific Ocean, usually around Christmas time. El Niño is due to a mass of warm water, usually located off Australia which moves eastward toward equatorial South America. El Niño develops every few years (observed on average every 4 years to a maximum of 7 years), and alters the weather in Australia, Africa, South Asia, and the tropical parts of the Americas. By understanding how winds and waves move in the tropical Pacific, scientists have been able to predict the El Niño phenomenon up to 1 year in advance. Similar phenomena are being studied in the Atlantic Ocean, where patterns seem to move much more slowly.

Besides being used to create global models, and in storm and weather forecasting, ocean data are also very important for day-to-day applications such as ship routing, oil production, and ocean fishing.

79.3.4.1 Ocean Colors

Ocean color data are critical for the study of global biogeochemistry and to determine the ocean's role in the global carbon cycle and the exchange of other critical elements and gases between the atmosphere and the ocean [35,36]. It is thought that marine plants remove carbon from the atmosphere at a rate equivalent to terrestrial plants, but knowledge of interannual variability is very poor. For most oceans, the color observed by satellite in the visible part of the spectrum varies with the concentration of chlorophyll and other plant pigments present in the water. Subtle changes in ocean color usually indicate various types and quantities of microscopic marine plants (i.e., phytoplankton are present in the water); the more phytoplankton present, the greater the concentration of plant pigments and the greener the water.

The launched (October 1997) Sea-viewing Wide Field-of-view Sensor (SeaWiFS), which is a part of MTPE, provides quantitative data on global ocean biooptical properties to the earth science community. SeaWiFS is a follow-on sensor to the Coastal Zone Color Scanner (CZCS), which ceased operations in 1986. See Figure 79.3 for a channel description of these two sensors; notice that all channels are concentrated in the (0.4, 0.7 μm) interval of the electromagnetic spectrum.

Other sensors for ocean color are the imaging spectrometer for ocean color applications MOS-IRS, launched on the Indian Remote Sensing Satellite IRS-P3 in March 1996, and the imaging spectrometer MOS-PRIRODA, launched aboard the Russian multisensor remote sensing module PRIRODA and docked to space station MIR in April 1996.

79.3.4.2 Ocean Dynamics

By studying ocean circulation and sea-level trends, scientists will be able to create global maps of ocean currents and of sea surface topography. Since sea surface height and sea-level variations are related to sea surface temperatures, the monitoring of mean sea levels enables the gathering of evidence that can measure global warming or El Niño-type events. For example, conditions related to El Niño may result in a change in sea surface height of 18 cm or greater [37].

TOPEX/Poseidon (T/P) is an important collaboration between USA/NASA and France/CNES. T/P uses radar altimetry to provide 10 day maps of the height of most of the ice-free oceans' surface. Circling the world every 112 min, the satellite gathers data for 3–5 years, and could be operational for 10 years. The T/P satellite was launched in August 1992 on an Ariane rocket. TOPEX measures the height of the ocean surface, as well as changes in global mean sea level. From these altimetry data, global maps of ocean topography are created, from which speed and direction of ocean currents are computed worldwide. Changes in mean sea level are monitored and currently are viewed mostly as related to natural ocean variability and not climate change. Climate change must be studied over a much longer time series of altimeter data. T/P also enables study of tides, wave geophysics, and ocean surface winds.

Sea winds are also being studied with scatterometers such as the NASA Scatterometer (NSCAT) and the soon to be launched EOS Scatterometer, SearWinds. These high-frequency radar instruments measure the reflected signals from the ocean surface to detect wind speed and direction.

ERS-1 is another satellite utilized to measure ocean dynamics. ERS-1 was launched in 1991 on a sun-synchronous, near-polar low-Earth orbit at an altitude of 780 km. ERS-1 orbits the Earth in 100 min and covers the entire planet in 3 days. Its payload consists of two specialized radars and one IR sensor. The Active Microwave instrument, consisting of a synthetic aperture radar and wind scatterometer, produces extremely detailed images of 100 km swath of the Earth's surface, with a spatial resolution of 20 m. The radar altimeter provides accurate range to sea surface and wave heights, and the along-track scanning radiometer constructs detailed pictures of the thermal structure of the seas and oceans from

surface temperature measurements at accuracy of <0.5 °C. ERS-1 images are also utilized for land applications where the instruments need to "look through" the cloud cover.

The study of sea ice with passive and active microwave sensors is also very important and additional reading in this topic can be found in Refs. [14,38].

79.3.5 Space Science Applications

Astronomical satellites have been developed to observe far distant objects that are usually beyond the range of ground-based instruments. They explore phenomena in the solar system, and beyond. Satellite observation of astronomical objects is also less sensitive to atmospheric interactions and can achieve higher accuracy than ground-based measurements. This section will give a brief description of the most important space science satellites.

The first astronomical satellite to be put into synchronous orbit was the International Ultraviolet Explorer (IUE) laboratory. IUE was launched in 1978 under a joint program involving NASA, ESA, and the United Kingdom. In more than 15 years of service, IUE gathered observations on >10,000 celestial objects. A program for coordinating its observations with those of the ROSAT satellite has been carried out under the title RIASS (Rosat-IUE All-Sky Survey). ROSAT, the Roentgen Satellite, is a joint collaboration between Germany, the United States, and the United Kingdom, and was launched in 1990. It is an x-ray observatory that carries two instruments, the x-ray telescope and the wide field camera.

The Infrared Astronomical Satellite (IRAS) is a joint project of the United States, the United Kingdom, and the Netherlands. The IRAS mission was intended to provide a survey of IR point sources (from 12 to 100 μm), but has also produced very high-quality image data. The Mid-Course Space Experiment (MSX), the Infrared Space Observatory (ISO), and the Space InfraRed Telescope Facility (SIRTF) are other examples of recently or soon-to-be launched sensors that provide an even finer resolution.

Hipparcos (High Precision Parallax Collecting Satellite) is an astronomy satellite launched in August 1989, with the purpose of determining the astrometric parameters of stars with unprecedented precision. After a life of 4 years, Hipparcos has produced two catalogs. The Hipparcos Catalogue provides position, parallax, and proper motion measurements with accuracy of 2 milliarcsec at 9 mag for over 120,000 stars. The Tycho Catalogue is the result of somewhat less precise astrometric measurements for some 1 million stars.

COBE, the Cosmic Origin Background Explorer developed by NASA, was launched in 1989. Designed to measure the diffuse IR and microwave radiation from the early universe, it carried three instruments: a Far Infrared Absolute Spectrophotometer (FIRAS), a Differential Microwave Radiometer (DMR), and a Diffuse Infrared Background Experiment (DIRBE). The first full-sky coverage was completed in 1990.

The Hubble Space Telescope (HST) is one of the most well-known astronomical satellites. It was built as a joint NASA/ESA project, and was launched in 1990 as a long-term space-based observatory. The heart of the system is a large reflector telescope 2.4 m in diameter. All the instruments on-board the HST use the light gathered by the reflector telescope. Current HST instruments are the Wide/Field Planetary Camera 2 (WFPC2), the Space Telescope Imaging Spectrograph (STIS), the Near-Infrared and Imaging Spectrograph (NICMOS), and the Faint Object Camera, FOC, provided by ESA. These different instruments can observe astronomical objects from UV to IR wavelengths. In 1993, the HST was serviced to correct a preliminary fault affecting the mirror with a corrective optical apparatus named COSTAR. Despite the preliminary mirror fault, and even more after correction, the HST has achieved much better results than those from observatories on Earth. Since it is located above the Earth's atmosphere (at 600 km), the HST produces highly detailed images of the stars and can detect objects beyond the range of ground-based instruments. Observations with the HST are scheduled as a space-based observatory according to worldwide astronomers' proposals.

The Advanced Satellite for Cosmology and Astrophysics (ASCA) is the product of a Japan/U.S. collaboration. Launched in 1993, this x-ray astronomy mission was still operational in 1997, and carries four large-area x-ray telescopes with arc minute resolution. ASCA data are being archived and can

be searched and retrieved online at the High Energy Astrophysics Science Archive Research Center, HEASARC. Gamma Ray Observatory (GRO) and the Advanced X-ray Astrophysics Facility (AXAF) are other examples of space sensors which operate in this spectrum range.

79.4 Management and Interpretation of Satellite Data

Satellite sensors gather the electromagnetic energy reflected or emitted from Earth (or any other planetary) surface features. This energy is then converted into a digital representation that is visualized by a user and interpreted either visually or with a computer. This section summarizes some preliminary ideas on how the digital representation is formed and the basic types of data processing necessary before any further interpretation of the data. For more details on the processing of remote sensing data, see Refs. [39–42].

79.4.1 Fundamental Data Levels

After transmission from the satellites, raw data are usually processed, calibrated, archived, and distributed by a ground-based data system. Most of NASA satellite data products are classified in the following data levels [7]:

- *Level* 0 *data* are the reconstructed raw instrument data at full resolution.
- *Level* 1*A data* are reconstructed, time-reference raw data, with ancillary information including radiometric and geometric coefficients.
- *Level* 1*B data* are corrected Level 1A data (in sensor units).
- *Level* 2 *data* are derived geophysical products from Level 1 data, at the same resolution and location, for example, atmospheric temperature profiles, gas concentrations, or winds variables.
- *Level* 3 *data* correspond to the same geophysical information as Level 2, but mapped onto a uniform space–time grid.
- *Level* 4 *data* are model output or results from analysis of lower-level data.

79.4.2 Image Restoration

Ideally, the scene as viewed and recorded by a sensor would be an exact rendering of the features within the sensor's viewing extent, represented as a spectral curve indicating the amount of energy reflecting/radiating for each point in a scene for a range of given wavelengths. From an engineering standpoint, this is impossible, however, because each image is discretized into a finite number of pixels. Variability defines nature, so each pixel will map into a region of the scene that contains a number of features, each producing its own unique spectral curve. The spectral signature recorded for a pixel is a function of these features and their relative sizes within the region covered by the pixel. The spectral response itself is also discretized into a finite number of bandwidths, where each bandwidth covers a small continuous band of the spectrum. The sensor records for each pixel the amount of energy observed for each band. This number itself, referred to as a Digital Number (DN), must be represented in a finite amount of computer memory, such as 8 bits, meaning that each band records activity as a whole number ranging from 0 to 255.

In practice, a number of events outside human control affect the quality of the observation, such as atmospheric scattering, variations in sun angle, high albedo, and instrument errors. Depending on the application, it may be desirable to correct for the presence of thin clouds within an image. The process of image restoration attempts to control and correct for these conditions [42].

Electromechanical effects due to the instrument itself can be discovered due to their periodic nature (such as caused by the repeated motion of a push broom, or the revolving of a mirror, or the physical process of gathering calibration points). A Fourier transform applied to an image from a sensor undergoing periodic interference exhibits strong noise spikes. A filter can then be used to remove the offending data.

Unfortunately, this also removes any good data that happens to fall at the same frequency, although normally this is but a small portion of the data. Data outages and instrument recorder failures appear as streaks in the image parallel with the scanline, and can be discovered by comparing the respective readings of the pixels in the surrounding scanlines of the image.

To account for the atmospheric effects of Rayleigh and aerosol scattering, an estimate of the portion of the signal that is due to the atmosphere is computed and subtracted from the recorded value. The reflectance of water in the near-IR region of the spectrum should be effectively zero, so the value to subtract for the near-IR band corresponds to the reading of the sensor observed over clear open water. To compute values to be subtracted for each of the other spectral components, a histogram should be formed for each band of a number of sample readings over clear open water. The lowest reading in each band is then used as an estimate of the value to subtract from each pixel to account for the atmospheric effect. In addition, information derived from TOVS, balloon readings, or the atmospheric correction software 5S can be useful in dealing with atmospheric effects.

79.4.3 Data Compression

Data compression is one of the most important tools to overcome the problems of data transmission, storage, and dissemination [43]. Data compression methods are usually classified as either lossless or lossy. With a lossless data compression scheme, the original data can be reconstructed exactly without any loss; in a lossy compression scheme, original data are reconstructed with a degree of error. For transmission from the satellite to the ground station, a lossless data compression must be utilized. For browsing purposes, lossy compression enables quick searches through large amounts of data. A compression scheme is also characterized by its compression ratio, that is, the factor by which the amount of information which represents the data is reduced through compression. For earth science data, lossless compression schemes provide compression ratios up to 2 or 3, while lossy techniques can reduce the amount of information by a factor of 20 or more without degrading the visual quality of the data.

Among the lossless compression methods, the Joint Photographic Experts Group (JPEG) developed a lossless compression method that is based on a predictor, an entropy encoder for prediction error, and an entropy code specifier. Another lossless compression scheme is the Rice algorithm, which can adapt to data of any entropy range. It is based on a preprocessor that spatially decorrelates the data, followed by a variable length encoder. This algorithm gives some of the best compression ratios among all lossless methods, and has been implemented on VLSI chips at NASA.

JPEG has also developed a lossy method based on the Discrete Cosine Transform (DCT). Other methods such as vector quantization or wavelet compression provide either lossless or lossy compressions. In a vector quantization technique, a dictionary of representative vectors is also called a codebook, and all data are encoded relative to the codebook. In this method, the one-time encoding step is computationally expensive but the decoding step at the user end is fast and efficient. Vector quantization is also utilized in a progressive scheme for "quick look"/browsing purposes. In a subband/wavelet compression method, signals are decomposed using quadrature mirror or wavelet filters [44]. Most energy is contained in the low-frequency subbands and high compression ratios can be obtained by compressing the high-frequency information.

For more information or references on data compression techniques, see Ref. [43].

79.4.4 Image Registration

In studying how the global environment is changing, programs such as Mission to Planet Earth [7] or the New Millennium program [45] involve the comparison, fusion, and integration of multiple types of remotely sensed data at various temporal, radiometric, and spatial resolutions. Results of this integration can be utilized for global change analysis, as well as for the validation of new instruments or of new data analysis. The first step in this integration of multiple data is registration, either relative image-to-image registration or absolute georegistration, to a map or a fixed coordinate system. Another case of

image registration is coregistration of multiple bands of one sensor. When the detectors of each spectral band have different spatial locations on the satellite's focal plane, there could be misregistration between each band's raw image [46,47].

Currently, the most common approach to image registration is to extract a few outstanding characteristics of the data, which are called *control points* (CPs), *tie points*, or *reference points*. The CPs in both images (or image and map) are matched by pair and used to compute the parameters of a geometric transformation. Most available systems follow this registration approach; and because automated procedures do not always offer the needed reliability and accuracy, current systems assume some interactive choice of the CPs. But such a point selection represents a repetitive, labor-, and time-intensive task that becomes prohibitive for large amounts of data. Also, since the interactive choice of control points in satellite images is sometimes difficult, too few points, inaccurate points, or ill-distributed points might be chosen, thus leading to large registration errors. A previous study [48] showed that even a small error in registration can have a large impact on the accuracy of global change measurements. For example, when looking at simulated 250 m spatial resolution MODIS data, a 1 pixel misregistration can produce 50% error in the computation of the NDVI. So, for reasons of speed and accuracy, automatic registration is an important requirement to ease the workload, speed up the processing, and improve the accuracy in locating a sufficient number of well-distributed accurate tie points.

Automatic image registration methods can be classified into two types: those that follow a human approach, by first extracting control points, and those that take a more global approach. Among the first methods, the most common features utilized as control points are the centers of gravity of regions—with or without region attributes such as areas, perimeters, ellipticity criteria, affine-invariant moments, and interregions distances. More recently, features extracted from wavelet decomposition have also been utilized, such as maxima and minima of wavelet coefficients, high-interest points, or local curvature discontinuities. A few methods utilize Delaunay triangulation methods to progressively increase the number of accurate control points. For the methods that do not match individual pairs of control points, the transformation is either found by correlation or by optimization, in the spatial or in the frequency domain. When in the spatial domain, correlation or optimization is performed either in the original data or on edge gradient data. Other methods propose a global image matching of edge segments or vectors linking feature points. Some recent research has also focused on the use of wavelets for global image registration. More complete surveys of image registration methods can be found in Refs. [47,49,50].

79.4.5 Dimension Reduction

The first step in analyzing multichannel data is to reduce the dimension of the data space. It is particularly important when the analysis method requires a training step, for example, supervised classification (see next section). The main issue in this case has often been referred as "the Curse of Dimensionality" [51]. If the original data have a large number of bands (e.g., for hyperspectral data), theoretical studies have shown that a very large training set should be utilized; but using a large training set deteriorates the estimation of the kernel density. To solve this problem, various dimension reduction schemes enable to perform classification in a smaller-dimensional subspace. Since the information contained in multiple channels is often redundant, it is possible to decorrelate spectrally the channels and reduce the number of channels to be analyzed without losing any information. Principal Component Analysis (PCA) and Projection Pursuit are the most common techniques for dimensionality reduction. For more information on these methods, refer to Refs. [39–42].

79.4.6 Data Mining

One objective of the NASA-initiated Mission to Planet Earth is to gather sufficient data to enable scientists to study the Earth as a dynamic system, resulting in a better understanding of the interactions between humans, the atmosphere, and the biosphere [8]. The episodic nature of most interesting events

would cause them to be missed if the data were not being gathered continuously. Comprehensive data sets allow scientists to construct and evaluate complex models of many Earth-related processes. But currently, due to computation and time constraints, only a small percentage of the data gathered by remote sensing are actually viewed by an individual user. Data-gathering missions tend to be multidisciplinary, so different aspects of the data sets are pertinent to different researchers.

Data mining can be defined as the process by which data content is automatically extracted from satellite data, enabling a scientist to query the data holdings based on high-level features present within an image [52]. Given the projected large volumes of data, it is not feasible to rely solely on conventional data management paradigms. Standard methods of segmenting images are inadequate as standalone techniques for image recognition, regardless of the speeds of processing, because there are no general methods for automatically assigning meaningful semantics to any homogeneous regions that are isolated. Metadata derived directly from the image header are not rich enough to enable robust querying of a database in most instances, but limit a user to retrieving all images at a given latitude/longitude during some time period, for example, regardless of the image quality or the unique features existing due to some unexpected set of circumstances. New approaches based on techniques such as image classification (described in the next section) are now feasible, due to the phenomenal increases in computing speed, the availability of massively parallel architectures, and the breakthroughs in signal processing.

An example of data mining is the browsing of 15 years of TOVS data with two complete coverages per day, which would require looking through 10,958 scenes per attribute. Of the several products generated, the scientists are primarily interested in browsing those with some given resolution. After locating a browse product that seems to indicate an interesting structure or phenomenon, a scientist might then retrieve this data temporally, or any supporting data set for further analysis. Scientists using the TOVS data sets desire a more intelligent form of querying so they can quickly and easily find relevant data sets that are pertinent to their research. Certain TOVS observations are more "interesting" than others, and the definition of "interesting" is a combination of objective fact and subjective opinion. Data mining is applicable here to aid in evolving a retrieval heuristic based on an individual scientist's definition of "interestingness." In one approach, the scientist could prepare a representative set of images that are labeled as positive or negative instances of "interesting," and a machine learning system (e.g., neural network, genetic algorithm) could perhaps be trained to classify the remaining images in the TOVS data set according to this definition. In a second approach, the scientists could be asked to identify explicitly structural features within the images that make them interesting, and image processing routines could then be applied to detect images with these features. Over time, a scientist could provide feedback to the heuristic classifier to improve its performance. Both approaches require that the underlying representation language (structures, bin size, spatial and temporal relationships) be robust and flexible enough to permit an appropriate level of expression.

79.4.7 Classification

Image classification is the task of developing a statistical model that labels every potential point in some multidimensional space. A *parametric classifier* assumes that data are described by some underlying parameterized probability density function (PDF). A training set of representative data from the domain is then used to supply appropriate values. For example, if a Gaussian or normal distribution is assumed, then the means, standard deviations, and joint-covariance matrix can be computed from the training data. A *nonparametric* or *unsupervised classifier* is typically used when there is insufficient knowledge about the type of underlying PDF for the domain. Self-organizing classifier models, such as certain kinds of neural networks, are also considered nonparametric classifiers when they make no *a priori* assumptions about any PDF.

In a *statistical* or *supervised classifier*, knowledge about the distribution of the data is utilized to assign a label to an unclassified pixel. Using "ground-reference data," a training set of known points is created. A prototype vector can then be calculated as the mean of all samples for each of the classes.

Assuming a Gaussian distribution in each of the channel readings for a given class, the standard deviation for each class is computed based on the sample. Then, the lowest distance from the given feature prototypes to an unclassified point determines the class of this incoming point. As simple and elegant as this approach might appear, in actuality its utility is limited. Features are not so discernible from a random labeling of an image. Thus, although this algorithm is inaccurate, it is consistent in its mislabelings. The algorithm's deterministic nature and underlying use of a continuous function combine to produce predictable behavior. In general, this algorithm labels all points similarly if they fall within the same neighborhood in the feature space.

Other parametric classifiers follow the *Maximum Likelihood Decision Rule*, which allows the construction of discriminant functions for the purposes of pattern classification. For more details on this technique, refer to Ref. [53].

The classifiers discussed earlier are, by definition, required to assign an unclassified pixel to the one nearest class. No measurement of the distance to that class or proximity to other classes is recorded, and no information on the confidence of the labeling is provided. *Fuzzy classifiers*, on the other hand, are not obligated to pigeonhole a pixel into a single class. Instead, the pixel is assigned a degree of membership for each possible class. Intuitively, and indeed for mathematical tractability, the pixel's memberships must sum to one, and the degree of membership for a given class must be between 0 and 1, inclusively. Two examples of fuzzy classifiers are given as follows. The *Fuzzy Nearest Neighbor* nonparametric classifier places an unclassified vector in the dominant class of its k-closest training vectors. If no class has an outright majority, then distances to the nearby vectors for each class which tied are summed and the unclassified vector is placed in the class with the minimum sum. The *Fuzzy Decision Tree Classifier* utilizes a decision tree as the data structure that encapsulates knowledge of what to do given a set of conditions. See Ref. [52] for more information on this method. Although this algorithm is conceptually simple, it is only recently that it has become computationally feasible due to the need to search the tree for each unclassified pixel to locate the nearest path. The search algorithm can also be sped up by running the algorithm on a parallel architecture such as a Single Instruction Multiple Data (SIMD) machine.

Many researchers have investigated the use of *neural networks* for automatically extracting metadata from images [54]. Many different neural network models have been considered, but with respect to performance accuracy, the backpropagation training technique has shown to be the best classifier [55]. The *backpropagation algorithm* is the most common method for training a neural network, and is the backbone of much of the current resurgence of research into neural nets [56]. With respect to pattern recognition, backpropagation can be considered to be a nonparametric technique for estimation of *a posteriori* probabilities.

79.4.8 Accuracy Assessment

A measurement derived solely from satellite imagery is of questionable use unless the technique employed for computing that measurement on those data has been validated. A technique that appears to work accurately on satellite imagery over some given location at some given time may perform abysmally on data from the same sensor at another location, or for the same location at another time. The reasons for this are many: through the course of a year, the sun angle changes causing different lighting conditions; from pass to pass, the viewing angle of the instrument can be different; with seasonal changes, surface reflectance varies due to weather conditions and the alteration of land cover as crops appear in different stages; atmospheric conditions fluctuate; and the sensor and spacecraft themselves age and possibly perform differently.

The key factor in any accuracy assessment of remote sensing data is the method and source used for determining what the satellite sensor is actually viewing. This ground reference data is gathered independent of the remote sensing data itself. There are several sources that can be construed as ground reference data, and each source has its own degree of accuracy. The most obvious is an actual site visit to the

area of interest. What is observed, also known as "ground truth," is recorded and compared to the digital rendition of the same spatial extent. This approach usually has a high degree of accuracy, but it is often prohibitively expensive. Depending on the time between the on-site ground reference gathering and the imaging of the area, the validity of the ground reference data may be lessened due to anthropomorphic or natural influences. The shorter the life of the feature being measured, the more difficult it is to find or gather meaningful time-critical ground reference data. If ground reference data are not available, it may be possible to perform photointerpretation with some degree of success. This itself depends on the knowledge of the photointerpreter, and the availability and suitability of a display device for viewing the image data and recording the photointerpreter's assessment. Another approach is to compare the digital image with other sources of ground reference data such as air photos or appropriate reference maps, provided the feature of interest is detectable using those sources. The degree of correspondence between the ground reference data and the measurement derived from the sensor data can then be compared for accuracy. In the worst case, the lack of adequate/accurate ground reference data requires using an unsupervised clustering approach that is usually less accurate but much cheaper to produce.

79.5 Future in Satellite Imaging and Sensing

Success of future earth and space science missions depends on increasing the availability of data to the scientific community who will be interpreting space-based observations, and on favoring interdisciplinary research for the analysis and the use of this data. One of the main challenges in the future of satellite imaging and sensing will be to handle, archive, and store all of these data in a way that can be easily accessible and retrieved by anyone who needs to use them. Systems such as the EOS Data and Information System (EOSDIS) [57,58] will require that over 1 terabyte per day be collected and processed into several levels of science data products within several hours after observation. After 15 years, the estimated amount of collected, processed, analyzed, and stored data will equal about 11,000 terabytes. Also at NASA, efforts are underway to design an advanced information system, based on an object-oriented database, with the express purpose of developing, incorporating, and evaluating state-of-the-art techniques for handling EOS-era scientific data challenges [59].

Another challenge will be to analyze this tremendous amount of data, and to find out new ways to fuse, integrate, and visualize this data. In particular, research in fast computational capabilities, such as field programmable gate arrays (FPGAs), will be of great importance.

On the other hand, the wide distribution of satellite data to the general public will be facilitated by regional distribution systems such as the Regional Application Centers (RACs) [60,61], whose goal is to provide local users, such as industry, agriculturalists, urban planners, regional communities, with local and "on-time" information about regional applications.

Satellite imaging and sensing is a field with a history of more than two decades, but is still in full expansion. The future in satellite imaging and sensing will see developments in several areas. The next millennium will see an explosion of commercial satellite systems and the profusion of satellite data, which will have economic and sociopolitical implications. As of this writing, over 30 commercial Earth sensing satellites are either being planned or being built. MTPE and EOS will generate unprecedented amounts of diverse resolution data. The future will also see the development of locally directed satellite systems, in answer to specific applications for specific areas of the planet. Telecommunications will also be a large part of the space market. In space, after the large success of the Mars Pathfinder mission, exploration of distant planets will see a flourishing of distant satellite systems providing unprecedented amounts of data to analyze regarding other planets' surface features, atmospheric, and magnetic properties. The understanding of other planets will also enable scientists to learn more about the Earth comparatively to other planets such as Mars, and to build a comprehensive data set to aid in planning future missions. The Mars Global Surveyor is an example of such as a mission; it will map the entire planet Mars by taking high-resolution pictures of the surface. The future might see a 10 year NASA program that will send pairs of Surveyor-like orbiters and Pathfinder-like landers to Mars every 26 months.

In order to gather novel and interesting data, this type of mission will need an increasing amount of on-board processing that will perform mission planning, image processing and understanding, as well as data compression and fusion. The design of systems including on-board processing will require new computational capabilities, such as reconfigurable hardware and parallel processors, as well as new developments in intelligent systems. In the near future, satellite imaging and sensing is a field that will produce unprecedented information about the Earth, its environment, and our solar system.

Acknowledgments

The authors would like to thank William J. Campbell for his support and for his useful comments upon reviewing our paper, Bob Mahoney for providing the spectral libraries used to generate Figure 79.2, and all the anonymous or nonanonymous authors of Web pages that we consulted during the research part of this endeavor. In particular, the online remote sensing tutorial by N.M. Short, edited by J. Robinson at the URL http://code935.gsfc.nasa.gov/Tutorial/TofC/Coverpage.html, and the list of selected links on remote sensing compiled by U. Malmberg and found at the URL http://www.ssc.se/rst/rss/index.html were very useful.

References

1. T. M. Lillesand and R. W. Kiefer, *Remote Sensing and Image Interpretation*, 2nd edn., John Wiley & Sons, New York, 1987.
2. J. B. Campbell, *Introduction to Remote Sensing*, 2nd edn., The Guilford Press, 1996.
3. A. R. Cracknell and L. W. B. Hayes, *Introduction to Remote Sensing*, Taylor & Francis Group, New York, 1991.
4. R. Greeves, A. Anson, and D. Landen, *Manual of Remote Sensing*, American Society of Photogrammetry, Falls Church, VA, 1975.
5. P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, Super-resolved surface reconstruction from multiple images, Technical Report FIA-94-12, NASA/Ames Research Center, Artificial Intelligence Branch, October 1994.
6. T. E. Bell, Harvesting Remote-Sensing Data, *IEEE Spectrum*, 32(3), 24–31, 1995.
7. G. Asrar and R. Greenstone (eds.), *1995 MTPE EOS Reference Handbook*, EOS Project Science Office, Code 900, NASA/Goddard Space Flight Center.
8. *1993 EOS Reference Handbook*, G. Asrar and D. Dokken (eds.), available from the Earth Science Support Office, Document Resource Facility, Washington, DC, 20024.
9. *SCOPE: SCenario for Observation of Planet Earth*, Publication of the National Space Development Agency of Japan, NASDA, 1995.
10. K. P. Corbley, Multispectral Imagery: Identifying more than meets the eye, *Geo Info Systems*, 38–43, June 1997.
11. A. J. Krueger, The global distribution of total ozone: TOMS Satellite measurements, *Planetary Space Sci.*, 37(12), 1555–1565, 1989.
12. S. Muller and A. J. Krueger, Analysis and comparison of ozone maps obtained by TOMS and TOVS during the Map/Globus 1983 Campaign, *Planetary Space Sci.*, 35(5), 539–545, 1987.
13. A. J. Krueger, Nimbus-7 Total Ozone Mapping spectrometer (TOMS) data during the GAP, France, Ozone Intercomparisons of June 1981, *Planetary Space Sci.*, 31(7), 773–777, 1983.
14. C. L. Parkinson, *Earth from Above. Using Color-Coded Satellite Images to Examine the Global Environment*, University Science Books, Sausalito, CA, 1997.
15. A. M. Thompson, The oxidising capacity of the Earth's atmosphere: Probable past and future changes, *Science*, 256, 1157–1165, 1992.
16. M. Schoeberl, J. Pfaendtner, R. Rood, A. Thompson, and B. Wielicki, *Atmospheres Panel Report to the Payload Panel, Palaeogeography, Palaeoclimatology, Palaeoecology* (Global and Planetary Change Section) 98, 9–21, Elsevier Science Publishers B.V., Amsterdam, 1992.

17. J. Susskind, J. Rosenfield, and D. Reuter, Remote sensing of weather and climate parameters from HIRS2/MSU on TIROS-N, *J. Geophys. Res.*, 89(D3), 4677–4697, 1984.
18. J. Simpson (ed.), TRMM: The Satellite Mission to measure tropical rainfall: Report of the Science Steering Group, NASA Publication, August 1988.
19. K. B. Kidwell, *NOAA Polar Orbiter Data Users Guide*, National Oceanic and Atmospheric Administration, December 1991.
20. J. R. G. Townshend (ed.), Improved global data for land applications. A proposal for a new high resolution data set, Global Change Report No. 20, Report of the Land Cover Working Group of IGBP-DIS, 1992.
21. J. R. G. Townshend and C. O. Justice, Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations, *Int. J. Remote Sensing*, 9, 187–236, 1988.
22. TREES, Tropical Ecosystem Environment Observations by Satellites. Strategy Proposal 1991–1993, Commission of the European Communities, Joint Research Centre, Institute for Remote Sensing Applications.
23. D. Skole and C. J. Tucker, Tropical deforestation and habitat fragmentation in the Brazilian Amazon: Satellite data from 1978 to 1988, *Science*, 260, 1905–1910, 1993.
24. C. J. Tucker, B. N. Holben, and T. E. Goff, Intensive Forest Clearing in Rondonia, Brazil, as detected by satellite remote sensing, *Remote Sensing Environ.*, 15, 255–261, 1984.
25. J. P. Malingreau, C. J. Tucker, and N. Laporte, AVHRR for Monitoring Global Tropical Deforestation, *Int. J. Remote Sensing*, 10(4–5), 855–867, 1989.
26. M. Ehlers, Integrating remote sensing and GIS for environmental monitoring and modeling: Where are we? *Geo Info Syst.*, 36–43, July 1995.
27. T. Cary, A world of possibilities: Remote sensing data for your GIS, *Geo Info Syst.*, 38–42, September 1994.
28. N. M. Short, *The Landsat Tutorial Workbook: Basics of Satellite Remote Sensing*, Scientific and Technical Information Branch, National Aeronautics and Space Administration, Washington, DC, 553 p, 1982.
29. D. L. Williams and A. Jenetos, Landsat-7 Science Working Group Report, July 1993.
30. Nicholas M. Short, P. D. Lowman, Jr., S. C. Freden, and W. A. Finch, Jr., *Mission to Earth: Landsat Views of the World, Scientific and Technical Information Office*, National Aeronautics and Space Administration, Washington, DC, 1976.
31. D. L. Evans, E. R. Stofan, T. D. Jones, and L. M. Godwin, Earth from Sky, *Sci. Am.*, 271(6), 70–75, December 1994.
32. N. M. Short, *Geomorphology from Space: A Global Overview of Regional Landforms*, Scientific and Technical Information Branch, National Aeronautics and Space Administration, Washington, DC, 1986.
33. D. E. Smith, R. Kolenkiewicz, P. J. Dunn, S. M. Klosko, J. W. Ronnins, M. H. Torrece, R. G. Williamson, E. C. Pavlis, N. B. Douglas, and S. K. Fricke, *Lageos Geodetic Analysis*, SL7.1, NASA Technical Memorandum 104549, September 1991.
34. S. C. Cohen and D. E. Smith, LAGEOS Scientific Results, *J. Geophys. Res.*, 90, 9217–9220, 1985.
35. S. B. Hooker, W. E. Esaias, G. C. Feldman, W. W. Gregg, and C. R. McClain, Volume 1, An overview of SeaWiFS and ocean color, in SeaWiFS Technical Report Series, S. B. Hooker (ed.), NASA Technical Memorandum 104566, Vol. 1, July 1992.
36. EOS-Ocean Color: Availability of the global data set, *Trans. Geophys. Union*, 70(23), June 1989.
37. TOPEX/Poseidon: Decoding the Ocean, French Space Agency/CNES Report, December 1993, available from Centre National d'Etudes Spatiales, 2 Place Maurice Quentin, 75039 Paris Cedex 01, France.
38. D. K. Hall and J. Martinec, *Remote Sensing of Ice and Snow*, Chapman & Hall, London, U.K., 1985.
39. P. H. Swain and S. M. Davis, *Remote Sensing: The Quantitative Approach*, McGraw-Hill, New York, 1978.
40. P. M. Mather, *Computer Processing of Remotely Sensed Images*, paperback edition, John Wiley & Sons, Chichester, 1989.
41. B. Jähne, *Digital Image Processing. Concepts, Algorithms and Scientific Applications*, Springer Verlag, New York, 1991.

42. J. G. Moik, Digital processing of remotely sensed images, NASA Publication SP-431, 1979.
43. J. C. Tilton and M. Manohar, Earth science data compression issues and activities, *Remote Sensing Rev*., 9, 271–298, 1994.
44. S. Mallat, A Theory for multiresolution signal decomposition, *IEEE Pattern Anal. Machine Intelligence*, PAMI-11(7), 674–693, 1989.
45. E. K. Casani, The New Millennium Program: Positioning NASA for the Ambitious Space and Earth Science Missions of the 21st Century, Albuquerque, NM, JPL Technical Report, October 1995.
46. J. Le Moigne, N. El-Saleous, and E. Vermote, Iterative edge- and wavelet-based image registration of AVHRR and GOES satellite imagery, *Image Registration Workshop, IRW97*, NASA/GSFC, Greenbelt, November 20–21, 1997.
47. J. Le Moigne, W. J. Campbell, and R. F. Cromp, An automated parallel image registration technique of multiple source remote sensing data, *IEEE Trans. Geosci. Remote Sensing*, 40(8), pp. 1849–1864, August 2002.
48. J. Townshend, C. O. Justice, C. Gurney, and J. McManus, The impact of misregistration on change detection, *IEEE Trans. Geosci. Remote Sensing*, 30, 1504–1060, 1992.
49. L. Brown, A survey of image registration techniques, *ACM Comput. Surv*., 24(4), pp. 325–376, 1992.
50. L. M. G. Fonseca and B. S. Manjunath, Registration techniques for multisensor remotely sensed imagery, *J. Photogrammetry Eng. Remote Sensing*, 62, 1049–1056, 1996.
51. D. W. Scott, The curse of dimensionality and dimension reduction, in *Multivariate Density Estimation: Theory, Practice, and Visualization*, John Wiley & Sons, New York, Chapter 7, pp. 195–217, 1992.
52. R. F. Cromp and W. J. Campbell, Data mining of multidimensional remotely sensed images, invited paper in *Proceedings of 2nd International Conference on Information Knowledge Management*, Washington, DC, pp. 471–480, November 1993.
53. H. C. Andrews, *Introduction to Mathematical Techniques in Pattern Recognition*, Wiley-Interscience, New York, 1972.
54. W. J. Campbell, S. E. Hill, and R. F. Cromp, Automatic labeling and characterization of objects using artificial neural networks, *Telematics Informatics*, 6(3–4), 259–271, 1989.
55. S. R. Chettri, R. F. Cromp, and M. Birmingham, Design of neural networks for classification of remotely sensed imagery, *Telematics Informatics*, 9(3/4), 145–156, 1992.
56. J. Hertz, A. Krogh, and R. Palmer, *Introduction to the Theory of Neural Computation*, Addison-Wesley, Redwood City, CA, 1991.
57. *EOS Data and Information System* (*EOSDIS*), NASA, Washington, DC, available from the Earth Science Support Office, Document Resource Facility, Washington, DC, May 1992.
58. DAAC/DADS Internal Design Documents, edited by Goddard DAAC, Code 902.2, NASA Goddard Space Flight Center.
59. R. F. Cromp, W. J. Campbell, and N. M. Short, Jr., An intelligent information fusion system for handling the archiving and querying of terabyte-sized spatial databases, *International Space Year Conference on Earth and Space Science Information Systems*, American Institute of Physics, Pasadena, CA, 1992.
60. W. J. Campbell, N. M. Short, Jr., P. Coronado, and R. F. Cromp, Distributed Earth Science Validation Centers for Mission to Planet Earth, *8th International Symposium, ISMIS'94*, Charlotte, NC, October 1994.
61. W. J. Campbell, P. Clemens, J. Garegnani, R. F. Cromp, and P. Coronado, Applying information technologies to facilitate information access and regional development, *Proceedings of the Workshop Technology 2007*, Boston, MA, September 1997.

80

Humidity and Moisture Measurement

Gert J.W. Visscher
Institute of Agricultural and Environmental Engineering

Water and water vapor can be found everywhere. Because of the asymmetrical distribution of their electric charge, water molecules are easily adsorbed on almost any surface, where they are present as a mono- or multimolecular layer of molecules. Water vapor in the air or any other gas is generally called *humidity*; in liquids and solids, it is usually designated as *moisture*. The determination of humidity and moisture, as in prediction of floods, fog, conditions for the appearance of plant diseases, etc., is of great economic importance. Stored foodstuffs or raw materials may dry up at low humidity or get moldy at high humidity. In many industrial processes, the measurement of moisture and humidity is important for the maintenance of the optimum conditions in manufacturing. Humidity and moisture content can be expressed in a number of ways, and the number of methods for measuring them is even greater. An engineer whose main concern is to avoid condensation no matter where in his system will be interested in the *dew point* of the gas flow. A chemist may be interested in the mere *quantity* of water vapor, whereas in a printing office or a storage room, the *relative humidity* is of more importance.

Water vapor is one of the constituent gases of the Earth's atmosphere, the total pressure P of which is, according to Dalton's law, the sum of the partial pressures. This means that

$$P = P_{N_2} + P_{O_2} + P_{H_2O} + P_{\text{other gases}} \tag{80.1}$$

Like other gases, water vapor can be considered to behave as an ideal gas, except near saturation. In average environmental conditions, water can also be present in the liquid and solid phase, the reason to

speak of water vapor rather than of water gas. An empty space in equilibrium with a flat water (or ice) surface can, at a given temperature, hold a well-defined maximum quantity of water vapor [1–3]. When this *saturation vapor pressure* is reached, any further addition of water vapor results in condensation. In the presence of air molecules at atmospheric pressure, the saturation vapor pressure is about 0.4% higher (expressed by the so-called enhancement factor [4]). The saturation vapor pressure for water is about 611 Pa at 0 °C, 2339 Pa at 20 °C, and 7383 Pa at 40 °C. So, one can say that the average water vapor pressure in the Earth's atmosphere around us ranges from about a half to a few percent of the barometric pressure.

There are several ways to express humidity:

1. The *vapor pressure* is that part of the total pressure contributed by the water vapor.
2. The *absolute humidity* (or vapor concentration or vapor density) is the mass of water vapor per unit of volume. Effects of temperature and pressure are, except near saturation, according to gas laws.
3. The *relative humidity* is the ratio of the actual vapor pressure and the saturation vapor pressure at the prevailing temperature. It is usually expressed as a percentage. Since the maximum water vapor pressure depends on temperature, the relative humidity (r.h.) also depends, at given water content, on temperature. At constant temperature and given water content, the r.h. is, according to the equation for P mentioned earlier, dependent on total pressure.
4. The *dew-point temperature* is the temperature to which a gas must be cooled, at constant pressure, to achieve saturation. When the condensate is ice, it is called *frost point*. It is, with unchanging composition of the gas, independent of temperature. It changes with pressure since P_{H_2O} is proportional to P. Of course, condensation will occur if saturation vapor pressure is reached.
5. The *mixing ratio* is the mass of water vapor per unit mass of dry gas, usually expressed in grams per kilogram. If the ratio is related to a unit mass of humid air, it is called the *specific humidity*.
6. The *mole fraction* is the ratio of the number of moles of a component to the total number of moles present.

Conversions between the different parameters used to be cumbersome. They are now becoming standard since the introduction of the microprocessor.

Concentrations of water in a *liquid* or a *solid* are normally given in kg/kg. Except in soil physics, volumetric units are rarely used. The expression *equilibrium relative humidity* (e.r.h.) refers to a condition where there is no net exchange of water vapor between a moisture-containing material (paper, medicines, foodstuffs, tobacco, seeds, etc.) and its environment. It is the equivalent for *water activity*, a_w, used in the fields of biology or food technology, generally expressed as a ratio rather than a percentage (i.e., 0.6 instead of 60%).

It is probably difficult to find a material that is inert to water molecules and with which it would be impossible, with some physical method, to measure the presence of water. Water molecules change the length of organic materials, the conductivity and weight of hygroscopic materials and chemical absorbents, and in general the impedance of almost any material. Water absorbs infrared as well as ultraviolet radiation. It changes the color of chemicals, the refractive index of air and liquids, the velocity of sound in air or electromagnetic radiation in solids, and the thermal conductivity of gases as well as that of liquids and solids. More fundamentally, the water content can be measured by removing the water (vapor) from the sample and measuring the change of weight (or the change of pressure in a gas). Other fundamental principles are the evaporation from a water surface into the stream of sample gas (*psychrometer*) and the cooling of the gas sample until condensation is detected. Microwave absorbance, the measurement of capacitance, and nuclear magnetic resonance (NMR) have found application in the measurement of moisture in liquids and solids.

After an engineer has decided which parameter has to be measured, he has to realize a few things [5]:

1. What is the minimum range of operation required? Overspecification can be expensive. Besides, instruments suitable for drying processes at high temperatures and, at the same time, trace detection of water vapor in dry gases do not exist.

2. Unlike a temperature sensor, a humidity sensor can, at least in air, essentially not be shielded from its direct environment. The question of how to measure humidity cannot be separated from the measuring problem in question (contamination, condensation, etc.). A question might be Is there a danger of condensation before or after the period of real measurement?
3. The accuracy that may be expected in the field of hygrometry is considerably lower than that in other fields of measurement. One should be careful not to ask or to expect accuracies better than 2% or 3% r.h. or 0.5 °C in dew point. Before deciding to use a certain method or to buy a particular instrument, it is very useful to go once more through things like humidity as well as temperature and pressure range, possible contaminants in the process, and accuracy and response time really needed. After this, a consideration of frequency of service and calibration, cost of sensor replacement, etc. has to be made. At the end, it is sensible to find a supplier or manufacturer that is willing to think along with the customer. Table 80.1 provides a selective listing of the methods, ranges, and manufacturers for measuring humidity and moisture. The methods are discussed later.

TABLE 80.1 Methods for Measuring Humidity and Moisture

Method	g, l, s[a]	Range	Manufacturer[b]
Mechanical (hair)	g	0%–100% r.h.	Lambrecht, Thies, Haenni, Jumo, Sato, Casella, Pékly et Richard
Condensation dew point	g	−80°C/+100°C dp	General Eastern, Michell Instr., EG&G, E+H[3] MBW, Protimeter, Panametrics
Dry and wet bulb	g	10%–100% r.h.	Lambrecht, Thies, Haenni, ASL, Jenway, Casella, Ultrakust, IMAG-DLO
Lithium chloride	g	−45°C/+95°C dp	Honeywell, Jumo, Lee Engineering, Siemens, Philips, Weiss
Polymer (capacitance)	g, s	0%–100% r.h.	Vaisala, Rotronic, Testo, Hycal, Panametrics, Novasina, EE Elektronik, Chino, Lee Integer
Electrical (others)	g, s	0%–100% r.h.	PCRC, General Eastern, Rotronic, Chino, Elmwood, Shinyei Kaisha
Thermal conductivity	g	0–130 g/m^3	Shibaura Electronics Co. Ltd
Al_2O_3/silicon	g.1	−80°C/+20°C dp	E+H, Gen. Eastern, Panametrics, Michell Instr., MCM, Shaw
Phosphorous pentoxide	g	0.5–10,000 ppm	Anacon, Beckman, DuPont
Crystal oscillator	g	0.02–1000 ppm	DuPont
Infrared absorbance	g, l	0–50 ppm up to 65°C dp	Siemens, H&B, ADC, Anacon, Kent, Horiba, Sieger, Beckman, Li-Cor
Infrared reflectance	s	0.02%–100%	Anacon, Infrared Engin., Moisture Systems Corp., Pier Electronic, Zeltex, Bran & Luebbe, Bühler
NMR	l, s	0.05%–100%	Oxford Anal. Instr., Bruker
Neutron moderation	s	>0.5%	Kay Ray, Berthold, Nuclear Ent.
Microwave attenuation	s	0%–85%	Mütec, Scanpro, Kay Ray Inc., BFMRA
TDR	s	0%–100%	IMKO GmbH, Campbell Sci. Inc., Soil Moisture, Tektronix
FD	s	0%–100%	ΔT Devices, IMAG-DLO, VITEL, Troxler

Abbreviations: E+H, Endress+Hauser; MBW, MBW Elektronik AG; ASL, Automatic Systems Laboratories Ltd; PCRC, Physical and Chemical Research Corporation; MCM, Moisture Control and Measurement Ltd; H&B, Elsag Bailey Hartmann & Braun; ADC, Analytical Development Company; IMAG-DLO, Institute for Agricultural and Environmental Engineering.

[a] g, 1, s, gas/liquid/solid.

[b] This table is inevitably incomplete; a manufacturer not mentioned may deliver a high-quality instrument (see also Ref. [9]).

80.1 Gases

There are several methods of measuring humidity, the most important of which are described here. The scope of the present survey is limited. More can be found in the literature [6–10]. Reference [9] covers measurement in liquids and solids as well.

80.1.1 Gravimetric Method

The gravimetric method is the most fundamental way of accessing the amount of water vapor in a moist gas. In a gravimetric hygrometer, the water vapor is frozen out by a cold trap or absorbed by a chemical desiccant and weighed, while the volume or the mass of dry gas is measured directly. Since the result of a measurement gives the average value over an extended time, the instrument is used in combination with a humidity generator, capable of producing a gas of constant humidity. The method is used for primary standards of, among others, NIST in the United States, NPL in the United Kingdom, and NRLM in Japan. Achievable accuracies are approximately 0.1%–0.2% in mixing ratio, or 0.04 °C in the range of –35 °C to +50 °C dew point, increasing to 0.08 °C at +80 °C and 0.15 °C at –75 °C. The operation of such a standard requires high skill and sophisticated hardware.

80.1.2 Precision Humidity Generators

For less elaborate calibration work, a precision humidity generator is preferred. The best ones are reported to have accuracy comparable to that of a gravimetric standard [11]. Therefore, such a generator may be considered as a primary standard as well. Three practical methods to produce an atmosphere of known humidity are described by Hasegawa [12]: the two-flow, the two-temperature, and the two-pressure method. Briefly, in the first method, a test chamber is fed by two streams of air; one being dry, and the second one saturated with water at a known temperature. The resulting humidity can be calculated from the two flow rates. The two-temperature method uses air that has been saturated with water vapor at a well-known temperature, after which the air is heated to a higher temperature. In the two-pressure method, air is saturated with water vapor at an elevated pressure, and expanded isothermally to a lower, normally atmospheric pressure. Both temperature and pressure of the saturator and the test chamber are measured accurately. In general, precision generators are not transportable, so intercomparisons have to be made with a transfer standard of high accuracy. A good, if not the only, choice is a standard mirror dew-point meter.

80.1.3 Condensation Dew-Point Hygrometer

The saturation vapor pressure in air increases with temperature (Figure 80.1). This means that the air under test can be cooled to a temperature where it is just saturated with water vapor. If this is done at constant pressure and specific humidity, the true dew-point temperature is obtained. In practice, a sample of the gas is usually drawn over a thermoelectrically cooled metal mirror. The mirror is cooled until dew or frost is detected, by optical means. In some cases, where the mirror is replaced by an inert substrate, the formation of dew is detected by electrical means or the use of surface acoustic waves. The temperature is maintained such that the thickness of the deposit is neither increasing nor decreasing. The highest accuracy of a transfer standard can be expected to be 0.03 °C to 0.05 °C in the range of –20 °C to +40 °C dew point. Industrial optical condensation dew-point hygrometers claim accuracies up to 0.2 °C, which may be true in case of a clean mirror; in practice, 0.5 °C is often more realistic. Advantages of this principle are its fundamental nature and the wide range dew points to 90 °C under ambient temperature can be measured. One of the disadvantages is the susceptibility of the mirror to contaminants, especially soluble salts. The sensor may measure the dew point of another condensable vapor if its dew point is above that of the water. A good control of the mirror temperature requires a temperature difference between dew point and ambient temperature of the sensing head. This means that at high relative humidities, the gas must be heated and measured outside the process stream. At temperatures

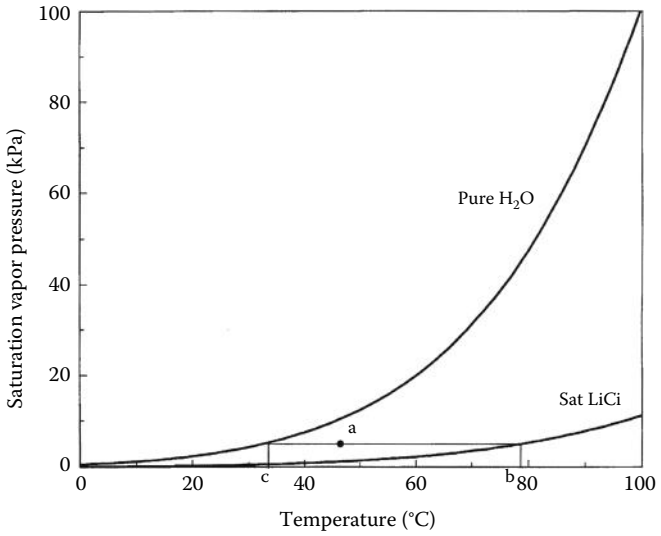


FIGURE 80.1 Saturation water vapor pressure above pure water and above a saturated solution of lithium chloride. In order to be in equilibrium with a gas sample at condition a, a saturated solution of LiCl has to be heated to temperature b. A free water surface must be cooled to a temperature b, as is the case in a mirror dew-point hygrometer.

below 0 °C, there may be supercooled water instead of ice on the mirror. Observations have shown that below –25 °C, the deposit will usually be ice. As the saturation vapor pressures over water and ice differ, one has to know the nature of the condensed layer; thus, for accurate measurements, a microscope is required. A rule of thumb is that the difference between water and ice on the mirror means a difference in dew point of one tenth of the temperature in degrees Celsius below zero, the dew point above water being the lower. Where low water content has to be measured, special attention should be given to the material and cleanliness of the pipes used. Stainless steel, polished at the internal surface, is to be preferred. The lower the moisture content, the more significant the effects are (Figure 80.2).

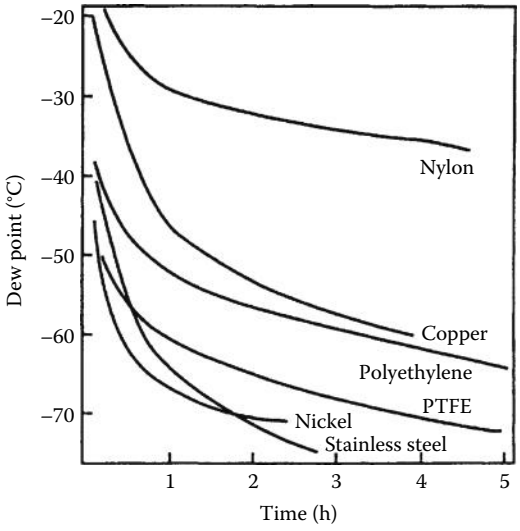


FIGURE 80.2 Illustration of the moisture given off by different tubing materials when flushed with very dry gas after being at ambient humidity. (From *A Guide to the Measurement of Humidity*, The Institute of Measurement and Control, London, U.K., 1996.)

80.1.4 Psychrometer

Two thermometers are ventilated with the gas of unknown humidity. One sensor, the dry bulb, measures the gas temperature t. The other sensor, the wet bulb, is surrounded by a wick saturated with pure water. The energy required to evaporate water into the air stream cools the wet bulb to a temperature t_w. The vapor pressure e in the sampled gas is calculated with the psychrometer equation:

$$e = e_w - A \cdot P\,(t - t_w) \tag{80.2}$$

where

e_w is the saturated vapor pressure at temperature t_w
P is the total atmospheric pressure
A is the psychrometer coefficient

A depends on ventilation speed, dimensions of the wet bulb, and radiative heat exchange between both dry and wet bulbs and their surroundings. The attractiveness of the psychrometric method lies in the fact that it is a direct and relatively simple method, with a theoretically strong basis. The accuracy of the method is determined by the accuracy of both dry and wet bulb sensors, the maintenance of a minimum ventilation speed, and a clean wick. With a ventilation speed of more than 3 m/s and a bulb diameter of 3–5 mm, A can, according to recent investigations, be assumed to be $(6.35 \pm 0.15) \times 10^{-4}\,°C^{-1}$, for Assmann-type psychrometers with a polished internal screen as well as sensors in a transverse air stream in a "black" radiation environment [13,14]. The radiation environment for axially ventilated psychrometers is of crucial importance. Especially when $t - t_w$ is measured directly the psychrometer is the preeminent instrument to measure near or at 100% r.h. The principle can be used up to a wet bulb temperature of 100 °C at atmospheric pressure; dry bulb temperature may exceed 100 °C (up to 165 °C, Ultrakust, Germany). Depending on the dry bulb temperature, the wick may have problems with water supply at strong evaporation. Frozen wet bulbs, already possible at ambient temperatures below 9 °C, can also lead to problems. Generally, this method adds water vapor to the atmosphere, which might be undesirable in specific applications.

80.1.5 Mechanical Hygrometers

Although mechanical hygrometers are losing ground, they are still widely used, mainly in room conditions. The principle relies on the elongation with r.h. of mainly human hair, textiles, or plastic fibers, the effect of which can be amplified mechanically to move a pen on a recorder. The best accuracy is 2%–3% r.h. (for hair, in the range between 35% and 95% r.h., if regenerated at regular intervals); in general, it is wise not to expect better than 5% r.h. In the case of hair, one must be aware of the fact that the hair may be in a state different from that during calibration. Hair exhibits a dry and a wet curve; the transition takes place below 35% r.h. Once at the dry curve, the instrument may read to 20% r.h. too high if it was calibrated at the wet curve as is usually done. If the instrument is kept overnight under a wet cloth, the wet curve calibration will be reestablished. The response time strongly depends on temperature: it ranges from a few minutes at 20 °C to 20 or 30 min at –10 °C. Temperature limits are –60 °C to +90 °C.

Although the initial costs of a mechanical hygrometer are low, the long-range costs of calibration and maintenance are considerable.

80.1.6 Lithium Chloride Dew-Point Meter

Addition of a hygroscopic soluble salt to pure water decreases the equilibrium saturation vapor pressure above the solution (*Raoult's law*). In Figure 80.1, this change is illustrated for a saturated lithium chloride solution, which has an equilibrium relative humidity of about 11% (see Table 80.2). A gas sample at condition *a* is in equilibrium with a saturated LiCl solution of a (higher) temperature *b*. (It is also in

TABLE 80.2 Equilibrium Relative Humidities (%) Over Some Saturated Salt Solutions

	Temperature (°C)			
	10	20	30	40
Salt	Relative Humidity (%)			
LiCl	11.3	11.3	11.3	11.2
$MgCl_2$	33.5	33.1	32.4	31.6
$Mg(NO_3)_2$	57.4	54.4	51.4	48.4
NaCl	75.7	75.5	75.1	74.7
KCl	86.8	85.1	83.6	82.3
K_2SO_4	98.2	97.6	97.0	96.4

equilibrium with pure water at a lower temperature *c*, note the similarity with a mirror dew-point meter.) The application of this principle led to a simple and effective sensor. A fabric sleeve over a bobbin with a bifilar winding of inert electrodes is coated with a dilute solution of lithium chloride. As the bobbin is heated by an alternating current, its resistance increases sharply at the point where the surface begins to dry out, the heating stops, the sensor begins to cool, attracts water vapor, etc., until an equilibrium temperature has been reached. This is measured by, for example, a resistance thermometer sensor in the bobbin. The sensor is simple, rugged, and relatively cheap and, after some calamity, it can be reactivated by recoating it. The dew-point range goes from −40 °C to +90 °C, with a claimed accuracy of about 0.5 °C. This estimate is often too optimistic because the influence of ambient temperature cannot be neglected. Moreover, LiCl has hydrates with one, two, three, and five molecules of water per molecule of salt. In the dew-point regions between 34 °C and 41 °C, the ambiguity leads to a possible error in that area of up to 1.5 °C in dew point. Below −12 °C, the error can rise to 3.5 °C. Low flow rates can cause stratification around the sensing surface, high rates can cool it too much, resulting in too high and too low readings, respectively. Flow rates somewhere between 0.05 and 1 m/s are generally recommended. Response times of commercially available sensors are on the order of minutes. Disadvantages include the lower limit lies at 11% relative humidity, the sensor is washed out by accidental immersion, and the power supply should not be turned off accidentally. From an ionic standpoint, the sensor can be considered precontaminated and therefore, according to the manufacturer [5], relatively insensitive to contamination. Another source [9] reports a number of gases, like sulfur vapors, ammonia, high concentrations of carbon dioxide, chlorine, hydrogen sulfide, and condensable hydrocarbons that could attack the lithium chloride sensor.

80.1.7 Electric Relative Humidity Sensors

Talking about numbers, the category of electric r.h. sensors is certainly by far the greatest part. The sensors are generally small, fast responding, do not dissipate heat, and can be used in confined spaces. Until 1975, the rights in this field were almost exclusively claimed by two resistance types: the Dunmore and the Pope sensor. The Dunmore sensor uses a dilute lithium chloride solution in a polyvinylacetate binder on an insulating substrate, with the danger of washout at saturation. The resistance of the sensor, measured between a bifilar grid, is a function of the r.h. of the surrounding air. This also applies to the Pope sensor, where a polystyrene substrate itself is the sensitive part of the sensor, after treatment with sulfuric acid. This sensor is less sensitive to washout, and has a wider range (15%–99% r.h.) than the Dunmore type. Modern bulk sensors, based on the measurement of the change in resistance, are manufactured by, for example, Shinyei Kaisha and Elmwood Sensors. Since 1970 a lot of work has been done in the development of capacitive sensors, where the stability, the influence of temperature, the temperature range, and the susceptibility for condensation conditions could be greatly improved. Since the 1980s, this sensor type has surpassed the mentioned resistance types. Basically, this type of sensor

consists of a thin polymer layer between two electrodes of various materials. The types differ mainly in polymer type and electrode material, resulting in sensors with different characteristics to withstand pollution and temperatures up to 190 °C. A drawback that still seems not completely under control is the fact that a polymer swells at high humidities, causing an undesirable shift in capacitance of the sensor.

80.1.8 Aluminum Oxide Hygrometers

Briefly, a sheet or wire of aluminum is anodized, producing a thin layer of water-sensitive pores. Subsequently, a conductive, water-permeable gold film is deposited over it. The radius of the pores is such that the sensor is specific for water molecules, and the amount absorbed is directly related to the dew point of the gas. The dew-point range goes from −110 °C to 20 °C, measured as a change in capacitance between the aluminum base and the gold electrode. The sensor works up to high pressures (30 MPa). The response time depends on the dew point, ranging from seconds at 10 °C to minutes below −40 °C. Stated accuracies are ±1 °C at higher dew points to ±3 °C at −100 °C. The sensors show slow drift, and recalibration at least twice a year is generally recommended. Where chemical attack of the aluminum can be expected, silicon capacitive type sensors may be an alternative. A commercially available type of MCM (Moisture Controls & Measurement Ltd., UK) is temperature-controlled, and has a very short response time: less than 15 s to a level of 1 ppm. Both aluminum and silicon sensors can withstand immersion in water.

80.1.9 Coulometric Method

In an electrolytic hygrometer, water in a sample stream is quantitatively absorbed by a phosphorous pentoxide layer. At the moment water molecules are present, the probe becomes electrically conducting. With a dc voltage over the sensor, water is electrolyzed and, according to Faraday's law, a current with a well-defined magnitude occurs (1 mA = 0.0935 μg H_2O/s). For the measurement of the mixing ratio, the flow through the cell must be measured accurately. Further calibration is, in principle, not necessary. At higher flow rates and moisture content (3000–10,000 ppm by volume), the water vapor may not be absorbed quantitatively, so this method is especially suited for low water contents, beginning at 1 ppm. The response time is about 1 min. Typical measurement uncertainties are 10% at 1 ppm, decreasing to 5% at higher values.

80.1.10 Crystal Oscillator

Even lower concentrations (down to 0.02 ppm by volume) can be detected with a hygroscopic coating on the surface of a quartz crystal. The resonant frequency of the crystal is a function of the mass of the coating—in other words, the moisture content of the gas. The crystal is alternately exposed to dry and humid gas, and the shift in frequency is measured. The response time is 1 min. Typical measurement uncertainty is 1 ppm or 5%. The instrument is relatively expensive (DuPont).

80.1.11 Infrared Methods

Like any heteroatomic gas, water vapor absorbs radiation in the infrared region. So, if the gas is led through the optical path between an infrared source and a detector, there is a reduction of the transmitted radiation. The source can be dispersive, that is, generated by a monochromator, or nondispersive, wideband radiation, generated by a heated tungsten or nichrome wire. In the monochromatic mode, the transmittance ratio at two different wavelengths is measured. In the nondispersive (NDIR) method, usually another path through a reference gas is taken. The detector can be gas-filled, solid-state, or pyroelectric, with adequate optical or gas filters in the right wave bands. A popular detector was and still is the Luft type. A Luft detector has two radiation-absorbing chambers containing the specific gas of interest, separated by a thin membrane. Since the detector is a dynamic device, the radiation paths are chopped, allowing the measurement of the change in capacitance. This method allows

measurement over a wide range, from ppm level up to saturation. It can be used in corrosive gases, the concentration of which can be measured simultaneously by the use of another detector. The instrument must be calibrated at regular intervals; typical measurement uncertainty is 1%. Response times to <1 s are possible, depending on the attainable refreshment time in the measuring chamber. Since the absorption depends on the number of atoms, transmittance depends on pressure. The method used to be relatively expensive, even compared with a mirror dew-point meter, but developments in the field of inexpensive semiconductor detectors and better optical filters have reduced prices. Further developments are going on in the direction of fiber-optic probes.

80.1.12 Miscellaneous Methods in Air

- The zirconia cell, acting as a battery on the presence of oxygen ions, is generally used in a mixture of air and steam. The sensor can be mounted directly in a hot gas stream (600 °C/1700 °C). Acceptable accuracies are reached at dew points of 70 °C or higher.
- The frequency of a signal, generated by an acoustic source, depends on the mixing ratio. Measurements in gases up to 250 °C, to dew points of about 70 °C (0–30 vol%, in special cases to 100 vol%, Mahlo).
- The Lyman alpha hygrometer uses the 121 nm emission line of hydrogen, which is strongly absorbed by water vapor. Extremely small response times (milliseconds) can be obtained. The instrument is relatively expensive.
- The difference in thermal conductivity of dry and wet air allows the measurement of absolute humidity with two sensitive temperature sensors in a Wheatstone bridge (Shibaura).

80.2 Liquids and Solids

80.2.1 Gravimetric

Drying of a material at a controlled temperature and taking the difference in weight before and after drying is the most fundamental method, greatly improved by microprocessor-based instruments. The assumption is that the loss in weight is caused by water only and that no other volatile components have been removed. Another problem can arise if water present by surface adsorption or if crystal water is removed.

80.2.2 Karl Fischer Method

A chemical method for the determination of water in solids and organic solvents is the Karl Fischer method. The Karl Fischer reagent is composed of iodine, sulfur dioxide, pyridine, and methanol. Addition of this reagent to water causes a chemical reaction in which, in excess of the other components, 1 mole of iodine is used for each mole of water. The reagent is added in a controlled way to a mixture of reagent and sample, while current between two electrodes is measured. If all the water has been used, a sudden change in the current through the mixture is observed.

80.2.3 Infrared Techniques

The infrared technique is applied to liquids and solids as well, in the wavelength bands of 1.45, 1.94, and 2.95 μm. For liquids, the transmission mode described is used, leading to particular problems because of the small optical path needed. Even smaller transmittance paths would be necessary for solids, one of the reasons to choose reflectance from the surface as a measure for water content. The surface has to be representative for the material in question. The system needs to be calibrated for each material. Concentrations from 0.02% to 100% can be measured. In case of specular reflection, the method cannot be used. A relatively new development is the attenuated total reflectance (ATR) crystal method. The crystal is inserted

in the liquid, which may be opaque or semiopaque, or slurry. A beam with two wavelengths, for reference and measurement entering the crystal, is reflected on the internal surface, penetrating the solution one-half wavelength at each reflection point. The amplitude of the reflected signal decreases at each reflection. The number of reflections is determined by the length of the crystal and the angle of incidence, so the sensitivity can be changed. The measuring range is claimed 0%–100% water in solutions; measuring water in emulsions requires some precautions. The response time is negligible.

80.2.4 Microwave Absorbance

Microwave absorbance is generally used in materials with a more or less constant composition, apart from the moisture content. Microwave radiation from a low-power, solid-state generator is absorbed by the sample and detected by a solid-state detector. The commonly used frequencies where water strongly absorbs are 1–2 and 9–10 GHz, the latter being less dependent on the composition of the material. Operating ranges from 1% to 70% of water are mentioned, with achievable accuracies of ±0.5% of water. The attenuation is influenced by bulk density, bulk material, and temperature. The path between source and detector should not contain any metallic material.

80.2.5 Nuclear Magnetic Resonance

Hydrogen atoms in the field of a permanent magnet are allowed, according to quantum mechanics, to have some defined orientations in that field. To shift an atom from one orientation to another requires a defined amount of energy, dependent on the strength of the magnetic field. If electromagnetic radiation at the right frequency is applied, resonance of the hydrogen atoms occurs, and a loss in frequency power can be detected. It is specific for all hydrogen atoms, so interference with liquids other than water in the sample can be expected. Temperature and flow must be controlled. Magnetic materials must be avoided. Measuring ranges of 0.05%–100% have been reported [9].

80.2.6 Neutron Moderation

This method is, like NMR, specific to hydrogen atoms. Neutrons of high energy are slowed by nuclei of hydrogen atoms. The main components are a detector of slow neutrons, next to a source of fast neutrons. The measuring range goes from 0.5%. The sensor can be made very rugged. The measured volume is a sphere of up to tens of centimeters in diameter. The method is dependent on the bulk density of the material, but largely independent of the properties of the material being analyzed [9]. The method is not suitable for foodstuffs. It suffers from necessarily severe government regulations and low acceptance by the public and the operator.

80.2.7 Time Domain Reflectometry

Time domain reflectometry (TDR) measures the propagation velocity of electrical pulses, mainly between 1 MHz and 1 GHz. This method is well established, especially for the measurement of water content in soil. With certain limitations, it is proven suitable for automatic installations. The main disadvantage is the complexity of the data analysis [15].

80.2.8 Frequency Domain

The general features of the frequency domain (FD) technique are comparable to those of TDR; however, there are some important differences. The dielectric properties of the material (soil, concrete, grain, oil, etc.) can be measured at a single frequency. The ability to choose a whole range of frequencies makes the FD sensor suitable for spectroscopic measurements. A sensor of this type has recently been developed by IMAG-DLO, Wageningen (The Netherlands). It is simple and inexpensive, and available for many purposes. The sensor uses a single application-specific integrated circuit (ASIC) and is suitable for the measurement of other properties of materials as well [15].

80.2.9 Measurement of Thermal Conductivity

The thermal conductivity of a material is related to the amount of water it contains. Heat pulses are supplied by a needle probe, and the cooling of the needle after ending the pulse is measured. It is a simple and inexpensive method that needs calibration for the material in which it is going to be used.

80.2.10 Water Activity or Equilibrium Relative Humidity

A material enclosed in a measuring chamber is, after some time, in equilibrium with its environment. The moisture content of the material can be derived from the so-called adsorption isotherms for that specific material, which must be determined experimentally. The method is used for many materials, such as foodstuffs, chemicals, grains, seeds, etc., with electric humidity sensors described earlier. In cases where the water potential of living material like potatoes and leaves must be measured, thus at very high relative humidities, the junction of a small thermocouple is cooled by an electric current and, after the current has been turned off, used as an unventilated wet bulb psychrometer. It is an excellent method that requires skill.

80.3 Formulae

A relatively simple equation for the calculation of the saturation vapor pressure $e_w(t)$ in the pure phase with respect to water is the *Magnus formula*:

$$\ln e_w(t) = \ln 6.112 + \frac{17.62t}{(243.12 + t)} \tag{80.3}$$

where

t is the temperature in °C (on ITS-90)
$e_w(t)$ is in pascals

Equation 80.3 covers the range between −45 °C and +60 °C, with a maximum standard deviation of 0.3%. Over ice, Equation 80.3 changes to

$$\ln e_i(t) = \ln 6.112 + \frac{22.46t}{(272.62 + t)} \tag{80.4}$$

covering the range between −65 °C and 0 °C with a standard deviation of <0.5%.

Equation 80.3 can easily be converted for the calculation of the dew point t_d:

$$t_d = \frac{[243.12 \ln(e/6.112)]}{[17.62 - \ln(e/6.112)]} \tag{80.5}$$

where e is the saturation vapor pressure e_w at dew-point temperature t_d. The standard deviation in the range mentioned is <0.02 K. The *frost point* t_f can be calculated from the following equation:

$$t_f = \frac{[272.62 \ln(e/6.112)]}{[22.46 - \ln(e/6.112)]} \tag{80.6}$$

where e is the saturation vapor pressure e_i at frost-point temperature t_f, in the range mentioned within a standard deviation of 4 mK.

More accurate formulae, given by Hyland and Wexler, updated by Sonntag can be found in Ref. [7]. The uncertainties mentioned earlier, given by Sonntag [7], are valid for the pure water system. In the presence of air, the uncertainties can only be maintained if the water vapor enhancement factor is taken into account.

80.4 Calibration

This section explains a few words about calibration, a most important and too often forgotten subject in humidity and moisture measurement. The gravimetric method has already been described, as well as some principles of humidity generators, which could be used as standards. Another method is the humidity chamber, with both temperature and humidity control, where an instrument is compared with a calibrated standard instrument like a mirror dew-point meter or a psychrometer. Saturated (or diluted) salt solutions can be used to create a constant relative humidity in a confined space. Some values are listed in Table 80.2 (from Ref. [16]). The user should have knowledge of some critical factors like temperature, etc. for the application of this method. Permeation tubes are used where a repeatable, accurate, and low concentration flow is required. Unlike with temperature or pressure sensors, it is, except at the highest national level, not realistic to expect a reference or standard for humidity with an accuracy of a factor 5 or 10 times better than the required accuracy for a measuring instrument. A good calibration guarantees traceability to a higher standard. This means, in general, that the calibration should be performed by a national or accredited laboratory. A rule of thumb for the intervals of calibration of humidity meters is as follows: condensation dew-point meters and psychrometers require calibration once every 1–2 years; electric relative humidity sensors need calibration every 6–12 months; and less stable types like aluminum oxide sensors must be calibrated every 6 months or sooner if desired [8]. A comprehensive treatment of calibration of hygrometers and attainable accuracies has been given by Wiederhold [10].

80.5 Developments

- Fiber optics is beginning to find their way in hygrometry, resulting in explosion-proof models. Changes in refractive index by the use of micropores are used in an instrument recently developed by Ultrakust (Germany). Fiber lengths go to 1000 m. The measuring range is 0% to 20% r.h. in a dew-point range of –90 °C to 20 °C.
- A rapidly developing field is that of the monolithic integrated circuit sensors. Hycal (United States) manufactured a capacitive relative humidity sensor, while a recent commercial development in Germany reports the use of a monolithic integrated sensor instead of a mirror in a miniature dew-point meter, with an accuracy of better than 0.5 °C. Since the use of porous silicon allows the development of smart sensors, developments in that area will undoubtedly continue [17].
- The progress in the theoretical background of many water-related processes, made possible by an advanced FD sensor mentioned earlier, is of great importance for future development of simple, cheap, stable, and reliable sensors that can be used for many purposes.

References

1. R. W. Hyland and A. Wexler, Formulations for the thermodynamic properties of the saturated phases of H_2O from 173.15 K to 473.15 K. *ASHRAE Trans*., 89(2A), 500–519, 1983.
2. A. Wexler, Vapor pressure formulation for ice. *J. Res. Natl. Bureau Standards*, 81A, 5–20, 1977.
3. Anonymous, *2009 ASHRAE Handbook Fundamentals SI Edition*. American Society of Heating, Refrigerating and Air Conditioning Engineers, Inc., Atlanta, GA, 2009, download http://www.gobookee.net/ashrae-handbook-fundamentals/, accessed on August 23, 2013.

4. L. Greenspan, Functional equations for the enhancement factors for CO_2-free moist air. *J. Natl. Bureau Standards*, 80A, 41–44, 1976.
5. Anonymous, Handbook, Selecting humidity sensors for industrial processes. General Eastern Instruments Corporation, 1982.
6. F. C. Quinn, The most common problems in moisture and humidity measurement control, *Proceedings of 1985 International Symposium Moisture and Humidity*, Washington, DC, April 15–18, pp. 1–5, 1985.
7. D. Sonntag, Advancements in the field of hygrometry. *Meteor. Zeitschrift, N.F.*, 3, 51–66, 1994 [Review].
8. Anonymous, *A Guide to the Measurement of Humidity*. The Institute of Measurement and Control, London, U.K., 1996.
9. K. Carr-Brion, *Moisture Sensors in Process Control*. Elsevier Applied Science Publishers, London, U.K., 1986.
10. P. R. Wiederhold, *Water Vapor Measurement*. Marcel Dekker, New York, 1997 (software included).
11. M. Stevens and S. A. Bell, The NPL standard humidity generator: An analysis of uncertainty by validation of individual component performance. *Meas. Sci. Technol.*, 3, 943–952, 1992.
12. S. Hasegawa, National basis of accuracy in humidity measurements. *Proceedings 1985 International Symposium Moisture and Humidity*, Washington, DC, pp. 15–19, April 15–18, 1985.
13. R. G. Wylie and T. Lalas, Detailed determination of the psychrometer coefficient for the wet cylinder in a transverse air stream and an analysis of its accuracy. Technical Paper No. 7 (CSIRO Division of Applied Physics, Sidney, Australia), 1981.
14. G. J. W. Visscher, Standard psychrometers: A matter of (p)references. *Meas. Sci. Technol.*, 5, 1451–1461, 1995.
15. M. A. Hilhorst and C. Dirksen, Dielectric water content sensors: Time domain versus frequency domain. *Proceedings Symposium TDR Environmental, Infrastructure and Mining Applications*, pp. 143–153, March 1994.
16. L. Greenspan, Humidity fixed points of binary saturated aqueous solutions. *J. Natl. Bureau Standards*, 81A(1), 89–96, 1977.
17. G. M. O'Halloran, M. Kuhl, P. J. Trimp, and P. J. French, A humidity sensitive capacitor based on a porous silicon dielectric. *Proceedings of 1996 National Sensor Conference*, Delft, The Netherlands, March 20–21, 1996.

81

Ventilation Measurement

L. Basano
Università di Genova

P. Ottonello
Università di Genova

81.1 Ventilation

This chapter deals mainly with *spirometry*, i.e., with the measurement of volumes and flows associated with respiration. Spirometric tests (which embody useful information about parameters related to pulmonary function) are often used for diagnostic purposes in conjunction with other measurements. Figure 81.1 and Table 81.1 show spirometric quantities of clinical interest [1].

The classification of spirometric instruments may be based on different criteria. For example, one may consider *open-circuit* devices (the subject takes a full inspiration or expiration before connecting to the meter) and *closed-circuit* devices (the subject remains connected to the apparatus during one or several respiratory cycles). Another distinction concerns whether the instrument is a *portable* one and is mainly used for monitoring purposes or is a *diagnostic* one whose purpose is to provide an accurate value that may be compared with a reference value. As for the nature of the physical functions more directly investigated, a subdivision into *volume measurements* and *flow measurements* is important, even though, using some care, it is sometimes possible to shift from one class to the other by time-differentiation or time-integration procedures. Another technique deals with the direct evaluation (as a function of frequency) of the complex impedance of the respiratory system viewed as a suitable network of resistors, capacitors, and inductors. This method has been the subject of considerable investigation in the last two decades, due to the advent of computers and their use in the spectral analysis of signals. This topic will be discussed in the final section.

As a general comment about measurements associated with ventilation, it may be said that there are few other fields in applied science where a correct *operation procedure* is virtually as important as the reliability of the measuring instruments themselves. See for example a study conducted on nearly 6000 patients affected by some airflow obstruction (Ref. [2] and references therein, especially n. 26).

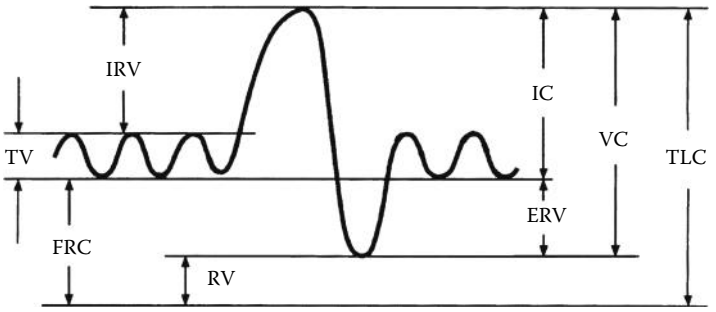


FIGURE 81.1 Spirometric trace illustrating the definitions of some significant quantities commonly evaluated in spirometry (see also Table 81.1).

TABLE 81.1 Common Spirometric Quantities

Vital capacity	VC
Inspiratory capacity	IC
Inspiratory reserve volume	IRV
Expiratory reserve volume	ERV
Tidal volume	TV
Functional residual capacity	FRC
Residual volume	RV
Total lung capacity	TLC

81.1.1 Volume Measurements

Here, gas volumes associated with the respiratory process are the main target of investigation; the principal instruments that have been used so far in the clinical routine and in research activity are [3]

1. *Spirometer*: An expandable chamber whose volume is monitored during inspiration or expiration. The subject is instructed to blow into a conduit communicating with the chamber; the latter may consist of a bell, a piston, or more often a bellows (as in the once ubiquitous Vitalometer®).
2. *Turbine meter*: Based on the principle that air blown through the inlet produces the rotation of a turbine connected to a revolution counter.
3. *Impedance plethysmograph*: Based on the measurement of resistance (*strain gage plethysmograph*) or of inductance (*inductance plethysmograph*); in both cases, the impedances of an elastic coil wrapped around the subject's chest and one wrapped around the subject's abdomen are monitored during respiration.
4. *Total-body plethysmograph*: A kind of sealed telephone booth inside which the subject sits; the pressure inside the box is sensed and converted to volume values.

81.1.2 Flow Measurements (Pneumotachometry)

In this case, the airflow through the upper airways is directly measured; in principle, the flow could be evaluated by time differentiating the spirometer records, but this method would suffer from limitations due to nonlinearity effects and to errors related both to poor frequency response and to some hysteresis of the volume meter. The reverse procedure is in fact more common: since they are normally connected to an electronic integrating device, flowmeters can be employed as volume-measuring instruments as well. Calibration is normally performed by discharging a syringe (of known capacity) through the flowmeter; the response of the latter is time integrated to check whether it corresponds to the calibration volume. It is convenient to repeat the calibration procedure a few times, with the

syringe discharged each time at a different speed, in order to detect possible effects due to deviations from linearity of the flow sensor.

A useful summary of standard specifications governing the performance of several types of pneumotachometers (PTM) together with a concise description of the more common PTM types can be found in Ref. [4]. Conventional devices for measuring respiratory flow are

1. *Linear resistance PTM* (LRPTM): Evaluates the pressure difference generated by the (laminar) airflow across a fixed hydrodynamic resistance. This procedure is based on the Poiseuille equation and is analogous to evaluating current by measuring voltage across a known resistor and using Ohm's law. The resistive element may consist either of a bundle of tubelets (Fleisch type) or of a wire mesh screen (Lilly type).
2. *Hot-wire PTM* (HWPTM): A very thin wire heated by an electric current, cooled by the flowing gas; the rate at which heat is conveyed away from the wire depends on the fluid flow; a variant of this instrument uses a heated film in lieu of the heated wire.
3. *Ultrasonic PTM* (UPTM): Based on the principle that the speed of a beam of ultrasounds exchanged between a pair of transducers is increased or decreased as it propagates through the moving air; the variations of the time of flight of the ultrasonic beam is related to the average speed of the flowing air.
4. *Vortex-shedding PTM* (VSPTM): Basically measures the frequency at which vortices are generated in the wake of a suitably shaped obstacle (the "bluff body") exposed to the flow.

81.1.3 Respiratory Impedance Measurements

In recent years, considerable interest has been devoted to the impedance of the respiratory system, a parameter that contains information about the morphology of our breathing apparatus and plays a role analogous to the impedance of an electrical circuit. *Respiratory impedance* may be obtained by connecting the patient's upper airways to an alternating pressure generator and by evaluating the ratio of the measured pressure to the measured flow; this can be done at several excitation frequencies, using a method that is referred to as the forced oscillation technique.

81.2 Instrumentation: Principles and Description

81.2.1 Measurements of Volume

81.2.1.1 Spirometer

A widely used spirometer is the bell type (see Figure 81.2): the volume of the lungs is monitored by the position of a light cylindrical bell (possibly equipped with counterweights) connected to the patient's mouth. The bell spirometer is rather cumbersome but, due to its great reliability, is an ideal tool for calibration purposes as well as for comparing data recorded at different centers. In its turn, the bell spirometer needs periodic calibration at least every 3 months; this may be performed by means of a 3 L syringe equipped with electronic volume readout [5]. The syringe must be accurate to within 25 mL and the spirometer should be able to measure volumes of at least 8 L with an accuracy of at least ±3% or 0.050 L, whichever is greater, with flows between zero and 14 L/s [6].

81.2.1.2 Turbine Meter

Its operating principle is quite straightforward. The revolutions of the turbine wheel, whose speed is proportional to the flow, are counted by electronic sensors (optoelectronics or Hall effect–based) and processed to give volume or flow values. The main limitation of the turbine meter is that low flow values are underestimated since a greater fraction of the air slips past the wheel as the flow rate decreases; it also displays poor frequency response and can be used only in unidirectional flows. Although these features

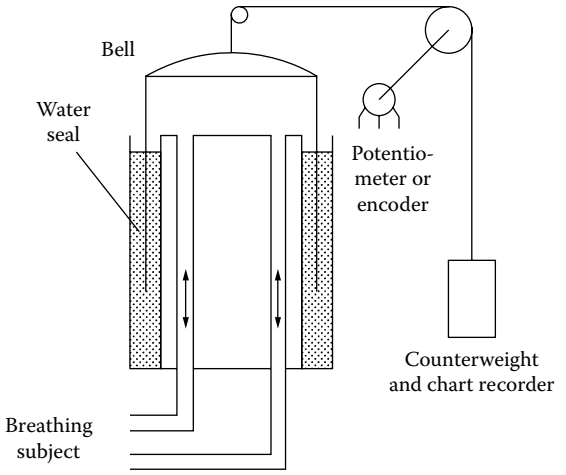


FIGURE 81.2 Schematic view of the classical bell spirometer.

make the device unsuitable for accurate laboratory measurements, it is used extensively to monitor the ventilation of patients in intensive care and in portable instruments, due to its overall (sensor and pulse processing) simplicity and low cost.

81.2.1.3 Inductance Plethysmograph

Like the spirometer, the PTM, and the total-body plethysmograph, this technique allows for noninvasive monitoring of human ventilation but it does not require a connection to the airways. Two elastic belts with serpentine inductors encircle the torso; the first around the rib cage and the other around the abdomen. Each inductor is part of the resonant tank circuit of a free-running oscillator (resonant frequency from 0.2 to 1 MHz) whose output is optically coupled to the demodulator circuit (for the purpose of electrical isolation of the subject from the equipment, usually operated by the ac power line). Ventilation causes the inductor cross-section areas to vary changing their inductances and thus varying the frequency of the two oscillators (FM modulation).

81.2.1.4 Strain Gage Plethysmograph

Its principle of operation is similar to that of the inductance plethysmograph; in this case, the sensing element consists of a strain gage whose resistance is varied as the gage is stretched and released during respiration.

81.2.1.5 Total-Body Plethysmograph (Constant-Volume Box)

This is a sealed chamber that allows the measurement of thoracic gas volume (TGV) by exploiting the pressure–volume relation of a fixed quantity of ideal gas. A schematic illustration of the method for a constant-volume plethysmograph is shown in Figure 81.3. The balloon-shaped object represents the subject's respiratory system, the volume of which varies during breathing. In the course of the measurement operation (which lasts only a few seconds), the shutter appearing in Figure 81.3 must remain closed in order that the alveolar pressure may be identified with the pressure measured at the mouth. In the following, the subscript A (for alveolar) denotes the gas in the lungs and the subscript C denotes the gas in the chamber.

For moderate respiratory acts, the air contained in the lungs at a given time can be treated as an approximately isothermal gas (temperature 37 °C, water vapor saturated) while the air in the chamber is more appropriately treated [1] as an adiabatic gas. By differentiating the ideal gas equation for the isothermal air in the lungs,

$$V_A = -P_A \frac{\Delta V_A}{\Delta P_A}$$

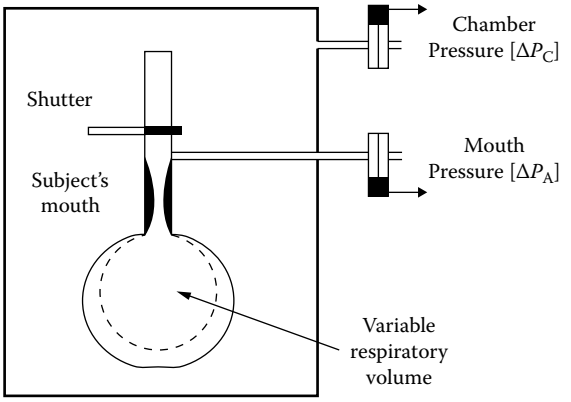


FIGURE 81.3 The respiratory system of a patient sitting inside an airtight chamber is schematized as an expandable balloon. The quantities of air inside and outside the balloon are separately constant: only their volumes and pressures change as the subject attempts to breathe against the closed shutter.

while for the adiabatic air in the chamber,

$$\Delta V_C = -V_C \frac{\Delta P_C}{\gamma P_C}$$

where γ is the usual ratio between molar heat capacities: $\gamma = c_p/c_v$.

Since the system is globally closed ($\Delta V_A = -\Delta V_C$), the lung volume V_A is related to the alveolar pressure P_A (which is taken to be the barometric pressure minus the water vapor pressure at 37 °C) and to the differential pressures ΔP_C and ΔP_A (equal to the chamber pressure fluctuation and to the alveolar pressure fluctuation, respectively) by the formula:

$$V_A = kP_A \frac{\Delta P_C}{\Delta P_A}$$

The coefficient k can be obtained through a calibration procedure that creates sinusoidal volume variations by means of a reciprocating pump. The pressure changes appearing on the right-hand side of the last equation may be measured by sending the outputs of the two sensors of Figure 81.3 to a data acquisition module installed in a host computer, typically a personal computer.

81.2.1.6 Correcting to Standard Conditions

Measurements of gas volumes make reference to lung BTPS (body temperature and pressure, saturated with water vapor) conditions. Volume measurements (V_A) made at ATP (ambient temperature and pressure) conditions should be corrected to BTPS conditions (V_B).

When bell or bellows spirometers are employed, the following formula [1] is often used:

$$V_B = \frac{310.2(P_B - P_W)}{(273.2 + t)(P_B - 6.3)} V_A \tag{81.1}$$

where

t is inside gas temperature (°C)
P_B is barometric pressure (kPa)
P_W is saturated water vapor pressure (kPa) of the gas

The temperature t of the air inside the spirometer should be measured accurately during each breathing maneuver (Table 3 in Ref. [7]).

81.2.2 Measurements of Flow

81.2.2.1 Linear Resistance PTM

The standard LRPTM contains the following elements (Figure 81.4a and b):

1. A fixed resistive load
2. A differential pressure sensor
3. Electronic instrumentation for processing and displaying the output of the pressure sensors

The relationship between flow rate and pressure difference should be linear within the range of useful flow rates; the maximum flow value, determined by the onset of turbulence, and the linearity range are given by the manufacturer's specifications.

It is good practice to calibrate the PTM periodically, by connecting it in series with a bell spirometer, and then use it without changing the geometry of the immediately adjacent tubing. In fact, accurate flow rate measurements are possible only if information about the flow pressure characteristics over the useful flow range is available; furthermore, this information should have been obtained in the same conditions in which the flowmeter will be used [8].

Figure 81.4a depicts the Fleisch PTM, whose resistive element consists of a bundle of capillary tubes. The function of the heater is to avoid the condensation of water vapor inside the tubes; the two chambers of the differential pressure sensor are connected at the two ends of the resistive element, whose resistance is lower than 0.1 kPa s/L (1 cm H_2O s/L). Since typical spirometric flows, for quiet breathing and slow maneuvers, are of the order of 1 L/s, they generate pressure differentials of about 1 cm H_2O. Considerably stronger flows, exceeding 10 L/s, may be produced by adults in some maneuvers; in these cases, Fleisch PTMs of larger dimensions should be employed in order to keep the spirometric flow within the linearity regime.

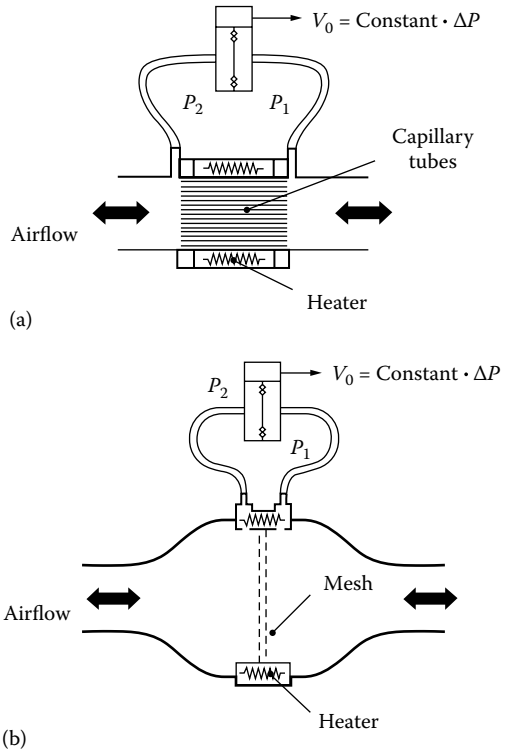


FIGURE 81.4 (a) The Fleisch (capillary) pneumotachograph and (b) the Lilly (wire-mesh screen) pneumotachograph. In both sensors, the task of the heater is to reduce water vapor condensation.

Figure 81.4b illustrates the Lilly PTM, whose resistive element consists of a wire mesh and is of the order of 0.05 kPa s/L (0.5 cm H_2O s/L) to which pressure differentials of the same order as the Fleisch are associated. The Lilly PTM is much less exposed than the Fleisch to the risk of nonlinearity at higher flow rates.

The linear resistance PTMs, in conjunction with an appropriate differential pressure transducer, an amplifier, and an analog or, more often, digital integrator, form the most widely used instrument for the combined measurement of flow and volume. In fact, recent recommendations on the standardization of lung function tests are limited to Fleisch and Lilly tachometers only [1].

81.2.2.2 Hot-Wire PTM

The HWPTM works on the principle that the cooling rate of a heated wire depends on the speed at which the surrounding fluid is flowing [9–11]. The rate at which thermal energy is lost by the wire may be written as

$$Q = hS(T_w - T) \tag{81.2}$$

where

Q is the heat transfer rate
S is the surface area of the wire
T_w is the wire mean temperature
T is the temperature of the fluid surrounding the wire
h is the transfer coefficient between sensor and fluid and is defined by Equation 81.2

The speed v of the fluid is linked to h by an important empirical relation whose analytical form has withstood the test of time; however, owing to the complexity of the formula and to the number of physical quantities involved in it, it is convenient to make direct reference to Ref. [11] where all necessary details can be found.

It is wise to remember, in fact, that from the operational point of view the last word on these instruments is always entrusted to a sound calibration procedure.

For spirometric purposes, it is recommended to use the HWPTM in the *constant-temperature* mode of operation, according to the diagram of Figure 81.5.

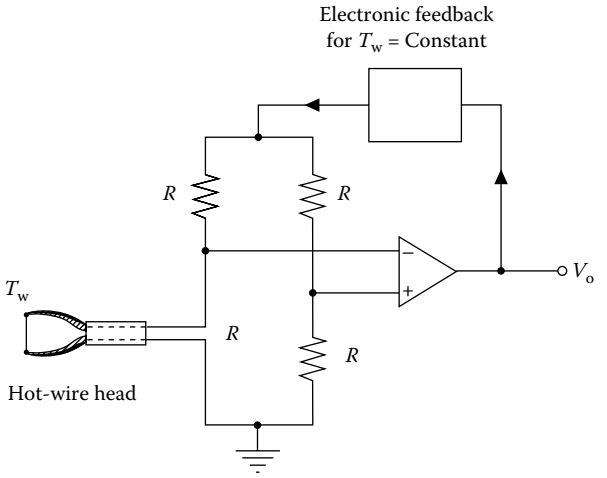


FIGURE 81.5 Block diagram of the constant-temperature hot-wire flowmeter. The electronic feedback system allows one to maintain the sensing wire temperature constant.

Using a suitable feedback system, the current is adjusted so as to maintain the wire temperature (hence, its resistance) at a constant value through the entire range of velocities allowed by the flowmeter. Once this condition is satisfied, the output voltage value is a measure of the heat transfer rate. The constant-temperature design has several advantages, e.g., the wire is protected against burnout and it is easy to compensate for environment temperature change either by using a temperature-sensitive element in the bridge or by measuring the actual temperature and correcting the results.

Accurate measurements require periodic calibration procedures (for example, using the syringe mentioned in connection with the spirometer) because dirt accumulating on the wire may significantly alter the heat transfer coefficient; it is easy, however, to remove dirt thoroughly from the sensor by manually superheating the wire for a few seconds or by immersing the sensor head into a suitable cleaning liquid. The calibration procedure should be repeated at intervals comparable to those of the LRPTM.

The setup in Figure 81.5 measures only unidirectional flows.

81.2.2.3 Hot-Film PTM

More recently, hot-film PTMs have been developed which can be used in place of the hot wire, the rest of the PTM remaining unchanged. The sensing element is deposited onto a nonconducting support (quartz) by vacuum sputtering, which ensures uniform thickness (about 0.1 μm) of the sensing element (platinum or nickel). In these sensors, heat can be conducted through the substrate and lost by convection to the ambient gas. The length-to-diameter ratio of the sensor is smaller; consequently, the temperature distribution along the sensor is less uniform. These drawbacks, which are more theoretical than real since calibration is always necessary, are largely compensated for by a lower fragility and sensitivity to particulate contamination.

81.2.2.4 Ultrasonic PTM

The basic working principles of UPTMs [10] are (1) sound is sped up or slowed down as it propagates through a moving medium and (2) the back and forth transit time of a sound signal is related to flow velocity and turns out to be largely independent of the acoustic velocity of the medium. UPTMs can be classified according to whether they use *pulsed* or *continuous* ultrasound signals.

In the first case, the transmitter is driven by a short pulse of sine waves; the round-trip transit time of individual pulses or sequences of pulses is measured.

In the second case, a continuous ultrasonic signal is transmitted along a closed path and either the phase shift or the frequency shift is measured. As a specific example, a pulsed transit time UPMT (Figure 81.6) is described.

Two piezoelectric crystal transducers are recessed into the wall of a conduit at an angle θ to the flow axis. Provisions are made in order to reduce the level of acoustical and electromagnetic

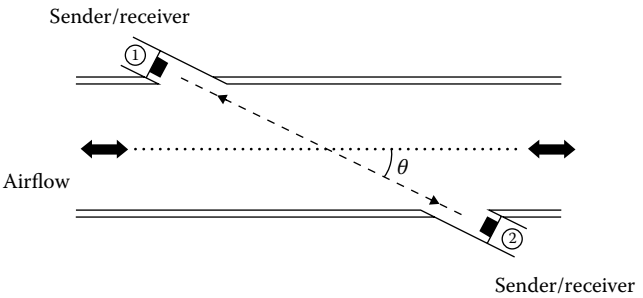


FIGURE 81.6 Transit time ultrasonic pneumotachograph. Owing to the large Q factor of the resonators, the central resonant frequencies of the two piezoelectric crystals should closely match.

external interferences. Moreover, moderate heating (at about 40 °C) prevents water from condensing on their surfaces. The two devices are made to function alternately as transmitter and receiver of a short burst of 50–200 kHz sound wave.

With no gas flowing through the conduit, the time required for sound transmission in either direction is the same. When gas is flowing, the two times are, respectively,

$$t_{12} = \frac{L}{(c + v \cdot \cos\theta)} \quad t_{21} = \frac{L}{(c - v \cdot \cos\theta)}$$

where

L is the transmitter–receiver distance

c is the speed of sound

The gas velocity v turns out to be independent of the actual value of c (which in turn depends on the gas type and on the working conditions and can range from around 200 to nearly 1000 m/s); an expression for v is easily obtained from the last equations by assuming $c^2 \gg v^2$

$$v \approx \frac{L}{2\cos\theta} \cdot \frac{\Delta t}{t_a^2} \tag{81.3}$$

where

$\Delta t = t_{21} - t_{12}$

$t_a = (t_{21} + t_{12})/2$

The flowmeter measures the average flow velocity of the gas along the path of the ultrasound beam.

In using the UPTM, it should be remembered that the working frequency is very close to the resonant frequencies of the two crystals which should be closely matched: small differences would cause a strong reduction of the instrument sensitivity owing to the sharpness of the resonance curves.

81.2.2.5 Vortex-Shedding PTM

When the Reynolds number is sufficiently high, vortices of regular periodicity are borne downstream of a bluff body located in the fluid flow (see Figure 81.7).

The frequency at which these eddies are generated is a linear function of flow velocity and depends weakly on the dimensionless Strouhal number [12,13]. Different types of VSPTMs use different methods for measuring the vortex frequency; to this end, ultrasonic beams and optical fibers are widely employed.

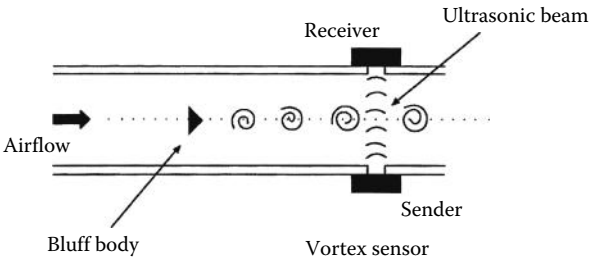


FIGURE 81.7 A vortex-shedding flowmeter equipped with an ultrasound vortex sensor. Vortices generated by the bluff body are sensed by the ultrasonic beam located downstream. In the optical model, the functions of vortex generation and counting are both taken on by an optical fiber.

Ultrasonic: An acoustic beam (Figure 81.7) is sent across the flow by an ultrasound transmitter located on one side of the tube (downstream of the bluff body); an ultrasound receiver is located on the opposite side of the tube. In this case, eddies generated by the bluff body act as modulators of the sound wave intensity. By doubling the setup [13], the ultrasonic VSPTM can also be used for bidirectional flows.

Optical fiber: In this model, an optical fiber is stretched across the tube and acts both as the bluff body and the vortex frequency sensor [14]. The light from an LED enters the fiber at one end, the other end being coupled to a photodiode. The intensity of the light propagating through the fiber is modulated by the mechanical vibrations of the fiber produced by the eddies moving with the fluid and fluctuates in step with the vortex generation frequency.

Both in the ultrasonic and the optical fiber VSPTM, the fluctuating intensity of the beam (respectively, sound and light) is processed to output a flow indication.

The main limitation of all types of VSPTMs is a "blind zone" at low fluid flows corresponding to the absence of vortices at low Reynolds numbers. In compensation, these devices are quite insensitive to the thermodynamic conditions of the gas and to the presence of particulate matter. Similarly to the case of HWPTMs, the last word on VSPTMs is entrusted to a sound calibration procedure.

81.2.2.6 Measurement of Volume Using a PTM Plus an Integrator

Another kind of plethysmograph can be built by combining a PTM and a digital integrator that calculates volume as follows:

$$V(\tau) = \int_0^{\tau} \dot{v}(t)dt \tag{81.4}$$

where $\dot{v}(t)$ is the instantaneous flow. The problems exhibited by this kind of instrument are the same as those of the PTM around which it is built, since the process of integration is invariably digital; moreover, correction to BTPS requires considerably more care [1]. Finally, when a heated PTM is employed, correction to BTPS is often done by assuming instantaneous thermalization of the gas within the tachometer (this means that the gas passing through the instrument assumes completely and without delay the same temperature of the tachometer [1]); this assumption may be quite unrealistic, especially in the case of the wire mesh screen type. The correction should thus be different for different types of PTMs.

81.2.3 Forced Oscillation Technique

The forced oscillation technique (FOT) is a modern tool [15] used to gain information on the structural and mechanical properties of the respiratory system by measuring how the latter responds to an externally imposed excitation. When linearity is assumed, well-known correlation techniques (borrowed from conventional signal analysis) can be exploited to measure the system complex impedance $Z_R(f)$, which is defined as the ratio of complex input pressure difference to complex output flow (i.e., the ratio of amplitudes and the difference of phases) and is equal to the inverse of the system transfer function $H_R(f)$.

Depending on the part of the respiratory system actually investigated, i.e., on the positions where the excitation pressure difference is applied and the flow is measured, different values for $Z_R(f)$ can be obtained:

- *Input impedance*: Both the excitation pressure (measured relatively to the pressure on the body) and flow are measured at the mouth.
- *Transfer impedance*: The excitation pressure is measured at the mouth, whereas flow is measured at the thorax (or vice versa) by means of a body plethysmograph.

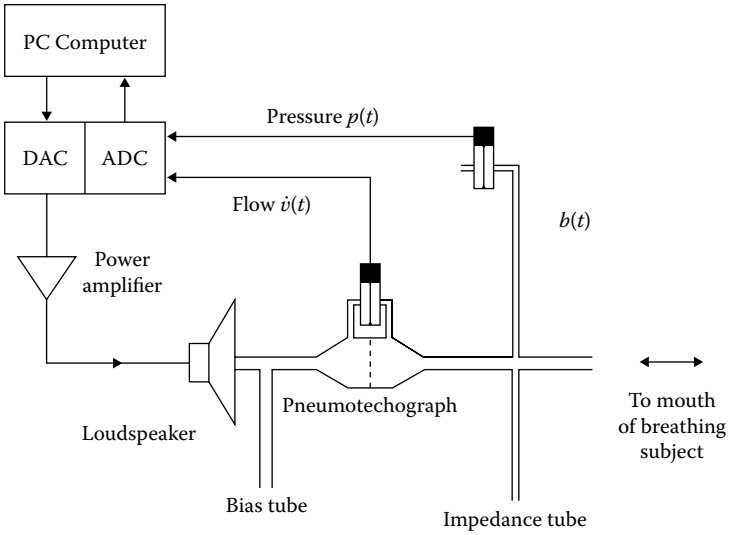


FIGURE 81.8 Experimental setup for the computer-based measurement of respiratory input impedance.

An experimental setup for the measurement of respiratory input impedance is shown in Figure 81.8. A suitable software provides

- Generation of a pseudorandom sequence (typically obtained by summing the first 25 harmonics of 2 Hz with random phases) which, after digital-to-analog conversion and proper power amplification, drives a large (20–30 cm diameter) loudspeaker
- Acquisition of two blocks of pressure and flow samples: 2 × 4096 data are collected in a 16 s run at a sampling rate of 256 samples/s
- Processing of data in order to obtain the desired $Z_R(f)$ function in the typical interval 2–48 Hz

The bias tube allows a flow of fresh air to be drawn through the system to minimize rebreathing of expired air. The impedance tube (2 m long, 2 cm diameter) allows spontaneous breathing; moreover, because its electrical equivalent is an inductor, it acts like a shunt for the low-frequency breathing components, thus reducing their level of disturbance on the pressure and flow measurements (noise), without affecting in a significant way the imposed pressure input (signal).

The respiratory apparatus is modeled as a system whose input is a random noise pressure $p(t)$ and whose output is a flow signal $\dot{v}(t)$, both corrupted by a breathing disturbance $b(t)$ uncorrelated to $p(t)$. Any text on random signal analysis [16] may provide useful information on the mathematical quantities employed in this section.

Using the following Fourier transform (FT) quantities,

- $P(f)$ = FT of the random noise pressure input $p(t)$
- $B(f)$ = FT of the subject's breathing noise $b(t)$

$$\dot{V}(f) = \text{FT of the linear system output } \dot{v}(t)$$

we have (Figure 81.9)

$$\dot{V}(f) = P(f) \cdot \frac{1}{Z_R(f)} + B(f) \tag{81.5}$$

A valid estimate of $Z_R(f)$ cannot be obtained from this relation because of the presence of the large breathing disturbance. According to the theory of signal processing, a nonbiased estimate of $Z_R(f)$ can be obtained by using autospectra and cross spectra of input [$p(t)$] and output [$\dot{v}(t)$] signals; if $G_{PP}(f)$ and

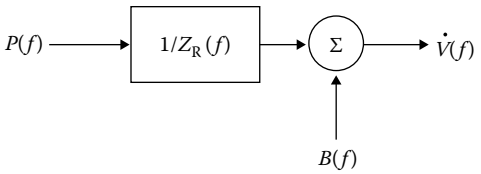


FIGURE 81.9 The respiratory system modeled in the frequency domain.

$G_{\dot{V}\dot{V}}(f)$ are the autopower spectra of pressure and flow, and $G_{\dot{V}P}(f)$ their cross power spectrum, the estimated impedance $Z_M(f)$ becomes

$$Z_M(f) = \frac{G_{PP}(f)}{G_{\dot{V}P}(f)} = \frac{G_{PP}(f)}{1/Z_R(f) \cdot G_{PP}(f) + G_{BP}(f)} \tag{81.6}$$

which becomes almost coincident with the true value $Z_R(f)$ when a long time average allows sensible reduction of the contribution of the $G_{BP}(f)$ term.

Moreover, since the systematic bias depends mainly on the presence of the breathing noise on both pressure and flow signals, all efforts should be made to reduce this disturbance as early as possible, i.e., at the sensor level. This is achieved by using the impedance tube and by keeping the value of the pneumotachograph impedance as low as possible. Further reduction is usually obtained by high-pass analog filtering (2 Hz corner frequency) before the ADC and/or by digital filtering before FFT calculations. It is worth noting that the choice of a pseudorandom excitation makes possible the use of a digital comb filter (a multiple narrow band pass, i.e., a band pass for each 2 Hz harmonic) [17].

When the FOT is used, it is necessary to ascertain that the coherence [16] between input and output, expressed by the γ^2 function defined as follows, is sufficiently large:

$$\gamma^2(f) = \frac{|G_{\dot{V}P}(f)|^2}{G_{\dot{V}\dot{V}}(f) \cdot G_{PP}(f)} \tag{81.7}$$

As a rule of the thumb, the calculated $Z_R(f_0)$ (the impedance for a particular value of frequency) is accepted only when the corresponding $\gamma(f_0)$ is larger than 0.9–0.95.

Actually, the value of $\gamma(f)$ is the only form of control used; at any rate, some care should be taken to interpret the coherence function as an index for the reliability of respiratory impedance data [18,19]. Obviously, higher values of the coherence function can be achieved by increasing the pressure input amplitude, but this may lead to an increase of nonlinearity effects [20].

In conventional setups, flow is evaluated by measuring the pressure drop ΔP that develops across the pneumotachograph, whose impedance is known and usually small with respect to the subject's input impedance Z_R. Since these two impedances are in series, the differential pressure ΔP measured by the pressure sensor is small compared with the pressure P applied to both sides of the transducer. Accurate measurements then require the use of highly symmetric sensors [21], where symmetry is expressed by the value of the common-mode rejection ratio, CMRR = 20 log$P/\Delta P$. In addition to this, however, other sources of error (such as the finite impedance of pressure sensors [22]) are usually present which render the correction procedure ineffective.

All these problems can be globally overcome by the following dynamic calibration procedure [23] that requires two simple preliminary measurements and the availability of a reference impedance Z_{REF}.

If Z_M^{∞} denotes the impedance measured when the measuring system is occluded, ($Z_R = \infty$), Z_M^{REF} denotes the impedance measured when a known reference impedance is used ($Z_R = Z_{REF}$) and Z_M denotes the impedance measured when the subject is connected to the measuring device, then the subject's corrected impedance is given by

$$Z_R = Z_{REF} \frac{1/Z_M^{REF} - 1/Z_M^{\infty}}{1/Z_M - 1/Z_M^{\infty}} \tag{81.8}$$

TABLE 81.2 Complex Values of Calibration Impedance

	Impedance (10^5 Pa s/m^3)		
Frequency (Hz)	Real	Imaginary	Modulus
0	3.09	0.00	3.09
5	3.11	1.11	3.30
10	3.14	2.17	3.82
15	3.20	3.22	4.54
20	3.28	4.26	5.37
25	3.37	5.28	6.26
30	3.47	6.29	7.18
35	3.58	7.28	8.12
40	3.70	8.26	9.05
45	3.82	9.23	9.98
50	3.94	10.17	10.91

Z_M^{∞} and Z_M^{REF} need not be evaluated each time a new subject is connected to the device, but only when some physical change (e.g., length variations of connecting tubes, replacements of sensors, etc.) has occurred since the last time the instrument was used.

81.2.3.1 Reference Impedance

The availability of a reference impedance is recommended [24]

1. To correct the measured impedance using the procedure described earlier
2. To compare measurements obtained by different devices, different techniques, and/or different groups

For these purposes, a compact calibrator [25] has been proposed which displays the following features:

- It is simply reproducible.
- Its impedance value is not too far from typical human values.
- It is tractable from the mathematical point of view.
- Its lowest resonances fall outside the usual working frequency range.

The calibrator consists of a bundle of 30 tubelets (each 20.0 cm long and of 2.00 mm inside diameter) and its complex impedance is given in Table 81.2.

81.3 Future Perspectives

Measurements in ventilation strongly depend on the availability of reliable sensors and detectors (pressure, temperature, optical and ultrasonic beams, etc.). The use of these components has been rising steadily in the last decade and this trend is expected to continue in the near future. The foreseeable consequences of this trend on the development of dedicated instrumentation may be

1. The increasing use of transducers of various kinds in all spirometric measurements
2. The universal use of personal computers (and specialized software) to process and display data
3. The growing importance of calibration procedures, because the accurate response of a particular transducer is usually reproducible but not always theoretically predictable; on the other hand, this is fundamental to all measurement systems

Defining Terms

Plethysmography: Method employed to evaluate volume variations, with particular reference to human organs.

Pneumotachography: Methodology for measuring gas flows.

Respiratory impedance: Complex ratio between input pressure difference and induced flow at a particular frequency, when the respiratory apparatus is viewed as a complex network of linear components (resistive, inductive, capacitive).

Spirometry: The measurement of volume changes of the ventilatory system (lungs and chest wall) usually inferred from the movement of gas to and from the system.

Ventilation: Physical interaction of a body of air with the respiratory system, either through spontaneous or mechanically assisted breathing.

References

1. P. H. Quanjer, G. J. Tammeling, J. E. Cotes, O. F. Pedersen, R. Peslin, and J. C. Yernault, Lung volumes and forced ventilatory flows, *Eur. Resp. J.*, 6: 5–40, 1993.
2. R. A. Wise, J. Connett, K. Kurnow, J. Grill, L. Johnson, R. Kanner, P. Enright, and the Lung Health Study Group, Selection of spirometric measurements in a clinical trial, the Lung Health Study, *Am. J. Respir. Crit. Care Med.*, 151: 675–681, 1995.
3. T. F. McAinsh, Ed., *Physics in Medicine & Biology Encyclopedia*, Oxford, U.K.: Pergamon Press, 1986.
4. D. I. Plaut and J. G. Webster, Ultrasonic measurement of respiratory flow, *IEEE Trans. Biomed. Eng.*, BME-27: 549–558, 1980.
5. W. S. Linn, J. C. Solomon, H. Gong, Jr., E. L. Avol, W. C. Navidi, and J. M. Peters, Standardization of multiple spirometers at widely separated times and places, *Am. J. Respir. Crit. Care Med.*, 153: 1309–1313, 1996.
6. American Thoracic Society, Standardization of Spirometry 1994 Update, *Am. J. Respir. Crit. Care Med.*, 152: 1107–1136, 1995.
7. L. R. Johnson, P. L. Enright, H. T. Voelker, and D. P. Tashkin, Volume spirometers need automated internal temperature sensors, *Am. J. Respir. Crit. Care Med.*, 150: 1575–1580, 1994.
8. K. E. Finucane, B. A. Egan, and S. V. Dawson, Linearity and frequency response of pneumotachometers, *J. Appl. Physiol.*, 32: 121–126, 1972.
9. N. H. Cook and E. Rabinowicz, *Physical Measurement and Analysis*, Reading, MA: Addison-Wesley, 1963.
10. L. Marton and C. Marton, Eds., *Methods of Experimental Physics. Fluid Dynamics*, Vol. 18, New York: Academic Press, 1981.
11. L. M. Fingerson, Thermal anemometry, current state, and future directions, *Rev. Sci. Instrum.*, 65: 285–300, 1994.
12. D. F. White, A. E. Rodely, and C. L. McMurtrie, The vortex shedding flowmeter, in R. B. Dowdell, Ed., *Flow—Its Measurement and Control in Science and Industry*, Pittsburgh, PA: Instrument Society of America, pp. 967–974, 1974.
13. Ch. Buess and W. Guggenbuhl, Ultrasonic airflow meters in medical application, *Proc. Annu. Conf. IEEE Eng. Med. Biol. Soc.*, 13–16, 1987.
14. S. Webster, R. McBride, J. S. Barton, and J. D. C. Jones, Air flow measurement by vortex shedding from multimode and monomode optical fibers, *Meas. Sci. Technol.*, 3: 210–216, 1992.
15. E. D. Michaelson, E. D. Grassman, and W. R. Peters, Pulmonary mechanics by spectral analysis of forced random noise, *J. Clin. Invest.*, 56: 1210–1230, 1975.
16. J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, New York: Wiley & Sons, 1986.

17. R. Farré and M. Rotger, Filtering the noises due to breathing in respiratory impedance measurements, *Eur. Respir. Rev*., 1: 196–201, 1991.
18. E. Oostveen and A. Zwart, Reliability of the coherence function for rejecting respiratory impedance data, *Eur. Respir. Rev*., 1: 218–221, 1991.
19. H. Franken, J. Clément, and K. P. van de Woestijne, Systematic and random errors in the determination of respiratory impedance by means of the forced oscillation technique: A theoretical study, *IEEE Trans. Biomed. Eng*., BME-30: 642–651, 1983.
20. M. Rotger, R. Peslin, R. Farré, and C. Duvivier, Influence of amplitude, phases and frequency content of pseudorandom pressure input on impedance data and their variability, *Eur. Respir. Rev*., 1: 178–182, 1991.
21. R. Peslin, P. Jardin, C. Duvivier, and P. Begin, In-phase rejection requirements for measuring respiratory input impedance, *J. Appl. Physiol: Respir. Environ. Exercise Physiol*., 56: 804–809, 1984.
22. R. Farré, R. Peslin, D. Navajas, C. Gallina, and B. Suki, Analysis of the dynamic characteristics of pressure transducers for studying respiratory mechanics at high frequencies, *Med. Biol. Eng. Comput*., 27: 531–537, 1989.
23. R. Farré, D. Navajas, R. Peslin, M. Rotger, and C. Duvivier, A correction procedure for the asymmetry of differential pressure transducers in respiratory impedance measurements, *IEEE Trans. Biomed. Eng*., BME-36: 1137–1140, 1989.
24. M. Cauberghs and K. P. van de Woestijne, Calibration procedure of the forced oscillation technique, *Eur. Resp. Rev*., 1: 158–162, 1991.
25. L. Basano and P. Ottonello, A calibrator for measurements of respiratory impedance, *Meas. Sci. Technol*., 6: 982–987, 1995.

82

Meteorological Measurement

John D. Garrison
San Diego State University

Stephen B.W. Roeder
San Diego State University

82.1 Introduction

Meteorological measurements are measurements of the physical properties of the atmosphere. These measurements are made at all elevations in the troposphere and the stratosphere. For measurements made at elevations above ground or tower level, instruments can be carried aloft by balloons, rockets, or airplanes. Ground radar is used to detect the presence of water in the form of droplets or ice crystals at all elevations and the winds associated with them. Lidar (optical radar) of selected wavelengths is used to detect the presence and amount of aerosols and other constituents of the atmosphere and to determine cloud height. Instruments on satellites measure properties of the atmosphere at all elevations.

Quantities measured are temperature, pressure, humidity, wind speed, wind direction, visibility, the presence and amount of precipitation, cloud amount, cloud opacity, cloud type, cloud height, broadband solar (or shortwave) radiation, long-wave radiation, ultraviolet radiation, and net radiation, sunshine duration, turbidity, and the amounts of trace gases such as NO, NO_2, SO_2, and O_3. Some of the methods and instruments used to measure a number of these variables are discussed in this chapter in the sections on pressure, temperature, humidity and moisture content, and air pollution monitoring. Additional information of interest is in the sections on resistive sensors, inductive sensors, capacitive sensors, satellite navigation, and radiolocation, and the sections under signal processing.

Meteorological measurements are made at individual sites, at several or many sites forming a local network, or at much larger networks. Much of the emphasis now is on global networks covering the entire northern and southern hemispheres. Individuals and groups can make measurements for their own purposes or they can use data provided by the various weather services. Weather service data are stored in archives that can cover many years of measurements. The U.S. National Climatic Data Center, Asheville, NC, has archived data produced from measurements at U.S. weather stations and other National Oceanic and Atmospheric Administration (NOAA) measurement sources, including satellites. It also has non-U.S. data. These data can be purchased. The website where information concerning NOAA data can be obtained is (http://www.ncdc.noaa.gov/). The U.S. National Renewable Energy

Laboratory has information on solar radiation and wind at website (http://www.rredc.nrel.gov/). Data at websites can often be retrieved by anonymous File Transfer Protocol (FTP).

Fabrication of meteorological instruments is usually done by companies specializing in these instruments. It is generally not economically feasible for individuals to fabricate their own instruments, unless their particular application cannot use commercially available instruments. The problem usually reduces to the determination of which commercial instruments to purchase. This determination depends on cost, durability, accuracy, maintenance requirements, ease of use, and the form of the output signal. A number of instrument manufacturers and distributors market complete weather stations.

The instruments used for meteorological measurements are fabricated for use in the special environment of the atmosphere. This environment varies with the latitude and longitude of the site and the elevation above the ground and above sea level. A common requirement is that the instruments be protected from adverse conditions that can cause errors in the measurement of the meteorological variables. For proper operation, some instruments (e.g., spectroradiometers) need to be in a temperature-controlled environment. The sun's heating can cause errors. Solar radiation can cause weathering of the instruments and shorten their useful life. Moisture from precipitation or dew can affect measurements adversely and also cause weathering and corrosion of the instruments. Blowing dust or sand can cause weathering of the instruments and affect the operation of mechanical parts. Insects, birds, and ice can also affect instruments adversely. Some gaseous constituents of the atmosphere can be corrosive. Packaging of the sensors and housing of the instruments in enclosures can protect them, but this protection must not interfere with the measurement of the meteorological variable. Packaging a sensor and putting the sensor in protective housing generally increases the response time of the sensor to changes in the meteorological variable it is measuring. Solar heating can be reduced by covering the sensor, its protective packaging, or housing with a white coating. Generally, housing or enclosures used to protect the instruments should be well ventilated. Sometimes, a fan is used to draw air through the housing or enclosure to reduce solar heating and make the air more representative of the outside air. Loss of measurement caused by loss of electric power can be avoided by having backup power from batteries or motor generators.

Another common requirement of meteorological measuring instruments is that they be calibrated before installation in the field. This is usually done by the manufacturer. For some applications, it may be important that the calibrations are traceable to National Institute of Standards and Technology (NIST) standards. The calibration should be checked routinely. Some instruments are constructed to be self-calibrating.

Generally, there is no reason to measure many of the meteorological variables to high precision. The temperature, humidity, and wind, for example, can vary in relatively short distances and sometimes in relatively short times by amounts that are large compared to the accuracy of the measuring instruments. This is especially true near ground level. The exact value of a meteorological variable at a particular site has little meaning unless the density of measuring sites is very high. A mesoscale network with high density of measuring sites over a relatively small area might be of interest for ecological studies. Larger-scale or global networks have widely separated sites. For these networks, it is important to install all of the instruments in a standard manner so as to reduce the effect of local fluctuations on the variation of the meteorological variables from site to site. Usually, this is done by placing them on a level, open area away from surrounding buildings and other obstructions and at a fixed distance above the ground. The ground should be drained ground and not easily heated to high temperature by the sun. In comparing temperatures, one must be aware of the "heat island effect" of large cities; furthermore, an increase in the degree of urbanization of a given site over time may affect the interpretation of the temperature trends observed.

The current trend is toward automatic measurement of meteorological variables at unattended sites with automatic data retrieval and computer processing and analysis of the data. Large-scale networks consisting of many stations covering a large area (e.g., the Northern Hemisphere) are used for regional, national, and global weather forecasting.

82.2 Measurement of the Atmospheric Variables

82.2.1 Temperature

The mean temperature of the atmosphere for each hour of the day at a particular site has a fairly regular annual and a diurnal variation when this mean temperature is an average over many years for each hour and day of the year. The temperature at a given site and time is a superposition of the mean temperature and the fluctuations from this mean temperature caused by current cloud and wind conditions, the past history of the air mass passing over the site, and interannual variations that are not yet well understood.

More detailed sources of information on temperature measurement include Refs. [1–5]. Common methods of measuring atmospheric temperature include electric resistance thermometer, thermistors, bimetallic strips, and liquid in glass thermometer.

82.2.1.1 Electric Resistance Thermometer

The variation of the resistance of a metal with temperature is used to cause a variation in the current passing through the resistance or the voltage across it. The electric circuit used for the measurement of temperature can utilize a constant current source, and temperature is determined from the voltage across the resistance after the circuit is calibrated. Alternatively, a constant voltage source can be used with the current through the resistance determining the temperature. Once the instrument is calibrated, it is generally expected to keep its calibration as long as electric power is supplied to the instrument. Commonly, platinum electric resistance thermometer (RTD) thermometers are made of a fixed length of fine platinum wire or a thin platinum film on an insulating substrate. The variation of the resistance as a function of temperature is approximately linear over the range of temperature found in meteorological measurements. The quadratic correction term is quite small. The accuracy and reproducibility of the measurements and the ease of using an electric signal for transmission of data from remote unmanned sites make electric resistance thermometers desirable for meteorological applications. Platinum is the best metal to use. With careful calibration and good circuit design, platinum resistance thermometers can measure temperature to a small fraction of a degree, much better than the accuracy needed for meteorological measurements.

82.2.1.2 Thermistors

Thermistors usually consist of an inexpensive mixture of oxides of the transition metals. The log of their resistance varies inversely with temperature. The change in resistance with temperature can be 10^3–10^6 times that of a platinum resistance thermometer. Their change in resistance with temperature is used to determine temperature in the same manner as metal resistance thermometers. They are somewhat lower in cost and are somewhat less stable than platinum resistance thermometers.

82.2.1.3 Bimetallic Strips

Bimetallic strips are normally used for casual monitoring of inside and outside temperatures at dwellings and office buildings and for heating and cooling controls. The accuracy is generally about ±1°C. They are low in cost.

82.2.1.4 Liquid in Glass Thermometer

These are a well-known method of measuring temperature. They are normally used for casual monitoring of inside and outside temperatures at dwellings or office buildings. These thermometers are more difficult to read than meter or dial readings of temperature, do not lend themselves to electric transmission of their readings, and are easily broken. Their cost can be low.

82.2.2 Pressure

One standard atmosphere of pressure corresponds to 1.01325×10^5 Pa (N/m^2) (14.6960 pounds per square inch, 1.01325 bars, 1013.25 mbars, 760.00 mm Hg, or 29.920 in. Hg). This is approximately

the mean atmospheric pressure at sea level. Atmospheric pressure at sea level usually does not deviate more than ±5% from one standard atmosphere. Atmospheric pressure decreases with altitude. Altitude measurements in airplanes are based on air pressure measuring instruments called altimeters. At about 5500 m (18,000 ft), the atmospheric pressure is half its sea level value. The following instruments are used to measure atmospheric pressure.

82.2.2.1 Mercury Manometer

Originally, barometric pressure was measured with a mercury manometer. This is a tube, about 1 m in length, filled with mercury and inverted into an open dish of mercury. The height of the column of mercury that the external pressure maintains in the tube is a measure of the external air pressure. Hence, one standard atmosphere is 760 mm Hg. While accurate, this device is awkward and has been replaced for general use.

82.2.2.2 Aneroid Barometer

It consists of a partially evacuated chamber that can expand or contract in response to changing external pressure. The evacuated chamber is often a series of bellows, so that the expansion and contraction occurs in one dimension. Basic aneroid barometers, which are still in use, have a mechanical linkage to a pointer giving a reading on a dial calibrated to read air pressure. High-quality mechanical barometers can achieve an accuracy of 0.1% of full scale. Aneroid barometers can also give electronic readout and eliminate the mechanical linkage; this is more the standard for serious meteorological measurements. In one method, a magnet attached to the free end of the bellows is in proximity to a Hall effect probe. The Hall probe output is proportional to the distance between the magnet and the Hall probe.

Barometric pressure is also measured with an aneroid type of device that consists of a rigid cylindrical chamber with a flexible diaphragm at its end. A capacitor is created by mounting one fixed plate close to the diaphragm and a second plate mounted on the diaphragm. As the diaphragm expands or contracts, the capacitance changes. Calibration determines the pressure associated with each value of capacitance. A range of 800 to 1060 mbars with an accuracy of ±0.3 mbars for ground-based measurements is typical. Setra Corporation produces this type of instrument for the U.S. National Weather Service ASOS network, the latter produced by AAI Systems Management Incorporated. The ASOS network is discussed later. Measurement of pressure is also discussed elsewhere in this chapter.

82.2.3 Humidity

Instruments that determine the density or pressure of water in vapor form in the atmosphere generally measure either relative humidity or dew-point temperature. The pressure of water vapor just above a liquid water surface when the vapor is in equilibrium with the liquid water is the saturated vapor pressure of the water. This saturated vapor pressure increases with the temperature and equals atmospheric pressure at the boiling temperature of water. Relative humidity is the ratio of the vapor pressure in air to the saturated vapor pressure at the temperature of the air. Relative humidity is usually expressed in percent, which is this ratio times 100. The dew-point temperature is the temperature to which the air must be lowered, so the vapor pressure in the air is the saturated vapor pressure with the relative humidity at 100%. Knowledge of water vapor density is used in weather prediction and in global climate modeling. It also affects light transmission through the atmosphere. Relative humidity is an important meterological variable. The temperature dew-point difference is an indicator of the likelihood of fog formation and can be used to estimate the height of clouds. More detailed sources of information on humidity measurement include Refs. [6–9]. Three common methods of measuring the vapor density in the atmosphere are given in the following sections.

82.2.3.1 Chilled Mirror Method

Chilled mirror instruments for measuring the dew-point temperature are not sold by most instrument companies. A chilled mirror instrument developed by Technical Services Laboratory is used in the U.S. National Weather Service ASOS network discussed later. It has a mirror cooled by a solid-state

thermoelectric cooler (using the Peltier effect) until water vapor in the air just starts condensing on the mirror. This condensation is detected using a laser beam reflecting from the mirror. When the reflected beam is first affected by the condensed water vapor, the temperature of the mirror is the dew-point temperature. The mirror temperature is controlled to remain at the dew-point temperature by an optic bridge feedback loop. The mirror is a nickel chromium surface plated on a copper block. The temperature of the block is measured to ±0.02% tolerance by a platinum resistance thermometer imbedded in the block. An identical platinum resistance thermometer measures ambient air temperature. Outside air is drawn through the protective enclosure surrounding the instrument by a fan, so that the effect of solar heating on the measured values of the dew-point temperature and ambient temperature is negligible and so that outside air is tested. The dew-point temperature and ambient temperature are measured between −60 °C and +60 °C to an accuracy of 0.5 °C rms. Dew-point errors are somewhat larger below 0 °C. To avoid errors that might arise from deterioration of the reflective properties of the mirror, the mirror should be inspected periodically, particularly in dirty or salty environments. This method is of higher cost than other methods of measuring the amount of water vapor in the atmosphere.

82.2.3.2 Thin Film Polymer Capacitance Method

The capacitance is formed with a thin polymer film as dielectric placed between two vapor-permeable electrodes. Water vapor from the air diffuses into the polymer, changing the dielectric constant of the dielectric and thus the capacitance. The capacitance can be measured electrically by comparison to fixed capacitance reference standards. The measured value of the capacitance is related to the relative humidity by calibration. Instruments using these capacitive sensors can measure relative humidity between 0% and 100% at temperatures between about −40 °C and +60 °C to about ±2% of relative humidity. These sensors can be made very small for incorporation into integrated circuits on silicon chips [9]. They are low in cost. Usually, instruments measuring humidity also measure temperature separately. The circuits used to measure relative humidity using a thin polymer capacitance yield an electric output signal (often 0–5 V), which lends itself to remote transmission of the relative humidity.

82.2.3.3 Psychrometric Method

Errors are introduced if the water is contaminated, if the water level in the reservoir supplying water to the wick becomes low, or if the reservoir runs dry. In extremely dry environments, it can be difficult to keep the wick wet, while salty environments can change the wet bulb reading. Accuracy is affected by air speed past the wet bulb. Because of these disadvantages, psychrometers have generally been replaced by more convenient methods of measuring humidity.

82.2.4 Wind Speed, Wind Direction, and Wind Shear

82.2.4.1 Anemometer

Weather stations commonly employ a three-cup anemometer. This consists of a vertical axis rotating collar with three vanes in the form of cups. The rotation speed is directly proportional to wind speed. Figure 82.1 shows an instrument of this type. An alternative to the cup anemometer is a propeller anemometer in which the wind causes a propeller to rotate. There are several ways to obtain an electrical signal indicating the speed: a magnet attached to the rotating shaft can induce a sinusoidal electrical impulse in a pickup coil; a Hall effect sensor can be used; or the rotating shaft can interrupt a light beam, generating an electric pulse in a photodetector. Rotating anemometers can measure wind velocities from close to 0 up to 70 m/s (150 mph).

82.2.4.2 Ultrasonic Wind Sensor

This sensor has no moving parts. Wind speed determination is as follows. An ultrasonic pulse emitted by a transducer is received by a nearby detector and the transit time calculated. Next, the transit time is

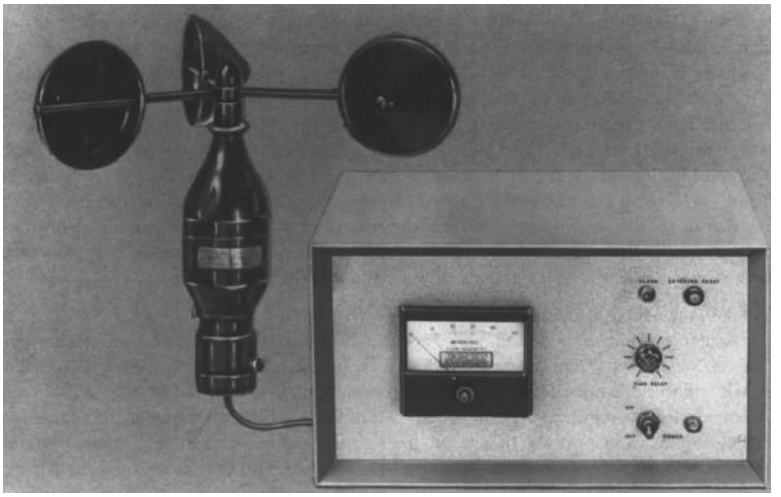


FIGURE 82.1 A cup-type anemometer for measuring wind speed. (Courtesy of Kahl Scientific Instrument Corporation.)

measured for the return path. In the absence of wind, the transit times are equal; but in the presence of wind, the wind component along the direction between the transmitter and receiver affects the transit time. Three such pairs, mounted 120° apart, enable calculation of both the wind speed and direction. Heaters in the transducer heads minimize problems with ice and snow buildup. The absence of moving parts eliminates the need for periodic maintenance.

82.2.4.3 Wind Direction

Wind direction sensors are generally some variant of the familiar weather vane. Sensitivity is maintained by constructing the weather vane to rotate on bearings with minimal resistance. Electronic readout can be achieved using a potentiometer (a "wiper" contact connected to the vane slides over a wire-wound resistor). The resistance between the contact and one end of the wire resistor indicates the position of the vane. Alternative methods of readout include optical and magnetic position sensors. Positional accuracy is ±5%.

82.2.4.4 Combination Wind Speed and Direction Sensor

A combination wind speed and direction sensor can be made in which a propeller anemometer is mounted on a weather vane. The vane keeps the propeller device pointed into the wind. Alternatively, two propeller anemometers, rigidly mounted in a mutually perpendicular arrangement, can be used to determine direction and magnitude of the horizontal wind simultaneously. Rotating anemometers and weather vanes are susceptible to ice and snow buildup and can be purchased with heaters. They need periodic maintenance.

82.2.4.5 Wind Shear

Wind shear occurs when wind direction and/or strength changes significantly over a short distance. This can occur in horizontal or vertical directions or sometimes in both. Measurement of wind shear conditions is particularly important at airports. Wind shear is determined by comparing readings made at the center of the airfield with measurements made at the periphery. An automated system to perform this function, entitled Low Level Wind Shear Alert System (LLWAS), is found at some airports. Wind shear can also be detected by Doppler radar. Doppler radar used by the U.S. National Weather Service is discussed later.

82.2.5 Precipitation

Precipitation measuring instrumentation includes devices that measure the presence of precipitation (precipitation sensors), those that determine the quantity of precipitation, those that measure rate of precipitation, and those that measure both quantity and rate.

82.2.5.1 Precipitation Presence Sensors

These sensors usually consist of two electric contacts in close proximity. Moisture causes electric conduction that is detected by a circuit monitoring conductance. A typical application consists of a circuit board consisting of a grid of two arrays of strips separated by small gaps. If the surface of the detector is heated, then only current precipitation will be detected and dew will not form to affect the measurement.

82.2.5.2 Rain Gages

These instruments measure amount of rainfall. A simple rain gage can consist of a cylinder, a funnel, and an inner collection tube of much smaller diameter than the funnel for amplification of the height of rain accumulation. The height of the water column in the inner tube is converted to total rainfall. Typical graduations on the tube enable determining rain accumulation to 0.025 cm (0.01 in.) of rain.

A tipping-bucket rain gage enables the measurement of both volume and rate of rainfall. A large funnel concentrates the precipitation, which is directed into one of two small buckets. When that bucket fills, it tips out of the way and empties, closing a switch to record the event, and another empty bucket moves into its place. Typical tipping-bucket gages respond to each 0.025 cm (0.01 in.) of rain. In conditions of snow and freezing rain, tipping-bucket rain gages can be equipped with heaters on the funnel to reduce snow and ice to water. The internal components are also heated to prevent refreezing. Reported accuracy for tipping-bucket rain gages is ±0.5% at 1.2 cm/h (0.5 in./h). Frise Engineering Company produces a tipping-bucket rain gage used in the ASOS network discussed later.

Highest accuracy rain gages that collect and concentrate precipitation should have their collection surfaces made of a plastic with a low surface tension for water. This minimizes losses from surface wetting.

Rain gages exist that do not rely on collection methods. Optical rain gages utilize an infrared beam. Drops falling through this beam induce irregularities in the beam that can be interpreted in terms of precipitation rate. This type of sensor is used for the precipitation identification or present weather sensor used in the ASOS network discussed later.

82.2.6 Solar Radiation

The mean annual intensity of solar radiation above the atmosphere (extraterrestrial solar radiation) continues to be measured to obtain a more precise value and to look for variations in the sun's energy output. It is called the solar constant. The solar constant is close to 1367 W/m^2. The intensity of solar radiation above the atmosphere varies approximately sinusoidally over the year with amplitude of close to 3.3% of the solar constant and a maximum near the first of January. This variation arises from the variation of the distance of the Earth from the sun. The sun has a spectral distribution that is roughly that of a blackbody at 5777 K with a peak of the spectrum at a wavelength of about 500 nm. Solar radiation is attenuated by scattering and absorption in the atmosphere. Attenuation is greater at wavelengths corresponding to absorption bands of certain gases in the atmosphere. On a clear day near noon, the solar intensity at the Earth's surface can be as high as 1000 W/m^2. Figure 82.2 shows a typical sea level solar spectrum with the sun about 48° away from the vertical (see, for example, Refs. [10,11]). This is called an air mass 1.5 spectrum, because the distance the radiation travels through the atmosphere is 1.5 times the distance when the sun is vertical. About 99% of the spectrum is in the range of wavelengths shown by the figure. Measurements of the solar constant and the solar spectrum are scientific measurements made with specialized instruments. Additional information can be found in Iqbal [11] and Coulson [12].

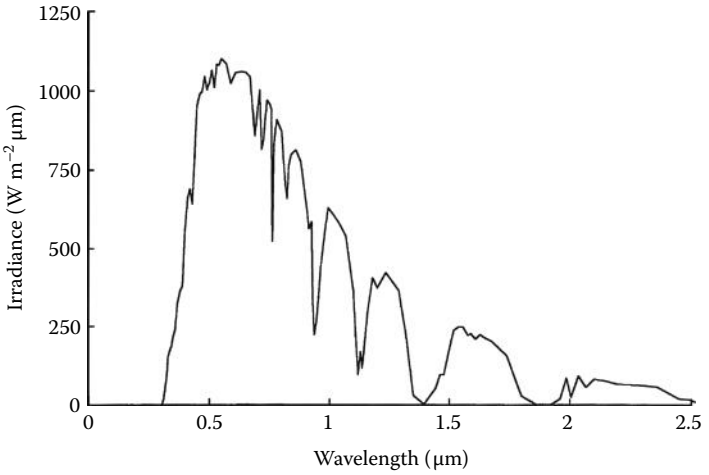


FIGURE 82.2 The intensity of solar radiation as a function of wavelength for a path length through the atmosphere of 1.5 times the vertical path length. The dips in the spectrum are molecular absorption bands.

Solar radiation instruments for general use in the field measure direct radiation from the sun, total or global radiation coming from the sky hemisphere, and diffuse or sky radiation (global radiation with the direct radiation removed). Solar radiation measuring instruments can be broadband instruments that measure the combined solar intensity (irradiance) at all wavelengths, or they can be spectrally selective instruments that measure the intensity at different wavelengths or in different wavelength bands. Only the much more common broadband instruments are discussed here. The instruments used for everyday measurement are field instruments. Field instruments are first class if they are of higher quality and provide greater accuracy and reliability (at higher cost).

Direct radiation is radiation coming directly from the sun without scattering in the atmosphere. Instruments measuring direct solar radiation usually include radiation coming from the sky out to an angular distance of about 3° away from the center of the sun's disk. They are called pyrheliometers. The radiation coming from clear sky near the sun, rather than from the solar disk, is the circumsolar radiation. This radiation can be subtracted from the pyrheliometer measurement for a more precise determination of direct radiation. This correction is often not made for routine measurements. The clear sky correction is calculated using the angular distribution of the intensity of the circumsolar radiation [13,14].

Instruments measuring global radiation are installed in a level position with a plane sensor facing up toward the sky. These instruments measure solar radiation coming from the whole sky hemisphere. They are called pyranometers. Global radiation measuring instruments should be sited in a level elevated area with no obstructions obscuring the sky hemisphere.

Diffuse radiation is the radiation coming from the sky hemisphere with the direct radiation subtracted. Pyranometers are used for measuring diffuse solar radiation and should be mounted in the same manner as pyranometers used for measurement of global radiation. They have an occulting (shade) disk or shadow band to prevent direct solar radiation from reaching the radiation sensor. The measurement of diffuse radiation involves correcting the pyranometer measurement for the part of the sky radiation shielded from the sensor by the occulting disk or shadow band. For clear skies, the occulting disk correction is calculated using the angular distribution of intensity of the circumsolar radiation. Corrections for partially cloudy and cloudy skies depend on the particular cloud conditions. Corrections for the shadow band are often determined by temporarily replacing the shadow band with an occulting disk when the sky is clear and when it is overcast. Corrections for measurements under other sky conditions can be determined by interpolation. The shadow band correction is discussed by LeBaron et al. [15]. The occulting disk

must have a tracking system to make the disk follow the sun over the sky. The shadow band removes solar radiation received from a narrow swath of the sky along the path the sun follows during the day. The shadow band must be adjusted regularly during the year as the path of the sun changes over the seasons. The more common solar radiation measuring instruments include the pyranometer and pyrheliometer.

82.2.6.1 Pyranometer

The sensor is usually a thermopile, consisting of a number of thermocouples in series, with alternate junctions heated by the sun. The unheated junctions are near ambient temperature. This is sometimes arranged by putting the unheated junctions in thermal contact with a white surface. Heating by the sun is accomplished by placing the junctions in contact with a matte black surface of high heat conductivity or by a black coating on the junctions. The blackened surface has a constant high solar absorptance (usually ~99%) over the solar spectrum. A constant high solar absorptance over the solar spectrum is important. The solar spectrum at the surface of the Earth varies with the time of day and year and the amount of clouds, because of the spectrally dependent scattering and absorption of solar radiation by the atmosphere. An absorbing surface whose absorption of solar radiation varies with wavelength will cause the sensor to have a different sensitivity for different wavelengths of the solar spectrum. Some less expensive pyranometers use a silicon photovoltaic sensor (solar cell) to measure solar radiation. These sensors have zero sensitivity above about 1.2 μm and the spectral response below 1.2 μm is not constant. This limits the accuracy of measurements of solar intensity with photovoltaic sensors. Instruments with a thermopile sensor can use a combination of thermopiles and resistors to compensate for the variation of the output of a single thermopile with temperature. The hemispherical windows of pyranometers are usually made of a special glass which transmits solar radiation of wavelengths between about 0.3 and 2.8 μm. This includes ~99% of the solar intensity. The absorbing surface must have a cosine response as a function of angle away from the normal to the surface (Lambert law response), and a flat response as a function of azimuth around the normal to the absorbing surface, for the global radiation to be measured correctly. The degree to which the pyranometer response is linear follows the cosine law, is temperature-, spectrum-, and azimuthally independent, and determines whether the instrument is a first-class instrument. Figure 82.3 shows a Kipp and Zonen (The Netherlands) first-class pyranometer.

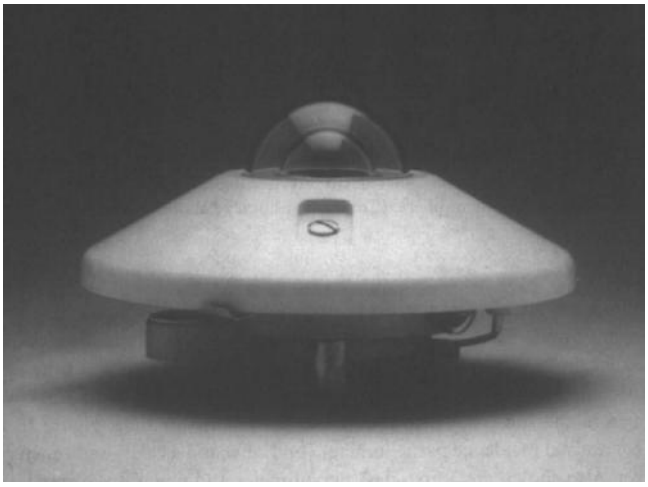


FIGURE 82.3 A class 1 pyranometer used for the measurement of global and diffuse solar radiation. (Courtesy of Kipp & Zonen Division of Enraf-Nonius Corporation.)

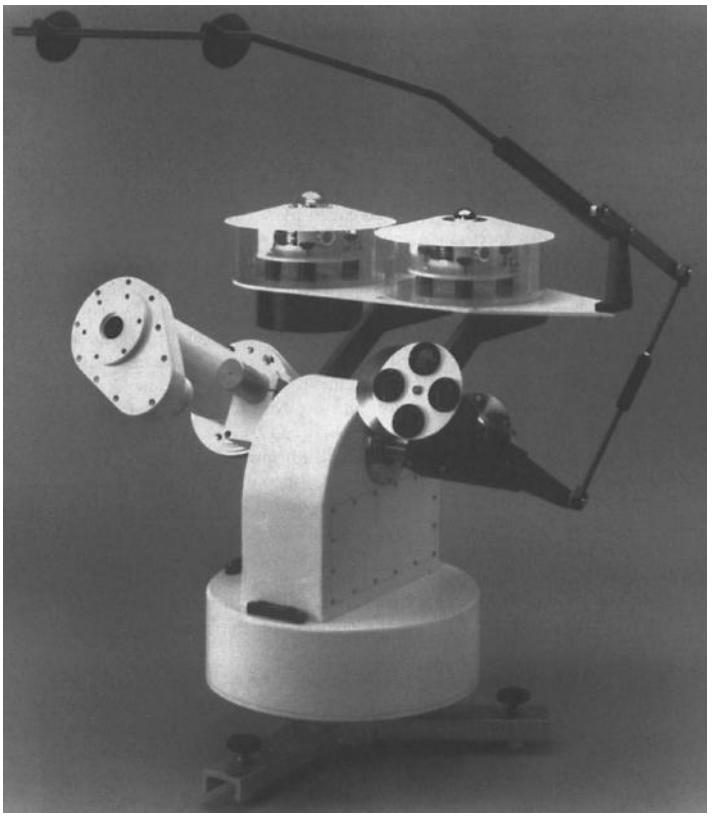


FIGURE 82.4 An automatic solar tracker with normal incidence pyrheliometer and cavity radiometer mounted on it. The tracker is coupled to two shade disks that shield direct sunlight from a first-class pyranometer and precision infrared radiometer facing the sky hemisphere. (Courtesy of Eppley Laboratories.)

82.2.6.2 Pyrheliometer

The field pyrheliometers usually have temperature-compensated thermopile sensors with flat spectral response and linear output. The sensor is placed at the bottom of an internally blackened, diaphragmed, collimator tube that limits the angular acceptance for solar radiation to be in the range of about 5° or 6° total acceptance angle. The pyrheliometer is mounted on an equatorial mount tracker that keeps the direct radiation from the sun parallel to the axis of the collimator tube. Figure 82.4 shows an Eppley Laboratory (United States) normal incidence pyrheliometer (on right) and a cavity radiometer (on left) mounted on a solar tracker. Also shown are a first-class pyranometer and precision infrared radiometer shielded from direct radiation from the sun by shade disks coupled to the tracker.

82.2.7 Visibility

Visibility in meteorology is a measure of how far one can see through the atmosphere.

82.2.7.1 Visibility Sensors

Measurement of forward scattering of light by aerosols in a given sample volume is used to determine visibility. A pulsed beam of near-infrared light (typically 850–880 nm wavelength) generated by an infrared-emitting diode is projected into the ambient air. A detector (usually a silicon photodiode), placed between 1 and 2 m away and oriented between 30° and 40° off the axis of the pulsed beam, samples the scattered light. The intensity of scattering is proportional to the atmospheric extinction coefficient.

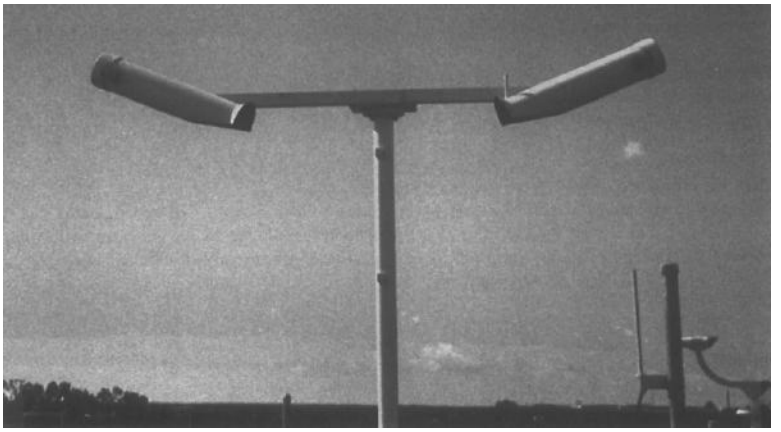


FIGURE 82.5 The xenon flash instrument that determines atmospheric visibility at an ASOS system station. This system is produced by AAI Systems Management Incorporation.

The extinction coefficient (k) is related to visual range (VR) by VR = 5/k (km). VR measurable by such a system lie between about 0.3 and 60 km. Figure 82.5 shows the visibility sensor used in the ASOS network. A xenon flash lamp source is used with this system.

82.2.7.2 Transmissometers

These measure the attenuation of a light beam to determine visibility.

82.2.8 Lidar Measurements

Lidar is optical radar in which a laser emits pulses of electromagnetic radiation. Scattered radiation that returns to an optimally tuned detector at the laser site for observation is called echos. The time from emission of the pulse to return of the echo determines the distance of the scatterer. The echos are sampled at all times between pulses, and thus all distances up to the maximum range. This provides information about the atmosphere along the pulse path. The strength of the echo is connected with the density, size, and shape of the scattering particles. Laser pulses are linearly polarized. Spherical scatterers return echos that have the same linear polarization. Nonspherical scatterers depolarize the echos. Thus, polarization of the echo also provides information concerning the scatterers. Using high spectral resolution lidar (HSRL), the scattering by air molecules can be separated from the scattering by aerosols. Differential absorption lidar (DIAL) measures the concentration of gaseous species in the atmosphere. Stephens [16] discusses the HSRL and DIAL methods. Zhao et al. [17] present an application of the DIAL method to the determination of O_3 concentration in the atmosphere. Because of their high power output, ruby, dye, neodymium, yttrium–aluminum–garnet (YAG), and CO_2 lasers are commonly used for lidar. The use of lidars in the observation of aerosols injected into the stratosphere by volcanic eruption, and their subsequent decay has become quite common (see Post et al. [18] and Jager [19] and other articles in the same journal issue).

82.2.9 Clouds

82.2.9.1 Observer Estimation

For many years, observers at weather stations have estimated the amount, type, and sometimes the opacity of low clouds, middle clouds, and high clouds. Estimation is usually every hour or similar period with the estimate in tenths or eighths (octa) of the sky covered. Satellite, radar, and ceilometer data are replacing observer estimation. Observer estimates have been archived.

TABLE 82.1 Companies Marketing Meteorological Instrumentation

Apex Instruments, Inc.
204 Technology Park Lane, Fuquay Varina, NC 27526
Phone: 919-557-7300, 800-882-3214 (toll free), Fax: 919-557-7110
http://www.apexinst.com/

Campbell Scientific, Inc.
815 W. 1800 N., Logan, UT 84321-1784
Phone: 435-753-2342, Fax: 435-750-9540
http://www.campbellsci.com/

Columbia Weather Systems, Inc.
2240 NE Griffin Oaks St., Suite 100, Hillsboro, OR 97124-7135
Phone: 503-629-0887, 866-625-8620 (toll free), Fax: 503-629-0898
http://www.columbiaweather.com/

Davis Instrument Mfg. Co. (T, P, RH, W, PR)[a]
4701 Mount Hope
Drive Baltimore, MD 21215
Phone: (800) 548-9409
http://www.davisnet.com/

Electric Speed Indicator (T, P, W)
12234 Triskett Rd., Dept. T-16
Cleveland, OH 44111-2519
Phone: (216) 251-2540
http://www.electricspeedindicator.com/

Eppley Laboratory, Inc.
12 Sheffield Ave. (SOL)
P.O. Box 419, Newport, RI 02840
Phone: (402) 847-1020, Fax: (401) 847-1031
http://www.eppleylab.com/

Frise Engineering Co. (PR)
2800 Sisson Street
Baltimore, MD 21211
Phone: (410) 235-8524

Hycal Division, Honeywell Corp. (T, RH)
9650 Telstar Avenue
El Monte, CA 91731
Phone: (818) 444-4000
http://sensing.honeywell.com/

Kahl Scientific Instrument Corp. (T, P, RH, W, PR, SOL, VIS)
P.O. Box 1166
737 South Main Street
El Cajon, CA 92022
Phone: (619) 444-2158
http://www.kahlsico.com/

Kipp & Zonen Division (SOL)
Enraf-Nonius Co., P.O. Box 507, L 2600 AM Delft, Roentgenweg 1, NL 2624 BD Delft, The Netherlands
Phone: 011 31 15 269 8500
390 Central Ave., Bohemia, NY 11716
Phone: (800) 229-5477
http://www.kippzonen.com/

Lockheed-Martin Corp. (Weather Surveillance Doppler Radar)
365 Lakeville Road
Greatneck, NY 11020
Phone: (516)574-1404
http://www.lockheedmartin.com/

Qualimetrics, Inc. (T, P, RH, W, PR, SOL, VIS, CEIL)[a]
1165 National Drive
Sacramento, CA 95134
Phone: (800) 806-6690
http://www.qualimetrics.com

Rainwise, Inc. (T, P, RH, W, PR, SOL)[a]
P.O. Box 443
Bar Harbor, ME 04609
Phone: (800) 762-5723
http://www.rainwise.com/

R. M. Young, Co. (T, P, W, RH, PR)
2801 Aero Park Dr.
Traverse City, MI 49686
Phone: (616) 946-3980
http://www.rmyoung.com/

Safety Flag Co. of America
82 Hadwin Street, Central Falls, RI 02863
Phone: 401-722-0900, 800-556-7584 (toll free),
Fax: 401-722-0912, 800-822-0450 (toll free)
http://www.safetyflag.com/

Setra Systems, Inc.
159 Swanson Rd., Boxborough, MA 01719
Phone: 978-263-1400, 800-257-3872 (toll free),
Fax: 978-264-0292
http://www.setra.com/

Vaisala Oy, PL 26 FIN-00421, (T, P, RH, W, PR, SOL, VIS, CEIL, RADS)[a]
Helsinki, Finland
100 Commerce Way
Woburn, MA 01801-1068
Phone: (617) 933-4500
http://www.vaisala.com/

Warren-Knight Instrument, Co.
2045 Bennett Dr., Philadelphia, PA 19116-30
Phone: 215-464-9300, Fax: 215-464-9303
http://www.warrenknight.com/

Note: T, Temperature; P, pressure, RH, relative humidity; W, wind; PR, precipitation; SOL, solar radiation; VIS, visibility; LID, lidar; CEIL, ceilometer; RADS, radiosonde.

[a] Complete weather station available.

82.2.9.2 Ceilometers

Ceilometers measure the altitude of clouds. Most ceilometers are based on lidar technology using an infrared pulsed laser diode, usually GaAs. Laser pulses are reflected back (echos) to a receiver, usually a Si diode detector. The round-trip time is converted into cloud height. Common ceilometers can measure cloud base altitudes to 4,000 m (12,000 ft). Recent instruments developed by Vaisala can measure cloud bases up to 23,000 m (75,000 ft) with an error of ±1%. The Vaisala ceilometer is used in the ASOS network discussed later. Maintenance includes cleaning of the window. Heaters and blowers respond to automatic sensing units to clear precipitation from the window and control instrument temperature.

Companies producing and marketing meteorological instruments are presented in Table 82.1. For information on additional companies marketing meteorological instruments in the United States, see *Thomas Register of American manufacturer's* and *Thomas Register Catalog File* found in most major libraries. For manufacturers in other countries, search library catalog files under "manufacturers."

82.3 U.S. Weather Service Facilities

The U.S. National Weather Service, the U.S. Federal Aviation Administration, and the U.S. Department of Defense have joined in an upgrade of weather instrumentation and analysis for U.S. weather stations. Other national weather services are upgrading their facilities. The new U.S. National Weather Service system has four components to obtain weather data throughout the United States and its possessions:

1. The Automated Surface Observing Systems (ASOS): These systems, produced by AAI Systems Management Incorporated, are installed at over 850 sites. Each system usually measures visibility, surface temperature, dew-point temperature, pressure, wind speed and direction, visibility, cloud height, precipitation identification, freezing rain, and precipitation accumulation. Instruments used for each ASOS site are discussed earlier with other meteorological measuring instruments. Figure 82.6 shows the arrangement of instruments at an ASOS site.
2. Doppler weather surveillance radar (NEXRAD—WSR-88D): Doppler radar are produced by Lockheed-Martin Corporation and installed at ~160 sites in the U.S. Doppler radar surveys the weather out to about 300 km from the radar. The standard radar scans angles above the horizon up to about 6° in six ~1° intervals with a 360° azimuth. Rain scans go up to about 20°. Radar echos are reflection of the pulsed radar microwave signal from water drops in the form of rain or cloud. Ice crystals and snow reflect radar pulses with lower echo strength. Doppler radar can detect short-lived, possibly catastrophic events such as tornados, downbursts, and flash floods in real time. The Doppler shift of the echo determines the radial component of the wind velocity at the height and distance of the source of the reflected echo. The strength of the echo provides information on the precipitation rate. Powerful computers and sophisticated computer programs are used to analyze the radar signals. Additional discussion of the use of radar for weather observations can be found in Stephens [16]. NEXRAD is discussed in Refs. [20–23], for example.
3. Radiosonde upper air sounding: The northern hemisphere has a network of about 700 radiosonde sites, including the U.S. network, where soundings are generally made twice daily. The northern hemisphere is largely decoupled from the southern hemisphere for weather development. Radiosondes provide information on wind velocity, temperature, and humidity at all elevations up to 30–40 km. A temperature sensor, humidity sensor, and pressure sensor are carried aloft by balloon at 0000 and 1200 UTC (±1 h). The wind velocity at different altitudes is measured by determining the radiosonde position at all times during the flight. The horizontal position of the radiosonde can be obtained by OMEGA, LORAN, or VLF global positioning systems. Additional information on global positioning is found in Section 76.3 (Chapter 76 of this book) on satellite navigation and radiolocation. Pressure can indicate the altitude of the radiosonde. Radar or GPS can measure the radiosonde position in three dimensions. Radiotheodolites can track the radiosonde to receive the data transmitted by radio transmitter. The radiosonde manufactured by Viz

FIGURE 82.6 A view of an ASOS system sited at an airport near San Diego. The instrument of Figure 82.5 is seen in the foreground. A stand for the freezing rain sensor is not used. The next instrument is the ceilometer. Behind that is a metal box containing connections to the sensors and instruments to transmit the data to a remote collection and processing station. Along the same line of instruments beyond this is the precipitation identification or present weather sensor. This is followed along the same line by the sensor measuring ambient and dewpoint temperatures and a tipping-bucket rain gage. To the north of the line of instruments is a pole with a cup anemometer, weather vane, flashing red lights, and lightning rod at the top.

Manufacturing Co. called Bsonde is used by the U.S. National Weather Service. The instruments are mounted on a lightweight white styrofoam container. The temperature sensor is a thin, 50 mm long rod thermistor painted white to minimize heating by solar radiation. It is mounted outside the container. It measures temperature to ±0.5 °C. The container has an air flow duct for ventilation of the relative humidity sensor mounted inside. The relative humidity sensor consists of an insulating strip coated with a film of carbon that measures relative humidity from 0% to 100% to ±5%. The pressure sensor is a nickel, C-span aneroid pressure sensor coupled mechanically to a 180 contact baroswitch providing the output pressure signal. Pressure is measured from 1060 to 5 mb with an accuracy of ±0.5 mb. Sensor signals are sent to the receiving station by an amplitude modulated transmitter operating from 1660 to 1700 MHz. The Vaisala Inc. radiosonde sensors used by the U.S. National Weather Service differ from Bsonde by use of a capacitance temperature sensor, a capacitance with polymer dielectric humidity sensor, and a capacitive aneroid pressure sensor.

4. Satellite remote sensing: Meteorological variables are deduced from measurements of the electromagnetic radiation coming from the atmosphere and the surface of the Earth using satellite-based sensors. Usually, three or more different detector wavelength bands are used. The interpretation is difficult because the electromagnetic radiation can come from many levels of the atmosphere or the Earth's surface, all at different temperatures, and may have undergone multiple scattering in the atmosphere and Earth reflection. The energy coming from the atmosphere and the Earth's

surface is mostly reflected and scattered solar radiation for wavelengths up to about 4 μm. Longer wavelength energy is mostly associated with thermal emission from the Earth and its atmosphere. Thermal emission from the atmosphere and the surface of the Earth consists of overlapping blackbody spectra of different temperatures. Absorption spectra characteristic of various gases in the atmosphere and their temperatures are superimposed on the overlapping blackbody spectra. Visible wavelengths from about 0.4 to 0.7 μm are useful for determining cloud coverage and cloud type. Infrared wavelengths from about 6 to 7 μm can be used to determine water vapor density. Infrared in the range from about 10 to 12 μm can be used to look at high clouds. It is also useful for night observation. Satellite measurements show the motion and development of the clouds and storm systems. More detailed discussion is available [16,24–27].

References

1. McGee, T. D., *Principles and Methods of Temperature Measurement*, New York: John Wiley & Sons, 1988.
2. Nicholas, J. V., *Traceable Temperatures: An Introduction to Temperature Measurement and Calibration*, New York: John Wiley & Sons, 1994.
3. Quinn, T. J., *Temperature*, New York: Academic Press, 1983.
4. Schooley, J. F., *Thermometry*, Boca Raton, FL: CRC Press, 1986.
5. Wilson, R. E., Temperature, in Ross, S. D. (Ed.), *Handbook of Applied Instrumentation*, Malabar, FL: Robert E. Krieger Publishing, 1982.
6. Hickes, W. F., Humidity and dew point, in Ross, S. D. (Ed.), *Handbook of Applied Instrumentation*, Malabar, FL: Robert E. Krieger Publishing, 1982.
7. Moisture and Humidity Measurement and Control in Science and Industry, *Proceedings of the 1985 International Symposium on Moisture Humidity, Washington, DC, April 15–18, 1985*, Research Triangle Park, NC: Instrument Society of America.
8. Silverthorne, S. V., Watson, C. W., and Baxter, R. D., Characterization of a humidity sensor that incorporates a CMOS capacitance measuring circuit, *Sensors Actuators*, 19, 371–383, 1989.
9. Marvin, C. F., *Psychrometric Tables for Obtaining the Vapor Pressure, Relative Humidity, and Temperature of the Dew Point from Readings of the Wet- and Dry-Bulb Thermometers*, Washington, DC: U.S. Government Printing Office, 1941.
10. Hulstrom, R., Bird, R., and Riordan, C., Spectral solar irradiance data sets for selected terrestrial conditions, *Solar Cells*, 15, 365–391, 1985.
11. Iqbal, M., *An Introduction to Solar Radiation*, New York: Academic Press, 1983.
12. Coulson, K. L., *Solar and Terrestrial Radiation: Methods and Measurements*, New York: Academic Press, 1975.
13. Zerlaut, G. A., Solar radiation measurements: Calibration and standardization efforts, in Boer, K. W. and Duffie, J. A. (Eds.), *Advances in Solar Energy*, Vol. 1, Boulder, CO: American Solar Energy Society, 1983.
14. Major, G., *Circumsolar Correction for Pyrheliometers and Diffusometers*, WMKO/TD-NO. 635, Geneva, Switzerland: World Meteorological Organization, 1994.
15. LeBaron, B. A., Michalsky, J. J., and Perez, R., A simple procedure for correcting shadowband data for all sky conditions, *Solar Energy*, 44, 249–256, 1990.
16. Stephens, G. L., *Remote Sensing of the Lower Atmosphere*, Oxford, U.K.: Oxford University Press, 1994.
17. Zhao, Y., Howell, J. N., and Hardesty, R. M., Transportable lidar for the measurement of ozone concentration and aerosol profiles in the lower troposphere, *SPIE International Symposium on Optical Sensing Environmental Monitoring, SPIE Proceedings*, Atlanta, GA, 310–320, October 11–14, 1993.
18. Post, M., Grund, C., Langford, A., and Proffitt, M., Observations of Pinatubo Ejecta over Boulder, Colorado by lidars of three different wavelengths, *Geophys. Res. Lett.*, 19, 195–198, 1992.

19. Jager, H., The Pinatubo eruption cloud observed by lidar at Garmisch-Partenkirchen, *Geophys. Res. Lett*., 19, 191–194, 1992.
20. *Next Generation Weather Radar: Results of Spring 1983 Demonstration of Prototype NEXRAD Products in an Operational Environment*, NEXRAD Joint System Program Office, Government Publication R400-N49, September 1984.
21. *Next Generation Weather Radar Product Description Document*, NEXRAD Joint System Program Office, Government Publication R400-PD-202, December 1986.
22. *A Guide for Interpreting Doppler Velocity Patterns*, NEXRAD Joint System Program Office, Government Publication R400-DV-101, October 1987.
23. Heiss, W. H., McGrew, D. L., and Sirmans, D., Nexrad: Next generation weather radar (WSR-88D), *Microwave J*., 33, 79–80, 1990.
24. Burroughs, W. J., *Watching the World's Weather*, Cambridge, U.K.: Cambridge University Press, 1991.
25. Carleton, A. M., *Satellite Remote Sensing in Climatology*, Boca Raton, FL: CRC Press, 1991.
26. Houghton, J. T., Taylor, F. W., and Rodgers, C. D., *Remote Sounding of Atmospheres*, Cambridge, U.K.: Cambridge University Press, 1984.
27. Scorer, R. S., *Cloud Investigation by Satellite*, New York: Halsted Press, 1986.

83

Robotics and Sensors: Environmental Applications

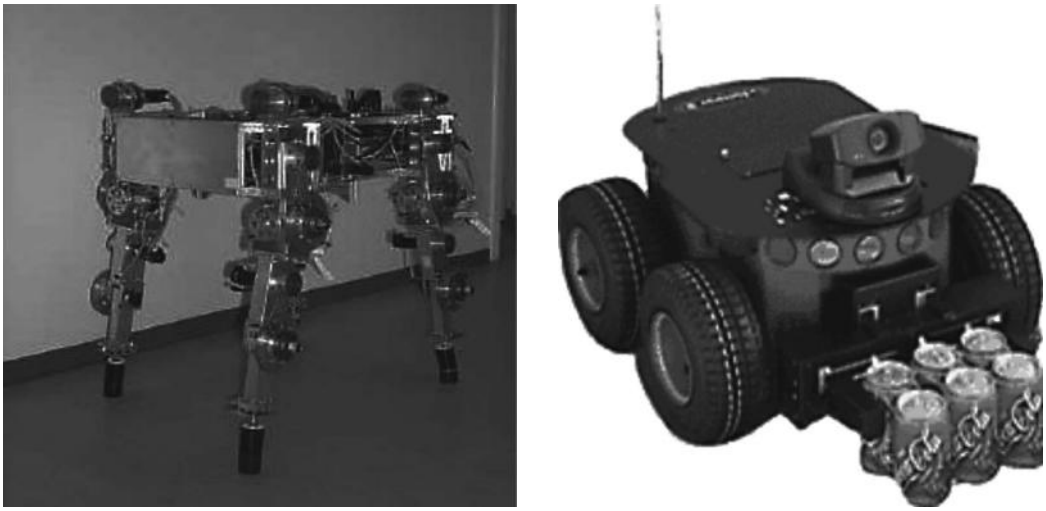
Gourab Sen Gupta
Massey University

83.1 Introduction

A robot is an automatically operated machine that replaces human effort, though it may not resemble human beings in appearance or perform functions in a humanlike manner. It is capable of being programmed in advance and is designed to execute one or more tasks repeatedly, with speed and precision.

Robots have changed the way we work, play, live, and unfortunately fight wars. Robots invaded the workplace many decades ago, initially for factory automation. They are increasing their presence in the home at a very rapid pace, primarily for assisted living. Wars are being fought using robots on the ground, above and below the waters, and in the air. In the next decade, the world will witness the largest growth of robots in the service industry. From the days of industrial automation using monstrous robots, the world has advanced to micro and nano robots traversing the veins of a human body to deliver drugs.

What makes the robots so capable and versatile as they are today? Will they ever be able to attain the full functionality, intelligence, and versatility of human beings? Or is it a wishful thinking? What will be the breakthrough technology that will enable the robots to make that quantum jump in their capabilities?



(a) (b)

FIGURE 83.1 Robots employing different types of locomotion methods: (a) a quadruped robot and (b) a wheeled robot.

For successful completion of tasks, robots have to perceive the world around them, the workspace in which they operate. At the heart of this perception are the inputs from a gamut of sensors. Accurate measurement of physical parameters and fusion of sensory data has a profound influence on the accuracy of the perception model. While a lot of energy and resources are still being expended for research into robot locomotion and actuators for motion, it is the advancement in sensors and measurement technology that will catapult the robots to the next level of versatility and acceptance. Miniaturization of sensors and precision measurement will be the flavor of research in the next decade.

Robots come in many different forms and sizes. Based on the method used for the locomotion of a robot, it can be classified as wheeled robot (including track robots) or legged robot. Legged robots may be two-legged humanoid or multilegged. Hybrid robots combine the aspects of wheeled and legged locomotion. Yet there are others types which cannot be classified as either wheeled or legged such as hovercrafts, underwater floating robots, and robots employing snakelike movements. A few different types of robots are shown in Figure 83.1.

Sensors are a very important element of a closed-loop robot control system (Figure 83.2). Sensors provide the robot the "awareness" of its surroundings—what's ahead, around, and "out there"? Sensors augment the capabilities of a robot to interact with the environment. For example, a robot lawn mower can "see" grass in its workspace. Sensors also give the robots the capability to goal seek such as by finding colored objects. Robots also need to employ sensors for protection and self-preservation. Localization in structured and unstructured environments is an important task of an autonomous robot for which sensors are required.

Sensors and measurement techniques increase the versatility of robots. Nature is a rich source of inspiration for the development of sensing and measurement systems. Sensors can be categorized based on nature—exteroceptive, proprioceptive, and interoceptive sensors.

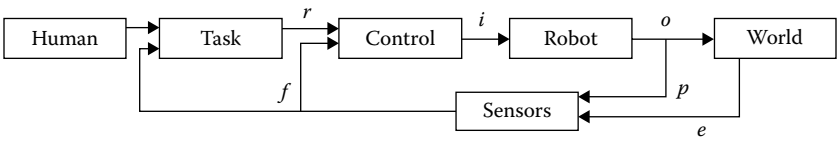


FIGURE 83.2 System blocks of a generic closed-loop control system of a robot.

Exteroceptive sensors deal with the external world and are used to find where is something, how does it look, and how big is it? Common exteroceptive sensors are camera, laser rangefinder, haptic sensors, etc. *Proprioceptive sensors* deal with self, for example, is the robot perfectly horizontal, where is it, how much is the joint bent, which way it is facing, how fast is it turning, etc. Common exteroceptive sensors are inclinometer, GPS, encoders, flex sensors, compass, and gyroscope. *Interoceptive sensors* deal with self, although without conscious perception, such as the amount of battery charge left for which a voltmeter may be used as the measurement device.

As shown in Figure 83.2, outputs from exteroceptive sensors (e) and proprioceptive sensors (p) are frequently combined to generate the feedback signal (f) for the robot controller. The feedback signal from the sensors is compared with the reference signal (r) in the control block to generate the inputs (i) for the robot.

83.2 Sensors for General Robotic Systems

Whole gamuts of sensors are used in robotic systems, from simple to very complex and sophisticated. These are detailed in this section.

83.2.1 Position and Velocity Encoders

Encoders measure displacement, both rotary and linear, and are used in odometry of a robot. The position information may also be used to calculate velocities. An encoder is an electromechanical device that can monitor position, motion, or direction of motion. It can be constructed with sensing mechanisms which can be mechanical, magnetic, or optical. There are two basic kinds of encoder information: absolute and incremental. Absolute encoders provide a code, which indicates the exact position of the shaft/wheel. Incremental encoders measure how far the wheel has moved and are normally used for measuring robot wheel velocity.

83.2.1.1 Magnetic Encoders

A magnetic encoder consists of a rotating gear (disk) made of ferrous metal and a magnetic pickup that contains a permanent magnet and the sensing element. The gear, which is mounted on the rotating shaft, has precisely machined teeth that provide the code pattern. As the disk rotates, these teeth disturb the magnetic flux emitted by the permanent magnet, causing the flux field to expand and collapse. These changes in the field are sensed by the sensing element, which generates a corresponding digital or pulse signal output. Two kinds of magnetic pickups exist—*Hall effect pickups* use a semiconducting sensing element that relies on the Hall effect to generate a pulse for every gear tooth that passes the pickup; *variable reluctance pickups* use a simple coil of wire in the magnetic field. As the gear teeth pass by the pickup and disturb the flux, they cause a change in the reluctance of the gear/magnet system. This induces a voltage pulse in the sensing coil that is proportional to the rate of flux change. Figure 83.3 shows the construction of a magnetic encoder.

83.2.1.2 Optical Encoders

The most popular type of encoder is the optical encoder, which consists of a rotating disk, a light source, and a photo detector (light sensor) as shown in Figure 83.4. The disk, which is mounted on the rotating shaft, has coded patterns of opaque and transparent sectors. As the disk rotates, these patterns interrupt the light emitted onto the photo detector, generating a digital or pulse signal output.

The encoding disk is made from glass for high-resolution applications (11 to >16 bits) and plastic (Mylar) or metal for applications requiring more rugged construction (resolution of 8 to 10 bits).

83.2.2 Flex or Bend Sensors

Flex or bend sensors are used to measure bend or angle. They are particularly of use in anthropomorphic robotic arms or exoskeleton devices used for control of robotic arms. These sensors are made of resistive

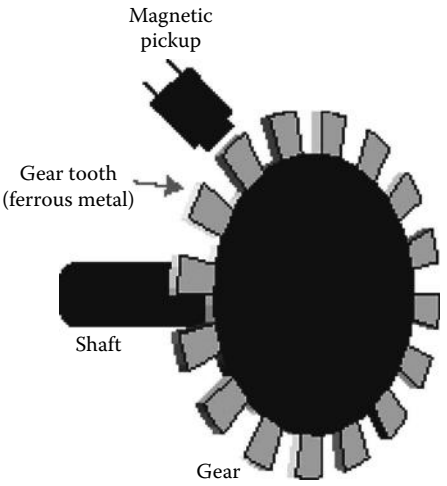


FIGURE 83.3 Magnetic encoder.

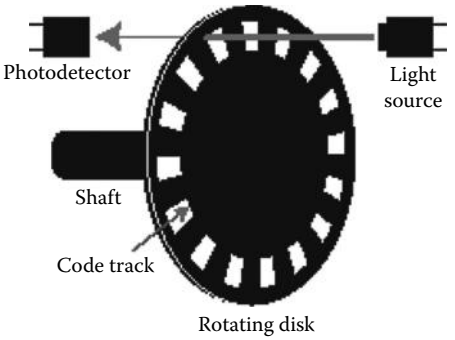


FIGURE 83.4 Optical encoder.

FIGURE 83.5 A resistive bend sensor.

strips as shown in Figure 83.5. The resistance changes when the strip is bent. While they are easy to integrate in a control system, they are however not very accurate.

83.2.3 Accelerometers

These sensors measure acceleration (as long as there is no rotation). They are frequently used to sense orientation, acceleration, vibration shock, falling, and for high-precision control of linear actuators of a robot and for joint feedback control. Unit of measurement is gram. A MEMS accelerometer can also be used as a tilt sensor. Available accelerometers can measure acceleration in 2-axis or 3-axis to detect magnitude and direction of the acceleration as a vector quantity. The output of the accelerometer can be analog signals or digital signals, usually on a serial interface. Often they are available commercially on a PCB breakout board, as shown in Figure 83.6, which makes it very easy to integrate with the rest of the electronic system of a robot.

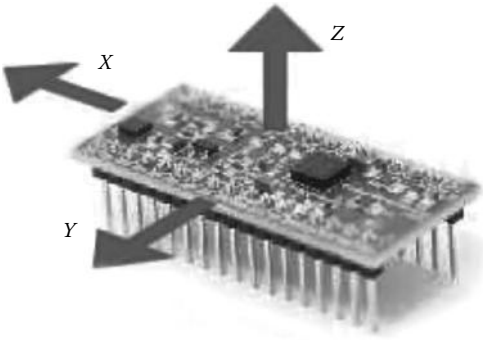


FIGURE 83.6 Three-axis accelerometers.

Apart from being incorporated in robots, micromachined and MEMS accelerometers are increasingly present in portable electronic devices (phones, tablet PC, cameras) and video game controllers (WII Console), to detect the position of the device or provide for game input.

There are many different ways to make an accelerometer. Accelerometers built using piezoelectric effect contain microscopic crystal structures that get stressed by accelerative forces, which cause a voltage to be generated. Another technique used is to sense the change in capacitance. If you have two microstructures next to each other, they have a certain capacitance between them. If an accelerative force moves one of the structures, then the capacitance will change. Additional circuitry is required to convert from capacitance to voltage. Commercially available accelerometer chips combine the processing circuitry on the chip itself and provide a voltage signal proportional to the acceleration. Other methods used are piezoresistive effect, hot air bubbles, and light.

83.2.4 Gyroscopes

A gyroscope has the capability of measuring the rate of rotation around a particular axis (angular rate or rotational speed). The unit of measurement is degree/second. It can be used for navigation on unmanned aerial vehicles (UAV), and radio-controlled helicopters. A gyroscope is essential for the stability of flying robots and is used to measure yaw, pitch, and roll of the robot. A difference between a gyroscope and an accelerometer is that a gyroscope gives an indication of the angular rate whereas an accelerometer measures the linear acceleration.

83.2.5 Electronic Compass

A digital compass sensor measures direction and bearing of a robot or autonomous mobile platforms. Often, the four cardinal (N, E, S, W) directions and the four intermediate (NE, NW, SE, SW) directions are measured. These sensors are essential for robots that navigate an unknown terrain. It uses a magnetic field sensor and a tilt sensor for compensation of inclination.

83.2.6 Two-Axis Tilt Sensors (Inclinometers)

A tilt sensor or inclinometer measures inclination, i.e., angles of slope (or tilt), single or dual axis. It measures elevation or inclination of an object with respect to gravity. Inclinometers measure both inclines (positive slopes, as seen by an observer looking upward) and declines (negative slopes, as seen by an observer looking downward). It is also known as a tilt meter, slope gauge, gradient meter, gradiometer, level gauge, declinometer, and pitch and roll indicator. Figure 83.7 shows an inclinometer.

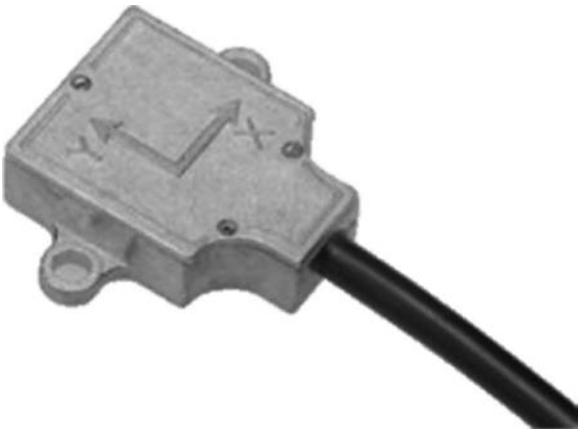


FIGURE 83.7 An inclinometer.

The output of these sensors can be analog signals or serial digital data. Common sensor technologies for inclinometers are MEMS, accelerometer, liquid capacitive, electrolytic, gas bubble in liquid, and pendulum.

83.2.7 Range Finders

For navigation and obstacle avoidance, robots need to detect the distance of objects from them. Range-finding sensors are used to measure range or distance. There are several technologies available to make such sensors, namely sensors based on infrared (IR) signals, ultrasonic signals, and lasers beams. The sensor has a transmitter–receiver pair. An IR light, sound, or laser beam is sent by the transmitter and reflected off the target, the distance of which is being measured. The time it takes for the signal to be returned to the receiver is a measure of how far the object is from the source.

83.2.8 GPS Sensors

More and more robots are being equipped with global positioning system (GPS) modules. The module can determine the latitude, longitude, and altitude of a receiver on earth by computing the time difference for signals from different satellites to reach the GPS receiver. The calculation to ascertain the location is based on a simple arithmetic theory known as trilateration. Especially in outdoor operations, where the GPS module has a direct line of sight to four or more orbiting satellites, GPS helps to localize the robot while navigating an unknown terrain. An added advantage of the GPS module is that it works in any weather condition and also provides additional information such as time, speed, bearing, trip distance, distance to destination, sunrise, and sunset time. Most GPS modules implement the NMEA 0183 standard which is a simple ASCII, serial communications protocol that defines how data is transmitted in a "sentence."

83.2.9 Vision Sensors (Camera)

Cameras give the robot the sense of sight. Cameras are very popular exteroceptive sensors, especially with mobile robots, for perceiving the external world. The image captured by the camera is processed to extract features of the surrounding and identify objects. The hardware to process images has become very cheap, fast, and small in form factor. This has made it very easy and cost-effective to integrate cameras in robotic systems (Figure 83.8). Primarily, there are two types of cameras based on the output

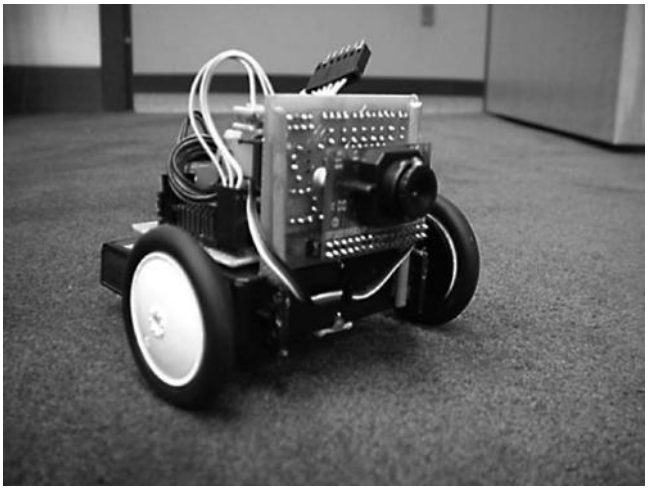


FIGURE 83.8 Use of camera in a robot.

signal—analog and digital. Digital cameras come with many different interfaces, namely USB, FireWire (IEEE 1394), TCP/IP, and SmartLink. Cameras with RF wireless transmitter built-in are also popular and can invariably be found on autonomous flying vehicles such as a Quadrotor and rescue robots.

83.2.10 Haptic (Tactile) Sensor

Haptic technology gives the robot the sense of touch by using tactile sensing which provides feedback by measuring forces and pressure. Research in the area of haptic sensing is in its infancy but is gaining momentum. Haptic sensors are employed in such scenarios where the robot has to manipulate objects. It helps the robot to discriminate between rigid and soft objects so that the gripping forces can be adjusted accordingly. Tactile sensors also give a feel of the object's profile (shape) which in turn helps the robot to decide on the optimal orientation for gripping. The integration of haptic sensors to refine information provided by vision sensors is a very promising approach in the development of autonomous robotic systems.

83.3 Sensors for a Humanoid Robot

A two-legged humanoid robot requires several types of sensors for balancing and gait generation. High accuracy and resolution of these sensors is of paramount importance for successful control of the humanoid. In this section, a couple of the sensors used in a typical humanoid robot are detailed. Primarily, there are three main types of sensors which are of utmost importance for humanoid control, namely foot pressure sensors, gyroscopes, and an Attitude Heading Reference System (AHRS).

83.3.1 Force Sensing Resistors

Along with a wide range of motion and acceleration sensors, force sensing resistors (FSRs) are very common devices in robotics applications. Due to their wide availability, ease of use, and low cost, designers have extensively relied on FSRs for force feedback in humanoid applications. FSRs are thin isometric force sensors whose resistance decreases, in a nonlinear way, with the force applied.

At least three commercial force sensors are currently available in the market—the Interlink FSR, the LuSense PS3, and the Tekscan FlexiForce sensor. The choice of the FSR is based on the application.

The FlexiForce sensors present better response in terms of linearity, repeatability, time drift, and dynamic accuracy, while the Interlink FSRs are more robust. If large changes in force are applied at a relatively high frequency, the Interlink or LuSense sensors should be selected, whereas if large slowly varying forces are applied infrequently for long durations, the Tekscan sensor should be selected. In a humanoid application, changes are short and changes in force are applied at a relatively high frequency; therefore, the Interlink FSRs are best suited. Due to nonconstant terrain, the ability to maintain a consistent and accurate area and position of the applied force will be the limiting factor in terms of sensor accuracy and precision.

The most basic FSR consists of two membranes separated by a thin air gap. The air gap is maintained by a spacer around the edges and by the rigidity of the two membranes. One of the membranes has two sets of interdigitated fingers that are electrically distinct, with each set connecting to one trace on a tail. The other membrane is coated with FSR ink. When pressed, the FSR ink shorts the two traces together with a resistance that depends on applied force.

The Interlink 402 FSR comprises of two polymer films: one with a conductive surface and the other with printed electrodes facing the first. Contact between the two surfaces causes the conductive layer to short circuit the printed electrodes, thereby reducing the electric resistance of the component. Typically, its resistance will drop from >1 MΩ to about 10 kΩ for an applied load of 100 (a force of roughly 1 N) to 10,000 g.

83.3.2 Three-Axis AHRS

To monitor the position of the body and pelvis, some form of sensor must be implemented. An AHRS consists of three gyroscopes, three accelerometers, and three magnetometers. These systems are currently used in intelligent gait humanoids. An on-board processor calculates the rate of change with respect to time outputting an absolute position in hexadecimal. The magnetometer is used as a horizontal reference to compute the rotational angle whereas the gravity is used for the roll and pitch.

Gyros are used to measure angular rates; accelerometers are used to determine initial attitude of the AHRS and to correct gyro drifts in the pitch and roll angles. The magnetometers essentially perform the same task where they correct gyro drift in the heading angle.

Gyro values are fed into the block named "numeric integration of the orientation equations" which is used to measure the rate of change in gyro angle. These values are fed into the block function "computation of correction signals" which reviews accelerometer and magnetometer values to produce and error value to be combined with the gyro value. Finally, an estimation of the heading pitch and roll angles is produced which combined with the other functional blocks is filtered using extended Kalman filtering. This filtering further increase the accuracy of the estimated orientation angles of the AHRS.

A major disadvantage of using MEMS systems for navigation/positioning is that they typically suffer from drift. Drift is the expression for an accumulated error in the position calculations. Because the system is continually adding minute changes to its previously calculated position, any error in measurement, however small, is accumulated resulting in a gradual drift effect.

To try and reduce the possibility of drift, there are several algorithms and filters that can be implemented. The CH 6 dm has a built-in extended Kalman filter (EKF) which through statistical models simulates the potential error in calculated value and actual value. EKF performance can be tuned by adjusting the process noise covariance matrix and the measurement noise covariance matrices (there are two measurement noise matrices—one for the accelerometers and other for the magnetometer).

The process noise matrix is used to specify how much the EKF "trusts" data from the gyros with respect to data from the magnetic sensors and accelerometers. The lower the values along the diagonal of the matrix, the more the rate gyros are trusted. Conversely, if the diagonal elements are large, the gyros are trusted less and the accelerometers and magnetometers are weighted more heavily.

83.4 Anthropomorphic Robotic Arm for Plant Health Monitoring Using RGB Color Sensor

Robots are increasingly finding applications in agriculture. Robots are employed to monitor the growth and health of plant tissues in a modified in vitro clonal plant propagation system. The color of the plant leaves is an indicator of its health. To monitor the color of plant material, and in particular to determine the RGB color content, a RGB color sensor that can interpret the color of an object can be used.

Colored objects that exhibit a particular color, as seen by humans, absorb certain wavelengths and reflect others back to the observer. Therefore, a red object would absorb all the green and blue wavelengths, but reflect back only the red wavelengths, thus giving its red appearance to the human. Currently, there are a number of color sensors available globally from numerous suppliers. Color sensors vary in price, size, sophistication, and the color space they use.

83.5 Sensors for Mobile Robotic Platforms in Environmental Applications

83.5.1 Greenhouse Applications

Various environmental factors influence the quality and productivity of plant growth in a greenhouse. Continuous monitoring of these environmental parameters gives valuable information to the grower to better understand how each factor affects the quality and the rate of plant growth, and how to maximize crop yield. The environmental parameters that are important are atmospheric temperature, humidity, carbon dioxide (CO_2), light intensity, and soil moisture. In this section, technologies for various environmental monitoring sensors have been reviewed.

83.5.2 Atmospheric Temperature Sensing

Temperature sensing technology is one of the most widely used sensing technologies in the world of sensors. It allows for the measurement of temperature in various applications and provides protection from excessive temperature excursions. Five different families of temperature sensors are available on the market. Each family of temperature sensors has its advantages and disadvantages. Depending on the application, one sensor may be more suitable than the other.

83.5.2.1 Thermocouples

A thermocouple is a junction between two wires of dissimilar metals. The point of contact between the wires generates a voltage that is proportional to the temperature. Thermocouples are suitable for measuring over a large temperature range, up to 2300 °C. They are less suitable for applications where smaller temperature differences need to be measured with high accuracy. For such applications, thermistors and resistance temperature detectors (RTDs) are more suitable. Applications include temperature measurement for kilns, gas turbine exhaust, diesel engines, and other industrial processes.

Advantages: Wide temperature range (200 °C–1300 °C), relatively cheap, highly accurate, minimal long-term drift, and fast response time.

Disadvantages: The relationship between the temperature and the thermocouple's output signal is not linear, low output signal (mV), vulnerable to corrosion, and calibration of thermocouples can be tedious and difficult.

83.5.2.2 Resistance Temperature Detectors

RTDs are widely used in many industrial applications such as air conditioning, food processing, and textile production, processing of plastics, microelectronics, and exhaust gas temperature measurement.

RTDs are basically temperature-sensitive resistor devices. The resistance increases with temperature. Most RTD elements consist of a length of fine-coiled wire wrapped around a ceramic or glass core.

Advantages: Linear over a wide temperature range, relatively accurate, good stability, and repeatability at high temperature (65 °C–700 °C).

Disadvantages: Low sensitivity, higher cost than thermocouples, and vulnerable to shock and vibration.

83.5.2.3 Thermistors

Similar to RTDs, thermistors are also temperature-dependent resistor devices. Thermistors are not as accurate or stable as RTDs but they are easier to wire, cost less, and almost all automation panels accept them directly. Thermistors are made of semiconductor materials with a resistivity that is especially sensitive to temperature.

Advantages: Highly sensitive, low cost, accurate over small temperature range, and good stability.

Disadvantages: Nonlinear resistance temperature characteristics, self-heating, and limited temperature operating range.

83.5.2.4 Integrated Circuit Temperature Sensors

In low-cost applications, most of the sensors stated earlier either are expensive or require additional circuits or components to be used. However, integrated circuit (IC) temperature sensors are complete, silicon-based sensing circuits with either analog or digital output. IC temperature sensors are often used in applications where the accuracy demand is low.

Advantages: Low cost, excellent linearity, and easy-to-read output.

Disadvantages: Limited temperature range, self-heating, fragile, and slightly less accurate when compared to other types.

83.5.3 Humidity Sensing

When it comes to humidity sensing technology, there are three types of humidity sensors: capacitive, resistive, and thermal conductivity humidity sensors.

83.5.3.1 Capacitive Humidity Sensors

Capacitive humidity sensors (CHS) are widely used in industrial, commercial, and weather telemetry applications. CHSs consist of a substrate on which a thin film of polymer or metal oxide is deposited between two conductive electrodes. The sensing surface is coated with a porous metal electrode to protect it from contamination and exposure to condensation. The substrate is typically glass, ceramic, or silicon. The changes in the dielectric constant of a CHS are nearly directly proportional to the relative humidity (RH) of the surrounding environment.

Advantages: Able to function in high-temperature environments (up to 200 °C), near linear voltage output, wide RH range, high condensation tolerance, reasonable resistance to chemical vapors and contaminants, minimal long-term drift, high accuracy, small size, and low cost.

Disadvantages: Limited sensing distance and sensor interface can be tedious and difficult.

83.5.3.2 Resistive Humidity Sensors

Resistive humidity sensors (RHS) measure the changes in electrical impedance of a hygroscopic medium such as conductive polymer, salt, or treated substrate. These sensors are suitable for use in control and display products for industrial, commercial, and residential applications. RHS consists of noble metal electrodes either deposited on a substrate by photoresist techniques or wire-wound electrodes on a plastic or glass cylinder.

Advantages: Faster response time than CHS, near linear voltage output, high accuracy, small size, low cost, and wide RH range.

Disadvantages: Lower operating temperature when compared to CHS, sensitive to chemical vapors, low tolerance against contaminants, and low condensation tolerance.

83.5.3.3 Thermal Conductivity Humidity Sensors

Thermal conductivity humidity sensors (TCHS) measure the absolute humidity by quantifying the difference between the thermal conductivity of dry air and that of air containing water vapors. These sensors are suitable for applications such as kilns for drying wood; machinery for drying textiles, paper, and chemical solids; pharmaceutical production; cooking; and food dehydration. TCHS consists of two matched negative temperature coefficient (NTC) thermistor elements in a bridge circuit; one is hermetically encapsulated in dry nitrogen and the other is exposed to the environment.

Advantages: Very durable, able to operate in high-temperature environments (up to 600°C), excellent immunity to many chemical and physical contaminants, high accuracy, and high condensation tolerance.

Disadvantages: Respond to any gas that has thermal properties different from those of dry nitrogen; this will affect the measurements.

83.5.4 Soil Moisture Sensing

Excessive moisture is undesirable in agriculture, houses, textiles, packaging materials, electronic appliances, dry food process, etc. Moisture detection is important in a number of different situations. For example, measurement of soil moisture is useful for minimizing the amount of irrigation water applied to grow plants and for optimizing plant growth. Excess moisture also creates "wet feet" in plants and destroys them. Because of the importance of the moisture content of materials, various techniques have been developed to measure it. This section outlines a number of soil moisture detection technologies that are available on the market and lists their advantages and disadvantages.

83.5.4.1 Frequency Domain Reflectometry Soil Moisture Sensor

Frequency domain reflectometry (FDR) is sometimes referred to as capacitance sensor. Soil sensor probes that use the FDR method of soil moisture measurement employ an oscillator to generate an electromagnetic signal that is propagated through the unit and into the soil. Part of this signal will be reflected back to the unit by the soil. This reflected wave is measured by the FDR probe, telling the user what the water content of the soil is.

Advantages: Highly accurate, fast response time, and inexpensive.

Disadvantages: Need to be calibrated for the type of soil they will be buried in.

83.5.4.2 Time Domain Reflectometry Soil Moisture Sensor

Time domain reflectometry (TDR) sensors propagate a pulse down a line into the soil, which is terminated at the end by a probe with wave guides (Figure 83.9). TDR systems measure water content of the soil by measuring how long it takes the pulse to come back.

Advantages: Highly accurate and fast response.

Disadvantages: Calibration can be tedious, difficult, and expensive.

FIGURE 83.9 Time domain reflectometry soil moisture sensors.

83.5.4.3 Gypsum Blocks

Gypsum blocks use two electrodes placed into a small block of gypsum to measure soil water tension. Wires connected to the electrodes are connected to either a portable handheld reader or a data logger. The amount of water in the soil is determined by the electrical resistance between the two electrodes within the gypsum block. More water present in the soil will reduce the resistance, while less water will increase it.

Advantages: Inexpensive and easy to install.

Disadvantages: Have to be replaced periodically and sensitive to the saline content of water.

83.5.4.4 Neutron Probes

Using neutron probes is another way to measure soil moisture content (Figure 83.10). A probe which is inserted in the ground emits low-level radiation in the form of neutrons. These neutrons collide with the

FIGURE 83.10 Use of neutron probes.

hydrogen atoms contained in water, which is detected by the probe. More the water content in the soil, the more neutrons are scattered back at the device.

Advantages: Very accurate and fast response.

Disadvantages: Expensive and users have to be registered with the government due to radioactive elements used to emit the neutrons.

83.5.5 Carbon Dioxide Sensing

Measuring carbon dioxide is important in monitoring indoor air quality and many industrial processes. Two types of CO_2 detectors are available to measure the CO_2 level in the surrounding environment—electrochemical CO_2 sensors and nondispersive IR (NDIR) CO_2 sensors.

83.5.5.1 Electrochemical CO_2 Sensors

The CO_2-sensitive element consists of a solid electrolyte formed between two electrodes, together with a heater substrate. By monitoring the changes in the electromotive force (EMF) generated between the two electrodes, it is possible to measure CO_2 gas concentration.

Advantages: Cheap, accurate, real-time sensing, high tolerance against contaminants, and small in size.

Disadvantages: Require a significant amount of power because they operate at high temperature.

83.5.5.2 NDIR CO_2 Sensors

NDIR sensors are spectroscopic sensors to detect CO_2 in a gaseous environment by its absorption characteristics. The key components are an IR source, a light tube, an interference (wavelength) filter, and an IR detector. The gas is pumped or diffused into the light tube and the electronics measures the absorption of the characteristic wavelength of light.

Advantages: High speed, real-time sensing, low power consumption, high contamination tolerance, and small in size.

Disadvantages: Carbon monoxide often coexists with CO_2 and absorbs a similar wavelength range as CO_2 which results in inaccurate readings. They are also very expensive.

83.5.6 Light Sensing

Light from the sun is responsible for nearly all life on the earth. Sunlight fuels the process of photosynthesis where plants convert carbon dioxide and water into carbohydrates. Plants use light in the range of 400–700 nm. This range is most commonly referred to as photosynthetically active radiation (PAR). Monitoring PAR is important to ensure that plants are receiving adequate light for photosynthesis. Typical applications include forest canopies, greenhouse monitoring, etc. PAR is also measured to estimate evaporation in bodies of water, as it plays a key role in surface water temperature. This section will present some of the popular light sensors on the market that can be used for environmental monitoring applications.

83.5.6.1 Photometric Sensors

Photometric sensors measure visible radiation or light as seen by the human eye.

Advantages: Highly sensitive, good stability, fast response time (10 μs), low temperature dependency, excellent linearity, and small size.

Disadvantages: Expensive and these sensors are mostly used to measure indoor lighting conditions. For environmental applications, PAR and solar radiation sensor are preferred.

83.5.6.2 Light-Dependent Resistor

Similar to photometric sensors, light-dependent resistors (LDRs) measure visible light as seen by the human eye. A LDR is basically a resistor; the internal resistance increases or decreases dependent on the level of light intensity impinging on the surface of the sensor.

Advantages: Very cheap, fast response, linear output, and small in size.

Disadvantages: Like photometric sensors, LDRs are mostly used to measure indoor lighting conditions.

83.5.6.3 Pyranometers

Pyranometers measure total solar radiation. These sensors are commonly used for agriculture, meteorology, and solar energy applications. The sensor is composed of a silicon photovoltaic detector mounted in a miniature head. The sensor output current is directly proportional to the level of solar radiation.

Advantages: Highly accurate, excellent linearity, good stability, and fast response time.

Disadvantages: Bulky and expensive.

83.5.6.4 Quantum Sensors

Quantum sensors measure the photosynthetic photon flux density (PPFD) of PAR. They are one of the most popular types of light sensors used in agriculture and environmental industries.

Advantages: Very sensitive, fast response, highly accurate, excellent linearity, good stability, and small in size.

Disadvantages: Expensive.

83.6 Biomimetic Sensor Design

Concepts from a nonengineering domain such as biology have sparked inspiration and innovation for a variety of technologies. Bacteria, plants, insects, mammals, reptiles, and the like have diverse forms, solving a variety of engineering functions. Elegant methods of sensing and communication are used in biological systems which may be considered as sophisticated adaptive systems. A biology-inspired branch of sensor research has emerged called biomimicry.

Biological systems offer exemplary methods of flight, imaging, sensing, adaptation to environment, and locomotion. Engineers have learned from these and created novel technologies. Biological domain provides inspiration at many levels, such as cellular, organism, and species. Nature has been developing biological sensors for billions of years. Lasting solutions have evolved to fulfill unique ecological niches, which make them ideal for study and imitation. Not only it is nature rich with sensing methods, but also it provides strategies to use these sensors. Biological sensors typically exhibit low-energy requirements, high sensitivity, and redundancy. They exhibit parallel sampling and processing of sensory information. Having tens or even hundreds of parallel receptor cells improves the signal-to-noise ratio through averaging. This also reduces the likelihood of error due to loss of or failure of a receptor organ. A great lesson from nature is redundancy; in most biological systems, there are many instances of redundancy.

Bibliography

Colestock, H., *Industrial Robotics: Selection, Design and Maintenance*, McGraw-Hill, New York, 2004.
Kurfess, T. R. ed., *Robotics and Automation Handbook*, CRC Press, Boca Raton, FL, 2005.
Nehmzow, U., *Mobile Robotics: A practical Introduction*, 2nd Ed., Springer, New York, 2003.
Nof, S. Y. ed., *Handbook of Industrial Robotics*, 2nd Ed., John Wiley & Sons, New York, 1999.
Spong, M. W., S. Hutchinson, and M. Vidyasagar, *Robot Modeling and Control*, John Wiley & Sons, New York, 2006.

IX

Signal Processing

84

Modulation

David M. Beams
The University of Texas at Tyler

84.1 Introduction

It is often the case in instrumentation and communication systems that an information-bearing signal may not be in an optimal form for direct use. In such cases, the information-bearing signal may be used to alter some characteristic of a second signal more suited to the application. This process of altering one signal by means of another is known as *modulation*; the original information is called the *baseband signal*, and the signal modulated by the baseband signal is termed the *carrier* (because it "carries" the information contained in the baseband signal). Recovery of the original information requires a suitable demodulation process to reverse the modulation process.

A prominent use of modulation techniques is found in communication systems. The extremely long wavelengths of electromagnetic waves at frequencies found in a typical audio signal make direct transmission impractical, because of constraints on realistic antenna size and bandwidth. Successful radio communication is made possible by using the original audio (baseband) signal to modulate a carrier signal of a much higher frequency and transmitting the modulated carrier by means of antennas of feasible size. Another example formerly in widespread use is that of modems used to transmit digital data through the telephone network. Digital data are not directly compatible with analog local subscriber connections, but audible signals modulated by digital data could be carried over local telephone lines. Instrumentation systems use modulation techniques for telemetry (where the distances may be on the order of centimeters for implanted medical devices to hundreds of millions of kilometers for deep-space probes), for processing signals in ways for which the original signals may be unsuited (such as avoiding $1/f$ noise in high-gain amplifiers by processing a high-frequency carrier modulated by a low-frequency signal), and for specialized amplification purposes (carrier and lock-in amplifiers).

Techniques that modulate the amplitude of the carrier are full-carrier *amplitude modulation* (AM), reduced- or suppressed-carrier double-sideband (DSB) AM, single-sideband (SSB) suppressed-carrier

modulation, vestigial-sideband (VSB) modulation, and on–off keying (OOK). Techniques that modulate the frequency or phase angle of the carrier include *frequency modulation* (FM), phase modulation (PM), frequency-shift keying (FSK), and phase-shift keying (PSK). Simultaneous variation of amplitude and phase are applied in quadrature AM (QAM). Each technique has its own particular uses. Full-carrier AM is used in radio broadcasting; VSB was formerly used in analog television broadcasting. DSB appears in instrumentation systems utilizing carrier amplifiers and modulating sensors, while SSB finds use in certain high-frequency radio communications. FM is used in broadcasting and point-to-point mobile communications. OOK is commonly used to transmit digital data in optical fiber links. FSK, PSK, and QAM are found in digital communications; analog QAM was formerly used to transmit chrominance (color) information in color television broadcasting. The emphasis of this particular chapter will be instrumentation systems; those interested principally in communications applications could begin by consulting Refs. [1–4].

Malmstadt et al. [5] list the purposes of modulation in instrumentation systems as using one signal to convey (carry) the information of another; of multiplexing; and of moving the spectrum of a signal to another band of frequencies to avoid noise or interference.

84.2 Generalized Modulation

Two assumptions are made at the outset: (1) the highest frequency present in the baseband signal is typically considerably less than the carrier frequency and (2) the results derived in the following section pertain to sinusoidal carriers but may be extended to other periodic carrier signals (such as square waves and triangle waves). The following equation gives a general expression as a function of time t for a modulated sinusoidal carrier signal of radian frequency ω_c:

$$f_s(t) = A_c(t)\cos[\omega_c t + \varphi(t)] \tag{84.1}$$

where

$f_s(t)$ is the modulated signal
$A_c(t)$ is the modulated signal's instantaneous amplitude
$\varphi(t)$ is the instantaneous modulation phase angle

Information may be carried by $f_s(t)$ by AM, *angle modulation*, or, in some cases, both. As will be shown later, both FM and PM may be seen as specific case forms of generalized angle modulation.

Equation 84.1 may be recast in an equivalent form:

$$f_s(t) = f_i(t)\cos(\omega_c t) - f_q(t)\sin(\omega_c t) \tag{84.2}$$

where

$f_i(t) = A_c(t)\cos[\varphi(t)]$
$f_q(t) = A_c(t)\sin[\varphi(t)]$

Equation 84.2 gives $f_s(t)$ as the sum of a cosinusoidal carrier term with time-varying amplitude $f_i(t)$ and a sinusoidal (quadrature) carrier term with time-varying amplitude $f_q(t)$; this formulation is known as the carrier-quadrature description of $f_s(t)$. The terms $f_i(t)$ and $f_q(t)$ are known, respectively, as the in-phase and quadrature components of $f_s(t)$. Carlson et al. [1] gives the Fourier transform of a signal in carrier-quadrature form:

$$F_s(\omega) = \frac{1}{2}[F_i(\omega - \omega_c) + F_i(\omega + \omega_c)] + \frac{j}{2}[F_q(\omega - \omega_c) + F_q(\omega + \omega_c)] \tag{84.3}$$

where $F_s(\omega) \leftrightarrow f_s(t)$, $F_i(\omega) \leftrightarrow f_i(t)$, and $F_q(\omega) \leftrightarrow f_q(t)$ are Fourier-transform pairs. The meaning of Equation 84.3 is that the spectrum of $F_s(\omega)$ contains the spectra of both $F_i(\omega)$ and $F_q(\omega)$,

each translated by $\pm\omega_c$. Time-varying modulation of the carrier in any sense causes energy to appear at frequencies (known as *sidebands*) other than the carrier frequency. Sidebands will be symmetrically distributed relative to the carrier in all but specialized cases (e.g., SSB) which are not found in instrumentation systems.

84.3 Amplitude Modulation

Doebelin [6] notes that interest in amplitude-modulated signals in instrumentation systems comes from two considerations:

1. Physical data are sometimes themselves amplitude-modulated.
2. AM may be intentionally introduced in some measurement systems to achieve one or more benefits.

AM in instrumentation systems usually takes the form of DSB suppressed-carrier AM. Examples of measurements using AM techniques will be given later.

84.3.1 DSB AM

AM applied to a sinusoidal carrier is described by

$$f_s(t) = A_c[k + \mu f_m(t)]\cos(\omega_c t) \tag{84.4}$$

where

A_c is the amplitude of the unmodulated carrier
k is the proportion of carrier present in the modulated signal
μ is the modulation index
$f_m(t)$ is the modulating baseband signal (presumed to be a real bandpass signal)
ω_c is the carrier radian frequency

The modulation index relates the change in amplitude of the modulated signal to the amplitude of the baseband signal. The value of k ranges from 1 in full-carrier AM to 0 in suppressed-carrier DSB AM. The instantaneous amplitude of the modulated signal is $k + \mu f_m(t)$, which may take on any positive value consistent with the dynamic range of the modulator and demodulating system. Phase reversal of the carrier occurs if the value of $k + \mu f_m(t)$ becomes negative. Recasting Equation 84.4 in carrier-quadrature form gives $f_i = A_c[k + \mu f_m(t)]$ and $f_q = 0$. The Fourier transform of this signal is

$$F_s(\omega) = \frac{A_c}{2}\left\{[k\delta(\omega - \omega_c) + k\delta(\omega + \omega_c)] + \mu[F_m(\omega - \omega_c) + F_m(\omega + \omega_c)]\right\} \tag{84.5}$$

where $\delta(\omega - \omega_c)$ and $\delta(\omega + \omega_c)$ are unit impulses at $+\omega_c$ and $-\omega_c$, respectively; these are the carrier components of $F_s(\omega)$. The frequency domain representation of $F_s(\omega)$ also contains symmetric sidebands about the carrier with the upper sideband arising from the positive-frequency component of $F_m(\omega)$ and the lower sideband from the negative-frequency component. A DSB AM signal thus has a bandwidth twice as large as that of the baseband signal.

Figure 84.1 represents a full-carrier AM signal resulting from modulation of a sinusoidal carrier by a sinusoidal modulating signal $f_m(t) = A_m \cos(\omega_m t)$, where A_m is the amplitude and ω_m is the radian frequency of the modulating signal. This waveform is said to show 100% sinusoidal AM because the smallest value of $k + \mu f_m(t)$ just reaches 0.

The spectra of the unmodulated sinusoidal carrier, the sinusoidal modulating signal, and the composite AM signal are shown in Figure 84.2 which shows the presence of symmetric sidebands spaced at $\pm\omega_m$ with respect to the carrier components (which themselves remain unchanged when modulation is applied).

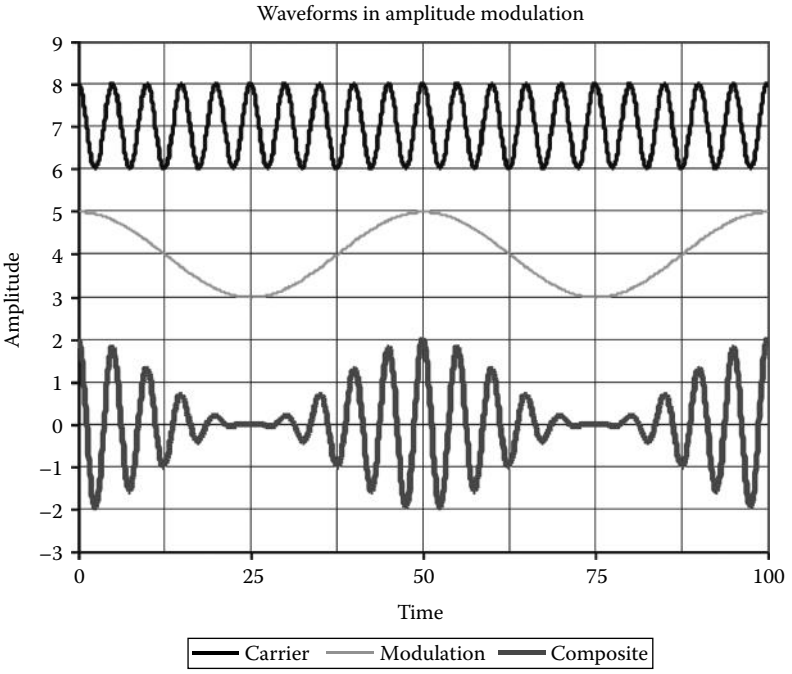


FIGURE 84.1 Full-carrier amplitude modulation showing 100% modulation of a sinusoidal carrier by a sinusoidal baseband modulating signal. The composite signal and the unmodulated carrier are in-phase throughout the modulation cycle. The time scale is arbitrary, and the carrier and modulation waveforms have been offset for clarity.

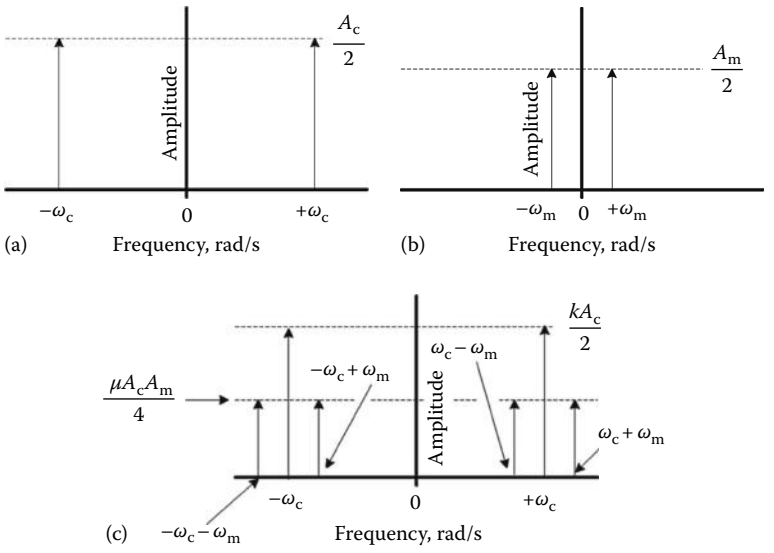


FIGURE 84.2 Double-sided spectra of (a) the unmodulated carrier, (b) the sinusoidal baseband signal, and (c) the amplitude-modulated signal.

The time-domain expression for a suppressed-carrier DSB signal is found by setting k to 0 in Equation 84.4. The phase of the suppressed-carrier DSB signal with respect to the unmodulated carrier undergoes phase reversal whenever $\mu f_m(t)$ goes negative, in contrast to full-carrier AM. The spectrum of a suppressed-carrier DSB signal is similarly found from Equation 84.5 by setting k to 0.

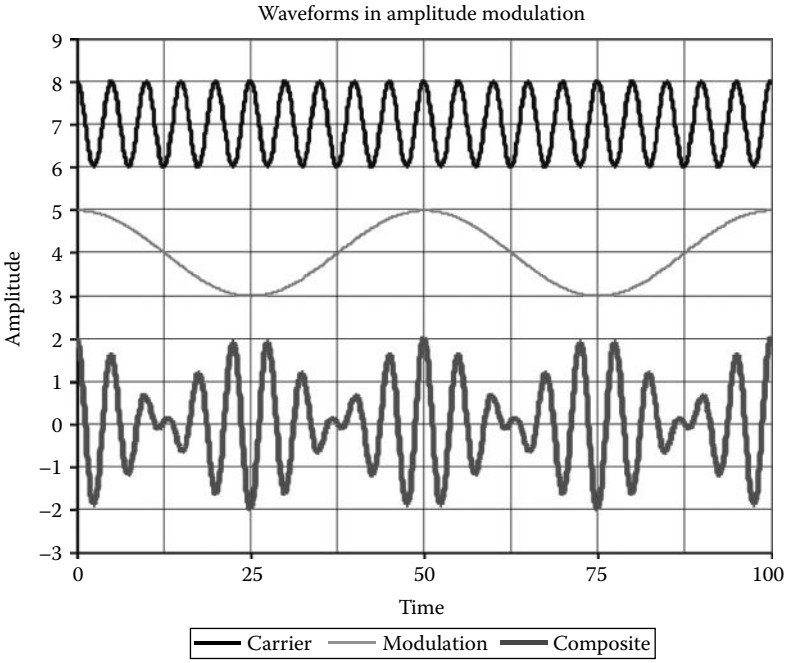


FIGURE 84.3 Suppressed-carrier double-sideband amplitude modulation of a sinusoidal carrier by a sinusoidal modulating signal; $k = 0$, $\mu = 2.0$, $A_c = 1.0$, and $A_m = 1.0$. The unmodulated carrier and the composite waveform are in phase on the positive half cycle of the modulating waveform, but the phase of the composite signal reverses with respect to the carrier on the negative half cycle. The time scale is arbitrary, and the carrier and modulation waveforms have been offset for clarity.

Figure 84.3 shows a time-domain representation of a suppressed-carrier DSB signal with the same sinusoidal baseband modulation as in Figure 84.1. The information in a full-carrier AM signal is found solely in the time-varying amplitude of the modulated signal, but the information carried by the suppressed-carrier DSB signal is found both in the amplitude and instantaneous phase of the modulated signal. (The phase of the DSB signal is inverted relative to the carrier when the baseband signal is negative and in phase when the baseband signal is positive.)

AM is a linear modulation technique; the sum of multiple AM signals produced from a common carrier by different baseband signals is the same as one AM signal produced by the sum of the baseband signals.

84.3.2 Generation of DSB AM Signals

Generation of AM signals in radio transmitters is frequently performed by applying the modulating waveform to the supply voltage to a nonlinear radio-frequency power amplifier with a resonant-circuit load as described by Carlson et al. [1]. Low-level AM may be achieved by direct multiplication of the carrier signal by $[k + \mu f_m(t)]$. Suppressed-carrier DSB signals often arise in instrumentation systems as the result of the use of an ac drive signal to a modulating sensor. Figure 84.4 shows examples of this. In Figure 84.4a, a sinusoidal drive voltage $v_{in}(t)$ is applied to the primary winding L_p of a linear variable differential transformer (LVDT). The magnitude of the output voltage $v_{out}(t)$ is proportional to the displacement $x(t)$ of the LVDT core from its center (null) position between the two secondary windings L_{x1} and L_{s2}; the output voltage is zero when the core is centered. The phase of $v_{out}(t)$ relative to $v_{in}(t)$ indicates the direction of displacement relative to the null position. In Figure 84.4b, an ac signal is applied to a resistive Wheatstone bridge. The differential output voltage $v_{out}(t)$ is zero when the bridge is balanced; the magnitude of $v_{out}(t)$ indicates the degree of unbalance in the bridge, and the phase of $v_{out}(t)$ relative

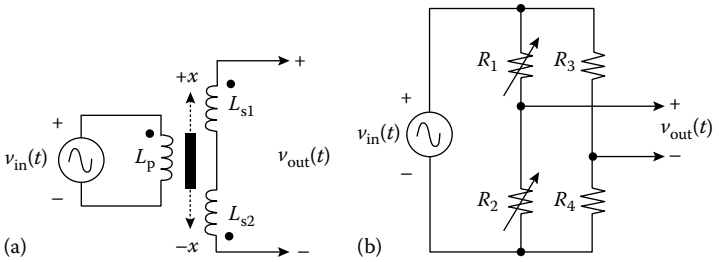


FIGURE 84.4 Balanced sensors producing suppressed-carrier DSB signals from ac excitation and physical measurands. In (a), an LVDT measures displacement x; in (b), a resistive Wheatstone bridge is unbalanced by a physical variable (e.g., mechanical strain).

to $v_{in}(t)$ indicates the direction of imbalance. In sensors like these, the modulating signal $f_m(t)$ is physical variable being measured. The use of ac excitation to a sensor is required for sensors containing reactive elements (inductances and/or capacitances); it may also be desirable in the case of resistive sensors (such as strain-gage bridges) requiring high-gain amplifiers. In these circumstances, amplification of an ac signal minimizes both $1/f$ noise and dc-offset problems associated with high-gain dc-coupled amplifiers.

84.3.3 Envelope Demodulation of DSB AM Signals

Full-carrier AM signals are readily demodulated by the simple envelope detector shown in Figure 84.5. The components of the RC low-pass filter are chosen such that $\omega_m << (1/RC) << \omega_c$. Envelope detection, however, cannot discriminate phase and is thus unsuitable for demodulation of signals in which phase reversal of the carrier occurs (such as reduced-carrier or suppressed-carrier signals). Synchronous demodulation is required for such signals.

84.3.4 Synchronous Demodulation of DSB AM Signals

Figure 84.6 shows three methods of synchronous demodulation of AM signals. In Figure 84.6a, the modulated signal $f_s(t)$ is multiplied by $A_d\cos(\omega_c t)$. In Figure 84.6b, the modulated signal $f_s(t)$ is gated by a square wave synchronous with $\cos(\omega_c t)$ so that the signal $f_s(t)$ is multiplied by a positive gain A during the positive half cycle of the carrier signal and by a negative gain $-A$ during the negative half cycle. Figure 84.6c shows a variation of this method in which the signal $f_s(t)$ is gated by a square wave such that it is amplified by a gain of A during the positive half cycle of $\cos(\omega_c t)$ and by 0 during the negative half cycle.

In the multiplying circuit of Figure 84.6a, the Fourier transform $F'_d(\omega)$ of the demodulated signal $f'_d(t) = f_s(t)A_d\cos(\omega_c t)$ is given by

$$F'_d(\omega) = \frac{A_d A_c k}{4}[\delta(\omega - 2\omega_c) + \delta(\omega + 2\omega_c) + 2\delta(\omega)] + \cdots$$

$$\cdots \frac{\mu A_d A_c}{4}[F_m(\omega - 2\omega_c) + F_m(\omega + 2\omega_c) + 2F_m(\omega)] \tag{84.6}$$

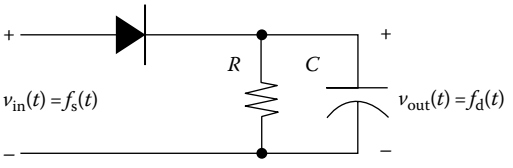


FIGURE 84.5 Simple envelope detector for full-carrier AM signals. The waveform $f_d(t)$ is the time-varying demodulated output.

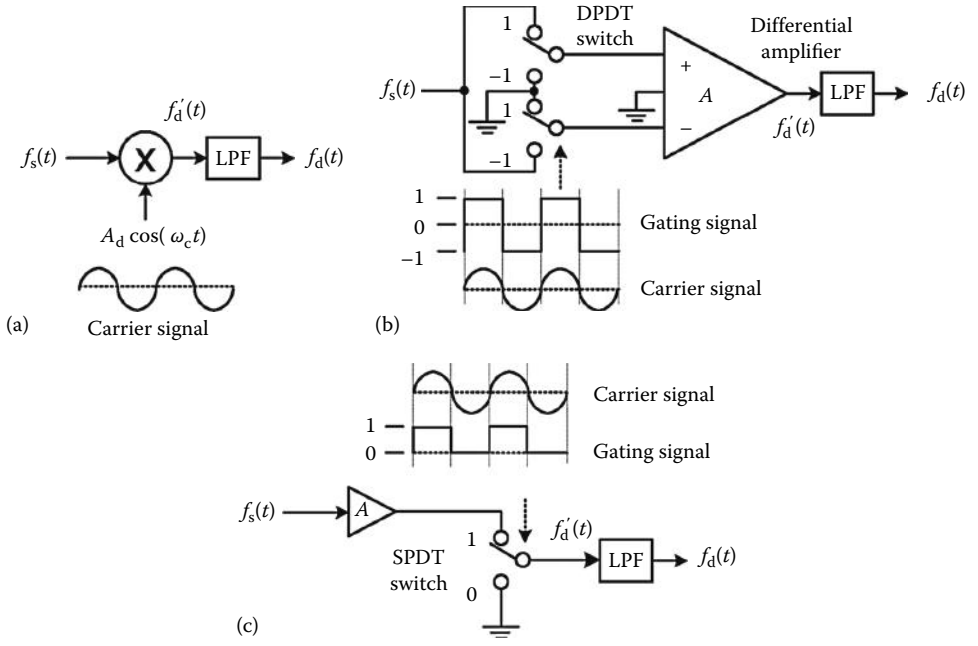


FIGURE 84.6 Synchronous detection methods for AM signals. The method of (a) is known as multiplying synchronous detection; those of (b) and (c) are methods of switching synchronous detection. The waveform $f_d'(t)$ is the time-varying demodulated output prior to filtering by the low-pass filter (labeled "LPF"). The waveform $f_d(t)$ is the filtered demodulated output. Switches are usually implemented with MOSFET transmission gates rather than mechanical switches.

It is assumed in Equation 84.6 that the input signal $f_s(t)$ is a generalized AM signal described by Equation 84.4. Suitable low-pass filtering after the multiplier removes the spectral components centered about $\pm 2\omega_c$. The Fourier transform $F_d(\omega)$ of the demodulated signal after filtering is

$$F_d(\omega) = \frac{A_d A_c k}{2}\delta(\omega) + \frac{\mu A_d A_c}{2} F_m(\omega) \tag{84.7}$$

Translating Equation 84.7 to the time domain gives

$$f_d(t) = \frac{A_d A_c}{2}[k + \mu f_m(t)] \tag{84.8}$$

The output of the multiplying synchronous detector consists of a dc term $A_d A_c k/2$ and the original baseband signal multiplied by a scale factor $\mu A_d A_c/2$. In the case of suppressed-carrier DSB ($k = 0$), the filtered detector output lacks the dc term and contains only a scaled version of the original baseband signal.

Figure 84.7 gives typical waveforms in a multiplying synchronous detector with a constant-amplitude sinusoidal signal. (This situation arises in balanced sensors when static measurand inputs are applied). It is of note that the output of the multiplier contains a dc term and a sinusoidal output at $2\omega_c$.

The gating action of the circuit of Figure 84.6b is equivalent to multiplying the original baseband signal by a square wave with levels $\pm A$. The Fourier-series representation of this square wave gating signal, $f_g(t)$, is

$$f_g(t) = \frac{4A}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)}\cos[(2n+1)\omega_c t] \tag{84.9}$$

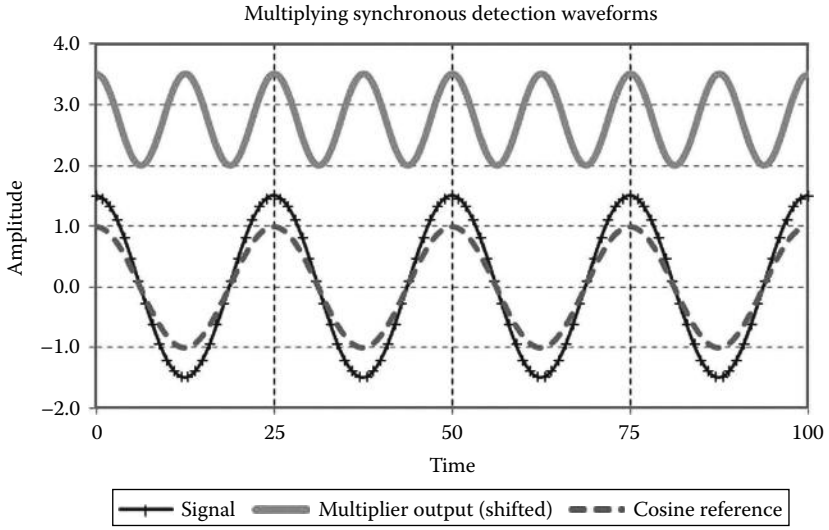


FIGURE 84.7 Waveforms in a multiplying synchronous demodulator when the signal to be demodulated is a constant-amplitude sinusoid in phase with the reference signal. The reference signal has been offset for clarity.

Multiplying an AM signal given by Equation 84.4 by the gating signal of Equation 84.9 and passing the product through a suitable low-pass filter gives

$$f_d(t) = \frac{2AA_c}{\pi}[k + \mu f_m(t)] \tag{84.10}$$

Comparison of Equation 84.10 with Equation 84.8 shows that synchronous detection of AM signals by multiplication or by switching produces similar results except for scaling factors.

Figure 84.8 shows typical waveforms in a switching-type synchronous demodulator of the type of Figure 84.6b when the input is a constant-amplitude sinusoid and the gain of the differential amplifier

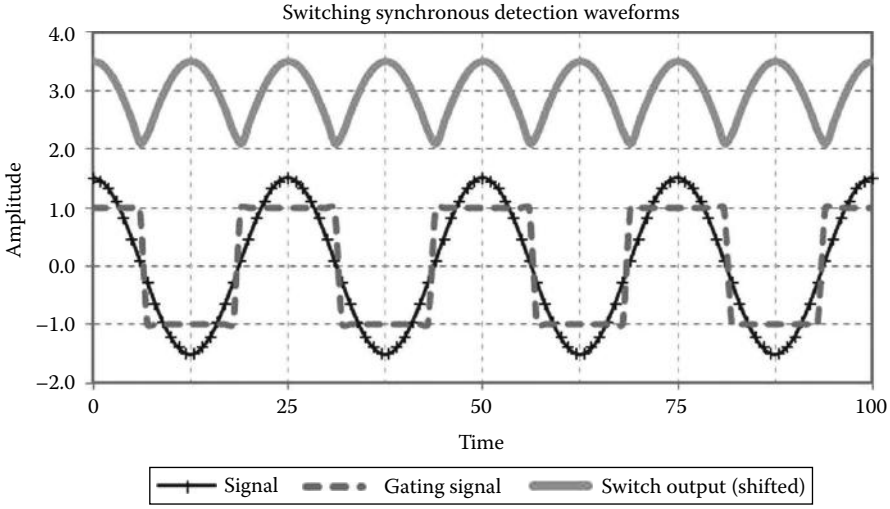


FIGURE 84.8 Waveforms in a full-wave synchronous demodulator when the input signal is a constant-amplitude sinusoid. The minimum value of the switch output waveform is 0; it has been offset for clarity.

is assumed to be 1. The output waveform of the switch has the form of a full-wave rectified sinusoid; this circuit is thus also known as a full-wave synchronous rectifier. The switch output waveform is periodic and nonsinusoidal with a fundamental frequency of $2\omega_c$. Low-pass filtering will extract the average (dc) value of the switch output while attenuating the time-varying components.

The gating function $f_g(t)$ of the circuit of Figure 84.6c is a square wave with levels A and 0. Its Fourier-series representation is

$$f_g(t) = \frac{A}{2} + \frac{2A}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)}\cos[(2n+1)\omega_c t] \tag{84.11}$$

Multiplication of this gating function with an AM signal described by Equation 84.4 and passing the product through a suitable low-pass filter gives

$$f_d(t) = \frac{AA_c}{\pi}[k + \mu f_m(t)] \tag{84.12}$$

Figure 84.9 shows typical waveforms in a switching-type synchronous demodulator of the type of Figure 84.6c when the input is a constant-amplitude sinusoid and the gain A of the amplifier is assumed to be 1. The output waveform of the switch has the form of a half-wave rectified sinusoid; this circuit is thus also known as a half-wave synchronous rectifier. The switch output waveform is periodic and non-sinusoidal with a fundamental frequency of ω_c.

Until this point, the implicit assumption has been made that the demodulating signal, whether sinusoidal (as in the case of the multiplying synchronous detector) or gating square wave (as in the case of switching synchronous detectors), is perfectly synchronized with the carrier of the modulated signal. However, perfect synchronization may not always be assumed. Multiplying a signal in carrier-quadrature form (as in Equation 84.2) by a demodulating signal $\cos(\omega_c t + \theta)$ and passing the result through a suitable low-pass filter gives

$$f_d(t) = \frac{f_i(t)}{2}\cos(\theta) + \frac{f_q(t)}{2}\sin(\theta) \tag{84.13}$$

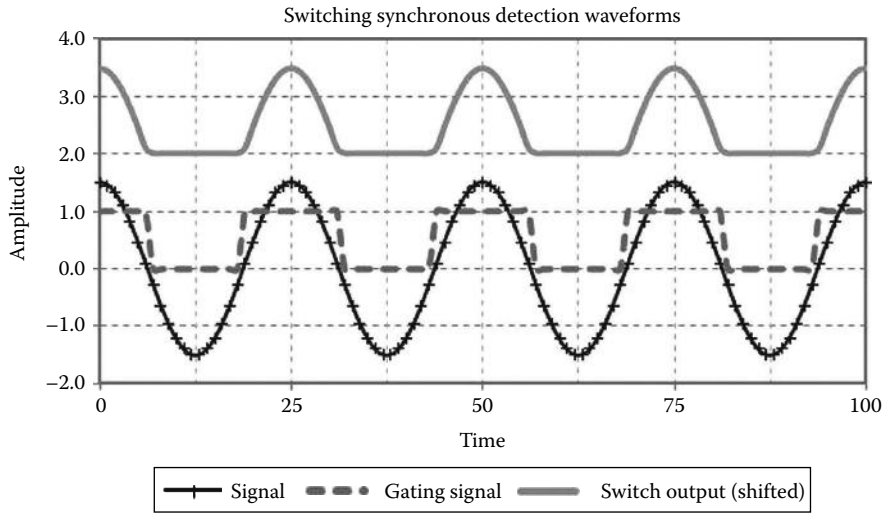


FIGURE 84.9 Waveforms in a half-wave synchronous demodulator when the input signal is a constant-amplitude sinusoid. The minimum value of the switch output waveform is 0; it has been offset for clarity.

Equation 84.13 is an important result; it demonstrates that both the magnitudes and the polarities of the components of the demodulated signal are functions of the synchronization in phase between the demodulating signal and the modulated carrier. Synchronous demodulation is thus often called phase-sensitive demodulation.

Balanced sensors with sinusoidal excitation (such as the LVDT and resistive Wheatstone bridge of Figure 84.4) produce modulated signals that are either in phase or inverted with respect to the carrier. In some measurements (e.g., electrical impedance), however, the phase between the modulated signal and the unmodulated carrier may take on a continuum of values, and information about the measurand may be contained in either or both the amplitude and phase of $f_s(t)$. A pair of phase-sensitive demodulators (PSD) may be employed to recover both amplitude and phase of $f_s(t)$; one uses a reference signal synchronous with $\cos(\omega_c t)$ while the other uses a reference signal synchronous with $\sin(\omega_c t)$. An example using half-wave synchronous rectifiers is shown in Figure 84.10.

Let $f_s(t)$ take the form $A_s\cos(\omega_c t + \theta)$ and be processed through the circuit of Figure 84.10. The in-phase gating signal $f_{GI}(t)$ has transitions synchronous with the zero crossings of $\cos(\omega_c t)$; quadrature gating signal $f_{GQ}(t)$ is synchronous with zero crossings of $\sin(\omega_c t)$. These gating signals may be expressed in terms of their Fourier series:

$$f_{GI}(t) = 0.5 + \frac{2}{\pi}\left[\cos(\omega_c t) - \frac{1}{3}\cos(3\omega_c t) + \frac{1}{5}\cos(5\omega_c t) - \frac{1}{7}\cos(7\omega_c t) + \cdots\right]$$

$$f_{GQ}(t) = 0.5 + \frac{2}{\pi}\left[\sin(\omega_c t) + \frac{1}{3}\sin(3\omega_c t) + \frac{1}{5}\sin(5\omega_c t) + \frac{1}{7}\sin(7\omega_c t) + \cdots\right] \tag{84.14a,b}$$

Multiplying $f_s(t)$ by the gating signals gives the switch output signals before filtering:

$$f_{I'}(t) = 0.5AA_s\cos(\omega_c t + \theta) + \cdots$$

$$\frac{AA_s}{\pi}\left\{\left[\cos(2\omega_c t + \theta) + \cos(\theta)\right] - \frac{1}{3}\left[\cos(4\omega_c t + \theta) + \cos(-2\omega_c t + \theta)\right] + \cdots\right\}$$

$$f_{Q'}(t) = 0.5AA_s\cos(\omega_c t + \theta) + \cdots \tag{84.15a,b}$$

$$\frac{AA_s}{\pi}\left\{\left[\sin(2\omega_c t + \theta) - \sin(\theta)\right] + \frac{1}{3}\left[\sin(4\omega_c t + \theta) - \sin(-2\omega_c t + \theta)\right] + \cdots\right\}$$

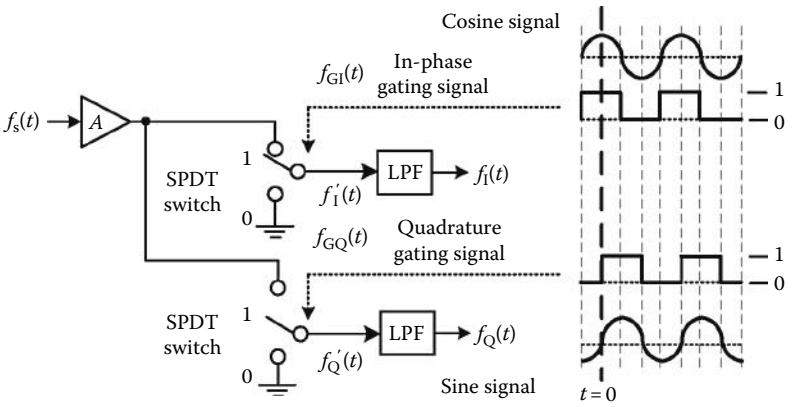


FIGURE 84.10 Phase-sensitive demodulator for arbitrary phase shifts between the unmodulated carrier and the modulated signal. Outputs $f_I(t)$ and $f_Q(t)$ are called the in-phase and quadrature outputs, respectively; the unfiltered signals at the outputs of the switches are designated as $f_I'(t)$ and $f_Q'(t)$.

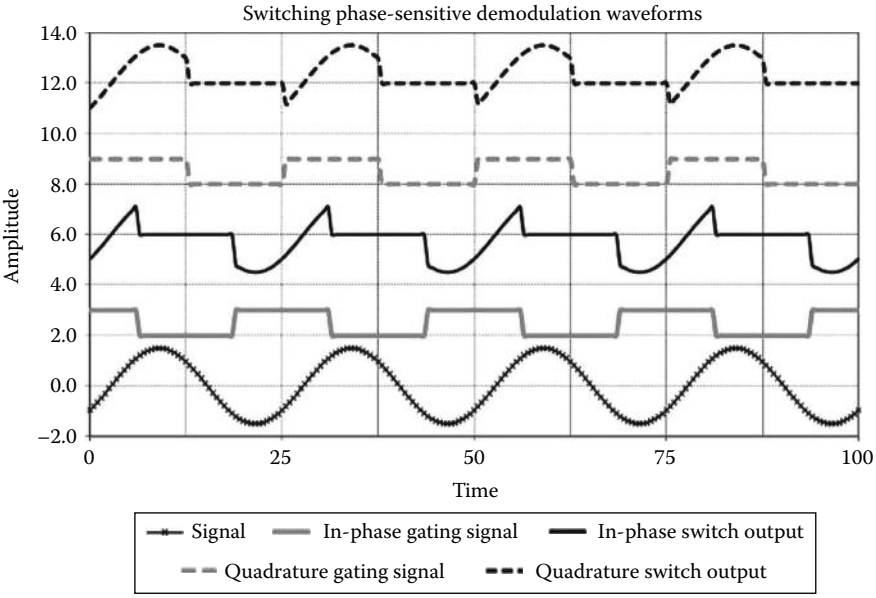


FIGURE 84.11 Example waveforms in the phase-sensitive demodulator of Figure 84.10. From bottom to top, the waveforms are the signal $f_s(t)$; the in-phase gating function $f_{GI}(t)$; the in-phase switch output prior to filtering $f_I'(t)$; the quadrature gating function $f_{GQ}(t)$; and the quadrature switch output prior to filtering $f_Q'(t)$. All waveforms except the signal waveform have been shifted vertically for clarity.

Equation 84.15 was derived using the trigonometric identities $\cos(a)\cos(b) = 0.5 \times [\cos(a + b) + \cos(a - b)]$ and $\cos(a)\sin(b) = 0.5 \times [\sin(a + b) - \sin(a - b)]$. Passing $f_I'(t)$ and $f_Q'(t)$ through low-pass filters having cutoff frequencies well below ω_c removes all but the dc components from the outputs $f_I(t)$ and $f_Q(t)$. The dc components are

$$f_I = \frac{AA_s}{\pi}\cos(\theta) \quad f_Q = -\frac{AA_s}{\pi}\sin(\theta) \tag{84.16}$$

The amplitude and phase of the modulated signal are found by the following equation:

$$A_s = \frac{\pi}{A}\sqrt{f_I^2 + f_Q^2} \quad \theta = \tan^{-1}\left(\frac{-f_Q}{f_I}\right) \tag{84.17a,b}$$

Figure 84.11 illustrates example waveforms in the PSD of Figure 84.10. In this case, the gain of the amplifier $A = 1$, the amplitude of the signal $A_s = 1.5$, and its phase θ relative to $\cos(\omega_c t)$ is –130. The calculated dc components of the switch output waveforms in Figure 84.11 are $f_I = -0.308$ and $f_Q = 0.367$. Using these values in Equation 84.17a,b gives a calculated amplitude of 1.505 and a phase angle of –130°.

Additional material on phase-sensitive demodulation in AM systems may be found in Pederson and Mayaram [7].

84.3.5 Applications of Phase-Sensitive Demodulation in Instrumentation Systems

It would be difficult to overstate the importance of phase-sensitive demodulation in instrumentation applications. There are numerous applications and developments described in the literature. For example, Munteanu et al. [8] describe the application of PSD to the measurement of electrical impedance.

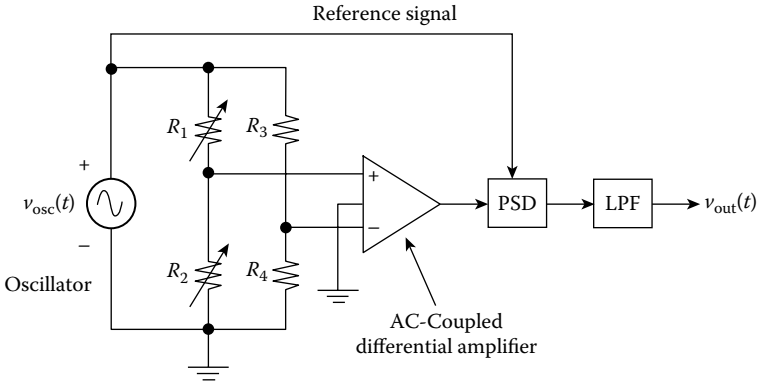


FIGURE 84.12 Carrier amplifier applied to measurement of imbalance in a resistive Wheatstone bridge. The oscillator provides a signal to operate the sensor, which itself may be measuring a number of possible measurands (e.g., mechanical strain). The block marked "PSD" is the phase-sensitive detector; "LPF" denotes a low-pass filter.

Syndenham [9] describes the use of PSD with LVDTs; Tarrant describes the use of PSD in optical systems, such as Faraday polarimeters and chopped optical systems [10]. The use of PSD to separate transformer voltage from signal voltage in electromagnetic flowmeters is described in [11]. A digital implementation of PSD for electrical capacitance tomography (ECT) applications is described in [12]. Ge and Lifeng [13] describe a digital PSD system for electric impedance tomography (EIT) applications. Tapson [14] describes a mixed-signal approach to PSD that provides benefits of precise bandwidth control, the ability to use square wave reference signals without problems of harmonic content, and elimination of need for a separate antialiasing filter that is required in a purely digital approach.

A particularly important application is the carrier amplifier. According to Pallàs-Areny and Webster [15], the carrier amplifier is a circuit that includes an oscillator circuit, ac-coupled amplification, phase-sensitive demodulation, and low-pass filtering. Figure 84.12 shows a carrier amplifier used for signal processing of the Wheatstone bridge of Figure 84.4. They further define a lock-in amplifier as a carrier amplifier in which the oscillator drives the system being measured rather than the sensor.

An example of a lock-in amplifier is the system for electrical-impedance plethysmography shown in Figure 84.13. In this case, a high-frequency ac current of constant magnitude is forced to flow through a part of the body. The current flow through the tissue produces a voltage which is picked up by sensing electrodes, amplified, and demodulated. Changes in perfusion in a limb or ventilation in the thorax provide small, but measurable, changes in electrical impedance $\Delta z(t)$ with respect to a much larger baseline impedance Z.

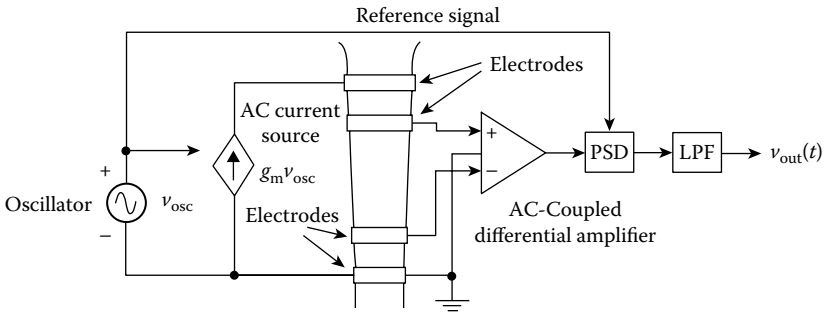


FIGURE 84.13 Lock-in amplifier employed for electrical-impedance plethysmography. The oscillator both controls the ac current source and provides a reference signal for the phase-sensitive demodulator.

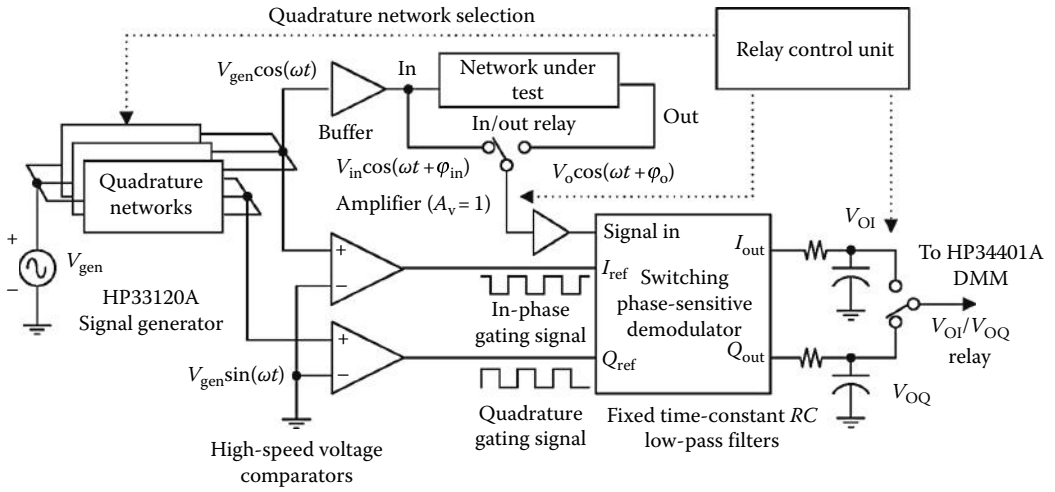


FIGURE 84.14 Lock-in amplifier methods using half-wave switching phase-sensitive demodulation in the Project TUNA instrument developed at the University of Texas at Tyler. The block labeled "Relay Control Unit" selects the appropriate quadrature network for the frequency of measurement, switches the input of the switching PSD between the input and output of the network under test, and switches the external DMM between the output voltage of the in-phase channel of the PSD (V_{OI}) and the output voltage of the quadrature channel (V_{OQ}).

As noted by Webster [16], high-frequency ac current passing through $Z + \Delta z(t)$ generates a full-carrier AM signal. This is unlike the signals produced by ac drive of balanced sensors, which are suppressed-carrier DSB signals. The full-carrier AM signal may be demodulated by a simple envelope detector. Phase-sensitive demodulation, however, gives superior performance because of its ability to reject interfering sources (e.g., pickup from 60 Hz power lines).

This ability to reject extraneous signals separated from the oscillator frequency allows carrier amplifiers to detect signals otherwise lost in noise or interfering signals. If the carrier oscillator frequency is ω_c and the bandwidth of the LPF extends from dc to an upper cutoff frequency ω_b, the carrier amplifier will reject all signals except those in the range $(\omega_c - \omega_b) < \omega < (\omega_c + \omega_b)$.

An example of lock-in amplifier methods coupled with quadrature detection is found in the Project TUNA (Texas Universal Network Analyzer) [17] and Project TUNA II [18] laboratory instruments developed for the electrical engineering laboratory at the University of Texas at Tyler. Both instruments use custom hardware and a LabVIEW (National Instruments, Austin, TX) virtual-instrument (VI) program. Both measure the frequency response (magnitude and phase shift) of the voltage gain of a linear network under sinusoidal, steady-state ac excitation, and both display these measurements in a Bode plot. Figure 84.14 shows the topology of the Project TUNA instrument. A signal generator (Agilent 33120A) generates a sinusoidal signal with frequency and amplitude controlled by the LabVIEW VI. The generator signal V_{gen} drives a quadrature network, an analog network that produces two sinusoidal outputs that maintain a 90° phase relationship over a wide range of frequencies. One output of the quadrature network is $V_{gen}\cos(\omega t)$ and the other output is $V_{gen}\sin(\omega t)$. The cosine signal is buffered and used to drive the network under test. Both the cosine and sine signals are passed through high-speed voltage comparators to produce square waves in-phase quadrature; these are the in-phase and quadrature gating signals of a half-wave switching phase-sensitive quadrature demodulator like that shown in Figure 84.10. The in-phase and quadrature output voltages of the PSD are read by an external digital multimeter (DMM) and these measurements are used to determine both the magnitude and the phase (relative to the cosine signal) of the signal input of the PSD. The magnitude and phase of both the input voltage $V_{in}\cos(\omega t + \varphi_{in})$ and output voltage $V_o\cos(\omega t + \varphi_o)$ of the network under test are measured, and the gain is computed and displayed. The switches of the PSD are implemented with ADG201HS

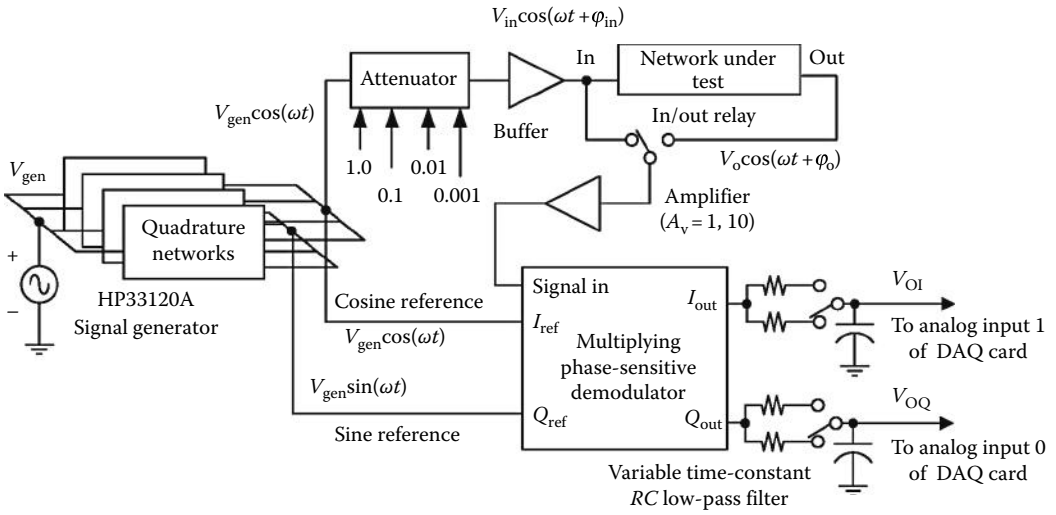


FIGURE 84.15 Lock-in amplifier methods using multiplying phase-sensitive demodulation in the Project TUNA II instrument developed at the University of Texas at Tyler. Improvements included a digitally controlled attenuator, a variable-gain amplifier, and variable time-constant low-pass filters.

high-speed analog switches (Analog Devices, Norwood, MA). A relay control unit, acting under control of the LabVIEW VI, activates the quadrature network appropriate to the frequency of the measurement, switches the input of the PSD to either the input or output of the network under test, and switches the DMM between the two outputs of the PSD.

The Project TUNA II instrument was an improved version of Project TUNA. Its chief differences were the use of a multiplying quadrature phase-sensitive detector, a variable-gain amplifier, and a variable attenuator under computer control. Analog Devices AD734 precision multipliers were used in the PSD. The topology of Project TUNA II is shown in Figure 84.15. The chief reason for the change to a multiplying PSD was to extend the maximal operating frequency of the instrument; charge injection in the analog switches of Project TUNA instrument limited the maximal usable frequency of Project TUNA to ~100 kHz.

Other topologies for lock-in amplifiers are also possible. The Analog Devices AD630 balanced modulator-integrated circuit makes a simple lock-in amplifier. It has on-chip precision resistances which allow the device to be used with a minimal number of external components. High-speed CMOS switches (such as the 74HC4053 or the ADG201HS mentioned in connection with Project TUNA) may be used in switching demodulators. Meade [19] provides detailed analysis of lock-in amplifiers; further information regarding lock-in amplifiers is found in [20,21].

84.4 Angle (Frequency and Phase) Modulation

In the preceding section, attention was paid to modulating the amplitude of a carrier signal. Information may be transmitted, however, by varying the phase angle in accordance with the baseband signal. Consider a signal of the form:

$$f_s(t) = A_c \cos[\omega_c t + \Delta\varphi \cdot x_m(t)] \tag{84.18}$$

The instantaneous phase of $f_s(t)$ is $\omega_c t + \Delta\varphi x_m(t)$; the phase relative to the unmodulated carrier is $\Delta\varphi x_m(t)$. The parameter $\Delta\varphi$ is the phase-modulation index (Carlson et al. [1]). The carrier is thus phase-modulated by $x_m(t)$. The instantaneous frequency of $f_s(t)$ is $\omega_c + \Delta\varphi[dx_m(t)/dt]$; phase modulation of the carrier also produces modulation of the carrier frequency unless $x_m(t)$ is a constant.

Assume that the modulating signal $x_m(t)$ is the time integral of the baseband signal $f_m(t)$. Then Equation 84.18 may be rewritten

$$f_s(t) = A_c \cos\left[\omega_c t + \Delta\varphi \cdot \int_0^t f_m(\tau)d\tau\right] \tag{84.19}$$

The instantaneous frequency $\omega_s(t)$ of $f_s(t)$ is $\omega_c + \Delta\varphi f_m(t)$. The instantaneous deviation in frequency in $\Delta\varphi f_m(t)$ is proportional to the baseband signal; the carrier is thus frequency-modulated by $f_m(t)$. This method of generating a frequency-modulated signal using a phase modulator controlled by the time integral of the baseband signal is known as indirect FM [22].

It is also possible to directly modulate the frequency of the carrier; in the case of the so-called direct FM, the expression for instantaneous frequency becomes

$$\omega_s(t) = \omega_c + \Delta\omega \cdot f_m(t) \tag{84.20}$$

Consider a baseband signal of the form $f_m(t) = A_m\cos(\omega_m t)$. Substituting this into Equation 84.19 gives

$$f_s(t) = A_c \cos\left[\omega_c t + \Delta\varphi \cdot A_m \int_0^t \cos(\omega\tau)d\tau\right] \tag{84.21}$$

Carrying out the integration in Equation 84.21 gives

$$f_s(t) = A_c \cos\left[\omega_c t + \frac{\Delta\varphi \cdot A_m}{\omega_m}\sin(\omega_m t)\right] = A_c \cos\left[\omega_c t + \beta\sin(\omega_m t)\right] \tag{84.22}$$

where β (called the modulation index) has replaced $\Delta\varphi A_m/\omega_m$. Figure 84.16 shows the time-domain representation of an FM signal and the original baseband signal. In this particular waveform, the unmodulated carrier frequency is 10 times the frequency of the sinusoidal baseband signal and the modulation index is 4.0.

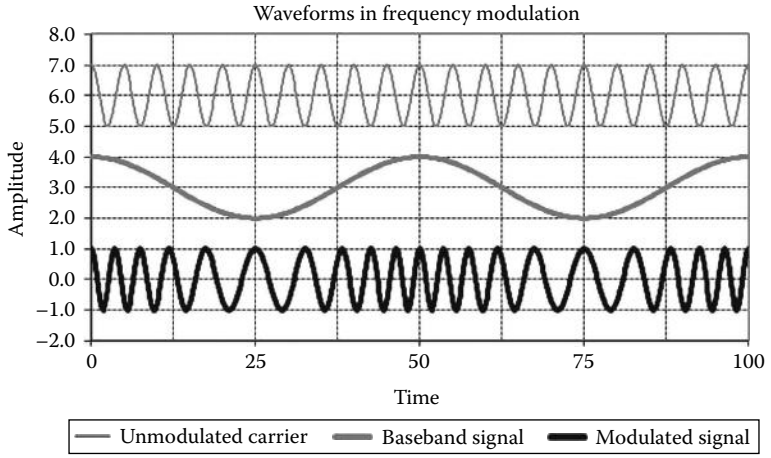


FIGURE 84.16 Frequency-modulated waveform. A_c = 1.0, A_m = 1.0, $\Delta\varphi$ = 0.5. The time scale is arbitrary, and the carrier and modulation waveforms have been offset for clarity.

The carrier-quadrature form of Equation 84.22 is

$$f_s(t) = A_c\left\{\cos[\beta\sin(\omega_m t)]\cos(\omega_c t) - \sin[\beta\sin(\omega_m t)]\sin(\omega_c t)\right\} \tag{84.23}$$

Unlike AM, an FM signal contains both in-phase and quadrature components, and the amplitudes of both are nonlinear functions of the modulation index. The terms $\cos[\beta\sin(\omega_m t)]$ and $\sin[\beta\sin(\omega_m t)]$ are generally expanded in terms of Bessel functions (see Schwartz [3] for detailed analysis). The Bessel function expression of an FM signal is

$$f_s(t) = A_c\sum_{n=-\infty}^{\infty} J_n(\beta)\cos[(\omega_c + n\omega_m)t] \tag{84.24}$$

where $J_n(\beta)$ represents a Bessel function of the first kind of order n. Beyer [23] provides tables of $J_0(\beta)$ and $J_1(\beta)$ and gives formulae for computing higher-order Bessel functions. Amplitudes for the carrier and sidebands through the fifth for modulation of a sinusoidal carrier with a sinusoidal baseband signal for selected values of β are given in Table 84.1.

It is also noted that $J_{-n}(\beta) = (-1)^n J_n(\beta)$. The approximations $J_0(\beta) = 1$ and $J_1(\beta) = \beta/2$ are valid for small values of the modulation index ($\beta < 0.2$). A carrier with sinusoidal FM and a low modulation index will show sidebands spaced at $\pm\omega_m$ about the carrier. Thus narrowband FM would be indistinguishable from full-carrier AM on a spectrum analyzer (which shows the amplitudes of spectral components, but not their phase relationships). As modulation index increases, however, new sideband pairs appear at $\pm2\omega_m$, $\pm3\omega_m$, $\pm4\omega_m$, etc., as the higher-order Bessel functions become more significant. The amplitude of the carrier component of $f_s(t)$ varies with $J_0(\beta)$; the carrier, in fact, disappears completely for certain values of the modulation index. These characteristics are unlike AM in which the carrier component of the modulated signal is constant and in which only one sideband pair is produced for each spectral component in the baseband signal. FM is inherently a nonlinear modulation process; the sum of multiple FM signals derived from a single carrier with multiple baseband signals does not give the same result as frequency modulation of the carrier by the sum of the baseband signals. The spectrum of a phase-modulated signal is similar to that of an FM signal, but the modulation index of a phase-modulated signal does not vary with ω_m.

TABLE 84.1 Carrier and Sideband Amplitudes for Frequency Modulation with a Sinusoidal Baseband Signal

Modulation Index	Carrier	f_1	f_2	f_3	f_4	f_5
0	1.0	—	—	—	—	—
0.25	0.98	0.12	0.01	—	—	—
0.5	0.94	0.24	0.03	—	—	—
1	0.77	0.44	0.11	0.02	—	—
1.5	0.51	0.56	0.23	0.06	0.01	—
2.0	0.22	0.58	0.35	0.13	0.03	0.01
5.0	−0.18	−0.33	0.05	0.36	0.39	0.26

Sources: The data produced are compiled from Seymour, K., *The Electronics Handbook*, Whitaker, J.C., Ed., CRC Press, Boca Raton, FL, 1996; Carlson, A.B. et al., *Communication Systems*, 4th ed., McGraw-Hill, New York, 2002.

84.4.1 Generation of Phase- and Frequency-Modulated Signals

Variable-frequency signals may arise directly in instrumentation applications from sources like turbine-type flowmeters or Doppler velocity sensors. Direct FM signals may be generated by applying the baseband signal to a voltage-controlled oscillator; Sherwin and Regan [24] apply this technique in a system using the LM566 VCO to transmit analog information over 60 Hz power lines. In radio-frequency applications, the oscillator frequency may be varied by application of the baseband signal to a voltage-variable reactance (typically a varactor diode). Indirect FM may be generated by PM of the carrier with the baseband signal integrated with respect to time as in Equation 84.21; DeFrance [25] and Comer and Comer [26] give examples of phase-modulator circuits.

84.4.2 Demodulation of Phase- and Frequency-Modulated Signals

PM signals may be demodulated by phase-sensitive detectors. Limiter circuits may be employed to transform analog signals to digital waveforms prior to the phase detector. This has the advantage of making the output of the phase detector insensitive to variations in amplitude of the signal prior to the limiter. The use of digital waveforms in a PSD also makes linear the relationship of output voltage to phase difference.

Simple digital circuits like those of Figure 84.17 can serve to detect the phase difference between two digital signals. In Figure 84.17a, an exclusive OR (XOR) logic gate produces logic 1 at its output Y only if

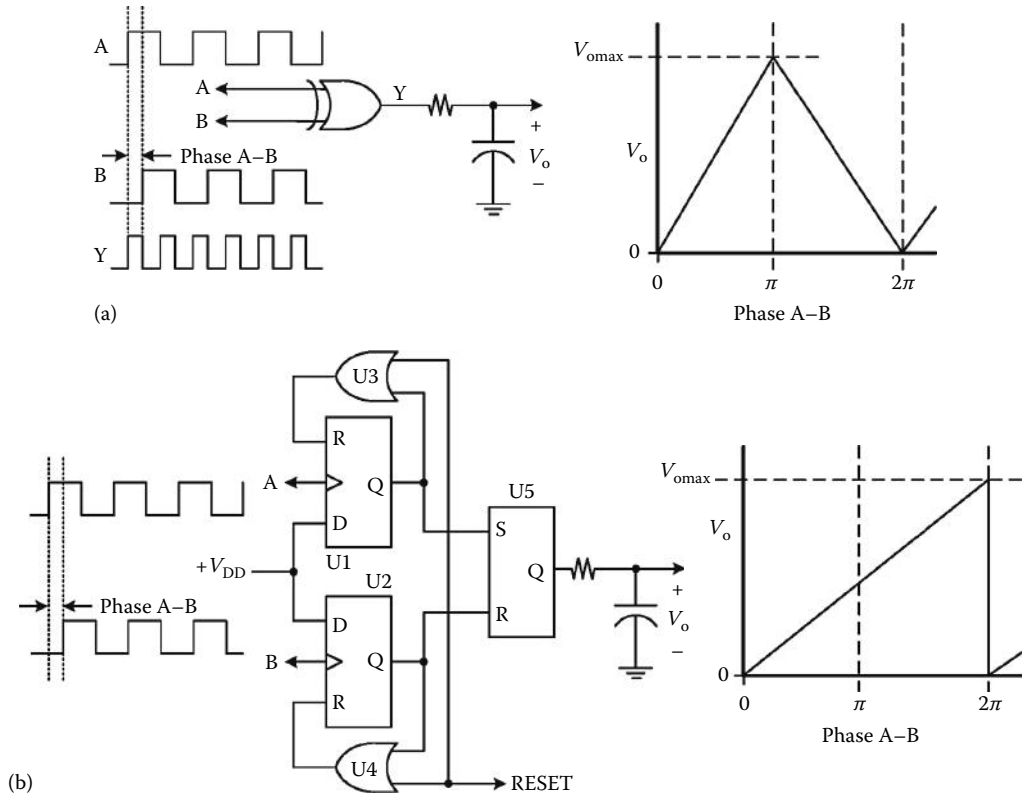


FIGURE 84.17 Examples of digital phase detectors. In both cases, the maximal output voltage V_{omax} is equal to the high-level output voltage of the logic devices. In each circuit, it is assumed that signals A and B have been hard-limited to suppress amplitude variations and to produce digital waveforms of the appropriate logic levels.

logic inputs A and B are different. The low-pass filter produces from the pulse train at output Y at voltage proportional to phase difference between signals A and B. However, the proportionality becomes negative for phase shifts between π and 2π radians. In Figure 84.17b, flip-flop U1 and OR-gate U3 produce a short logic 1 pulse on the rising edge of A; this pulse sets SR flip-flop U5. Flip-flop U2 and OR-gate U4 likewise produce a short pulse on the rising edge of B, causing U5 to be reset. The low-pass filter produces a voltage proportional to the phase difference between signals A and B. The RESET input may be used to force the initial states of U1 and U2.

Figure 84.18 shows examples of three common FM detectors. Figure 84.18a is a quadrature detector of the type frequently used in integrated-circuit FM receivers. Figure 84.18b is a frequency-to-voltage converter utilizing a monostable multivibrator, while Figure 84.18c shows a block diagram of a phase-locked loop (PLL).

In the quadrature detector of Figure 84.18a, a limiter circuit suppresses amplitude variations in the modulated signal and the amplitude-limited signal is passed to a PSD as its reference signal. The limiter output signal is also coupled (via a small capacitor C_1) to a resonant network consisting

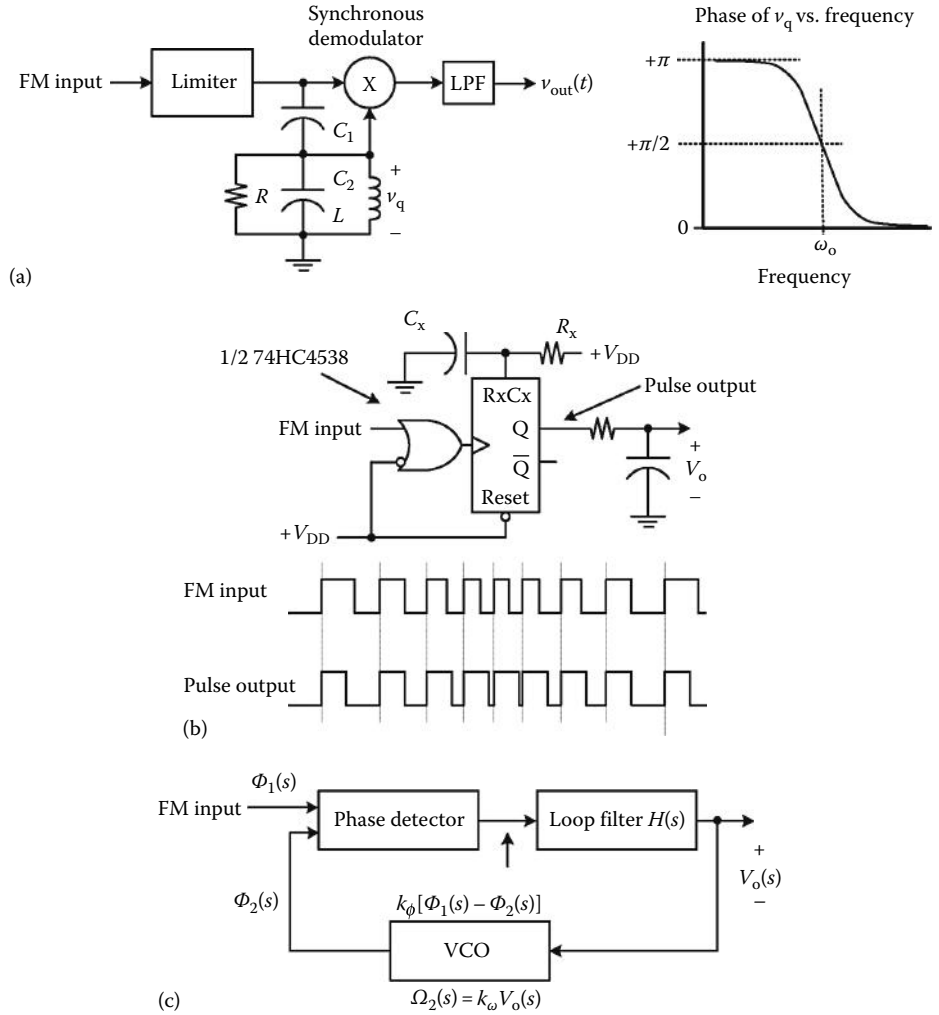


FIGURE 84.18 Examples of FM demodulators.

of L, R, and C_2. The phase of the voltage across the quadrature network relative to the limiter output, $\varphi(\omega)$, is given by

$$\varphi(\omega) = \pi - \tan^{-1} \frac{\omega L/R}{[1-(\omega/\omega_0)^2]} \tag{84.25}$$

where ω_0 is

$$\omega_0 = \frac{1}{\sqrt{L(C_1 + C_2)}} \tag{84.26}$$

This relationship is illustrated in the graph associated with Figure 84.18b. The variation in phase is nearly linear for frequencies close to ω_0. The baseband signal is recovered from the output of the phase-sensitive detector. The voltage across the quadrature network also varies with frequency, but there is generally sufficient voltage across the quadrature network to force the PSD into a nonlinear regime in which its output voltage is relatively insensitive to the amplitude of the voltage across the quadrature network.

Figure 84.18b shows a frequency-to-voltage converter consisting of a monostable multivibrator (one-shot) triggered on each rising edge of the digital input waveform; the low-pass filter recovers the baseband signal from the resulting pulse train.

Figure 84.18c shows FM demodulation by means of a PLL. The phase-comparator circuit produces an output voltage proportional to the difference in phase between the input signal and the output signal of the VCO. The output voltage of the phase comparator is filtered and used to drive the VCO; the filtered output voltage of the phase comparator is also the demodulated output signal.

The operation of the PLL may be explained in the following. The phase and frequency of the FM input signal are designated as $\varphi_1(t)$ and $\omega_1(t)$, respectively, and have Laplace transforms $\Phi_1(s)$ and $\Omega_1(s)$. The phase and frequency of the VCO output are represented by $\varphi_2(t)$ and $\omega_2(t)$; their Laplace transforms are $\Phi_2(s)$ and $\Omega_2(s)$. The phase detector produces an output voltage $v_\varphi(t)$ given by

$$v_\varphi(t) = k_\varphi[\varphi_1(t) - \varphi_2(t)] \tag{84.27}$$

The Laplace transform of $v_\varphi(t)$ is designated $V_\varphi(s)$ and is given by

$$V_\varphi(s) = k_\varphi[\Phi_1(s) - \Phi_2(s)] \tag{84.28}$$

This voltage is filtered by the loop filter having transfer function $H(s)$ and is applied to the control voltage input of the VCO. The VCO produces an output signal with a frequency proportional to the control voltage. The Laplace transform of the VCO frequency is

$$\Omega_2(s) = k_\omega k_\varphi H(s)[\Phi_1(s) - \Phi_2(s)] \tag{84.29}$$

Since $\Phi_1(s) = \Omega_1(s)/s$ and $\Phi_2(s) = \Omega_2(s)/s$, appropriate substitutions and rearrangement of Equation 84.29 gives

$$\Omega_2(s) = \left[\frac{k_\omega k_\varphi H(s)}{s + k_\omega k_\varphi H(s)}\right]\Omega_1(s) \tag{84.30}$$

If the input frequency to the PLL is a step function of magnitude ω_1, the Laplace transform of the VCO frequency is

$$\Omega_2(s) = \left[\frac{k_\omega k_\varphi H(s)}{s + k_\omega k_\varphi H(s)}\right]\frac{\omega_1}{s} \tag{84.31}$$

Application of the final-value theorem of the Laplace transform to Equation 84.31 gives $\omega_2 = \omega_1$. In other words, the output frequency of the VCO matches the input frequency. The VCO input voltage is ω_1/k_ω and is directly proportional to the input frequency. If the input signal is a carrier frequency-modulated by some baseband signal, the filtered phase-detector output follows that baseband signal. Taub and Schilling [2] give an extensive analysis of the PLL, as do Pederson and and Mayaram [7] and O'Dell [27].

The simplest loop filter simply applies the output of the phase detector to the input of the VCO and has $H(s) = 1$. If this were done, the transfer function $\Omega_2(s)/\Omega_1(s)$ would have a single-pole low-pass response with a cutoff frequency $\omega_{-3\text{dB}} = k_\omega k_\varphi$. The VCO output frequency tracks slowly varying input frequencies well, but the response to rapidly varying input frequencies is limited by the low-pass transfer function of the loop. Baseband spectral components which lie beyond the cutoff frequency of the loop are attenuated. A single-pole *RC* network is frequently used as the loop filter in a PLL. The transfer function of a PLL with such a loop filter is

$$\frac{\Omega_2(s)}{\Omega_1(s)} = \left[\frac{k_\omega k_\varphi \omega_\text{f}}{s^2 + s\omega_\text{f} + k_\omega k_\varphi \omega_\text{f}}\right] \tag{84.32}$$

where $\omega_\text{f} = 1/(RC)$ is the cutoff frequency of the *RC* low-pass filter. It is noted that the loop transfer function $\Omega_2(s)/\Omega_1(s)$ is a second-order system and thus may exhibit dynamics such as resonance and underdamped transient response, depending upon the values of the parameters k_φ, k_ω, and ω_f.

84.4.3 Examples

Two common applications in instrumentation involving FM signals are Doppler radar velocimetry and distance measurement using FM continuous-wave (FMCW) radar.

Doppler radar velocimetry measures relative velocity between an emitter of microwave energy and a reflecting object by taking advantage of the Doppler frequency shift. Figure 84.19 shows a simple Doppler radar velocity sensor. A microwave source (typically a MOSFET or Gunn diode) supplies energy to the feedpoint of a high-gain microwave antenna. Some of the emitted energy is

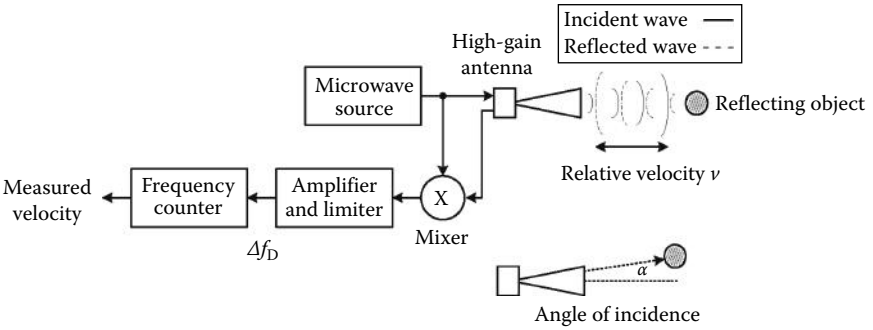


FIGURE 84.19 Microwave Doppler velocity sensor.

reflected from nearby objects, and the reflected signal is mixed with the emitted signal to produce the sum and difference of the transmitted and received frequencies. If the object and the microwave emitter are in relative motion, the reflected signal returns to the emitter shifted in frequency by the Doppler shift frequency Δf_D:

$$\Delta f_D = 2\left(\frac{v}{c}\right) f_c \cos(\alpha) \tag{84.33}$$

where

v is the relative velocity of the source and the reflecting object (m/s)
c is the speed of light in air (3.0×10^8 m/s)
f_c is the frequency of the microwave (carrier) source (Hz)
α is the angle of incidence between the source and the reflector

For K-band microwave sources (24 GHz) and head-on incidence ($\alpha = 0$), the Doppler frequency shift is 160 Hz/(m/s). One drawback of the system of Figure 84.19 is that it cannot distinguish whether the Doppler shift is positive (indicating that the source and reflector are approaching each other) or negative (indicating that the source and reflector are moving away from each other).

The Doppler velocity sensor uses a constant-frequency carrier; modulation of the carrier frequency takes place through nonelectronic means. The Doppler velocity sensor may be used to measure distance traveled through integration of the measured velocity, but it cannot measure distance or detect stationary objects. FMCW radar uses a frequency-modulated carrier and detects the difference between the instantaneous transmitted frequency and the received frequency. An example of an FMCW distance-measurement system is shown in Figure 84.20.

The frequency of the carrier is swept between its minimum and maximum frequency values f_{cmin} and f_{cmax} in a triangular pattern at a rate $\pm\Delta f_c/\Delta t$. The transmitted signal strikes the reflecting object

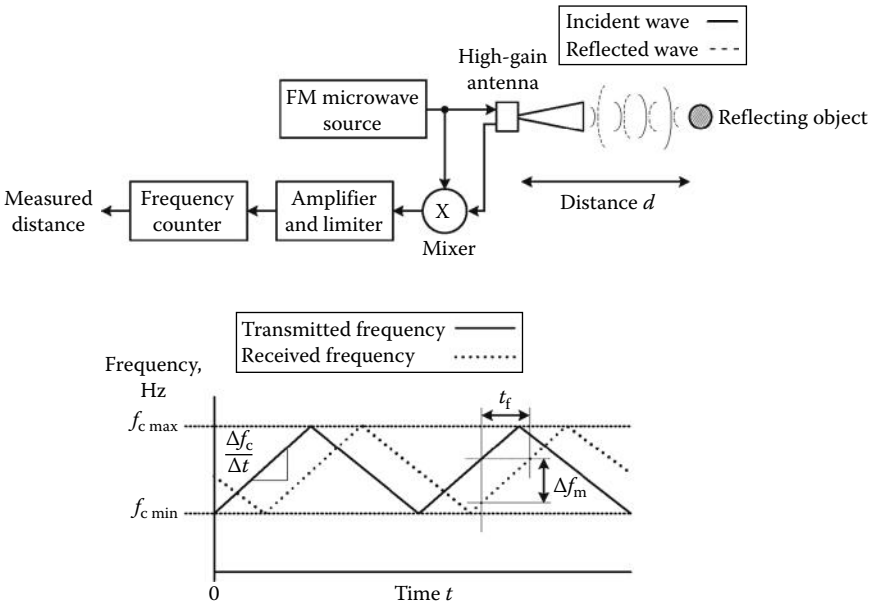


FIGURE 84.20 FMCW distance sensor. (The diagram of carrier frequency vs. time is adapted from Daniewicz, J.L. and Boyes, W.H., in Lipták, B.G., Ed., *The Instrument Engineer's Handbook. Process Measurement and Analysis*, 4th ed., Vol. 1, CRC Press, Boca Raton, FL, 2003.)

at a distance d from the source and returns to the source after a total time of flight t_f, which is the time required for the signal to traverse a distance of $2d$. During the time of flight, the carrier frequency will have changed by $(\pm\Delta f_c/\Delta t)t_f$, assuming that the carrier frequency sweep has not changed polarity during that time. Mixing the transmitted and received signals produces a measured difference frequency Δf_m:

$$\Delta f_m = t_f \frac{\Delta f_c}{\Delta t} \tag{84.34}$$

Since $t_f = 2d/c$ (where c is the speed of light in air), the distance may be expressed as

$$d = \frac{c\Delta f_m}{2\left(\Delta f_c/\Delta t\right)} \tag{84.35}$$

However, as Figure 84.20 makes evident, Equations 84.34 and 84.35 are valid only during those times that the frequency sweep rates of both the transmitted and reflected signals have the same polarity. FMCW distance sensors find wide use in material-handling applications (e.g., sensing the liquid level in a tank or reservoir).

Cohen et al. [29] in Figure 84.21 describe a novel use of a direct FM technique to perform noninvasive monitoring of human ventilation. Elastic belts with integral serpentine inductors encircle the chest and

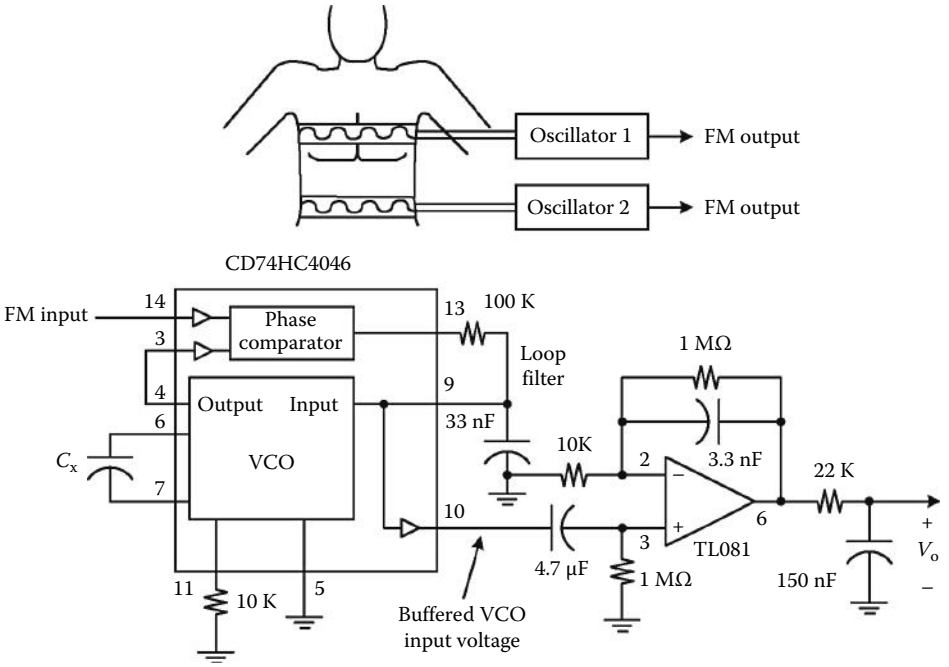


FIGURE 84.21 Direct FM techniques applied to the measurement of human ventilation. (Adapted from Cohen, K.P. et al., *Physiol. Meas.*, 15, 217, 1994.) One of two PLL demodulator circuits is shown. The capacitor C_x was chosen specifically for each demodulator depending upon the nominal frequency of each VCO. Pin numbers are shown adjacent to each integrated circuit. Optical coupling was used between the oscillator outputs and demodulator inputs to provide electrical isolation between the subject and the measurement system.

abdomen of the subject; these inductors are components of resonant tank circuits of free-running radio-frequency oscillators. Changes in chest and abdominal volume during respiration cause the inductors' cross-sectional areas to vary, changing their inductances and thus varying the frequencies of the oscillators. Figure 84.21 also shows one of the two PLL demodulator circuits. The PLLs utilize single-pole *RC* low-pass loop filters.

84.5 Components and Instrumentation

84.5.1 Integrated Circuits

Table 84.2 lists certain integrated circuits which may be used in application of the modulation techniques covered in this chapter. This is neither an exhaustive list of all useful types nor of all manufacturers. Most integrated circuits are available in a number of packages and performance grades; the prices given are indicative of the pricing of the least expensive versions purchased in small quantities. The AD532, AD633, AD734 four-quadrant multipliers may be used to make balanced modulator and multiplying demodulator circuits. The AD650 may be used as either a voltage-to-frequency or a frequency-to-voltage converter. The ADG201HS is the high-speed analog switch used as the switching element of the PSD of the Project TUNA instrument described earlier. Depending upon the maximal frequency, any number of analog switches from various sources may be suitable.

Lock-in amplifiers for general laboratory instrumentation purposes are available from several manufacturers, as listed in Table 84.3.

A partial list of instruments using modulation techniques is given in Table 84.4. A partial list manufactures of integrated circuits for modulation is given in Table 84.5 and manufacturers that make modulated sources and analyzers can be found in Table 84.6.

TABLE 84.2 Integrated Circuits Used in Modulation Applications

Designation	Function	Manufacturer
AD630JNZ	Balanced modulator	Analog Devices
AD532JHZ	Analog multiplier	Analog Devices
AD633JNZ	Analog multiplier	Analog Devices
AD734ANZ	Analog multiplier	Analog Devices
LM565	PLL	National Semiconductor
74HC4046	PLL	Various
AD650JNZ	V-F, F-V converter	Analog Devices
AD652JPZ	V-F converter	Analog Devices
AD654JNZ	V-F converter	Analog Devices
ADG201HSJNZ	High-speed analog switch	Analog Devices
74HC4538	Dual-retriggerable monostable	Various
NE564	PLL	ST Microelectronics

Note: This is not intended as a comprehensive or exhaustive list.

TABLE 84.3 Manufacturers of General-Purpose Lock-In Amplifiers

Manufacturer	Location	URL
EG&G Princeton Applied Research	Oak Ridge, TN	www.princetonappliedresearch.com
Stanford Research Systems	Sunnyvale, CA	www.thinksrs.com
Bentham Instruments	Reading, Berkshire, U.K.	www.bentham.co.uk

TABLE 84.4 Instruments Utilizing Modulation Techniques

Manufacturer	Model Number	Description
Hewlett-Packard	HP 53310A	Modulation domain analyzer
	HP 5371A	Frequency and time-interval analyzer
	HP 5372A	Frequency and time-interval analyzer
	HP 5373A	Modulation domain pulse analyzer
	HP 8780A	Vector signal generator
	HP 8782B	Vector signal generator
	HP 8981B	Vector modulation analyzer
	HP 11736B	I/Q modulation tutorial software
I/Q Tutor	HP 89410A	Vector signal analyzer
	HP 89440A	Vector signal analyzer
	HP 89441A	Vector signal analyzer
	HP 11715A	AM/FM test source
	HP 8901A	Modulation analyzer
	HP 8901B	Modulation analyzer
Tektronix	1720	Vectorscope
	1721	Vectorscope
	1725	Vectorscope
Rohde & Schwarz	DS 1200	TV demodulator
	SME-02	Signal generator
	SME-03	Signal generator
	SMHU-58	Signal generator
	SMT-02	Signal generator
	SMY-01	Signal generator
	SMY-02	Signal generator

TABLE 84.5 Companies that Make Integrated Circuits for Modulation

Analog Devices, Inc.
One Technology Way
Box 9106
Norwood, MA 02062
(617) 329-4700

Harris Corp. Semiconductor Products Division
Box 883
Melbourne, FL 37902
(407) 724-3730

Motorola, Inc., Semiconductor Products Sector
3102 N. 56th St.
Phoenix, AZ 85018
(602) 952-3248

National Semiconductor Corp.
2900 Semiconductor Dr.
Box 58090
Santa Clara, CA 95052-8090

Philips Components—Signetics
811 E. Arques
Sunnyvale, CA 94088
(408) 991-2000; (800) 227-1817

TABLE 84.6 Companies that Make Modulated Sources and Analyzers

Hewlett-Packard Co.
Test and Measurement Sector
Box 58199
Santa Clara, CA 95052-9943
(800) 452-4844

Tektronix Inc. Corporate Offices
26600 SW Parkway
Box 1000
Wilsonville, OR 97070-1000
(503) 682-3411; (800) 426-2200

Rohde & Schwarz Inc.
4425 Nicole Dr.
Lanham, MD 20706
(301) 459-8800

References

1. A. B. Carlson, P. B. Crilly, and J. C. Rutledge, *Communication Systems*, 4th ed. New York: McGraw-Hill, 2002.
2. H. Taub and D. L. Schilling, *Principles of Communication Systems*. New York: McGraw-Hill, 1971.
3. M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed. New York: McGraw-Hill, 1990.
4. W. Tomasi, *Advanced Electronic Communications Systems*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1992.
5. H. Y. Malmstadt, C. G. Enke, and S. R. Crouch, *Electronics and Instrumentation for Scientists*. Menlo Park, CA: Benjamin/Cummings Publishing Company, 1981.
6. E. O. Doebelin, *Measurement Systems Application and Design*, 4th ed. New York: McGraw-Hill, 1990.
7. D. O. Pederson and K. Mayaram, *Analog Integrated Circuits for Communication: Principles, Simulation and Design*, 2nd ed. Berlin, Germany: Springer, 2010.
8. R. Munteanu, G. Todoran, and R. Copindean, Measurement of circuit parameters R, X, G, B, through a method of phase sensitive demodulation. *Optimization of Electrical and Electronic Equipments, 1998. OPTIM '98. Proceedings of the 6th International Conference on*, Braşov, Romania, Vol. 3, pp. 727–730, May 14–15, 1998.
9. P. H. Syndenham, Electrical magnetic inductive processes, in W. Boyes, Ed., *The Instrumentation Reference Book*, 4th ed. Amsterdam, The Netherlands: Elsevier, 2010.
10. A. W. S. Tarrant, Polarimeters and detector techniques, in W. Boyes, Ed., *The Instrumentation Reference Book*, 4th ed. Amsterdam, The Netherlands: Elsevier, 2010.
11. J. G. Kopp, B. G. Lipták, and H. Eren, Magnetic flowmeters, in B. G. Lipták, Ed., *The Instrument Engineer's Handbook. Process Measurement and Analysis*, 4th ed., Vol. 1. Boca Raton, FL: CRC Press, 2003.
12. X. Zhang and H. Wang, Digital phase-sensitive demodulation in electrical capacitance tomography system, 2008. *WCICA 2008. 7th World Congress on Intelligent Control and Automation,* Chongqing, China, pp. 6730–6733, June 25–27, 2008.
13. K. Ge and R. Lifeng, FPGa-based digital phase-sensitive demodulator for EIT system, 2007. *ICEMI '07. 8th International Conference on Electronic Measurement and Instruments,* Xian, China, pp. 4-845–4-848, August 16, 2007–July 18, 2007.

14. J. Tapson, Mixed signal phase sensitive detection. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS),* Paris, France, pp. 1292–1295, May 30, 2010–June 2, 2010.
15. R. Pallàs-Areny and J. G. Webster, *Sensors and Signal Conditioning*, 2nd ed. New York: John Wiley & Sons, 2001.
16. J. G. Webster, Measurement of flow and volume of blood, in J. G. Webster, Ed., *Medical Instrumentation, Application and Design*, 4th ed. New York: John Wiley & Sons, 2009.
17. D. M. Beams, *Project TUNA: The Development of a LabVIEW Virtual Instrument as a Class Project in a Junior-Level Electronics Course*. Presented at the 2000 Annual Conference of the American Society for Engineering Education, St. Louis, MO, June 18–21, 2000.
18. D. M. Beams, *Project TUNA II—Bode Analyzer and Teaching Tool*. Presented at the 2006 Annual Conference of the American Society for Engineering Education, Chicago, IL, June 18–21, 2006.
19. M. L. Meade, *Lock-in Amplifiers: Principles and Applications*. London, U.K.: Peregrinus, 1984.
20. W. J. Witteman, *Detection and Signal Processing: Technical Realization* (Springer Series in Advanced Microelectronics). Berlin, Germany: Springer, 2010.
21. Bentham Instruments, Lock-in amplifiers. http://www.bentham.co.uk/pdf/F225.pdf, retrieved January 14, 2011.
22. K. Seymour, FM, *The Electronics Handbook*, J. C. Whitaker, Ed. Boca Raton, FL: CRC Press, 1996.
23. W. H. Beyer, *CRC Standard Mathematical Tables*, 28th ed. Boca Raton, FL: CRC Press, 1987.
24. J. Sherwin and T. Regan, FM remote speaker system, National Semiconductor Application Note AN-146, in *National Semiconductor Corp. 1991 Linear Applications Databook*, 1991.
25. J. J. DeFrance, *Communication Electronics Circuits*, 2nd ed. San Francisco, CA: Rinehart Press, 1972.
26. D. J. Comer and D. T. Comer, *Advanced Electronic Circuit Design*. New York: John Wiley & Sons, 2003.
27. T. H. O'Dell, *Circuits for Electronic Instrumentation*. Cambridge, U.K.: Cambridge University Press, 1991.
28. J. L. Daniewicz and W. H. Boyes, Radar, noncontacting level sensors, in B. G. Lipták, Ed., *The Instrument Engineer's Handbook. Process Measurement and Analysis*, 4th ed., Vol. 1. Boca Raton, FL: CRC Press, 2003.
29. K. P. Cohen, D. Panescu, J. H. Booske, J. G. Webster, and W. J. Tompkins, Design of an inductive plethysmograph for ventilation measurement, *Physiol. Meas*., 15, 217–229, 1994.

85

Filters

Rahman Jamal
National Instruments Germany

Robert Steer
Frequency Devices

85.1 Introduction

In its broadest sense, a filter can be defined as a signal-processing system whose output signal, usually called the *response*, differs from the input signal, called the *excitation*, such that the output signal has some prescribed properties. In more practical terms, an electric filter is a device designed to suppress, pass, or separate a group of signals from a mixture of signals according to the specifications in a particular application. The application areas of filtering are manifold, for example to band-limit signals before sampling to reduce aliasing, to eliminate unwanted noise in communication systems, to resolve signals into their frequency components, to convert discrete-time signals into continuous-time signals, to demodulate signals, etc. Filters are generally classified into three broad classes: *continuous-time*, *sampled-data*, and *discrete-time* filters depending on the type of signal being processed by the filter. Therefore, the concepts of signals are fundamental in the design of filters.

A *signal* is a function of one or more independent variables such as time, space, temperature, etc. that carries information. The independent variables of a signal can either be continuous or discrete. Assuming that the signal is a function of time, in the first case the signal is called continuous time and in the second, discrete time. A continuous-time signal is defined at every instant of time over a given interval, whereas a discrete-time signal is defined only at discrete-time instances. Similarly, the *values* of a signal can also be classified in either continuous or discrete.

In real-world signals, often referred to as analog signals, both amplitude and time are continuous. These types of signals cannot be processed by digital machines unless they have been converted into discrete-time signals. In contrast, a digital signal is characterized by discrete signal values that are defined only at discrete points in time. Digital signal values are represented by a finite number of digits, which are usually binary-coded. The relationship between a continuous-time signal and the corresponding discrete-time signal can be expressed in the following form:

$$x(kT) = x(t)_{t=kT}, \quad k = 0, 1, 2, \ldots \tag{85.1}$$

where T is called the sampling period.

Filters can be classified on the basis of the input, output, and internal operating signals. A continuous data filter is used to process continuous-time or analog signals, whereas a digital filter processes digital signals. Continuous data filters are further divided into *passive* or *active* filters, depending on the type of elements used in their implementation. Perhaps, the earliest type of filters known in the engineering community are *LC* filters, which can be designed by using discrete components like inductors and capacitors, or crystal and mechanical filters that can be implemented using *LC* equivalent circuits. Since no external power is required to operate these filters, they are often referred to as *passive* filters. In contrast, *active* filters are based on active devices, primarily *RC* elements, and amplifiers. In a sampled-data filter, on the other hand, the signal is sampled and processed at discrete instants of time. Depending on the type of signal processed by such a filter, one may distinguish between an *analog sampled-data* filter and a *digital* filter. In an analog sampled-data filter, the sampled signal can principally take any value, whereas in a digital filter the sampled signal is a digital signal, the definition of which was given earlier. Examples of analog sampled-data filters are switched capacitor (SC) filters and charge-transfer device (CTD) filters made of capacitors, switches, and operational amplifiers.

85.2 Filter Classification

Filters are commonly classified according to the filter function they perform. The basic functions are low pass, high pass, bandpass, and bandstop. If a filter passes frequencies from zero to its cutoff frequency Ω_c and stops all frequencies higher than the cutoff frequencies, then this filter type is called an ideal *low-pass filter*. In contrast, an ideal *high-pass filter* stops all frequencies below its cutoff frequency and passes all frequencies above it. Frequencies extending from Ω_1 to Ω_2 are passed by an ideal *bandpass filter*, while all other frequencies are stopped. An ideal bandstop filter stops frequencies from Ω_1 to Ω_2 and passes all other frequencies. Figure 85.1 depicts the magnitude functions of the four basic *ideal filter* types.

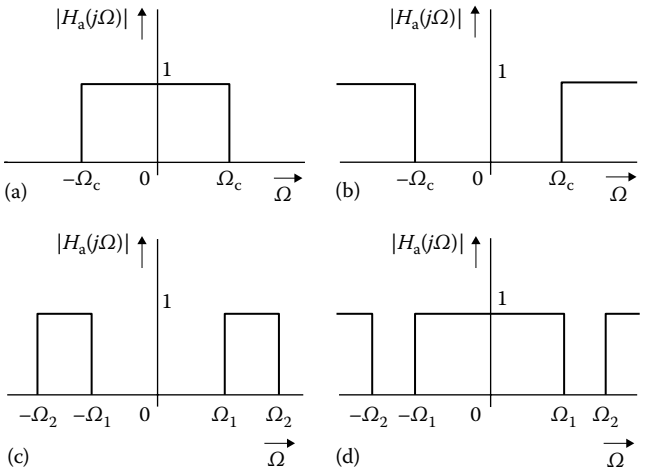


FIGURE 85.1 The magnitude function of an ideal filter is 1 in the passband and 0 in the stopband as shown for (a) low-pass, (b) high-pass, (c) bandpass, and (d) stopband filters.

So far, we have discussed ideal filter characteristics having rectangular magnitude responses. These characteristics, however, are physically not realizable. As a consequence, the ideal response can only be approximated by some nonideal realizable system. Several classical approximation schemes have been developed, each of which satisfies a different criterion of optimization. This should be taken into account when comparing the performance of these filters characteristics.

85.3 Filter Approximation Problem

Generally, the input and output variables of a linear, time-invariant, causal filter can be characterized either in the time domain through the convolution integral given by

$$y(t) = \int_0^t h_a(t-\tau)x(t)d\tau \tag{85.2}$$

or, equivalently, in the frequency domain through the transfer function

$$H_a(s) = \frac{Y(s)}{X(s)} = \frac{\sum_{i=0}^{N} b_i s^i}{1 + \sum_{i=0}^{N} a_i s^i} \Leftrightarrow H_a(s) = \frac{b_N}{a_N} \prod_{i=1}^{N} \left(\frac{s - s_{0i}}{s - s_{\infty i}} \right) \tag{85.3}$$

where

$H_a(s)$ is the Laplace transform of the impulse response $h_a(t)$ and $X(s)$
$Y(s)$ are the Laplace transforms of the input signal $x(t)$ and the output or the filtered signal $y(t)$
$X(s)$ and $Y(s)$ are polynomials in $s = \sigma + j\Omega$
The overall transfer function $H_a(s)$ is a real rational function of s with real coefficients

The zeroes of the polynomial $X(s)$ given by $s = s_{\infty i}$ are called the poles of $H_a(s)$ and are commonly referred to as the *natural frequencies* of the filter. The zeros of $Y(s)$ given by $s = s_{0i}$ which are equivalent to the zeroes of $H_a(s)$ are called the *transmission zeros* of the filter. Clearly, at these frequencies the filter output is zero for any finite input. Stability restricts the poles of $H_a(s)$ to lie in the left half of the s-plane excluding the $j\Omega$-axis, that is, $\mathrm{Re}\{s_{\infty i}\} < 0$. For a stable transfer function, $H_a(s)$ reduces to $H_a(j\Omega)$ on the $j\Omega$-axis, which is the continuous-time Fourier transform of the impulse response $h_a(t)$ and can be expressed in the following form:

$$H_a(j\Omega) = |H_a(j\Omega)|\, d^{j\theta(\Omega)} \tag{85.4}$$

where

$|H_a(j\Omega)|$ is called the magnitude function
$\theta(\Omega) = \arg H_a(j\Omega)$ is the phase function

The gain magnitude of the filter expressed in decibels (dB) is defined by

$$\alpha(\Omega) = 20\log|H_a(j\Omega)| = 10\log|H_a(j\Omega)|^2 \tag{85.5}$$

Note that a filter specification is often given in terms of its attenuation, which is the negative of the gain function also given in decibels. While the specifications for a desired filter behavior are commonly given

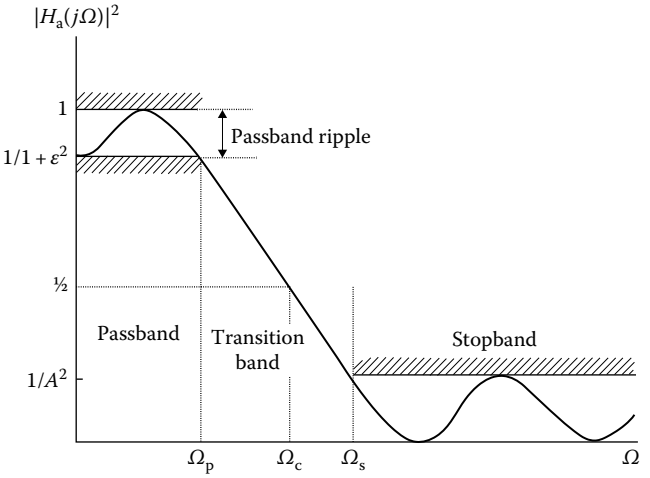


FIGURE 85.2 The squared magnitude function of an analog filter can have ripple in the passband and in the stopband.

in terms of the loss response $\alpha(\Omega)$, the solution of the filter approximation problem is always carried out with the help of the characteristic function $C(j\Omega)$ giving

$$\alpha(\Omega) = 10\log[1+|C(j\Omega)|^2] \tag{85.6}$$

Note that $\alpha(\Omega)$ is not a rational function, but $C(j\Omega)$ can be a polynomial or a rational function and approximation with polynomial or rational functions is relatively convenient. It can also be shown that frequency-dependent properties of $|C(j\Omega)|$ are in many ways identical to those of $\alpha(\Omega)$. The approximation problem consists of determining a desired response $|H_a(j\Omega)|$ such that the typical specifications depicted in Figure 85.2 are met. This so-called tolerance scheme is characterized by the following parameters:

- Ω_p Passband cutoff frequency (rad/s)
- Ω_s Stopband cutoff frequency (rad/s)
- Ω_c –3 dB cutoff frequency (rad/s)
- ε Permissible error in passband given by $\varepsilon = (10^{r/10} - 1)^{1/2}$, where r is the maximum acceptable attenuation in dB; note that $10\log(1)/(1+\varepsilon^2)^{1/2} = -r$
- $1/A$ Permissible maximum magnitude in the stopband, i.e., $A = 10^{\alpha/20}$, where α is the minimum acceptable attenuation in dB; note that $20\log(1/A) = -\alpha$

The *passband* of a low-pass filter is the region in the interval $[0, \Omega_p]$ where the desired characteristics of a given signal are preserved. In contrast, the *stopband* of a low-pass filter (the region $[\Omega_s, \infty]$) rejects signal components. The *transition* band is the *region* between $(\Omega_\times - \Omega_p)$ which would be 0 for an ideal filter. Usually, the amplitudes of the permissible ripples for the magnitude response are given in decibels.

The following sections review four different classical approximations: Butterworth, Chebyshev Type I, elliptic, and Bessel.

85.3.1 Butterworth Filters

The frequency response of an Nth-order Butterworth low-pass filter is defined by the squared magnitude function:

$$|H_a(j\Omega)|^2 = \frac{1}{1+(\Omega/\Omega_c)^{2N}} \tag{85.7}$$

It is evident from Equation 85.7 that the Butterworth approximation has only poles, i.e., no finite zeros and yields a maximally flat response around zero and infinity. Therefore, this approximation is also called maximally flat magnitude (MFM). In addition, it exhibits a smooth response at all frequencies and a monotonic decrease from the specified cutoff frequencies.

Equation 85.7 can be extended to the complex *s*-domain, resulting in

$$H_a(s)H_a(-s) = \frac{1}{1+(s/j\Omega_c)^{2N}} \tag{85.8}$$

The poles of this function are given by the roots of the denominator

$$s_k = \Omega_c e^{j\pi[1/2+(2k+1)/2N]}, \quad k = 0, 1, \ldots, 2N-1 \tag{85.9}$$

Note that for any N, these poles lie on the unit circle of radius Ω_c in the s-plane. To guarantee stability, the poles that lie in the left half-plane are identified with $H_a(s)$. As an example, we will determine the transfer function corresponding to a third-order Butterworth filter, i.e., $N = 3$.

$$H_a(s)H_a(-s) = \frac{1}{1+(-s^2)^3} = \frac{1}{1-s^6} \tag{85.10}$$

The roots of denominator of Equation 85.10 are given by

$$s_k = \Omega_c e^{j\pi[1/2+(2k+1)/6]}, \quad k = 0, 1, 2, 3, 4, 5 \tag{85.11}$$

Therefore, we obtain

$$\begin{aligned}
s_0 &= \Omega_c e^{j\pi 2/3} = -1/2 + j\sqrt{2}/2 \\
s_1 &= \Omega_c e^{j\pi} = -1 \\
s_2 &= \Omega_c e^{j\pi 4/3} = -1/2 - j\sqrt{3}/2 \\
s_3 &= \Omega_c e^{j\pi 5/3} = 1/2 - j\sqrt{3}/2 \\
s_4 &= \Omega_c e^{j2\pi} = 1 \\
s_5 &= \Omega_c e^{j\pi/3} = 1/2 + j\sqrt{3}/2
\end{aligned} \tag{85.12}$$

The corresponding transfer function is obtained by identifying the left half-plane poles with $H_a(s)$. Note that for the sake of simplicity we have chosen $\Omega_c = 1$:

$$H_a(s) = \frac{1}{(s+1)(s+1/2-j\sqrt{3}/2)(s+1/2+j\sqrt{3}/2)} = \frac{1}{1+2s+2s^2+s^3} \tag{85.13}$$

Table 85.1 gives the Butterworth denominator polynomials up to $N = 5$.

Table 85.2 gives the Butterworth poles in real and imaginary components and in frequency and Q.

TABLE 85.1 Butterworth Denominator Polynomials

Order (N)	Butterworth Denominator Polynomials of $H(s)$
1	$s + 1$
2	$s^2 + \sqrt{2}s + 1$
3	$s^3 + 2s^2 + 2s + 1$
4	$s^4 + 2.6131s^3 + 3.4142s^2 + 2.6131s + 1$
5	$s^5 + 3.2361s^4 + 5.2361s^3 + 5.2361s^2 + 3.2361s + 1$

TABLE 85.2 Butterworth and Bessel Poles

	Butterworth Poles				Bessel Poles (−3 dB)			
N	Re a	Im$(\pm j)b$	Ω	Q	Re a	Im$(\pm j)b$	Ω	Q
1	−1.000	0.000	1.000	—	−1.000	0.000	1.000	—
2	−0.707	0.707	1.000	0.707	−1.102	0.636	1.272	0.577
3	−1.000	0.000	1.000	—	−1.323	0.000	1.323	—
	−0.500	0.866	1.000	1.000	−1.047	0.999	1.448	0.691
4	−0.924	0.383	1.000	0.541	−1.370	0.410	1.430	0.522
	−0.383	0.924	1.000	1.307	−0.995	1.257	1.603	0.805
5	−1.000	0.000	1.000	—	−1.502	0.000	1.502	—
	−0.809	0.588	1.000	0.618	−1.381	0.718	1.556	0.564
	−0.309	0.951	1.000	1.618	−0.958	1.471	1.755	0.916
6	−0.966	0.259	1.000	0.518	−1.571	0.321	1.604	0.510
	−0.707	0.707	1.000	0.707	−1.382	0.971	1.689	0.611
	−0.259	0.966	1.000	1.932	−0.931	1.662	1.905	1.023
7	−1.000	0.000	1.000	—	−1.684	0.000	1.684	—
	−0.901	0.434	1.000	0.555	−1.612	0.589	1.716	0.532
	−0.623	0.782	1.000	0.802	−1.379	1.192	1.822	0.661
	−0.223	0.975	1.000	2.247	−0.910	1.836	2.049	1.126
8	−0.981	0.195	1.000	0.510	−1.757	0.273	1.778	0.506
	−0.831	0.556	1.000	0.601	−1.637	0.823	1.832	0.560
	−0.556	0.831	1.000	0.900	−1.374	1.388	1.953	0.711
	−0.195	0.981	1.000	2.563	−0.893	1.998	2.189	1.226

In the next example, the order N of a low-pass Butterworth filter is to be determined whose cutoff frequency (–3 dB) is $\Omega_c = 2$ kHz and stopband attenuation is greater than 40 dB at $\Omega_s = 6$ kHz. Thus the desired filter specification is

$$20\log|H_a(j\Omega)| \le -40, \quad \Omega \ge \Omega_s \tag{85.14}$$

or equivalently,

$$|H_a(j\Omega)| \le 0.01, \quad \Omega \ge \Omega_s \tag{85.15}$$

It follows from Equation 85.7

$$\frac{1}{1+(\Omega_s/\Omega_c)^{2N}} = (0.01)^2 \tag{85.16}$$

Solving the earlier equation for N gives $N = 4.19$. Since N must be an integer, a fifth-order filter is required for this specification.

85.3.2 Chebyshev Filters or Chebyshev I Filters

The frequency response of an Nth-order Chebyshev low-pass filter is specified by the squared-magnitude frequency response function:

$$|H_a(j\Omega)|^2 = \frac{1}{1+\varepsilon^2 T_N^2(\Omega/\Omega_p)} \tag{85.17}$$

where

$T_N(x)$ is the Nth-order Chebyshev polynomial

ε is a real constant <1 which determines the ripple of the filter

Specifically, for nonnegative integers N, the Nth-order Chebyshev polynomial is given by

$$T_N(x) = \begin{cases} \cos(N\cos^{-1}x), & |x| \le 1 \\ \cosh(N\cos h^{-1}x), & |x| \ge 1 \end{cases} \tag{85.18}$$

High-order Chebyshev polynomials can be derived from the recursion relation:

$$T_{N+1}(x) = 2xT_N(x) - T_{N-1}(x) \tag{85.19}$$

where

$T_0(x) = 1$
$T_1(x) = x$

The Chebyshev approximation gives an *equiripple* characteristic in the passband and is maximally flat near infinity in the stopband. Each of the Chebyshev polynomials has real zeros that lie within the interval (–1, 1) and the function values for $x\varepsilon[-1, 1]$ do not exceed +1 and –1.

The pole locations for Chebyshev filter can be determined by generating the appropriate Chebyshev polynomials, inserting them into Equation 85.17, factoring, and then selecting only the left half plane roots. Alternatively, the pole locations P_k of an Nth-order Chebyshev filter can be computed from the relation, for $k = 1 \rightarrow N$

$$P_k = -\sin\Theta_k \sin h\beta + j\cos\Theta_k \cos h\beta \tag{85.20}$$

where

$\Theta_k = (2k - 1)\pi/2N$
$\beta = \sin h^{-1}(1/\varepsilon)$

Note: P_{N-k+1} and P_k are complex conjugates and when N is odd there is one real pole at

$$P_{N+1} = -2\sin h\beta$$

For the Chebyshev polynomials, Ω_p is the last frequency where the amplitude response passes through the value of ripple at the edge of the passband. For odd N polynomials, where the ripple of the Chebyshev polynomial is negative going, it is the $[-1/(1 + \varepsilon^2)]^{(1/2)}$ frequency and for even N, where the ripple is positive going, it is the 0 dB frequency.

The Chebyshev filter is completely specified by the three parameters: ε, Ω_p, and N. In a practical design application, ε is given by the permissible passband ripple and Ω_p is specified by the desired passband cutoff frequency. The order of the filter, i.e., N, is then chosen such that the stopband specifications are satisfied.

85.3.3 Elliptic or Cauer Filters

The frequency response of an Nth-order elliptic low-pass filter can be expressed by

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \varepsilon^2 F_N^2(\Omega/\Omega_p)} \tag{85.21}$$

where $F_N(\cdot)$ is called the Jacobian elliptic function. The elliptic approximation yields an equiripple passband and an equiripple stopband. Compared with the same-order Butterworth or Chebyshev filters, the elliptic design provides the sharpest transition between the passband and the stopband. The theory of elliptic filters, initially developed by Cauer, is involved; therefore for an extensive treatment, refer to Ref. [1].

Elliptic filters are completely specified by the parameters ε, α, Ω_p, Ω_s, and N where ε is passband ripple, α is stopband floor, Ω_p is the frequency at the edge of the passband (for a designated

passband ripple), Ω_s is the frequency at the edge of the stopband (for a designated stopband floor), and N is the order of the polynomial.

In a practical design exercise, the desired passband ripple, stopband floor, and Ω_s are selected and N is determined and rounded up to the nearest integer value. The appropriate Jacobian elliptic function must be selected and $H_a(j\Omega)$ must be calculated and factored to extract only the left plane poles. For some synthesis techniques, the roots must be expanded into polynomial form.

This process is a formidable task. While some filter manufacturers have written their own computer programs to carry out these calculations, they are not readily available. However, the majority of applications can be accommodated by use of published tables of the pole/zero configurations of low-pass elliptic transfer functions. An extensive set of such tables for a common selection of passband ripples, stopband floors, and shape factors is available in Ref. [2].

85.3.4 Bessel Filters

The primary objectives of the preceding three approximations were to achieve specific loss characteristics. The phase characteristics of these filters, however, are nonlinear. The Bessel filter is optimized to reduce nonlinear-phase distortion, i.e., a maximally flat delay. The transfer function of a Bessel filter is given by

$$H_a(s) = \frac{B_0}{B_N(s)} = \frac{B_0}{\sum_{k=0}^{N} B_k s^k}, \quad B_k = \frac{(2N-k)!}{2^{N-k}k!(N-k)!} \quad k = 0, 1, \ldots, N \tag{85.22}$$

where $B_N(s)$ is the Nth-order Bessel polynomial. The overall squared-magnitude frequency response function is given by

$$|H_a(j\Omega)|^2 = 1 - \frac{\Omega^2}{2N-1} + \frac{2(N-1)\Omega^4}{(2N-1)^2(2N-3)} + \cdots \tag{85.23}$$

To illustrate Equation 85.22, the Bessel transfer function for $N = 4$ is given as follows:

$$H_a(s) = \frac{105}{105 + 105s + 45s^2 + 10s^3 + s^4} \tag{85.24}$$

Table 85.2 lists the factored pole frequencies as real and imaginary parts and as frequency and Q for Bessel transfer functions that have been normalized to $\Omega_c = -3$ dB.

85.4 Design Examples for Passive and Active Filters

85.4.1 Passive *R, L, C* Filter Design

The simplest and most commonly used passive filter is the simple, first-order ($N = 1$) RC filter shown in Figure 85.3. Its transfer function is that of a first-order Butterworth low-pass filter. The transfer function and −3 dB Ω_c are

$$H_a(s) = \frac{1}{RCs + 1} \quad \text{where } \Omega_c = \frac{1}{RC} \tag{85.25}$$

While this is the simplest possible filter implementation, both source and load impedance change the dc gain and/or corner frequency and its roll-off rate is only first order, or −6 dB/octave.

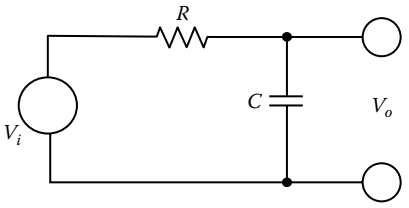


FIGURE 85.3 A passive first-order *RC* filter can serve as an antialiasing filter or to minimize high-frequency noise.

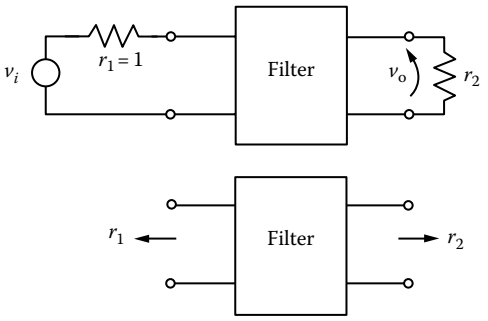


FIGURE 85.4 A passive filter can have the symbolic representation of a doubly terminated filter.

To realize higher-order transfer functions, passive filters use *R*, *L*, *C* elements usually configured in a ladder network. The design process is generally carried out in terms of a doubly terminated two-port network with source and load resistors R_1 and R_2 as shown in Figure 85.4. Its symbolic representation is given later.

The source and load resistors are normalized in regard to a reference resistance $R_B = R_1$, i.e.,

$$r_i = \frac{R_1}{R_B} = 1, \quad r_2 = \frac{R_2}{R_B} = \frac{R_2}{R_1} \tag{85.26}$$

The values of *L* and *C* are also normalized with respect to a reference frequency to simplify calculations. Their values can be easily scaled to any desired set of actual elements:

$$l_v = \frac{\Omega_B L_v}{R_B}, \quad c_v = \Omega_B C_v R_B \tag{85.27}$$

Low-pass filters, whose magnitude-squared functions have no finite zero, i.e., whose characteristic functions $C(j\Omega)$ are polynomials, can be realized by lossless ladder networks consisting of inductors as the series elements and capacitors as the shunt elements. These types of approximations, also referred to as *all-pole approximations*, include the previously discussed Butterworth, Chebyshev Type I, and Bessel filters. Figure 85.5 shows four possible ladder structures for even and odd *N*, where *N* is the filter order.

In the case of doubly terminated Butterworth filters, the normalized values are precisely given by

$$a_v = 2\sin\left(\frac{(2v-1)\pi}{2N}\right), \quad v = 1, \ldots, N \tag{85.28}$$

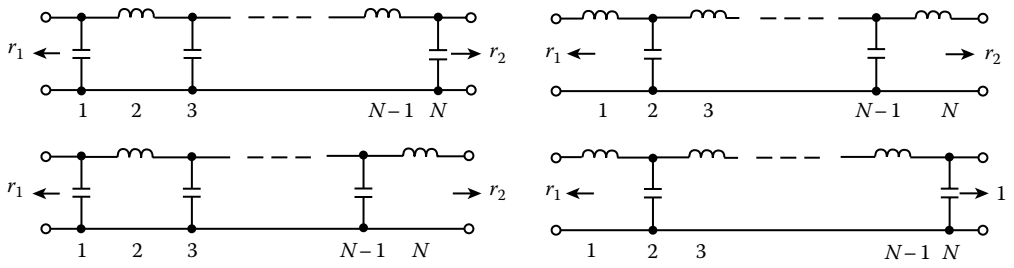


FIGURE 85.5 Even and odd N passive all-pole filter networks can be realized by several circuit configurations. N odd, above; N even, below.

where a_v is the normalized L or C element value. As an example, we will derive two possible circuits for a doubly terminated Butterworth low pass of order 3 with $R_B = 100\ \Omega$ and a cutoff frequency $\Omega_c = \Omega_B = 10$ kHz. The element values from Equation 85.28 are

$$l_1 = 2\sin\left(\frac{(2-1)\pi}{6}\right) = 1 \Rightarrow L_1 = \frac{R_B}{\Omega_c} = 1.59 \text{ mH}$$

$$c_2 = 2\sin\left(\frac{(4-1)\pi}{6}\right) = 2 \Rightarrow C_2 = \frac{2}{\Omega_c R_B} = 3.183 \text{ nF} \tag{85.29}$$

$$l_3 = 2\sin\left(\frac{(6-1)\pi}{6}\right) = 1 \Rightarrow L_3 = \frac{R_B}{\Omega_c} = 1.59 \text{ mH}$$

A possible realization is shown in Figure 85.6.

Table 85.3 gives normalized element values for the various all-pole filter approximations discussed in the previous section up to order 3 and is based on the following normalization:

1. $r_1 = 1$
2. All the cutoff frequencies (end of the ripple band for the Chebyshev approximation) are $\Omega_c = 1$ rad/s
3. r_2 is either 1 or ∞, so that both singly and doubly terminated filters are included

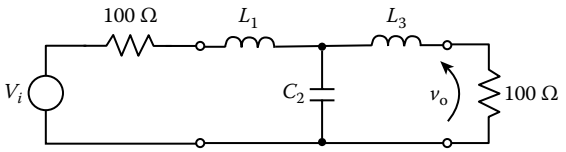


FIGURE 85.6 A third-order passive all-pole filter can be realized by a doubly terminated third-order circuit.

TABLE 85.3 Element Values for Low-Pass Filter Circuits

		$N = 2$, Element Number		$N = 3$, Element Number		
Filter Type	r_2	1	2	1	2	3
Butterworth	∞	1.4142	0.7071	1.5000	1.3333	0.5000
	1	1.4142	1.4142	1.0000	2.0000	1.0000
Chebyshev Type I	∞	0.7159	0.4215	1.0895	1.0864	0.5158
0.1 dB ripple	1	—	—	1.0316	1.1474	1.0316
Chebyshev Type I	∞	0.9403	0.7014	1.3465	1.3001	0.7981
0.5 dB ripple	1	—	—	1.5963	1.0967	1.5963
Bessel	∞	1.0000	0.3333	0.8333	0.4800	0.1667
	1	1.5774	0.4227	1.2550	0.5528	0.1922

The element values in Table 85.3 are numbered from the source end in the same manner as in Figure 85.4. In addition, empty spaces indicate unrealizable networks. In the case of the Chebyshev filter, the amount of ripple can be specified as desired, so that in the table only a selective sample can be given. Extensive tables of prototype element values for many types of filters can be found in Ref. [3].

The example given earlier, of a Butterworth filter of order 3, can also be verified using Table 85.3. The steps necessary to convert the normalized element values in the table into actual filter values are the same as previously illustrated.

In contrast to all-pole approximations, the characteristic function of an elliptic filter function is a rational function. The resulting filter will again be a ladder network but the series elements may be parallel combinations of capacitance and inductance and the shunt elements may be series combinations of capacitance and inductance.

Figure 85.6 illustrates the general circuits for even and odd N, respectively. As in the case of all-pole approximations, tabulations of element values for normalized low-pass filters based on elliptic approximations are also possible. Since these tables are quite involved, the reader is referred to Ref. [3].

85.4.2 Active Filter Design

Active filters are widely used and commercially available with cutoff frequencies from millihertz to megahertz. The characteristics that make them the implementation of choice for several applications are small size for low-frequency filters because they do not use inductors; precision realization of theoretical transfer functions by use of precision resistors and capacitors; high input impedance that is easy to drive and for many circuit configurations the source impedance does not effect the transfer function; low output impedance that can drive loads without effecting the transfer function and can drive the transient, switched capacitive, loads of the input stages of A/D converters and low (N+THD) performance for pre-A/D antialiasing applications (as low as −100 dBc).

Active filters use R, C, A (operational amplifier) circuits to implement polynomial transfer functions. They are most often configured by cascading an appropriate number of first- and second-order sections.

The simplest first-order ($N = 1$) active filter is the first-order passive filter of Figure 85.3 with the addition of a unity gain follower amplifier. Its cutoff frequency (Ω_c) is the same as that given in Equation 85.25. Its advantage over its passive counterpart is that its operational amplifier can drive whatever load that it can tolerate without interfering with the transfer function of the filter.

The vast majority of higher-order filters have poles that are not located on the negative real axis in the s-plane and therefore are in complex conjugate pairs that combine to create second-order pole pairs of the form:

$$H(s) = s^2 + \frac{\omega_p}{Q}s + \omega_p^2 \Leftrightarrow s^2 + 2as + a^2 + b^2 \tag{85.30}$$

where

$$p_1,\ p_2 = a \pm jb$$

$$\omega_p^2 = a^2 + b^2$$

$$Q = \frac{\omega_p}{2a} = \frac{\sqrt{(a^2 + b^2)}}{2a}$$

The most commonly used two-pole active filter circuits are the *Sallen and Key* low-pass resonator, the *multiple feedback* bandpass, and the *state variable* implementation as shown in Figure 85.7a–c. In the analyses that follow, the more commonly used circuits are used in their simplest form. A more comprehensive treatment of these and numerous other circuits can be found in Ref. [4].

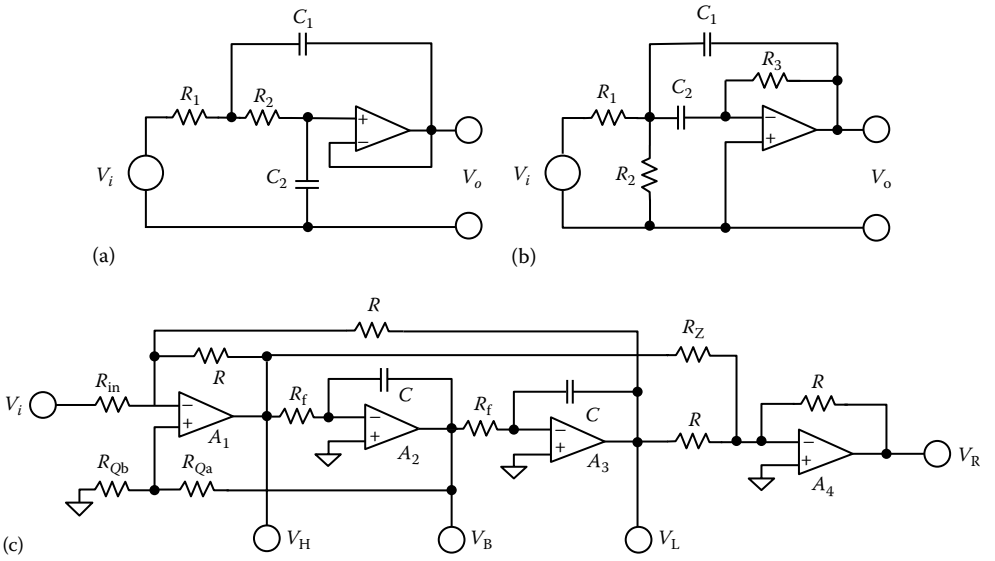


FIGURE 85.7 Second-order active filters can be realized by common filter circuits: (a) Sallen and Key low pass, (b) multiple feedback bandpass, and (c) state variable.

The Sallen and Key circuit of Figure 85.7a is used primarily for its simplicity. Its component count is the minimum possible for a two-pole active filter. It cannot generate stopband zeros and therefore is limited in its use to monotonic roll-off transfer functions such as Butterworth and Bessel filters. Other limitations are that the phase shift of the amplifier reduces the Q of the section and the capacitor ratio becomes large for high Q circuits. The amplifier is used in a follower configuration and therefore is subjected to a large common mode input signal swing which is not the best condition for low distortion performance. It is recommended to use this circuit for a section $Q < 10$ and to use an amplifier whose gain bandwidth product is greater than $100 f_p$.

The transfer function and design equations for the Sallen and Key circuit of Figure 85.7a are

$$H(s) = \frac{1/R_1R_2C_1C_2}{s^2 + (1/R_1C_2)s + 1/R_1R_2C_1C_2} = \frac{\omega_p^2}{s^2 + (\omega_p/Q)s + \omega_p^2} \tag{85.31}$$

from which obtains

$$\omega^2 = \frac{1}{R_1R_2C_1C_2}, \quad Q = \omega_pR_1C_2 = \sqrt{\frac{R_1C_2}{R_2C_1}} \tag{85.32}$$

$$R_1,\ R_2 = \frac{1}{4\pi f_pQC_2}\left[1 \pm \sqrt{1 - \frac{4Q^2C_2}{C_1}}\right] \tag{85.33}$$

which has valid solutions for

$$\frac{C_1}{C_2} \geq 4Q^2 \tag{85.34}$$

In the special case where

$$\begin{aligned} &R_1 = R_2 = R, \text{ then} \\ &C = 1/2\pi Rf_p, \quad C_1 = 2QC, \quad \text{and} \quad C_2 = C/2Q \end{aligned} \tag{85.35}$$

The design sequence for Sallen and Key low pass of Figure 85.7a is as follows.

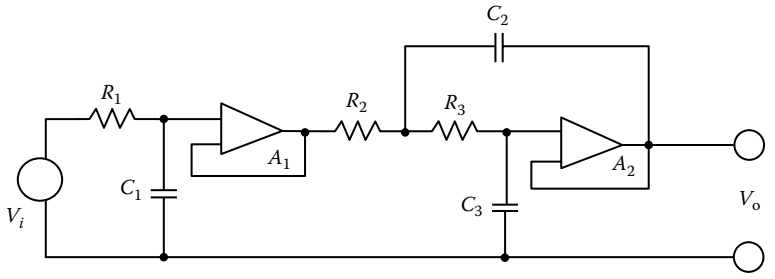


FIGURE 85.8 A three-pole Butterworth active can be configured with a buffered first-order *RC* in cascade with a two-pole Sallen and Key resonator.

For a required f_p and Q, select C_1 and C_2 to satisfy Equation 85.34. Compute R_1 and R_2 from Equation 85.33 (or Equation 85.35 if R_1 is chosen to equal R_2) and scale the values of C_1 and C_2 and R_1 and R_2 to desired impedance levels.

As an example, a three-pole low-pass active filter is shown in Figure 85.8. It is realized with a buffered single-pole *RC* low-pass filter section in cascade with a two-pole Sallen and Key section.

To construct a three-pole Butterworth filter, the pole locations are found in Table 85.2 and the element values in the sections are calculated from Equation 85.25 for the single real pole and in accordance with the Sallen and Key design sequence listed earlier for the complex pole pair.

From Table 85.2, the normalized pole locations are

$$f_{p1} = 1.000, \quad f_{p2} = 1.000, \quad \text{and} \quad Q_{p2} = 1.000$$

For a cutoff frequency of 10 kHz and if it is desired to have an impedance level of 10 kΩ, then the capacitor values are computed as follows:

For R_1 = 10 kΩ:

$$\text{From Equation 85.25, } C_1 = \frac{1}{2\pi R_1 r_{p1}} = \frac{1}{2\pi(10{,}000)(10{,}000)} = \frac{10^{-6}}{200\pi} = 0.00159\ \mu\text{F}$$

For $R_2 = R_3 = R$ = 10 kΩ:

$$\text{From Equation 85.35, } C = \frac{1}{2\pi R f_{p2}} = \frac{1}{2\pi(10{,}000)(10{,}000)} = \frac{10^{-6}}{200\pi} = 0.00159\ \mu\text{F}$$

from which

$$C_2 = 2QC = 2(0.00159)\ \mu\text{F} = 0.00318\ \mu\text{F}$$

$$C_3 = \frac{C}{2Q} = 0.5(0.00159)\ \mu\text{F} = 0.00795\ \mu\text{F}$$

The *multiple feedback* circuit of Figure 85.7b is a minimum component count, two-pole (or one-pole pair), bandpass filter circuit with user definable gain. It cannot generate stopband zeros and therefore is limited in its use to monotonic roll-off transfer functions. Phase shift of its amplifier reduces the Q of the section and shifts the f_p. It is recommended to use an amplifier whose open-loop gain at f_p is $>100Q^2H_p$.

The design equations for the *multiple feedback* circuit of Figure 85.4b are

$$H(s) = \frac{s/R_1C_1}{s^2 + s/R_3\left(1/C_1 + 1/C_2\right) + (R_1 + R_2)/R_1R_2R_3C_1C_2} = -\frac{s\omega_p H_p/Q}{s^2 + s\omega_p/Q + \omega_p^2} \tag{85.36}$$

when $s = j\omega_p$, the gain H_p is

$$H_p = \frac{R_3C_2}{R_1(C_1 + C_2)} \tag{85.37}$$

From Equations 85.36 and 85.37 for a required set of ω_p, Q, and H_p

$$R_1 = \frac{Q}{C_1H_p\omega_p}, \quad R_2 = \frac{Q}{\omega_p}\left(\frac{1}{Q^2(C_1 + C_2) - H_pC_1}\right), \quad R_3 = \frac{R_1H_p(C_1 + C_2)}{C_2} \tag{85.38}$$

For R_2 to be realizable,

$$Q^2(C_1 + C_2) \geq H_pC_1 \tag{85.39}$$

The design sequence for a *multiple feedback* bandpass filter is as follows.

Select C_1 and C_2 to satisfy Equation 85.39 for the H_p and Q required. Compute R_1, R_2, and R_3. Scale R_1, R_2, R_3, C_1, and C_2 as required meeting desired impedance levels.

Note that it is common to use $C_1 = C_2 = C$ for applications where $H_p = 1$ and $Q > 0.707$.

The *state variable* circuit of Figure 85.7c is the most widely used active filter circuit. It is the basic building block of programmable active filters and of SC designs. While it uses three or four amplifiers and numerous other circuit elements to realize a two-pole filter section, it has many desirable features. From a single input, it provides low-pass (V_t), high-pass (V_H), and bandpass (V_B) outputs and by summation into an additional amplifier (A_4) (or the input stage of the next section) a band reject (V_R) or stopband zero can be created. Its two integrator resistors connect to the virtual ground of their amplifiers (A_2, A_3) and therefore have no signal swing on them. Therefore, programming resistors can be switched to these summing junctions using electronic switches. The sensitivity of the circuit to the gain and phase performance of its amplifiers is more than an order of magnitude less than single amplifier designs. The open-loop gain at f_p does not have to be multiplied by either the desired Q or the gain at dc or f_p. Second-order sections with Q up to 100 and f_p up to 1 MHz can be built with this circuit.

There are several possible variations of this circuit that improve its performance at particular outputs. The input can be brought into several places to create or eliminate phase of inversions; the damping feedback can be implemented in several ways other than the R_{Qa} and R_{Qb} that are shown in Figure 85.7c and the f_p and Q of the section can be or adjusted independently from one another. The dc offset adjustment components can be added to allow the offset at any one output to be trimmed to zero.

For simplicity of presentation, Figure 85.7c makes several of the resistors equal and identifies others with subscripts that relate to their function in the circuit. Specifically, the feedback amplifier A_1 that generates the V_H output has equal feedback and input resistor from the V_L feedback signal to create unity gains from that input. Similarly, the "zero-summing" amplifier, A_4 has equal resistors for its feedback and input from V_L to make the dc gain at the V_R output the same as that at V_L. More general configurations with all elements included in the equation of the transfer function are available in numerous reference texts including Ref. [4].

The *state variable* circuit, as configured in Figure 85.7c, has four outputs. Their transfer functions are

$$V_{\mathrm{L}}(s) = -\frac{R}{R_1(R_{\mathrm{f}}C)^2}\left(\frac{1}{D(s)}\right) \tag{85.40a}$$

$$V_{\mathrm{B}}(s) = \frac{R}{R_1}\left(\frac{s/(R_{\mathrm{f}}C)}{D(s)}\right) \tag{85.40b}$$

$$V_{\mathrm{H}}(s) = -\frac{R}{R_1}\left(\frac{s^2}{D(s)}\right) \tag{85.40c}$$

$$V_{\mathrm{R}}(s) = \frac{R}{R_1(R_{\mathrm{f}}C)^2}\left(\frac{(R_z/R)s^2+1}{D(s)}\right) \tag{85.40d}$$

where

$$D(s) = s^2 + \frac{a}{R_{\mathrm{f}}C}s + \frac{1}{(R_{\mathrm{f}}C)^2} = s^2 + \frac{\omega_{\mathrm{p}}}{Q}s + \omega_{\mathrm{p}}^2 \quad a = \frac{R_{\mathrm{Qb}}}{(R_{\mathrm{Qa}}+R_{\mathrm{Qb}})}\left(2+\frac{R}{R_1}\right) \tag{85.41}$$

Note that the dc gain at the low-pass output is

$$V_{\mathrm{L}}(0) = -\frac{R}{R_1} \tag{85.42a}$$

from which obtains

$$\omega_{\mathrm{p}} = \frac{1}{R_{\mathrm{f}}C} \quad \text{and} \quad \frac{1}{Q} = \frac{R_{\mathrm{Qb}}}{(R_{\mathrm{Qa}}+R_{\mathrm{Qb}})}\left(2+\frac{R}{R_1}\right) \tag{85.42b}$$

The design sequence for the state variable filter of Figure 85.7c is to select the values of R_{f} and C to set the frequency ω_{p}, set the values of R_i for the desired dc gain, and adjust R_{Qa} and R_{Qb} for the desired Q and dc gain.

85.5 Discrete-Time Filters

A digital filter is a circuit or a computer program that computes a discrete output sequence from a discrete input sequence. Digital filters belong to the class of discrete-time linear time-invariant (LTI) systems, which are characterized by the properties of causality, recursibility, and stability, and may be characterized in the time domain by their impulse response and in the transform domain by their transfer function. The most general case of a discrete-time LTI system with the input sequence denoted by $x(kT)$ and the resulting output sequence $y(kT)$ can be described by a *set of linear difference equations with constant coefficients*:

$$y(kT) = \sum_{\mu=0}^{N} b_\mu x(kT-\mu T) - \sum_{\mu=1}^{N} a_\mu y(kT-\mu T) \tag{85.43}$$

where $a_0 = 1$. An equivalent relation between the input and output variables can be given through the convolution sum in terms of the impulse response sequence $h(kT)$:

$$y(kT) = \sum_{\mu=0}^{N} h(kT)x(kT - \mu T) \tag{85.44}$$

The corresponding transfer function is given by

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{\mu=0}^{N} b_\mu z^{-\mu}}{1 + \sum_{\mu=1}^{N} a_\mu z^{-\nu}} \Leftrightarrow H(z) = b_0 \prod_{\mu=1}^{N} \left(\frac{z - z_{0\mu}}{z - z_{\infty\mu}} \right) \tag{85.45}$$

where

$H(z)$ is the z-transform of the impulse response $h(kT)$

$X(z)$ and $Y(z)$ are the z-transform of the input signal $x(kT)$ and the output or the filtered signal $y(kT)$

As can be seen from Equation 85.44, if for at least one μ, $a_\mu \neq 0$, the corresponding system is recursive; its impulse response is of infinite duration—*infinite impulse response (IIR) filters*. If $a_\mu = 0$, the corresponding system is nonrecursive—*finite impulse response (FIR) filter*; its impulse response is of finite duration and the transfer function $H(z)$ is a polynomial in z^{-1}. The zeros of the polynomial $X(z)$ given by $z = z_{\infty i}$ are called the poles of $H(z)$ and are commonly referred to as the *natural frequencies* of the filter. The condition for the stability of the filter is expressed by the constraint that all the poles of $H(z)$ should lie inside the unit circle, that is $|z_{\infty i}| < 1$. The zeros of $Y(z)$ given by $z = z_{0t}$ which are equivalent to the zeros of $H(z)$ are called the *transmission zeros* of the filter. Clearly, at these frequencies the output of the filter is zero for any finite input.

On the unit circle, the transfer function frequency $H(z)$ reduces to the frequency response function $H(e^{j\omega T})$, the discrete-time Fourier transform of $h(kT)$, which in general is complex and can be expressed in terms of magnitude and phase

$$H(e^{j\omega T}) = |H(e^{j\omega T})| e^{j\theta(\omega)} \tag{85.46}$$

The gain function of the filter is given as

$$\alpha(\Omega) = 20\log_{10} |H(e^{j\omega T})| \tag{85.47}$$

It is also common practice to call the negative of the gain function of the attenuation. Note that the attenuation is a positive number when the magnitude response is <1.

Figure 85.9 gives a block diagram realizing the difference equation of the filter, which is commonly referred to as the *direct-form I* realization. Notice that the element values for the multipliers are obtained directly from the numerator and denominator coefficients of the transfer function. By rearranging the structure in regard to the number of delays, one can obtain the canonic structure called *direct-form II* shown in Figure 85.10, which requires the minimum number of delays.

Physically, the input numbers are samples of a continuous signal and real-time digital filtering involves the computation of the iteration of Equation 85.43 for each incoming new input sample. Design of a filter consists of determining the constants a_μ and b_μ that satisfies a given filtering requirement. If the filtering is performed in real time, then the right side of Equation 85.46 must be computed in less than the sampling interval T.

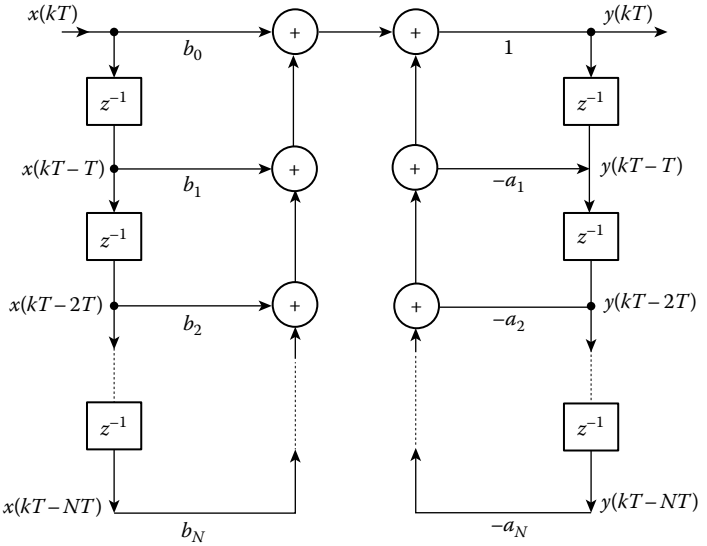


FIGURE 85.9 The difference equation of a digital filter can be realized by a direct-form I implementation that uses separate delay paths for the *X* and *Y* summations.

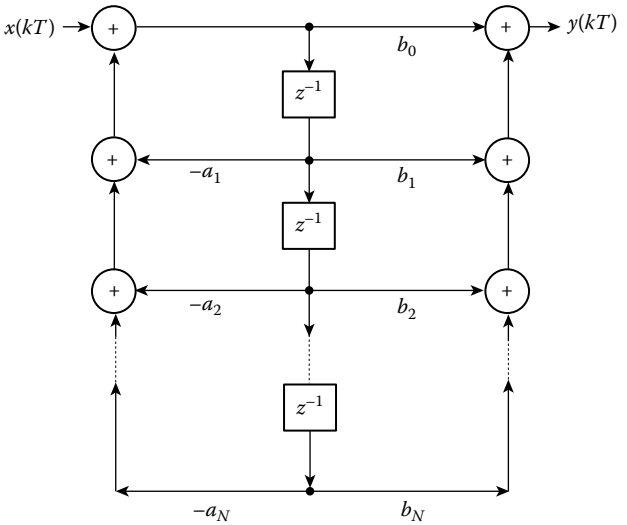


FIGURE 85.10 A direct-form II implementation of the difference equations minimizes the number of delay elements.

85.6 Digital Filter Design Process

The digital filter design procedure consists of the following basic steps:

1. Determine the desired response. The desired response is normally specified in the frequency domain in terms of the desired magnitude response and/or the desired phase response.
2. Select a class of filters (e.g., linear-phase FIR filters or IIR filters) to approximate the desired response.
3. Select the best member in the filter class.
4. Implement the best filter using a general-purpose computer, a DSP, or a custom hardware chip.
5. Analyze the filter performance to determine whether the filter satisfies all the given criteria.

85.7 FIR Filter Design

In many digital signal-processing applications, FIR filters are generally preferred over their IIR counterparts, because they offer a number of advantages compared with their IIR equivalents. Some of the good properties of FIR filters are a direct consequence of their nonrecursive structure. First, FIR filters are inherently stable and free of limit cycle oscillations under finite-word length conditions. In addition, they exhibit a very low sensitivity to variations in the filter coefficients. Second, the design of FIR filters with exactly *linear phase* (constant group delay) vs. frequency behavior can be accomplished easily. This property is useful in many application areas, such as speech processing, phase delay equalization, image processing, etc.

Finally, there exists number of efficient algorithms for designing optimum FIR filters with arbitrary specifications. The main disadvantage of FIR filters over IIR filter is that FIR filter designs generally require, particularly in applications requiring narrow transition bands, considerably more computation to implement.

An FIR filter of order N is described by a difference equation of the form

$$y(kT) = \sum_{\mu=0}^{N} b_\mu x(kT - \mu T) \tag{85.48}$$

and the corresponding transfer function is

$$H(z) = \frac{Y(z)}{X(z)} = \sum_{\mu=0}^{N} b_\mu z^{-\mu} \tag{85.49}$$

The objective of FIR filter design is to determine $(N \pm 1)$ coefficients given by

$$h(0), h(1), \ldots, h(N) \tag{85.50}$$

so that the transfer function $H(e^{j\omega T})$ approximates a desired frequency characteristic. Note that because Equation 85.47 is also in the form of a convolution summation, the impulse response of an FIR filter is given by

$$h(kT) = \begin{cases} b_\mu, & k = 0, 1, \ldots, N \\ 0, & \text{otherwise} \end{cases} \tag{85.51}$$

Two equivalent structures for FIR filters are given in Figure 85.11.

The accuracy of an FIR approximation is described by the following parameters:

δ_p Passband ripple
δ_s Stopband attenuation
$\Delta\omega$ Transition bandwidth

These quantities are depicted in Figure 85.12 for a prototype low-pass filter. δ_p and δ_s characterize the permissible errors in the passband and stopband, respectively. Usually, the passband ripple and stopband attenuation are given in decibels, in which case their values are related to the parameters δ_p and δ_s by

$$\text{Passband ripple (dB): } A_p = -20\log_{10}(1 - \delta_p) \tag{85.52}$$

$$\text{Stopband ripple (dB): } A_s = -20\log_{10}(\delta_s) \tag{85.53}$$

Note that due to the symmetry and periodicity of the magnitude response of $|H(e^{j\omega T})|$, it is sufficient to give the filter specifications in the interval $0 \le \omega \le \pi$.

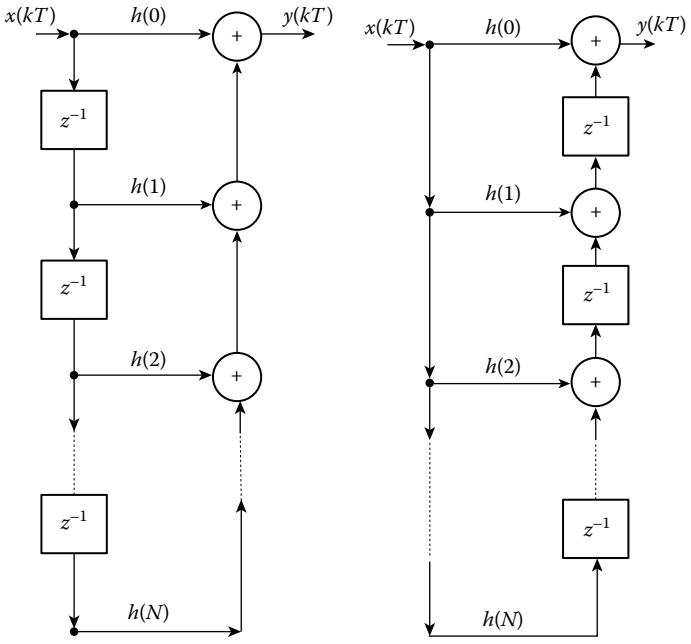


FIGURE 85.11 The sequence of the delays and summations can be varied to produce alternative direct-form implementations.

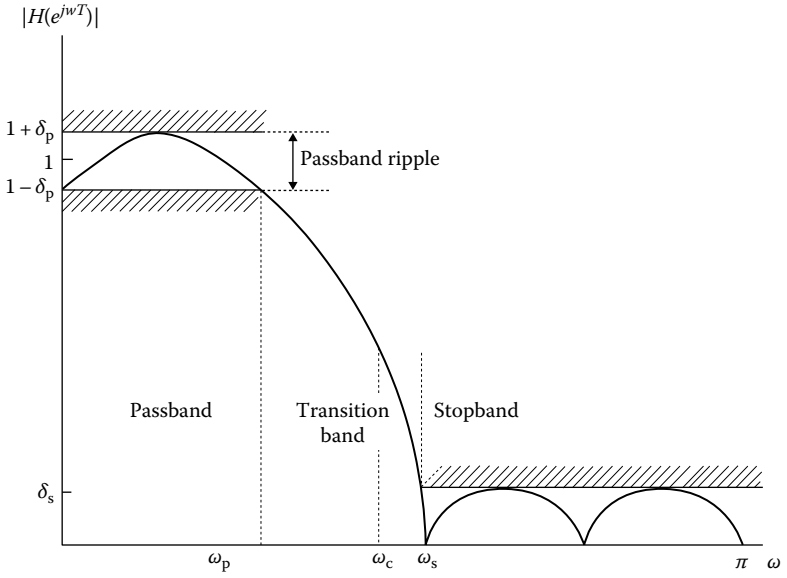


FIGURE 85.12 Tolerance limits must be defined for an FIR low-pass filter magnitude response.

85.7.1 Windowed FIR Filters

Several design techniques can be employed to synthesize linear-phase FIR filters. The simplest implementation is based on *windowing*, which commonly begins by specifying the ideal frequency response and expanding it in a Fourier series and then truncating and smoothing the ideal impulse response by means of a window function. The truncation results in large ripples before and after the discontinuity of

the ideal frequency response known as the Gibbs phenomena, which can be reduced by using a window function that tapers smoothly at both ends. Filters designed in this way possess equal passband ripple and stopband attenuation, i.e.,

$$\delta_p = \delta_s = \delta \tag{85.54}$$

To illustrate this method, let us define an ideal desired frequency response that can be expanded in a Fourier series:

$$H_d(e^{jwT}) = \sum_{k=-\infty}^{\infty} h_d(kT) e^{-jk\omega T} \tag{85.55}$$

where $h_d(kT)$ is the corresponding impulse response sequence, which can be expressed in terms of $H_d(e^{j\omega T})$ as

$$h_d(kT) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{jwT}) e^{jk\omega T} d\omega \tag{85.56}$$

The impulse response of the desired filter is then found by weighting this ideal impulse response with a window $w(kT)$ such that

$$h(kT) = \begin{cases} w(kT) h_d(kT), & 0 \le k \le N \\ 0, & \text{otherwise} \end{cases} \tag{85.57}$$

Note that for $w(kT)$ in the earlier given interval we obtain the rectangular window. Some commonly used windows are Bartlett (triangular), Hanning, Hamming, Blackmann, etc., the definitions of which can be found in Ref. [5].

As an example of this design method, consider a low-pass filter with a cutoff frequency of ω_c and a desired frequency of the form:

$$H_d(e^{j\omega T}) = \begin{cases} e^{-j\omega NT/2}, & |\omega| \le \omega_c \\ 0, & \omega_c < |\omega| \le \pi \end{cases} \tag{85.58}$$

Using Equation 85.56, we obtain the corresponding ideal impulse response:

$$h_d(kT) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} e^{-j\omega TN/2} e^{jk\omega T} d\omega = \frac{\sin[\omega_c(kT - TN/2)]}{\pi(kT - TN/2)} \tag{85.59}$$

Choosing $N = 4$, $\omega_c = 0.6\pi$ and a Hamming window defined by

$$w(kT) = \begin{cases} 0.54 - 0.46\cos(2\pi kT/N), & 0 \le k \le N \\ 0, & \text{otherwise} \end{cases} \tag{85.60}$$

we obtain the following impulse response coefficients:

$$\begin{aligned} h(0) &= -0.00748 \\ h(1) &= 0.12044 \\ h(2) &= -0.54729 \\ h(3) &= 0.27614 \\ h(4) &= -0.03722 \end{aligned} \tag{85.61}$$

85.7.2 Optimum FIR Filters

As mentioned earlier, one of the principal advantages of FIR filters over their IIR counterparts is the availability of excellent design methods for optimizing arbitrary filter specifications. Generally, the design criterion for the optimum solution of an FIR filter design problem can be characterized as follows. The maximum error between the approximating response and the given desired response has to be minimized, i.e.,

$$E(e^{j\omega T}) = W_{\mathrm{d}}(e^{j\omega T})\left\| H_{\mathrm{d}}(e^{j\omega T})\right| - \left|H(e^{j\omega T})\right\| \tag{85.62}$$

where

$E(e^{j\omega T})$ is the weighted error function on a close range X of $[0, \pi]$

$W_{\mathrm{d}}(e^{j\omega T})$ is a weighting function, which emphasizes the approximation error parameters in the design process

If the maximum absolute value of this function is less than or equal ε on X, i.e.,

$$\varepsilon = \max_{\omega \varepsilon X} \left| E(e^{j\omega T}) \right| \tag{85.63}$$

the desired response is guaranteed to meet the given criteria. Thus, this optimization condition implies that the best approximation must have an equiripple error function. The most frequently used method for designing optimum magnitude FIR filters is the Parks–McClellan algorithm. This method essentially reduces the filter design problem into a problem in polynomial approximation in the Chebyshev approximation sense as discussed earlier. The maximum error between the approximation and the desired magnitude response is minimized. It offers more control over the approximation errors in different frequency bands than is possible with the window method. Using the Parks–McClellan algorithm to design FIR filters is computationally expensive. This method, however, produces optimum FIR filters by applying time-consuming iterative techniques. A FORTRAN program for the Parks–McClellan algorithm can be found in the IEEE publication Programs for DSP in Ref. [6]. As an example of an equiripple filter design using the Parks–McClellan algorithm, a sixth-order low-pass filter with a passband $0 \le \omega \le 0.6\pi$, a stopband $0.8\pi \le \omega \le \pi$, and equal weighting for each band was designed by means of this program.

The resulting impulse response coefficients are

$$\begin{aligned} h(0) &= h(6) = -0.00596 \\ h(1) &= h(5) = -0.18459 \\ h(2) &= h(4) = 0.25596 \\ h(3) &= 0.70055 \end{aligned} \tag{85.64}$$

85.7.3 Design of Narrowband FIR Filters

When using conventional techniques to design FIR filters with especially narrow bandwidths, the resulting filter lengths may be very high. FIR filters with long filter lengths often require lengthy design and implementation times, and are more susceptible to numerical inaccuracy. In some cases, conventional filter design techniques, such as the Parks–McClellan algorithm, may fail the design altogether. A very efficient algorithm called the interpolated finite impulse response (IFIR) filter design technique can be employed to design narrowband FIR filters. Using this technique produces narrowband filters that require far fewer coefficients than those filters designed by the direct application of the Parks–McClellan algorithm. For more information on IFIR filter design, see Ref. [7].

85.8 IIR Filter Design

The main advantage of IIR filters over FIR filters is that IIR filters can generally approximate a filter design specification using a lower-order filter than that required by an FIR design to perform similar filtering operations. As a consequence, IIR filters execute much faster and do not require extra memory, because they execute in place. A disadvantage of IIR filters, however, is that they have a nonlinear-phase response. The two most common techniques used for designing IIR filters will be discussed in this section. The first approach involves the transformation of an analog prototype filter. The second method is an optimization-based approach allowing the approximation of an arbitrary frequency response.

The transformation approach is quite popular because the approximation problem can be reduced to the design of classical analog filters, the theory of which is well established, and many closed-form design methods exist. Note that this is not true for FIR filters, for which the approximation problems are of entirely different nature. The derivation of a transfer function for a desired filter specification requires the following three basic steps:

1. Given a set of specifications for a digital filter, the first step is to map the specifications into those for an equivalent analog filter.
2. The next step involves the derivation of a corresponding analog transfer function for the analog prototype.
3. The final step is to translate the transfer function of the analog prototype into a corresponding digital filter transfer function.

Once the corresponding analog transfer function for the analog prototype is derived, it must be transformed using a transformation that maps $H_a(s)$ into $H(z)$. The simplest and most appropriate choice for s is the well-known bilinear transform of the z-variable

$$s = \frac{2(1-z^{-1})}{T_d(1+z^{-1})} \Leftrightarrow z = \frac{1+(T_d/2)s}{1-(T_d/2)s} \tag{85.65}$$

which maps a stable analog filter in the s-plane into a stable digital filter in the z-plane. Substituting s with the right-hand side of Equation 85.63 in $H_a(s)$ results in

$$H(z) = H_a\left(\frac{2(1-z^{-1})}{T_d(1+z^{-1})}\right) \Rightarrow H(e^{j\omega T})\Big|_{z=e^{j\omega T}} = H_a\left(\frac{2j}{T_d}\tan\left(\frac{\omega T}{2}\right)\right) \tag{85.66}$$

As it can be seen from Equation 85.66, the analog frequency domain (imaginary axis) maps onto the digital frequency domain (unit circle) nonlinearly. This phenomenon is called frequency warping

and must be compensated in a practical implementation. For low frequencies, Ω and ω are approximately equal. We obtain the following relation between the analog frequency Ω and the digital frequency ω:

$$\Omega = \frac{2}{T_d}\tan(\omega T/2) \tag{85.67}$$

$$\omega = \frac{2}{T}\arctan(\Omega T_d/2) \tag{85.68}$$

The overall bilinear transformation procedure is as follows:

1. Convert the critical digital frequencies (e.g., ω_p and ω_s for low-pass filters) to the corresponding analog frequencies in the s-domain using the relationship given by Equation 85.67.
2. Derive the appropriate continuous prototype transfer function $H_a(s)$ that has the properties of the digital filter at the critical frequencies.
3. Apply the bilinear transform to $H_a(s)$ to obtain $H(z)$ which is the required digital filter transfer function.

To illustrate the three-step IIR design procedure using the bilinear transform, consider the design of a second-order Butterworth low-pass filter with a cutoff frequency of $\omega_c = 0.3\pi$. The sampling rate of the digital filter is to be $f_s = 10$ Hz, giving $T = 0.1$ s. First, we map the cutoff frequency to the analog frequency

$$\Omega_c = \frac{2}{0.1}\tan(0.15\pi) = 10.19 \text{ rad/s} \tag{85.69}$$

The poles of the analog Butterworth filter transfer function $H_a(s)$ are found using Equation 85.11. As explained earlier, these poles lie equally spaced in the s-plane on a circle of radius Ω_c:

$$H_a(s) = \frac{1}{s^2 + \sqrt{2\Omega_c s + \Omega_c^2}} \tag{85.70}$$

Application of the bilinear transformation

$$S = \frac{2(1 - z^{-1})}{0.1(1 + z^{-1})} \tag{85.71}$$

gives the digital transfer function

$$H(z) = \frac{0.00002 + 0.00004z^{-1} + 0.00002z^{-2}}{1 - 1.98754z^{-1} + 0.98762z^{-2}} \tag{85.72}$$

The earlier computations were carried out using Ref. [8], which greatly automates the design procedure.

85.8.1 Design of Arbitrary IIR Filters

The IIR filter design approach discussed in the previous section is primarily suitable for frequency-selective filters based on closed-form formulas. In general, however, if a design other than standard low pass, high pass, bandpass, and stopband is required, or if the frequency responses of arbitrary specifications are to be matched, in such cases it is often necessary to employ algorithmic methods implemented on computers. In fact, for nonstandard response characteristics, algorithmic

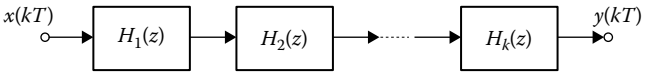


FIGURE 85.13 An IIR filter can be implemented by a cascade of individual transfer functions.

procedures may be the only possible design approach. Depending on the error criterion used, the algorithmic approach attempts to minimize the approximation error between the desired frequency response $H_d(e^{j\omega T})$ and $H(e^{j\omega T})$ or between the time-domain response $h_d(kT)$ and $h(kT)$. Computer software is available for conveniently implementing IIR filters approximating arbitrary frequency response functions [8,9].

85.8.2 Cascade-Form IIR Filter Structures

Recall that theoretically there exist an infinite number of structures to implement a digital filter. Filters realized using the structure defined by Equation 85.44 directly is referred to as direct-form IIR filters. The direct-form structure, however, is not employed in practice except when the filter order $N \leq 2$, because they are known to be sensitive to errors introduced by coefficient quantization and by finite-arithmetic conditions. Additionally, they produce large round-off noise, particularly for poles closed to the unit circle.

Two less-sensitive structures can be obtained by partial fraction expansion or by factoring the right-hand side of Equation 85.46 in terms of real rational functions of orders 1 and 2. The first method leads to *parallel connections* and the second one to *cascade connections* of corresponding lower-order sections, which are used as building blocks to realize higher-order transfer functions. In practice, the cascade form is by far the preferred structure, since it gives the freedom to choose the pairing of numerators and denominators and the ordering of the resulting structure. Figure 85.13 shows a cascade-form implementation, whose overall transfer function is given by

$$H(z) = \prod_{k=1}^{M} H_k(z) \tag{85.73}$$

where the transfer function of the kth building block is

$$H_k(z) = \frac{b_{0k} + b_{1k}z^{-1} + b_{2k}z^{-2}}{1 + a_{1k}z^{-2} + a_{2k}z^{-2}} \tag{85.74}$$

Note this form is achieved by factoring Equation 85.45 into second-order sections.

There are, of course, many other realization possibilities for IIR filters, such as state-space structures [8], lattice structures [10], and wave structures. The last is introduced in the next section.

85.9 Wave Digital Filters

It was shown earlier that for recursive digital filters the approximation problem can be reduced to classical design problems by making use of the bilinear transform. For wave digital filters (WDFs), this is carried one step farther in that the structures are obtained directly from classical circuits. Thus, to every WDF there corresponds an *LCR* reference filter from which it is derived. This relationship accounts for their excellent properties concerning coefficient sensitivity, dynamic range, and all aspects of stability under finite-arithmetic conditions. The synthesis of WDFs is based on the wave network characterization; therefore, the resulting structures are referred to as WDFs. To illustrate the basic idea behind the

theory of WDFs, consider an inductor L, which is electrically described by $V(s) = sLI(s)$. In the next step, we define wave variables $A_1(s)$ and $B_1(s)$ as

$$\begin{aligned} A_1(s) &= V(s) + RI(s) \\ B_1(s) &= V(s) - RI(s) \end{aligned} \tag{85.75}$$

where R is called the port resistance. Substituting $V(s) = sLI(s)$ in the earlier relation and replacing s in $A_1(s)$ and $B_1(s)$ with the bilinear transform given by Equation 85.65, we obtain

$$B(z) = \frac{(1-z^{-1})L - (1+z^{-1})R}{(1-z^{-1})L + (1+z^{-1})R} A(z) \tag{85.76}$$

Letting $R = L$, the earlier relation reduces to

$$B(z) = -z^{-1}A(z) \tag{85.77}$$

Thus an inductor translates into a unit delay in cascade with an inverter in the digital domain. Similarly, it is easily verified that a capacitance can be simulated by a unit delay and a resistor by a digital sink. Figure 85.14 shows the digital realizations of impedances and other useful one-port circuit elements.

To establish equivalence with classical circuits fully, the interconnections are also simulated by the so-called wave adaptors. The most important of these interconnections are series and parallel connections, which are simulated by series and parallel adaptors, respectively. For most filters of interest, only two- and three-port adaptors are employed. For a complete design example, consider Figure 85.15.

For a given LC filter, one can readily derive a corresponding WDF by using the following procedure. First, the various interconnections in the LC filter are identified as shown in Figure 85.15. In the next

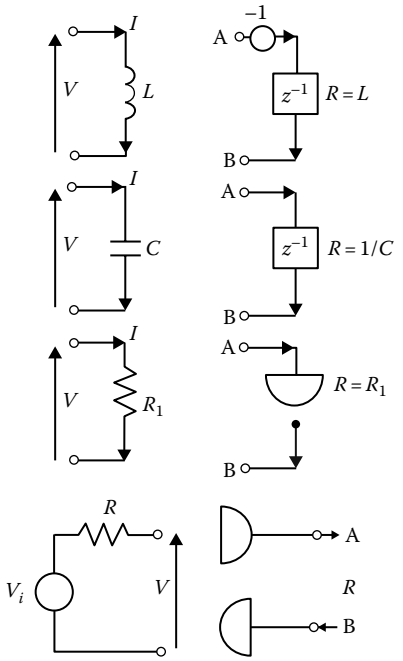


FIGURE 85.14 Digital filter implementations use functional equivalents to one-port linear filter elements.

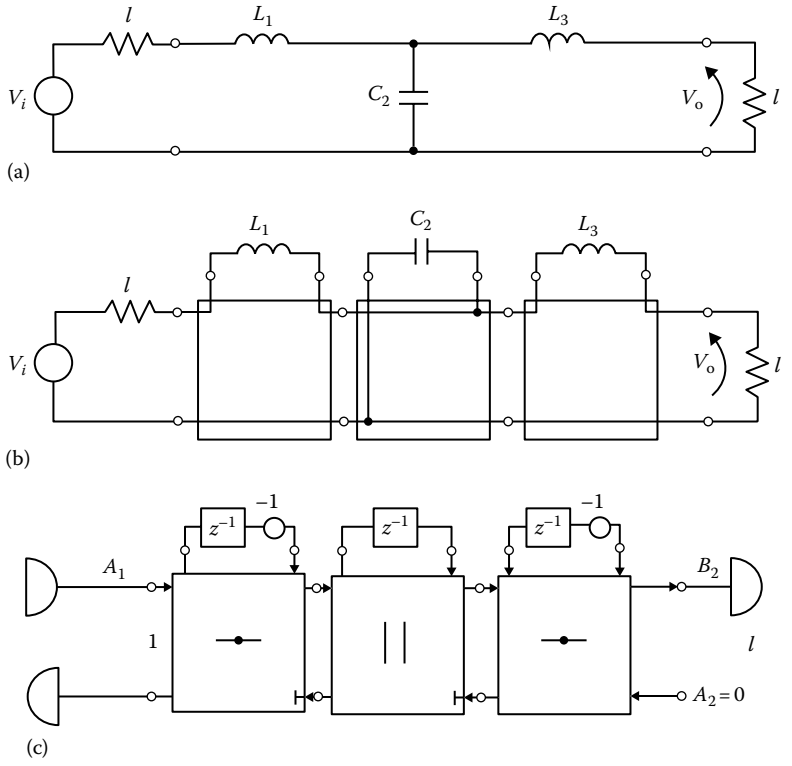


FIGURE 85.15 Digital wave filters establish equivalence with classical filter circuits by use of wave adapter substitutions: (a) LC reference low pass; (b) identification of wire interconnections; and (c) corresponding wave digital filter.

step, the electrical elements in the *LC* filter are replaced by its digital realization using Figure 85.15. Finally, the interconnections are substituted using adaptors. Further discussions and numerical examples dealing with WDFs can be found in Refs. [11–13].

85.10 Antialiasing and Smoothing Filters

In this section, two practical application areas of filters in the analog conditioning stage of a data acquisition system are discussed. A block diagram of a typical data acquisition system is shown in Figure 85.16, consisting of an *antialiasing filter* before the analog-to-digital converter (ADC) and a smoothing filter after the digital-to-analog converter (DAC).

For a complete discrete reconstruction of a time-continuous, band-limited input signal having the spectrum $0 \le f \le f_{max}$, the sampling frequency must be, according to the well-known Shannon's sampling theorem, at least twice the highest frequency in the time signal. In our case, in order to be able to represent frequencies up to f_{max}, the sampling frequency $f_s = 1/T > 2f_{max}$. The necessary band limiting to $f \le f_{max}$ of the input time-continuous signal is performed by a low-pass filter, which suppresses higher

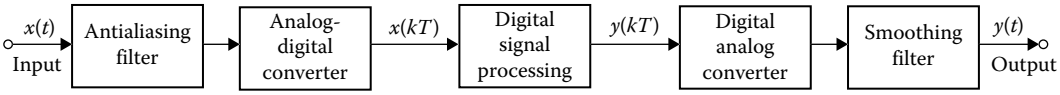


FIGURE 85.16 A data acquisition system with continuous-time inputs and outputs uses antialias prefiltering, an A/D converter, digital signal processing, a D/A converter, and an output smoothing filter.

spectral components greater than f_{max}. Violation of this theorem results in alias frequencies. As a result, frequency components above $f_x/2$, the so-called Nyquist frequency, appear as frequency components below $f_x/2$. Aliasing is commonly addressed by using antialiasing filters to attenuate the frequency components at and above the Nyquist frequency to a level below the dynamic range of an ADC before the signal is digitized. Ideally, a low-pass filter with a response defined by

$$H(j\Omega) = \begin{cases} 1, & |\Omega| \le \pi/T \\ 0, & |\Omega| \ge \pi/T \end{cases} \tag{85.78}$$

is desired to accomplish this task. In practice, a variety of techniques based on the principles of continuous-time analog low-pass filter design can be employed to approximate this "brick-wall" type of characteristic. Antialiasing filters typically exhibit attenuation slopes in the range from 45 to 120 dB/octave and stopband rejection from 75 to 100 dB. Among the types of filters more commonly used for antialias purposes are the Cauer elliptic, Bessel, and Butterworth. The optimum type of filter depends on which kinds of imperfections, e.g., gain error, phase nonlinearity, passband and stopband ripple, etc., are most likely to be tolerated in a particular application. For example, Butterworth filters exhibit very flat frequency response in the passband, while Chebyshev filters provide steeper attenuation at the expense of some passband ripple. The Bessel filter provides a linear-phase response over the entire passband but less attenuation in the stopband. The Cauer elliptic filter, with its extremely sharp roll-off, is especially useful as an antialiasing filter for multichannel digitizing data acquisition systems. However, the large-phase nonlinearity makes it more appropriate for applications involving analysis of the frequency content of signals as opposed to phase content or waveform shape.

Many considerations discussed earlier also apply to smoothing filters. Due to the sampling process, the frequency response after the digital-to-analog conversion becomes periodic with a period equal to the sampling frequency. The quantitization steps that are created in the DAC reconstruction of the output waveform and are harmonically related to the sampling frequency must be suppressed through a low-pass filter having the frequency response of Equation 85.78 also referred to as a smoothing or reconstruction filter. While an antialiasing filter on the input avoids unwanted errors that would result from undersampling the input, a smoothing filter at the output reconstructs a continuous-time output from the discrete-time signal applied to its input.

Consideration must be given to how much antialiasing protection is needed for a given application. It is generally desirable to reduce all aliasable frequency components (at frequencies greater than half of the sampling frequency) to less than the LSB of the ADC being used. If it is possible that the aliasable input can have amplitude as large as the full input signal range of the ADC, then it is necessary to attenuate it by the full 2^N range of the converter. Since each bit of an ADC represents a factor of 2 from the ones adjacent to it, and 20 log(2) = 6 dB, the minimum attenuation required to reduce a full-scale input to less than a LSB is

$$\alpha < -20N(6\text{ dB}) \tag{85.79}$$

where N is the number of bits of the ADC.

The amount of attenuation required can be reduced considerably if there is knowledge of the input frequency spectrum. For example, some sensors, for reasons of their electrical or mechanical frequency response, might not be able to produce a full-scale signal at or above the Nyquist frequency of the system and therefore "full-scale" protection is not required. In many applications, even for 16 bit converters that, in the worst case, would require 96 dB of antialias protection, 50–60 dB is adequate.

Additional considerations in antialias protection of the system are the noise and distortion that are introduced by the filter that is supposed to be eliminating aliasable inputs. It is possible to have a perfectly clean input signal which, when it is passed through a prefilter, gains noise and harmonic distortion

components in the frequency range and of sufficient amplitude to be within a few LSBs of the ADC. The ADC cannot distinguish between an actual signal that is present in the input data and a noise or distortion component that is generated by the prefilter. It is necessary that both noise and distortion components in the output of the antialias filter must also be kept within LBS of the ADC to ensure system accuracy.

85.11 SC Filters

SC filters, also generally referred to as analog sampled-data filters, provide an alternative to conventional active *RC* filters and are commonly used in the implementation of adjustable antialiasing filters. SC filters comprise switches, capacitors, and op amps. Essentially, an SC replaces the resistor in the more traditional analog filter designs. Because the impedance of the SC is a function of the switching frequency, one can vary the cutoff frequency of the SC filter by varying the frequency of the clock signal controlling the switching. The main advantage of SC filters is that they can be implemented in digital circuit process technology, since the equivalent of large resistors can be simulated by capacitors having small capacitance values.

When using SC filters, one must also be aware that they are in themselves a sampling device that requires antialias protection on the input and filtering on their outputs to remove clock feedthrough. However, since clock frequencies are typically 50–100 times f_c of the filter, a simple first or second *RC* filter on their inputs and outputs will reduce aliases and noise sufficient to permit their use with 12–14 bit ADCs. One also need to consider that they typically have dc offset errors that are large, vary with time, temperature, and programming or clock frequency. Interested readers may refer to Refs. [13,14].

85.12 Adaptive Filters

Adaptive filtering is employed when it is necessary to realize or simulate a system whose properties vary with time. As the input characteristics of the filter change with time, the filter coefficients are varied with time as a function of the filter input. Some typical applications of adaptive filtering include spectral estimation of speech, adaptive equalization, echo cancellation, and adaptive control, to name just a few. Depending on the application, the variations in the coefficients are carried out according to an optimization criterion and the adaptation is performed at a rate up to the sampling rate of the system. The self-adjustment capability of adaptive filter algorithms is very valuable when the application environment cannot be precisely described. Some of the most widely used adaptive algorithms are least-mean square (LMS), recursive least squares (RLS), and frequency domain, also known as block algorithm. The fundamental concept of an adaptive filter is depicted in Figure 85.17.

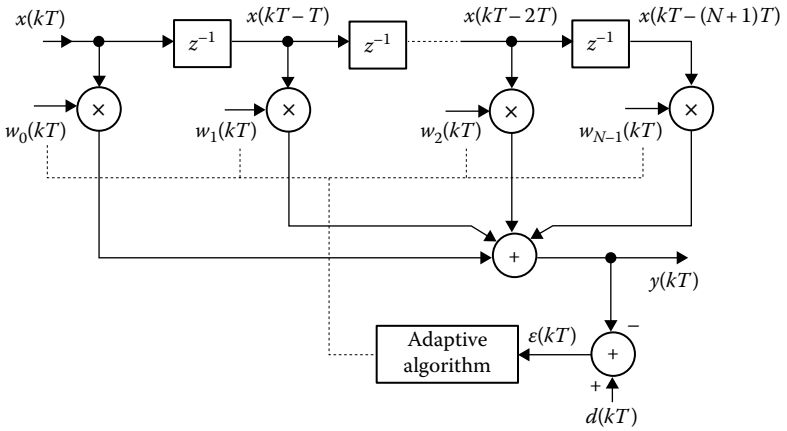


FIGURE 85.17 An adaptive filter uses an adaptive algorithm to change the performance of a digital filter in response to defined conditions.

An adaptive filter is characterized by the filter input $x(kT)$ and the desired response $d(kT)$. The error sequence $\varepsilon(kT)$ formed by

$$\varepsilon(kT) = \sum_{\mu=0}^{N-1} w_\mu(kT)x(kT - \mu T) \tag{85.80}$$

and $x(kT)$, …, $x(kT - T(N-1))$ serve as inputs to an adaptive algorithm that recursively determines the coefficients $w_0(kT + T)$, …, $w_{N-1}(kT + T)$. A number of adaptive algorithms and structures can be found in the literature that satisfies different optimization criteria in different application areas. For more detailed developments, refer to Refs. [1,5,15].

Defining Terms

Antialiasing filter: Antialiasing filters remove any frequency elements above the Nyquist frequency. They are employed before the sampling operation is conducted to prevent aliasing in the sampled version of the continuous-time signal.

Bandpass filter: A filter whose passband extends from a lower cutoff frequency to an upper cutoff frequency. All frequencies outside this range are stopped.

Equiripple: Characteristic of a frequency response function whose magnitude exhibits equal maxima and minima in the passband.

Finite impulse response (FIR) filter: A filter whose response to a unit impulse function is of finite length, i.e., identically zero outside a finite interval.

High-pass filter: A filter that passes all frequencies above its cutoff frequency and stops all frequencies below it.

Ideal filter: An ideal filter passes all frequencies within its passband with no attenuation and rejects all frequencies in its stopband with infinite attenuation. There are five basic types of ideal filters: low pass, high pass, bandpass, stopband, and all pass.

Infinite impulse response (IIR) filter: A filter whose response to a unit impulse function is of infinite length, i.e., nonzero for infinite number of samples.

Low-pass filter: A filter that attenuates the power of any signals with frequencies above its defined cutoff frequency.

Passband: The range of frequencies of a filter up to the cutoff frequency.

Stopband: The range of frequencies of a filter above the cutoff frequency.

Transition region: The range of frequencies of a filter between a passband and a stopband.

References

1. S. Mitra and J. Kaiser, *Handbook for Digital Signal Processing*, New York: John Wiley & Sons, 1993.
2. E. Christian and E. Eisenmann, *Filter Design Tables and Graphs*, New York: John Wiley & Sons, 1966.
3. R. Saal, *Handbuch zum Filterentwurf [Handbook of Filter Design]*, Frankfurt, German: Allgemeine Elektricitäts-Gesellschaft AEG-Telefunken, 1979.
4. C. Lindquist, *Active Network Design*, Long Beach, CA: Steward & Sons, 1977.
5. W. K. Chen, *The Circuits and Filters Handbook*, Boca Raton, FL: CRC Press, 1995.
6. DSP Committee, IEEE ASSP, Eds., *Programs for Digital Signal Processing*, New York: IEEE Press, 1979.
7. P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Englewood Cliffs, NJ: Prentice Hall, 1993.

8. Anonymous, *LabVIEW Digital Filter Design Toolkit*, Austin, TX: National Instruments, 2013, http://digital.ni.com/manuals.nsf/websearch/3707B0DA6E28618C86257B7200261CF0, accessed on August 23, 2013.
9. M. Cerna and R. Jamal, The design of digital filters using graphical programming techniques, *Proc. MessComp*, 232–238, 1995.
10. R. A. Roberts and C. T. Multis, *Digital Signal Processing*, Reading, MA: Addison-Wesley, 1987.
11. A. Antoniou, *Digital Filter: Analysis and Design*, New York: McGraw-Hill, 1979.
12. F. Taylor, *Digital Filter Design Handbook*, New York: Marcel Dekker, 1983.
13. A. Fettweis, Wave digital filters: Theory and practice, *Proc. IEEE*, 74, 270–327, 1986.
14. G. Temes and S. Mitra, *Modern Filter Theory and Design*, New York: John Wiley & Sons, 1973.
15. M. L. Honig and D. Messerschmitt, *Adaptive Filters—Structures, Algorithms and Applications*, Boston, MA: Kluwer Academic Publishers, 1984.
16. Edward P. Cunningham, *Digital Filtering*, New York: John Wiley & Sons, New York, 1995.
17. A. H. Gray, Jr. and J. D. Markel, Digital lattice and ladder filter synthesis, *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-23: 268–277, 1975.
18. M. Bellanger, *Digital Processing of Signals—Theory and Practice*, New York: John Wiley & Sons, 1988.
19. R. C. Dorf, and Z. Wan, Transfer functions of filters, in *The Electrical Engineering Handbook*, R. C. Dorf, Ed., Boca Raton, FL: CRC Press, 1997.
20. A. Zverev, *Handbook of Filter Synthesis*, New York: John Wiley & Sons, 1967.

86

Spectrum Analysis and Correlation

Ronney B. Panerai
University of Leicester

86.1 Introduction

Most sensors and instruments described in the previous chapters of this handbook can produce continuous measurements in time or sequential measurements at fixed or variable time intervals, as represented in Figure 86.1. The temporal patterns resulting from such measurements are usually referred to as *signals*. Signals can either be *continuous* or *discrete in time*. The main objective of *spectral analysis* is to provide an estimate of the distribution of signal power at different frequencies. Spectral analysis and correlation techniques are an aid to the interpretation of signals and to the systems that generate them. These methods are now widely used for the analysis and interpretation of measurements performed in medicine, geophysics, vibration analysis, communications, and several other areas.

Although the original concept of a *signal* involves measurements as a function of time (Figure 86.1), this term has been generalized to include measurements along other dimensions, e.g., distance. In addition, signals can have multiple dimensions—the instantaneous velocity of an airplane can be regarded as a four-dimensional signal since it depends on time and three spatial coordinates.

With the growing availability of signal-processing computer packages and dedicated instruments, most readers will perform spectral analysis and correlation at the "touch of a button," visualizing results on a screen or as a computer plot. These "black-box" systems are useful for saving time and money, but users should be aware of the limitations of the fundamental techniques and circumstances in which inappropriate use can lead to misleading results. This chapter presents the basic concepts of spectral analysis and correlation based on the *fast Fourier transform* (FFT) approach. FFT algorithms allow the most efficient computer implementation of methods to perform spectral analysis and correlation and have become the most popular option. Nevertheless, other approaches, such as parametric techniques, wavelet transforms (WT), and time-frequency analysis, are also available. These will be

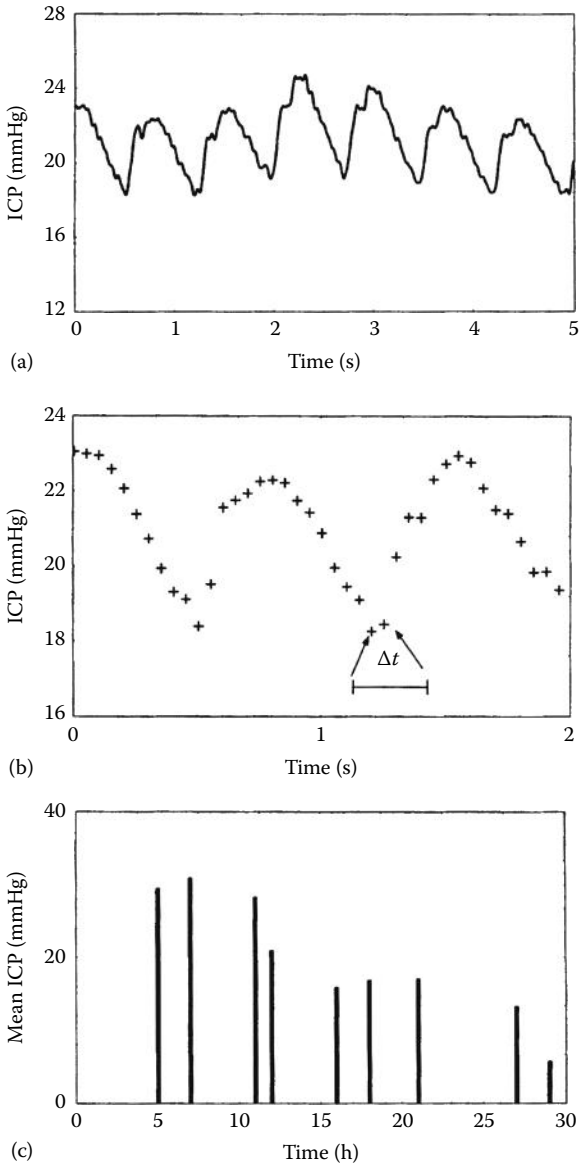


FIGURE 86.1 Examples of continuous and discrete-time signals. (a) Continuous recording of intracranial pressure in a head-injured patient, (b) intracranial pressure measurements obtained at regular intervals of 50 ms, and (c) nonuniformly spaced measurements of mean intracranial pressure over a period of 30 h following surgery.

briefly discussed and the interested reader will be directed to the pertinent literature for applications that might benefit from alternative approaches.

86.2 Fundamental Concepts

86.2.1 Spectral Analysis

Practical applications of spectral and correlation analysis are performed on discrete-time signals (Figure 86.1). These are obtained either from a sequence of discrete measurements or from the transformation of a continuous signal to digital format using an analog-to-digital converter (ADC).

When the latter is adopted to allow computer analysis of an originally continuous signal, two main characteristics of the ADC must be considered. The first is the number of bits available to represent each sample, as this will determine the resolution and accuracy of the sampled signal. The second important consideration is the *sampling interval* Δt (Figure 86.1). From the *Nyquist theorem* [1], the maximum value of Δt must be such that the *sampling frequency* $f_s - 1/\Delta t$ is at least twice the highest frequency of interest in the original signal. If this rule is not followed, spectral and correlation estimations might be considerably distorted by a phenomenon called *aliasing* [2]. Low-pass filtering before ADC is always recommended to limit the bandwidth of the continuous signal to allow the correct choice of f_s or Δt. In practice, the sampling frequency is usually much higher than the minimum required by the Nyquist theorem to provide a better visual representation of the sampled data.

Let x_n represent a discrete-time signal with samples at $n = 0, 1, 2, \ldots, N-1$. The Fourier theorem [1,2] states that it is possible to decompose x_n as a sum of cosine and sine waveforms of different frequencies using an appropriate combination of amplitude coefficients. Therefore,

$$x_n = a_0 + \sum_{k=1}^{N-1} a_k \cos\left(\frac{2\pi kn}{N}\right) + \sum_{k=1}^{N-1} b_k \sin\left(\frac{2\pi kn}{N}\right) \tag{86.1}$$

where $k = 1, 2, \ldots, N-1$ determines the frequency of each cosine and sine waveforms as $f_k = k/N\Delta t$. The corresponding coefficients are calculated from

$$a_0 = \frac{1}{N}\sum_{n=0}^{N-1} x_n \tag{86.2a}$$

$$a_k = \frac{1}{N}\sum_{n=0}^{N-1} x_n \cos\left(\frac{2\pi kn}{N}\right) \tag{86.2b}$$

$$b_k = \frac{1}{N}\sum_{n=0}^{N-1} x_n \sin\left(\frac{2\pi kn}{N}\right) \tag{86.2c}$$

Note that Equation 86.2a represents the mean value of x_n and that the argument $2\pi kn/N$ is the same for the *direct* (Equation 86.2) and *inverse* (Equation 86.1) *discrete Fourier transforms* (DFT).

From Euler's formula [3], it is possible to combine the cosine and sine terms to express the DFT in exponential form:

$$e^{j\theta} = \cos\theta + j\sin\theta \tag{86.3}$$

leading to

$$x_n = \sum_{k=0}^{N-1} c_k e^{j(2\pi kn/N)} \tag{86.4}$$

with

$$c_k = \frac{1}{N}\sum_{n=0}^{N-1} x_n e^{-j(2\pi kn/N)} \tag{86.5}$$

where c_k is now a complex value related to the original cosine and sine coefficients by

$$c_0 = a_0 \tag{86.6a}$$

$$c_k = a_k - jb_k \quad k = 1, 2, \ldots, N-1 \tag{86.6b}$$

A graphic representation of the a_k, b_k, or c_k coefficients for each value of k (or f_k) constitutes the frequency spectrum of x_n, expressing the relative contribution of different sinusoidal frequencies to the composition of x_n (Equation 86.4). Since c_k is complex (Equation 86.6b), a more meaningful physical interpretation of the spectrum is obtained with the *amplitude* and *phase* spectra, defined as

$$A_k = \left(a_k^2 + b_k^2\right)^{1/2} = |\, c_k \,| \tag{86.7a}$$

$$\theta_k = \tan^{-1}\left(-\frac{b_k}{a_k}\right) \tag{86.7b}$$

Figure 86.2 shows the amplitude (or magnitude) and phase spectra for the signal in Figure 86.1a, sampled at intervals Δt = 20 ms. The signal was low-pass-filtered at 20 Hz before ADC. The total duration is

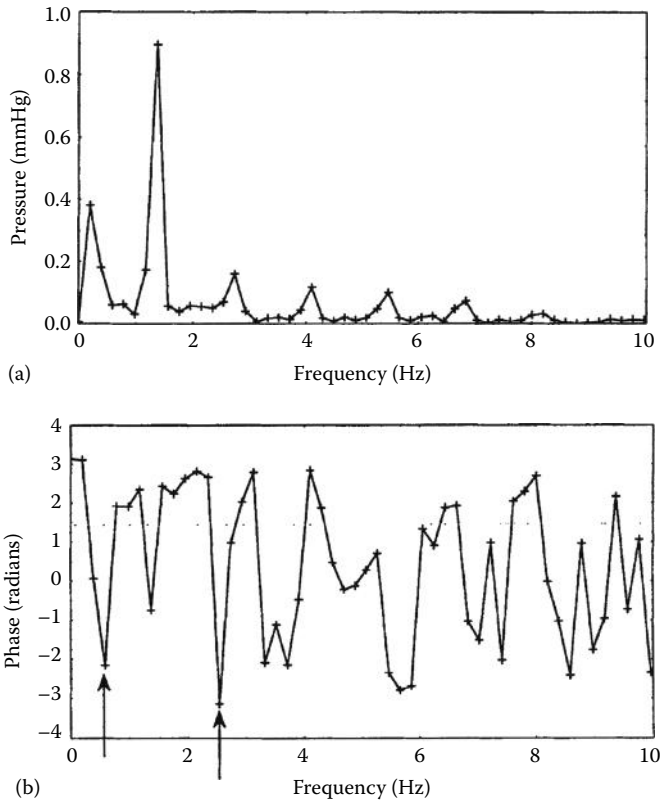


FIGURE 86.2 (a) Amplitude, and (b) phase spectra of the intracranial pressure signal represented in Figure 86.1a after analog-to-digital conversion with a sampling interval of 20 ms. The main peak in the amplitude spectrum corresponds to the frequency of the cardiac cycle in Figure 86.1a. Wraparound of the phase spectrum is apparent in the third and 13th harmonics (arrows). Both spectra have been plotted to 10 Hz only.

given by $T = N\Delta t = 5$ s, corresponding to $N = 250$ samples. Before calculating the spectral coefficients, the mean value of the complete record was removed (dc term) and any linear trends were removed by fitting a straight line to the data (detrending). As will be discussed later, it is also important to apply a window to the data, to minimize the phenomenon of leakage. For $k > N/2$, both spectra present symmetrical values. This can be easily demonstrated from the fact that cosine (and sine[2]) functions have even symmetry while sine has odd symmetry. From Equation 86.7, it follows that A_k and θ_k have *even* and *odd* symmetry, respectively [4]. Consequently, only half the spectral components ($k < N/2$) are required to give a complete description of x_n in the frequency domain.

The amplitude spectra indicate the combined amplitude of the cosine and sine terms to reconstruct x_n; the phase spectra reflect the relative phase differences (or time delays) between the sinusoidal waveforms to generate the temporal pattern of x_n. The amplitude spectra also reflect the signal power at different frequencies. For simplicity, the power spectrum can be defined as

$$P_k = A_k^2 = |c_k|^2 \tag{86.8}$$

Direct implementation of Equation 86.8, however, leads to spectral power estimates which are biased and inconsistent. More appropriate procedures for estimating the *power spectrum* (or *power density spectrum*) will be discussed later.

Parseval's theorem [5] demonstrates that the total signal energy can be computed in either time or frequency domain:

$$\frac{1}{N}\sum_{n=0}^{N-1} x_n^2 = \sum_{k=0}^{N-1} P_k \tag{86.9}$$

If x_n has zero mean, the left-hand side of Equation 86.9 is the biased estimator of signal variance [6]. Although most applications of spectral analysis concentrate on the characteristics of the amplitude or power spectra, it is important to bear in mind that the phase spectrum is also responsible for the temporal pattern of x_n. As an example, both the Dirac impulse function and white noise have a flat, constant amplitude (or power) spectra [6], it is the difference in the phase spectra which accounts for the different morphologies in the time domain.

Interpretation of the amplitude and phase spectra of both theoretical functions and sampled data is facilitated by taking into account several properties of the DFT (Equations 86.4 and 86.5), namely, *symmetry*, *linearity*, *shifting*, *duality*, and *convolution* [7]. To these, a very important property of Equations 86.1 and 86.4 must be added. Since cosine and sine functions are periodic, and exist for $-\infty < t < \infty$, Equations 86.1 and 86.4 will reconstruct x_n not only in the interval of interest ($0 \le t \le$ T) but also at all other multiple intervals $pT \le t \le (p + 1)T (p = 0, \pm 1, \pm 2, \ldots)$. As a consequence, spectral estimations obtained with the DFT inherently assume that x_n is *periodic* with period $T = T/\Delta t$. As discussed in the following sections, this property needs to be taken into account when performing spectral analysis with the DFT and FFT.

86.2.2 Correlation Analysis

The basic concept of the correlation coefficient, as a measure of the strength of linear relationship between two variables [6], can be extended to signal analysis with the definition of the *cross-correlation function* (CCF) as [5]

$$r_{xy}(p) = \frac{1}{N}\sum_{n=0}^{N-1} x_n y_{n-p} \quad p = 0, \pm 1, \pm 2, \ldots \tag{86.10}$$

where x_n and y_n are zero-mean, discrete-time signals defined in the interval $n = 0, 1, 2, \ldots, N-1$. For each value of p, the cross correlation is computed by shifting y_n by $p\Delta t$ and calculating the average product in Equation 86.10. If x_n and y_n are unrelated, the sum of positive and negative products will tend to zero. Conversely, y_n tends to follow x_n, but with a time delay D, $r_{xy}(p)$ will show a peak at $p = D/\Delta t$. This property of the CCF is illustrated in Figure 86.3. As noted by Bergland [8], cross correlation can be viewed as "one signal searching to find itself in another signal."

For $y_n = x_n$, $r_{xx}(p)$ becomes the *autocorrelation function* (ACF):

$$r_{xx}(p) = \frac{1}{N}\sum_{n=0}^{N-1} x_n x_{n-p} \tag{86.11}$$

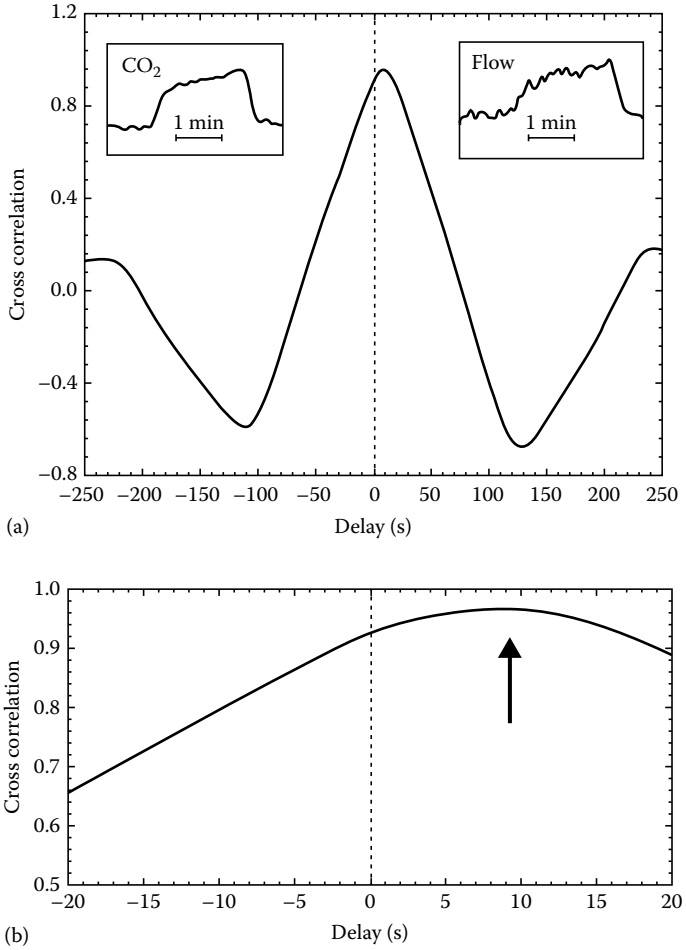


FIGURE 86.3 CCF between changes in arterial CO_2 and blood flow to the brain. Arterial CO_2 was estimated from end-tidal measurements and cerebral blood flow with Doppler ultrasound in the middle cerebral artery. (a) CCF and original signals (inserts). The cross-correlation value of ~1.0, observed at time delays near zero, reflects the similar temporal patterns between the two measurements. The negative cross correlations are obtained when either signal is shifted by approximately the duration of the plateau phase, which lasts 2 min. (b) Enlarging the scale around delay = 0 shows that the peak cross correlation occurs at 10 s, reflecting the time it takes for the flow to respond to the CO_2 change. (Data kindly provided by Dr. Joanne Dumville, Mr. A. Ross Naylor, and Prof. David H. Evans, University of Leicester, Leicester, U.K.)

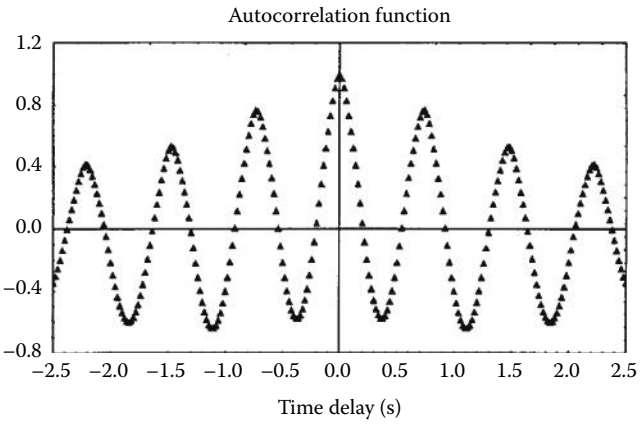


FIGURE 86.4 ACF of the discrete-time version of the signal in Figure 86.1a. The periodicity of the ACF reflects the quasi-periodic pattern of the intracranial pressure signal (Figure 86.1a).

and it is intuitive that the maximum value of $r_{xx}(p)$ occurs for $p = 0$ with

$$r_{xx}(0) = \frac{1}{N}\sum_{n=0}^{N-1} x_n^2 \tag{86.12}$$

which represents the signal variance or total energy. Therefore, for signals with unit standard deviation, the autocorrelation peak is equal to 1.

The Wiener–Khintchine theorem [9] demonstrates that the ACF and the power spectrum constitute a Fourier transform pair, that is,

$$S_k = \sum_{p=0}^{N-1} r_{xx}(p) e^{-j(2\pi kp/N)} \tag{86.13}$$

where S_k is usually called the autospectra of x_n. Equation 86.13 indicates that it is possible to estimate the power spectra from a previous estimate of the ACF. As a transform pair, the ACF can also be derived from the autospectra by substituting S_k for c_k in Equation 86.4.

From Equation 86.11, it is clear that $r_{xx}(p)$ has even symmetry, that is, $r_{xx}(+p) = r_{xx}(-p)$. This property is apparent in Figure 86.4, which shows the estimated ACF for the signal in Figure 86.1a. Another characteristic of ACF, which can be visualized in Figure 86.4, is the occurrence of secondary peaks reflecting the presence of an oscillatory component in x_n (Figure 86.1a).

86.3 Fast Fourier Transform

The FFT is not a single algorithm but rather a large family of algorithms which can increase the computational efficiency of the DFT. The main ideas behind the formulation of FFT algorithms are discussed later. A detailed description of the different algorithms that have been proposed is beyond the scope of this introduction; this can be found in Refs. [5–7,10–14].

For both software and hardware implementations of Equations 86.4 and 86.5, the computational efficiency is usually expressed by the number of complex multiplications and additions, or simply, by the *number of* operations [10]. Straight implementation of either Equation 86.4 or 86.5 leads to N^2 operations. Typically, FFT algorithms can reduce this number to $N\log_2 N$. For $N = 1024$, the FFT algorithm is 100 times faster than the direct implementation of Equation 86.4 or 86.5.

The essence of all FFT algorithms is the periodicity and symmetry of the exponential term in Equations 86.4 and 86.5, and the possibility of breaking down a transform into a sum of smaller transforms for subsets of data. Since n and k are both integers, the exponential term is periodic with period N. This is commonly represented by

$$W_N = e^{-j(2\pi/N)} \tag{86.14}$$

and Equation 86.5 can be written as

$$c_k = \frac{1}{N}\sum_{n=0}^{N-1} x_n W_N^{kn} \quad k = 0, 1, 2, \ldots, N-1 \tag{86.15}$$

In many applications, the terms W_N^{kn} are called *twiddle factors*. Assuming $N = 8$, calculation of the DFT with Equation 86.15 will require 64 values of w_8^{kn}. Apart from the minus sign, a simple calculation can show that there are only four different values of this coefficient, respectively: $1, j, (1+j)/\sqrt{2}$, and $(1-j)/\sqrt{2}$. Consequently, only these four complex factors need to be computed, representing a significant savings in number of operations.

Most FFT algorithms are based on the principle of *decimation-in-time*, involving the decomposition of the original time (or frequency) sequence into smaller subsequences. To understand how this decomposition can reduce the number of operations, assume that N is even. In this case, it is possible to show that Equation 86.15 can be written as [4,5,7,11]:

$$c_k = \frac{1}{N}\sum_{r=0}^{(N/2)-1} x_r^e \cdot W_{N/2}^{kr} + \frac{1}{N}\sum_{r=0}^{(N/2)-1} x_r^0 \cdot W_{N/2}^{kr} \tag{86.16}$$

where x_r^e and x_r^0 represent the even- and odd-order samples of x_n, respectively. Comparing Equations 86.15 and 86.16, it is clear that the latter represents two DFTs with dimension $N/2$, involving $2(N/2)^2$ operations rather than the N^2 operations required by Equation 86.15. This process of decimation-in-time can be carried out further to improve computational performance. In the general case, N can be decomposed into q factors:

$$N = \prod_{i=1}^{q} r_i = r_1 r_2 \ldots r_q \tag{86.17}$$

The number of operations required is then [6]

$$\text{Number of operations} = N\sum_{i=1}^{q} r_i \tag{86.18}$$

In the original algorithm of Cooley and Tukey [10], $r_i = 2$ and $N = 2^q$. In this case, the theoretical number of operations required would be $2Nq = 2N\log_2 N$. As pointed out in Ref. [6], further improvements in efficiency are possible because of the symmetry of the twiddle factors. The efficiency gain of most FFT algorithms using radix-2, i.e., $N = 2^q$ is

$$\text{Efficiency gain} = \frac{N^2}{N\log_2 N} = \frac{N}{\log_2 N} \tag{86.19}$$

For $N = 1024$, $q = 10$ and the efficiency gain is ~100. Specific applications might benefit from other decompositions of the original sequence. Cases of particular interest are radix-4 and radix-8 FFTs [14]. However, as shown by Rabiner and Gold [11] (p. 585), it is not possible to generalize the superiority of radix-8 over radix-4 algorithms.

In general, most FFT algorithms accept complex x_n sequences in Equation 86.5. By limiting x_n to the most common situation of real-valued signals, it is possible to obtain more efficient algorithms as demonstrated by Sorensen et al. [15]. Uniyal [16] performed a comparison of different algorithms for real-valued sequences showing that performance is architecture-dependent. For machines with a powerful floating point processor, the best results were obtained with Brunn's algorithm [17].

The application of FFT algorithms for spectral and correlation analysis is discussed in the following sections.

86.4 FFT Spectral Analysis

For some deterministic signals, x_n can be expressed by a mathematical function and the amplitude and phase spectra can be calculated as an exact solution of Equations 86.5 and 86.7. The same is true for the power spectra (Equations 86.8 and 86.13). Examples of this exercise can be found in many textbooks [1,2,4,5,7].

In most practical applications, there is a need to perform spectral analysis of experimental measurements, corresponding to signals which, in general, cannot be described by simple mathematical functions. In this case, the spectra have to be estimated by a numerical solution of Equations 86.5 through 86.8, which can be efficiently implemented on a digital computer with an FFT algorithm. For estimation of the power spectrum, this approach is often classified as *nonparametric*, as opposed to other alternatives which are based on parametric modeling of the data such as *autoregressive* methods [18]. Considerable distortions can result from applications of the FFT unless attention is paid to the following characteristics and properties of the measured signal and the DFT/FFT.

86.4.1 Limited Observation of Signal in Time

Limited observation of a signal x_n in time can be seen as the multiplication of the original signal x_n^∞ by a rectangular window of duration $T = N\Delta t$ as exemplified for a single sinusoid in Figure 86.5. The DFT assumes that x_n is periodic, with period T, as mentioned previously. Instead of a single harmonic at the frequency of the original sinusoid, the power spectrum estimated with the FFT will have power at other harmonics as indicated by the spectrum in Figure 86.5c. The spectral power, which should have been concentrated on a single harmonic (Figure 86.5c, dashed line), has "leaked" to neighboring harmonics and for this reason this phenomenon is usually called *leakage*. The morphology of the distorted spectrum of Figure 86.5c can be explained by the fact that the Fourier transform of a rectangular window function (Figure 86.5b) is given by a sin c function (sin x/x) which presents decreasing side lobes [1,2]. Multiplication in time corresponds to the convolution operation in the frequency domain. In the general case of signals comprising several harmonics, the sin c functions will superimpose and the resulting spectrum is then a distorted version of the "true" spectrum. As the individual *sine* functions superimpose to produce the complete spectrum, a *picket-fence* effect is also generated [8]. This means that spectral leakage not only adds spurious power to neighboring harmonics but also restricts the frequency resolution of the main spectral peaks. The effects of spectral leakage can be reduced by (1) increasing the period of observation and (2) multiplying the original signal x_n by a window function with a smooth transition as represented by the dashed line window in Figure 86.5b. The Fourier transform of a window function with tapered ends has smaller side lobes, thus reducing the undesirable effects leakage.

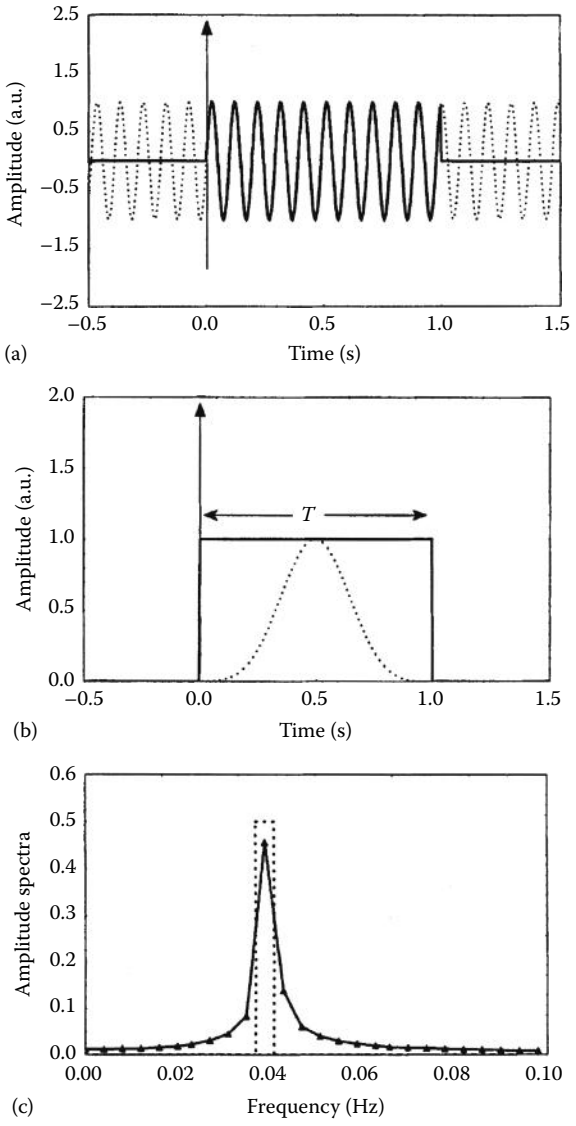


FIGURE 86.5 Effect of limited observation time T on the amplitude spectra of a sinusoidal component. (a) Observation of a single harmonic (dashed line) for a limited period of time T is equivalent to the multiplication for the rectangular function represented in (b). The Blackman–Harris window is also represented in (b) (dashed line). (c) Truncating a single harmonic produces spectral estimates smeared by *leakage* (solid line) as compared with the theoretical result (dashed line) with width equal to the frequency resolution ($f_r \approx 0.004$ Hz).

A large number of tapering windows have been proposed, as reviewed by Harris [19]. As an example, the four-term Blackman–Harris window, defined as

$$w_n = a_0 - a_1 \cos\left(\frac{2\pi n}{N}\right) + a_2 \cos\left(\frac{4\pi n}{N}\right) - a_3 \cos\left(\frac{6\pi n}{N}\right) \quad n = 0, 1, 2, \ldots, N-1 \tag{86.20}$$

produces side lobe levels of –92 dB if the a coefficients are chosen as $a_0 = 0.35875$, $a_1 = 0.48829$, $a_2 = 0.14128$, $a_3 = 0.01168$ [19]. Windows also play an important role in the sampling properties of power spectral estimates, as will be discussed later. Windowing attenuates the contribution of signal samples

at the beginning and end of the signal and, therefore, reduces its effective signal duration. This effect is reflected by the equivalent noise bandwidth (ENBW) defined as [19]

$$\text{ENBW} = \frac{\sum_{n=0}^{N-1} w_n^2}{\left[\sum_{n=0}^{N-1} w_n\right]^2} \tag{86.21}$$

For a rectangular window ENBW = 1.0 and for the Blackman–Harris window (Equation 86.20), the corresponding value is 2.0. The majority of other window shapes have intermediate values of ENBW [19].

86.4.2 Effects of "Zero Padding"

Most FFT algorithms operate with $N = 2^q$ samples, the choice of q is many times critical. Since frequency resolution is inversely proportional to N, in many circumstances a value of q leading to $2^q > N$ is preferable to the option of limiting the signal to $N' = 2^{q-1}$ samples with $N' < N$. The most common and simple way of extending a signal to comply with the $2q$ condition is by *zero padding*. For signals with zero mean and with first and last values around zero, this can be accomplished by complementing the signal with Q zeros to achieve the condition $N + Q = 2^q$. For signals with end points different from zero, these values can be used for padding. If initial and final values differ significantly, a linear interpolation from the last to the first point is also a practical option. However, with the application of windowing, most signals will have similar initial and final points and these can be used for zero padding. As discussed in the next section, zero padding has important applications for the estimation of correlation functions via FFT. For spectral analysis, it is relatively simple to demonstrate that adding Q zeros corresponds to over sampling the N point original spectrum with a new frequency resolution which is $(N + Q)/N$ times greater than the original resolution. Consequently, although zero padding does not introduce major distortions, it produces the false illusion of higher resolution than warranted by the available N measured signal samples.

86.4.3 Phase Spectrum Estimation

The use of Equation 86.7b to estimate the phase spectrum is fraught with a different kind of problem, resulting from the indetermination of the $\tan^{-1}$ function to discriminate between phase angles with absolute values greater than π. This problem is illustrated in Figure 86.2 showing that phase angles decrease continuously until reaching $-\pi$ and then "jump" to continue decreasing from the $+\pi$ value. This feature of the phase spectrum is called *wraparound*. Methods to "unwrap" the phase spectrum have been proposed [20], but a general satisfactory solution to this problem is not available. In some cases, the shifting property of the DFT [5,7] can be used to "rotate" the original signal in time, thus minimizing the slope of the phase spectrum and, consequently, the occurrence of wraparound.

86.4.4 Sampling Properties of Spectral Estimators

The most straightforward approach to computing the power spectrum is to use Equation 86.8. This method is known as the *periodogram* [5]. Application is limited to signals which are *stationary*, meaning stable statistical properties (such as the mean and the variance) along time. For measurements performed on *nonstationary* systems, such as speech or systems with time-varying parameters, other methods of spectral estimation are available and will be mentioned later. It is possible to demonstrate that when the period of observation T tends to infinity, Equation 86.8 gives an unbiased estimate of the power spectrum. In practice, due to finite values of T, the phenomenon of spectral leakage described earlier will lead to power spectral estimates which are *biased*.

The second inherent problem with the periodogram is the *variance* of the resulting spectral estimates. Assuming x_n to follow a Gaussian distribution, it follows that a_k and b_k will also be Gaussian because

Equation 86.2 represents a linear transformation. Since Equations 86.7a and 86.8 involve the sum of two squared Gaussian variates, P_k will follow a χ^2 distribution with two degrees of freedom [6]. In this case the mean and the standard deviation of the power spectral estimate will be the same, *independently of the frequency considered*. As a consequence, power spectral estimates obtained from Equation 86.8, using a simple sample x_n, should be regarded as highly unreliable. In addition, the variance or standard deviation of this χ^2 distribution does not decrease with increases in sample duration N. This indicates that the periodogram (Equation 86.8) is an *inconsistent* estimator of the power spectrum.

For a χ^2 distribution with m degrees of freedom, the coefficient of variation is given by

$$\mathrm{CV}\left[\chi_m^2\right] = \frac{\sqrt{2m}}{m} = \sqrt{\frac{2}{m}} \tag{86.22}$$

showing that it is possible to improve the reliability of power spectral estimates by increasing m. This can be achieved by replacing Equation 86.8 by [21]

$$\hat{P}_k = \frac{1}{L}\sum_{l=1}^{L} c_{k,l}^2, \quad k = 0, 1, 2, \ldots, N-1 \tag{86.23}$$

with L representing a number of separate samples x_n each with length $T = N\Delta t$. If only one record of x_n can be obtained under stationary conditions, it is possible to break down this record into L segments to obtain an improved estimate of the power spectrum with variance reduced by a factor of L. However, the spectral resolution, given by $f_r = 1/T$, will be reduced by the same factor L, thus indicating an inescapable compromise between resolution and variance.

A *modified periodogram* was introduced by Welch [21] consisting of the multiplication of x_n by a triangular, or other window shape, before computing the individual spectral samples with Equation 86.5. The application of a window justifies overlapping adjacent segments of data by as much as 50%. For a signal with a total duration of N samples, the combination of overlapping with segmentation (Equation 86.23) can lead to a further reduction of the spectral variance by a factor of 11/18.

Averaging L spectral samples as indicated by Equation 86.23 represents one approach to improve spectral estimation by means of *smoothing*. A similar effect can be obtained with the *correlogram*. Equation 86.13 indicates that it is possible to estimate the power spectrum from the ACF. Limiting the number of shifts of the ACF to $p \ll N$ becomes equivalent to smoothing the original spectrum by convolution with the Fourier transform of a Bartlett (triangular) window [22]. As discussed in the next section, the ACF can also be computed more efficiently with the FFT and it can be shown that in this case it involves a smaller number of numerical operations than the Welch method based on the periodogram [5].

86.5 FFT Correlation Analysis

Before considering the application of FFT algorithms to compute ACF and CCF, it is important to discuss their sampling properties using Equations 86.10 and 86.11 as estimators. Assuming that variables x_n and y_n are not defined outside the interval $0 \le n \le N-1$, it follows from Equation 86.10 that as p increases and the two functions "slide" past each other, the effective number of summed products is $N - |p|$ rather than N as implied by Equations 86.10 and 86.11. For this reason, these equations are often rewritten as

$$r_{xy}(p) = \frac{1}{N-|p|}\sum_{n=0}^{N-|p|-1} x_n y_{n-1}, \quad p = 0, \pm 1, \pm 2, \ldots \tag{86.24}$$

The main justification for this modification, however, is that Equations 86.10 and 86.11 lead to biased estimations of correlation functions while Equation 86.24 is *unbiased*.

Equation 86.24 normally assumes that x_n and y_n are standardized variables with zero mean and unit variance. If the mean values are different from zero, Equations 86.10 and 86.11 will produce distorted estimates with a "pyramid effect" due to the presence of the dc term. However, this effect is compensated for in Equation 86.24 and in this case the effect of the mean value is to add a constant term:

$$r_{xy}(p) = r'_{xy}(p) - m_x m_y \tag{86.25}$$

where $r'_{xy}(p)$ is the cross correlation of variables with mean values m_x and m_y, respectively.

Similarly to the DFT, Equations 86.10 and 86.11 involve N^2 operations and Equation 86.24 slightly less. Since the autocorrelation and the power spectra constitute a Fourier transform pair (Equation 86.13), the computation of correlation functions can also be sped up by means of an FFT algorithm. For the sake of generality, the cross *spectrum* of x_n and y_n can be defined as [6]

$$C_{xy}(f_k) = X(f_k)Y^*(f_k) \tag{86.26}$$

with $X(f_k)$ and $Y(f_k)$ representing the Fourier transforms of x_n and y_n, respectively. The generalized Wiener–Khintchine theorem then gives the CCF as

$$r_{xy}(p) = \sum_{k=0}^{N-1} C_{xy}(f_k) e^{-j(2\pi kp/N)} \tag{86.27}$$

Therefore, "fast" correlation functions can be computed [23] using the forward FFT to calculate $X(f_k)$ and $Y(f_k)$ and then the inverse FFT to obtain $r_{xy}(p)$ with Equation 86.27. Obviously, when ACF are being computed with this method, only one transform is necessary to obtain the autospectra (Equation 86.13) instead of the cross spectra (Equation 86.26).

When correlation functions are computed with the FFT, it is critical to pay attention again to the periodicity of the transformed variables as an intrinsic property of the DFT. When the two functions in either Equation 86.10 or 86.11 are displaced by p samples, for periodic functions there will be nonzero products outside the range $0 \le n \le N - 1$, thus leading to significant errors in the estimated ACF or CCF. In this case, the resulting estimates are called *circular* correlations [6]. This error can be avoided by zero padding the original signals from $n = N$ to $n - 2N - 1$ and computing the FFTs with $2N$ points. The resulting correlation functions will be noncircular and, in the range $0 \le p \le N - 1$ will agree with correlations computed with the original Equation 86.10 or 86.11. Finally, to remove bias the results of Equation 86.27 should also be multiplied by $NI(N - |p|)$ to agree with Equation 86.24.

86.6 Further Information

Software for FFT special analysis is available from multiple sources. Off-the-shelf software ranges from specialized packages for digital signal processing, such as DADiSP, to statistical packages which include FFT analysis of time series. Mathematical and engineering packages such as MATLAB® also include routines for FFT spectral and correlation analysis. For a review of available options, see Ref. [24]. For readers who want to implement their own software, FFT routines can be found in Refs. [4,13–15,18]. Additional references are [25–27].

Hardware implementations of FFT algorithms are common in areas requiring real-time spectral analysis as in the case of blood flow velocity measurement with Doppler ultrasound. For a review of hardware implementations, see Refs. [11,28]. Developments in this area follow the pace of change in VLSI technology [29].

One of the limitations of the FFT is the fact that frequency resolution is the inverse of the signal observation time. Improved resolution can be obtained with *parametric* methods of spectral analysis and their application is particularly relevant when only short segments of data are available or when it is necessary to discriminate between frequency harmonics which are closely spaced in the spectrum. Broadly speaking, parametric methods assume that the data follow spectral densities with a known pole–zero structure of variable complexity, characterized by a given *model order*. All-zero models correspond to the *moving average* structure while the all-pole version represents the *autoregressive* model. The general case is the autoregressive-moving average model (ARMA). For a comprehensive review of these methods, see Ref. [30]; further information and software implementations can be found in Refs. [18,27].

Nonstationary signals present a particular problem. In cases where the signal statistical properties change relatively slowly with time, it is possible to select short segments of quasi-stationary data and to use the DFT or parametric methods to estimate the spectra as mentioned previously. However, when these changes in systems parameters or statistical moments are fast in relation to the phenomenon under observation (e.g., speech or seismic data), this approach is not feasible because of the poor frequency resolution resulting from short observation times. Methods proposed to cope with signal nonstationarity often depend on the underlying cause of nonstationary behavior [9,31]. More general methods, known as *time-frequency distributions*, are now favored by most investigators [32]. The Wigner–Ville and Choi–Williams transforms are some of the more widely used of these time-frequency distributions. In each case the signal is described by a simultaneous function of time *and* frequency and hence is graphically represented by a three-dimensional plot having time and frequency as dependent variables.

A different approach to the analysis of nonstationary data is the application of *wavelets* [33]. This alternative also has advantages in the representation of fast transients and in applications requiring data compression and pattern classification. Similarly to the sine and cosine functions, which are the basis of Fourier analysis, wavelets are orthogonal functions which can be used to decompose and reconstruct signals using a finite set of coefficients obtained by a WT. The main difference between wavelets and sinusoids, however, is that the former are limited in time. In addition, the complete orthogonal set of wavelets can be obtained simply by expansion (or compression) and scaling of a single function, known as the *mother wavelet*. Because of their limited time duration, wavelets can provide a much more synthetic decomposition of fast transients, or sharp edges in image analysis, than it is possible to obtain with the DFT. Their property of expansion/contraction of a single mother wavelet can also overcome a major limitation of the DFT, that is, to allow good frequency resolution at both low and high frequencies. For applications of the WT and commercially available software, see Refs. [34,35].

Defining Terms

Analog-to-digital conversion: The process of converting a continuous signal to a discrete-time sequence of values usually sampled at uniform time intervals.

Autocorrelation function (ACF): A measure of longitudinal variability of a signal which can express the statistical dependence of serial samples.

Correlogram: Numerical calculation and graphical representation of the ACF or CCF.

Cross-correlation function (CCF): A measure of similarity between signals in the time domain which also allows the identification of time delays between transients.

Cross spectrum: The complex product of the power spectra of two different signals.

Decimation-in-time: The process of breaking down a time series into subsequences to allow more efficient implementations of the FFT.

Discrete Fourier transform (DFT): The usual method to obtain the Fourier series of a discrete-time signal.

Fast Fourier transform (FFT): Algorithm for the efficient computation of the DFT.

Periodogram: A family of methods to estimate the power spectrum using the DFT.

Power spectrum: The distribution of signal power as a function of frequency.

Signal: Continuous or discrete representation of a variable or measurement as a function of time or other dimension.

Stationarity: Property of signals which have statistical moments invariant with time.

Twiddle factors: Exponential term in the DFT whose periodicity allows repeated use and hence considerable savings of computation time in the FFT.

Wraparound: Overflow of phase spectral estimations above $|\pi|$ due to the uncertainty of the $\tan^{-1}$ function.

Zero padding: Extension of a signal with zeros, constant values, or other extrapolating functions.

References

1. B.P. Lathi, *Communication Systems*, New York: John Wiley & Sons, 1968.
2. R.N. Bracewell, *The Fourier Transform and Its Applications*, Englewood Cliffs, NJ: Prentice Hall, 1988.
3. E. Kreyszig, *Advanced Engineering Mathematics*, New York: John Wiley & Sons, 1962.
4. P.A. Lynn and W. Fuerst, *Introductory Digital Signal Processing with Computer Applications*, Chichester, U.K.: John Wiley & Sons, 1989.
5. J.G. Proakis and D.G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 2nd edn., New York: Macmillan, 1992.
6. J.S. Bendat and A.G. Piersol, *Random Data: Analysis and Measurement Procedures*, 2nd edn., New York: John Wiley & Sons, 1986.
7. A.V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing*, Englewood Cliffs, NJ: Prentice Hall, 1989.
8. G.D. Bergland, A guided tour of the fast Fourier transform, *IEEE Spectrum*, 6: 41–52, 1969.
9. M.B. Priestley, *Spectral Analysis and Time Series*, London, U.K.: Academic Press, 1981.
10. J.W. Cooley and J.W. Tukey, An algorithm for the machine computation of complex Fourier series, *Math. Computation*, 19: 297–301, 1965.
11. L.R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Englewood Cliffs, NJ: Prentice Hall, 1975.
12. E.O. Brigham, *The Fast Fourier Transform and Its Applications*, 2nd edn., Englewood Cliffs, NJ: Prentice Hall, 1988.
13. Digital Signal Processing Committee, *Programs for Digital Signal Processing*, New York: IEEE Press, 1979.
14. R.C. Singleton, An algorithm for computing the mixed radix fast Fourier transform, *IEEE Trans. Audio Electroacoust*., 17: 93–103, 1969.
15. H.V. Sorensen, D.L. Jones, M.T. Heideman, and C.S. Burrus, Real-valued fast Fourier transform algorithms, *IEEE Trans. Acoust. Speech Signal Proc*., 35: 849–863, 1987.
16. P.R. Uniyal, Transforming real-valued sequences: Fast Fourier versus fast Hartley transform algorithms, *IEEE Trans. Signal Proc*., 41: 3249–3254, 1994.
17. G. Brunn, *z*-Transform DFT filters and FFT, *IEEE Trans. Acoust. Speech Signal Proc*., 26: 56–63,1978.
18. S.M. Kay, *Modern Spectral Estimation*, Englewood Cliffs, NJ: Prentice Hall, 1988.
19. F.J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, *Proc. IEEE*, 66: 51–83, 1978.
20. J.M. Tribolet, A new phase unwrapping algorithm, *IEEE Trans. Acoust. Speech Signal Proc*., 25: 170–177, 1977.

21. P.D. Welch, The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Trans. Audio Electroacoust*., 15: 70–73, 1967.
22. R.B. Blackman and J.W. Tukey, *The Measurement of Power Spectra*, New York: Dover Publications, 1958.
23. T.G. Stockham, Jr., High-speed convolution and correlation, *1966 Spring Joint Computer Conference, AFIPS Proc*., 28: 229–233, 1966.
24. R. Braham, Math & visualization: New tools, new frontiers, *IEEE Spectrum*, 32(11): 19–36, 1995.
25. W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge, U.K.: Cambridge University Press, 1986.
26. D.M. Monro, Complex discrete fast Fourier transform, *Appl. Stat*., 24: 153–160, 1975.
27. S.L. Marple, Jr., *Digital Spectral Analysis with Applications*, Englewood Cliffs, NJ: Prentice Hall, 1987.
28. G.D. Bergland, Fast Fourier transform hardware implementations—An overview, *IEEE Trans. Audio Electroacoust*., 17: 104–108, 1969.
29. E. Bidet, D. Castelain, C. Joanblanq, and P. Senn, A fast single-chip implementation of 8192-complex point FFT, *IEEE J. Solid-State Circuits*, 30: 300–305, 1995.
30. S.M. Kay and S.L. Marple, Jr., Spectrum analysis—A modern perspective, *Proc. IEEE*, 69: 1380–1419, 1981.
31. J. Leuridan, H.V. Van der Auweraer, and H. Vold, The analysis of nonstationary dynamic signals, *Sound Vibration*, 28: 14–26, 1994.
32. L. Cohen, Time-frequency distributions—A review, *Proc. IEEE*, 77: 941–981, 1989.
33. I. Daubechies, Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math*., 41: 909–996, 1988.
34. A. Aldroubi and M. Unser, Eds., *Wavelets in Medicine and Biology*, Boca Raton, FL: CRC Press, 1996.
35. A. Bruce, D. Donoho, and H.Y. Gao, Wavelet analysis, *IEEE Spectrum*, 33: 26–35, 1996.

87

Applied Intelligence Processing

Peter H. Sydenham
University of South Australia

Rodney Pratt
University of South Australia

87.1 Introduction

For a variety of reasons, the signal that comes from the output of a sensor will usually need to be processed, as has been explained elsewhere. This may be needed for a single sensor alone, as found in the conversion from an electronic resistance temperature sensor into the equivalent digital display number.

Alternatively and increasingly so, as is now discussed, the outputs from several sensors need to be combined into a single signal to form a mapping of those sources from "many to one." An example of this would be when a set of sensors measuring many different variables is used to determine if a machine tool is correctly making a component.

There are now many ways that can be used to combine signals, each having its own features that make it the best to use in a given circumstance.

Signal-processing methods range from the well-proven and mostly used method using processing based on mathematical relationships that are very precise—these are explained elsewhere. These are generally executed using digital computation and are often referred to as digital signal-processing (DSP) methods.

Alternatively, it has now been shown conclusively that less quantitative methods can also be used to great effect despite their lack of complete formal mathematical formulation. These are here called the applied intelligence (AI) methods, a convention developed to distinguish man-made systems from the very broad, and oversold, use of the term *artificial intelligence* (AI).

Many seemingly different methods exist in the latter group but, as will be shown here, they are all part of a continuum of ways that range from use of subjective to exactly objective procedures. These are not well explained as a group because they are presented in the literature as different methods used in isolation of each other. This account shows how they all tie together, thus makes it easier to decide which is appropriate in a given application. They are particularly useful for handling highly nonlinear situations where algorithms cannot be realized.

While we appear to prefer the total objectivity of a mathematically formulated method of signal processing, it is now well proven that the AI methods often are better choices to use in terms of better speed of performance and often lower cost of processing. Often they are the only solution since the algorithmic approach cannot be deployed because of the lack of a suitable mathematical formulation or powerful enough processor to run the algorithm.

Signal processing in the modern instrument, therefore, will often make use of many different methods. This account is an introduction to the characteristics of the various forms and is written to assist selection. Space limitations prevent presentation of each kind in detail.

87.2 Overview of Algorithmic Methods

Traditionally the most popular method used to develop mapping models is that of mathematical modeling. The mathematical model is usually what is sought, as it provides the highest level of understanding about the subject and the most precise representation of the behavior. The major disadvantage of mathematical models is that they can quickly become so complex that implementation of these models in measurement systems is often impractical.

In this class, the single, or set of multiple, input signal(s) to the data processor is converted to the output form using tightly formulated mathematical description. This relationship is called the algorithm. Strict relationships hold; the relationship is said to be *formal*, meaning that for any given input the output will always be the same. The algorithm supports only one interpretation.

This method of signal processing is the most highly developed method and is certainly one to aim for because it is devoid of ambiguity. All will agree on how it will respond. It carries a comforting level of understanding and, thus, acceptance.

Algorithmic methods can be very accurate, traceable, and can be calibrated with relative ease and agreement. They are the basis of many instrumentation systems. The origin of their use in instrumentation goes back to the early days of computing using, first, mechanical computational machines (late 1800s to around 1930) and then analog electric devices (early 1900s to 1960s), all of which were mostly replaced by the use of digital computers commencing around 1950. All of these algorithmic methods of processing can be simplistically regarded as embodiments of a mathematical equation inside a suitable technological machine.

As the demanded complexity and performance requirements grew over time, so did the demands on the detail of the algorithm and the means to model it inside a computational machine. Mathematical description eventually reaches limits of definition as the models push the boundaries of mathematical methods and human development. Too often, this arises before adequate detail is able to be built into the model. The algorithm is then an inadequate model of the need.

As the algorithm increases in complexity, the processing power needed must be increased to maintain both fidelity and speed of processing.

Despite great advances being made in algorithm development and in computer power, the algorithmic methodology eventually encountered mathematical and technological barriers in many fields. The method is seen to not always be the best to use because of lack of an adequate algorithm or the high cost of computing.

In instrumentation, another factor also arises. Fast, detailed processing brings with it the need for increasing electrical bandwidth requirements in signal transmission. This increases implementation costs and also eventually reaches technological constraints.

87.3 Overview of Applied Intelligence Methods

Fortunately, the solutions that may overcome these limiting constraints in many circumstances were developing in other fields under the general name of AI (now called applied intelligence in engineering), as new forms of mathematics and in other fields, such as decision theory.

Principally, a key limitation of the algorithmic method is that its unforgiving level of formalism carries with it a depth of processing exactitude that is often not warranted.

Other methods have emerged that allow vaguely subjective, as opposed to tightly objective, processing to be applied to good effect.

These AI methods have gradually gained acceptance to the degree that many are now routinely used and are supported by dedicated applications software and electronic integrated circuitry.

At first, these many alternatives were seen to be isolated methods. Gradually, the literature has shown trends to merge them in pairs. Their use in a more widely mixed form is still limited. This account seeks to give a comprehensive appreciation of the commonly met AI processing methods by placing them into relative perspective.

It is interesting to contemplate that natural world computing in animals does not appear to make much use of algorithmic methods, but does make extensive use of the methods presented here in the AI class.

The paradigm invoked here is that experience has shown that informal methods based on knowledge-based systems (KBS) can produce mappings of many inputs to one by use of less than completely formal description.

The AI methods can yield surprisingly efficient solutions to previously unsolved needs. They often can outperform algorithmic methods or carry out a similar task with far less computing power. They are all associated with multiple input processing and can be applied to forming decisions from data supplied by sensors. Each situation has to be judged on the balance between use of computing effort and effective processing.

On the downside, they lack formality and thus may be very hard to calibrate and authenticate. They, not having adequate scientific foundation and a solid formal base of operation, are not easily accepted as "sound." They are often hard to comprehend by a second party, for their description is not always adequately documented or done to any agreed convention. As their principles vary widely, they must be well understood before application is developed.

For all of these negative factors, they often are able to provide "more performance for less cost" and thus will be increasingly adopted.

Their rising level of use should not suggest the algorithmic methods will become obsolete, but more that the instrument designer now has a much larger set of processing tools available.

87.4 Mapping, in General

The high-level purpose of most signal processing is to yield knowledge of a situation so that decisions can be made.

For example, consider a health-monitoring system installed on an aircraft engine. A set of sensors of different measurand types and locations is installed at various critical points on the engine—temperatures, pressures, flow rates, metal content of the lubricating oil, and more. The data from these are collected and transmitted to a central processor using a common digital bus. The many inputs then need to be combined in some way to decide such conditions as emergency states, when to change the oil, and engine efficiency. This combination is a "mapping of many to a few."

These are not always simple mappings, for there is no adequate algorithm available to give a mathematical description for such things as degradation of oil condition. However, human intuition can be used quite effectively to obtain answers—the human mind is very capable of carrying out such mapping functions. This form of natural processing makes use of what are technically called "heuristics"—but more commonly known as "rules of thumb."

Consider the question, "How could we decide, using an automated measurement system, when loaves being made in a bakery are satisfactory to sell?" As the way to decide, this almost all people asked would suggest that the weight, size, crustiness, appearance, and softness inside would be the parameters that must all be satisfactory (that is, lie within a small range of values for each) to be declared suitable. Weight and size are easily measured; the others are not for they are really heuristics, as is the choice of the set of parameters.

The thought process implemented here is that the designer starts with a desire to know something about a situation. Consider how we could automatically monitor the "risk of working in a hazardous place" in order to give an alarm at set levels. In this kind of situation, a study of the problem will lead to identification of key parameters. These parameters can each be assigned safety functions that express how each parameter varies with system changes. With this framework, it is then possible to set up a signal-processing system that continuously calculates the risk level. This form of solution is based on ideas embodied in the wide field of decision-making theory.

The heart of application of AI methods of signal processing in instrumentation lies with appreciation of decision theory methods.

A range of multivariable mappings methods using AI ideas have emerged. Those well established in instrumentation are

- Representational measurement theory and ways sets are mapped into other sets (whose usefulness is still emerging)
- Rule and frame representation of heuristic knowledge and ways they are used to form expert systems (ESs) and other KBSs
- Binary Boolean trees as crisp logical mappings, which is the foundation of the fuzzy logic method based on fuzzy set theory

Another class of AI methodology that does not fit the same sequence, yet includes powerful methods, is those that assist optimization of the mapping setup. There are two main methods in use:

1. Genetic algorithm (GA) and its use to optimize fuzzy logic and other multisensor setups
2. Artificial neural net, a mapping method that learns by itself, from experience, how to achieve an optimal mapping in a given situation that it has been taught to work in

87.5 Basics of Decision Theory

87.5.1 Rules about a Decision-Making Process

Before entering into the detail of the AI signal-processing methods, it is necessary to develop a foundation about the ways in which decisions can be made by computers using sensed information. It is not that well appreciated, but setting up a mapping-type signal-processing situation is actually implementing a decision-making information system. General appreciation of decision making can be found in the introductory work of Kaufmann [1].

Unlike the human brain decision maker which carries out smart thinking with ease to build a machine counterpart, an engineered object needs effective externalization of the process involved. This begins by developing appreciation of the basic rules that always apply about a decision-making situation. These have been summarized by Baker et al. [2] and condense to

1. There must be a clearly expressed criterion for making a judgment of the options available, which must be such that others will understand how the judgment was made.
2. All decisions will involve choosing alternative strategies to arrive at the best one to use. This will involve assigning score numbers to the selected parameters and deciding how to process the set of numbers.
3. A decision is made by a choice of competing alternatives in the face of given restraints. Decisions can only rarely be made in the general sense but will be made for a given set of circumstances. The complexity of a problem rises rapidly as the number of parameters rises.

4. The process used attempts to achieve some payoff as a value added or lost. It aims to rank the various mapping alternatives to advise the apparently best to use. Note that once a decision-making mapping is built, the information about alternatives is no longer available as it will only usually embed one set of parameters as a single process.
5. A decision matrix carrying the competing parameters results. A method of combining the matrix constituents is needed, and, again, there is no singularly definitive, absolutely correct way to process the matrix.

In setting up a signal-processing mapping, these rules will need to be addressed. They will be embedded in the software of the hardware processor as its operational strategy. Considerable creativity is needed by the designer of the processor, for much of the setup of decision-making methods requires subjective human interpretation in several steps of the process. Decision making is really only needed when there is no exact and obvious answer. The devices built to mimic the human process will never be perfect. There will be much debates about which are the best methods and parameters to use. Engineers must live with this situation and make machines that will make good decisions, that are as close to perfect as possible.

87.5.2 Extracting Parameters

The first step in setting up a decision mapping is to understand the need. That means researching it by observation and from literature on the topic. This sets up the specific knowledge base to allow one to progress to the next step.

Then comes the need to define the key parameters of the situation. There is no organized way to develop these. They arise from inventive and innovative thought processes that seem to be based on prior learning.

To streamline this intuitive step, it is useful to apply some ordered processes that assist in externalizing appropriate parameters. Three methods are now briefly described.

87.5.2.1 Slip Writing

A group of people familiar with the problem area are read a brief statement of the problem by a person independent from the problem. An example could be "What do you think are the main parameters that a person uses to decide if a loaf of bread is fresh?"

Without much time to reflect, the group is then asked to write down the key parameters that come to mind immediately as they work without talking about the problem as a group. They write down each parameter on a separate piece of paper, doing this as fast as ideas come to them. This only happens for a few minutes. The slips of paper are then collected and classified. The whole process takes around 10 min and is known as slip writing.

It will usually be found that there is common agreement about the majority of parameters with some quite unexpected ones also arising.

Slip writing is a good way to find consensus. It probes the mind well and can bring out appreciation of factors that open discussion might easily inhibit. It is important in this method to decouple the thoughts of each person during the process; otherwise, the real parameters may not be externalized because some people may exert influence on others of the group.

87.5.3 Brainstorming and Think Tanks

If participants are shown what others are thinking and are encouraged to debate issues, it is possible to gain consensus and also allow group participants the opportunity to help each other be innovative at the same time. This method works best when the participants are prepared to go into open discussion. Several similar processes are those known as brainstorming or carrying out a think-tank session.

Here a problem in need of solution is written down as a well-prepared statement by the session organizer. A team of experts, each covering the expected aspects of the problem area, are selected and sent the statement along with any supporting exhibits. Each person considers, over a few days, how he or she might contribute a solution.

The group is then assembled. The problem is first reviewed by the session leader and each person is then asked for ideas. As ideas are externalized they are recorded in very brief form—large sheets of butcher paper are suitable. These sheets must be readable by all in the group and be prepared instantly to keep up with the thoughts of the group.

It will be found that innovative ideas will arise as candidate solutions are put up and seen by others in the group. This method encourages group-driven inventiveness.

Gradually the group will settle on a few solutions that it feels have a good chance of succeeding. This list is then ordered in priority of likelihood of success. The session leader then writes up the outcomes, ready for further investigation.

87.5.4 Knowledge Trees

The final method to be described here for developing parameters of a decision, called knowledge trees, has the merit of ordering the relative place of parameters as well as encouraging inventiveness of solutions. It also provides a mapping structure. This procedure is based on age-old realization that we think problems through by breaking them down into ever smaller subproblems until we feel able to solve them. Overall solution is then very much a matter (but not entirely so in practice) of implementing the solution of all subproblems and combining them by a process called integration.

The need is first written down. For example, "How would we measure the quality of loaves of bread?" or in a shorter form "Is the loaf OK?"

This forms the top-level parameter of the tree given as Figure 87.1. Consideration of the situation at hand then leads to realization of the collection of parameters relevant to get this answer. These might be weight, size, crustiness, appearance, and softness. They may have been externalized by a group process, such as slip writing, or created by the individual.

Each branch on the tree is then visited to see how that parameter might be measured. Only one of these can be measured as it stands—it is reasonable to assume that weight can be measured with scales.

Size is not so easy to measure as it is expressed for there is inadequate definition. More thought will yield another level to the tree—length, width, and height—for this parameter. As linear measurements, these can also be measured with ease.

When building the branching downward, a thought process has decided how the parameters map upward and downward. Size dictates the mapping of three parameters into one, so there must also be a defined mapping model for that mapping.

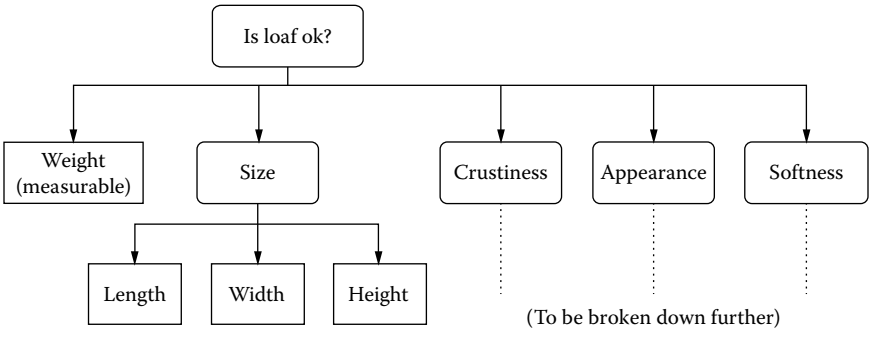


FIGURE 87.1 Knowledge trees allow facts and their relationships to be captured in pictorial form.

Note also that to branch downward, the thought process used has actually been driven by some heuristics. Each branching has been driven by rules of some kind—but more on that in the rule-based decision-making method covered later.

It is also easy to see why the other parameters could be measured with automated instrumentation. Softness could be assessed in terms of the squeeze factor, which is actually measurable as the compliance of the loaf at the center point of a side of the loaf. Appearance would map down the tree into color, texture, and graininess of the image. It is left to the reader to think up and draw a complete tree.

When all branch ends have been reticulated down to the point where they can be measured, the system mapping can be implemented with a human-made sensing system. The parameters are externalized and the mapping process is largely decided.

Use of tree-based thinking is a simple, yet powerful, way of keeping track of decision making for a complex situation. The recorded form also allows others to see the thought process used.

87.5.5 Two Examples of Decision Assistance Methods

87.5.5.1 Triangle of Pairs

Having now seen how to externalize parameters and how they might be interrelated using trees, we can move on to investigate how to set up a suitable decision-making process.

Knowing the parameters is not enough yet to design a multisensor mapping processor. The relative importance of the parameters is also a key factor. Furthermore, the tree is not the only way to combine sensor signals.

Two, of many, examples of decision assistance methods are now outlined to illustrate these points.

The first, the triangle of pairs (TOP) method, allows parameters to be ranked against others on the binary-only basis of which is preferred of each two compared. In the bread example, compare some of the parameters for their relative importance. Crustiness is preferred to weight. Softness is preferred to size and so on until all pairs have been considered. If all combinations are carried through and recorded as a matrix, a TOP results, as in Table 87.1.

Having formed a matrix of competing parameters, the next step is to decide how the matrix can be processed. This is where much debate can occur. For the TOP method, however, the binary nature of the choices means a simple count of first preferences gives the ordered preference of parameters—at least as that person assessed it!

TABLE 87.1 Triangle of Pair Assessment Is a Simple Way to Decide Which Choice to Make

Size	Crustiness	Appearance	Softness	Parameters
$\frac{\Rightarrow Si}{W}$	$\frac{C}{\Rightarrow W}$	$\frac{A}{\Rightarrow W}$	$\frac{\Rightarrow So}{W}$	Weight
	$\frac{\Rightarrow C}{Si}$	$\frac{A}{\Rightarrow Si}$	$\frac{\Rightarrow So}{Si}$	Size
		$\frac{\Rightarrow A}{C}$	$\frac{So}{\Rightarrow C}$	Crustiness
			$\frac{\Rightarrow So}{A}$	Appearance
Scores (number of times preferred)				
Si	*C*	*A*	*So*	*W*
2	2	1	3	2

Notes: This table gives the workings for the grading of bread as in the example of Figure 87.1. Si, size; Cr, crustiness; A, appearance; So, softness; W, weight.

We will see later how the idea is extended by giving parameters a varying degree of "preference" rather than the simple binary choice allowed here.

87.5.6 Utility Analysis

A more fully developed method for making decisions in complex and subjective situations is one called utility analysis (UA). This is a process that can be applied to find the usefulness of a design, piece of equipment, or any similar situation where one can externalize and prioritize a set of measurable parameters. Although mostly applied to decision making as a paper study, the process is amenable to the creation of a multisensor mapping processor.

Appreciation of this process is easily obtained by working through an example. Consider the need to set up a method for grading the quality of bread made in an automated bakery.

The first step is to decide the parameters that will affect the choice finally made. We select weight, size, and appearance as the key parameters to illustrate the method. (More parameters might be needed in a real situation.) We also decide that these are important according to the relative weighting ratios of 1.0:0.2:0.8.

Next utility curves must be set up, one for each parameter. These show how the usefulness of the parameter changes as the parameter ranges.

The simplest way to obtain these graphs is to use one's own intuition, but a better way is to make use of some form of consensus-forming procedure as discussed earlier. Figure 87.2 shows what these three functions might look like. As a guide to their construction, if the weight of the loaf is too low, it fails to comply with legal requirements and thus has zero utility as a product below the allowed uncertainty of weight. The size factor depends on the type of bread. Here it is assumed it is for sandwich making, in which case the size can be too small for sliced meats or too big for a toaster. Appearance can only be measured by mapping downward to a set of measurands; it is convenient to plot the function in more vaguely defined terms in this method.

Note that the weighting ratios have already been incorporated into the utility charts by setting their best values at the correct percentage.

It is necessary to reinforce the fact that the set of graphs is needed for the parameters, not for each case to be considered. With the charts setup, the selection process can begin.

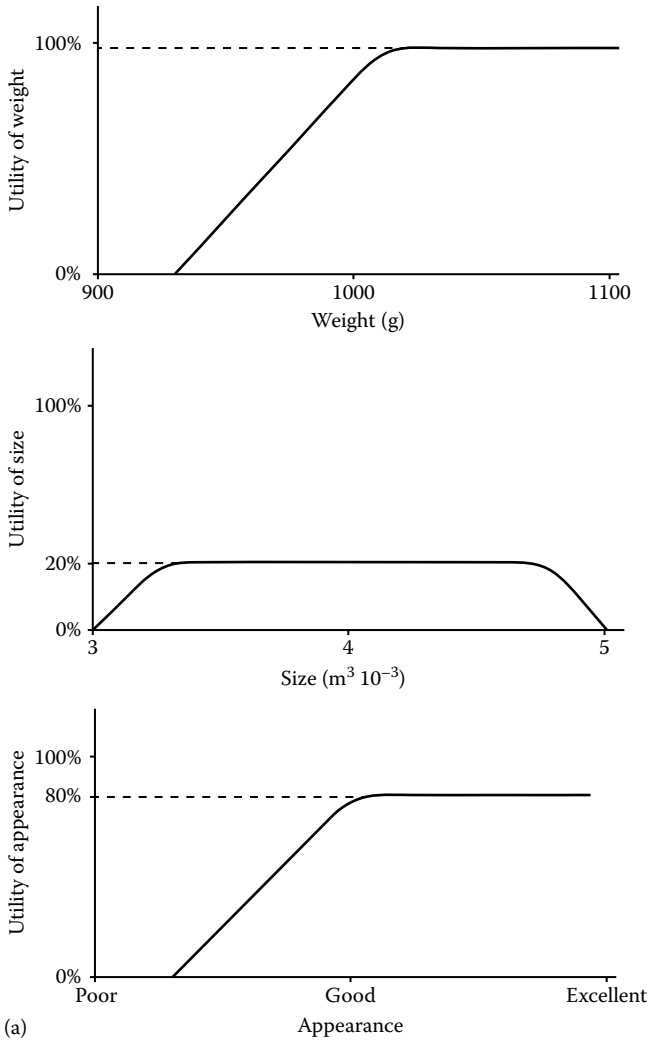
A matrix is now created, as also shown in Figure 87.2. The actual measured weight value (960 g) for loaf 1 (under automatic inspection) is compared with the graph of weight to yield a utility of 0.4. The size is done likewise, using the size graph to get 0.2, and the appearance sensor set tells us it is between poor and good to give us 0.4. Each loaf is subjected to the same process. The other two loaves have different sets of scores.

The combined usefulness of a given loaf is now to be decided by processing the set of numbers for that loaf. Here is where some difficulty arises, because there are many ways to combine the three scores for each loaf. One way often used is simply to sum the values, as is done in the example. Note that this can be satisfactory unless a zero or other unacceptable value arises, in which case more mapping processing is needed. Assume that an acceptable combined score has been determined to be 1.8 or higher (with the best, due to the weightings, 2.0).

Having carried out the processing of the matrix, we find that the first loaf is not acceptable and the other two are equally acceptable.

What we have done is to form an automatic measuring system that can make assisted and graded decisions. There is certainly a degree of human subjectivity inherent in the method, but by recording what is taking place the task can be automated and it can also be evaluated in the light of experience, correcting choices of parameters and weightings.

Other decision-making methods exist: they use elements of the earlier concepts in a variety of assemblages.



(a)

Object under Assessment \ Parameter	Weight (g)	Size (m^3 10^{-3})	Appearance (grade)	Score(Σ)
Loaf 1	(960)	(4)	(poor/good)	1.0
	0.4	0.2	0.4	
Loaf 2	(1000)	(3.5)	(good/excellent)	2.0
	1.0	0.2	0.8	
Loaf 3	(1100)	(4.9)	(excellent)	1.9
	1.0	0.1	0.8	
⋮	⋮	⋮	⋮	⋮
Loaf "*n*"	…	…	…	…

(b)

FIGURE 87.2 Utility analysis is a more-detailed way to automate loaf inspection with a set of sensors: (a) three possibilities of the utility functions for the example in Figure 87.1, and (b) scoring matrix for loaves passing the inspection point.

The earlier has been a very rapid introduction to decision-making methods. These are rarely taught as the foundation of the now popularized AI signal-processing methods given next. They are, however, the conceptual basis of the AI processes.

87.6 Principal AI Methods

87.6.1 Measurement Theory

Sensors observe a finite number of variables associated with real-world events. Measurement systems are required to convert the information from these sensors into knowledge. This process often involves what is known as mapping.

Depending on the loss or preservation of information, the mapping process can be classified as either transformation or abstraction.

If a rule exists that assigns every element of x in the set X (written $x \in X$) to precisely one element $y \in Y$, we say that a function exists that maps the set X to the set Y, and this is represented as Figure 87.3 and symbolically by

$$y = f(x)(x \in X)$$

A transformation defines a one-to-one mapping on the set of all elements of x into y, in which for all $x \in X$ there exists a unique value in Y; therefore, the inverse transformation is guaranteed, i.e., $x = f^{-1}(y)$.

An abstraction differs from a transformation in that it can map a number of elements $x_i \in X$ into the same y, or a many-to-one mapping; hence, no inverse operator can exist such that the inverse image of X can be uniquely retrieved from Y. Therefore, abstraction is usually characterized by a loss of information.

It is possible to define the set theoretical relationship as a formal mathematical model and then embed that in the signal processor. Although this is a clearly useful method of formalizing mappings, this approach is not commonly used at the time of writing but can be expected to find more favor in the future.

87.6.2 Rule- and Frame-Based Systems

In early stages of problem solving, we seem naturally to look to rules of thumb to get the solution started. Even the algorithmic methods start with this form of thinking, for one has to decide what the elements of the decision are and how they might be assembled into a strategy for implementation in the signal processor. As the rules are externalized, one's thought patterns also usually build knowledge trees that show the relationship between the rules.

Consideration of the knowledge supporting the structure of a tree will reveal that the decision needed about which way to branch as one moves through a tree is actually the implementation of a rule that has relevance at that junction. Rules link parameters together.

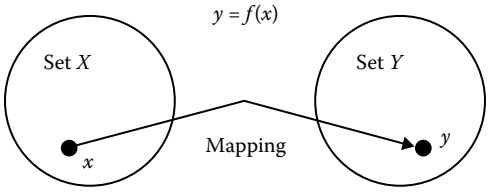


FIGURE 87.3 Pictorial representation of the set theoretical mapping process.

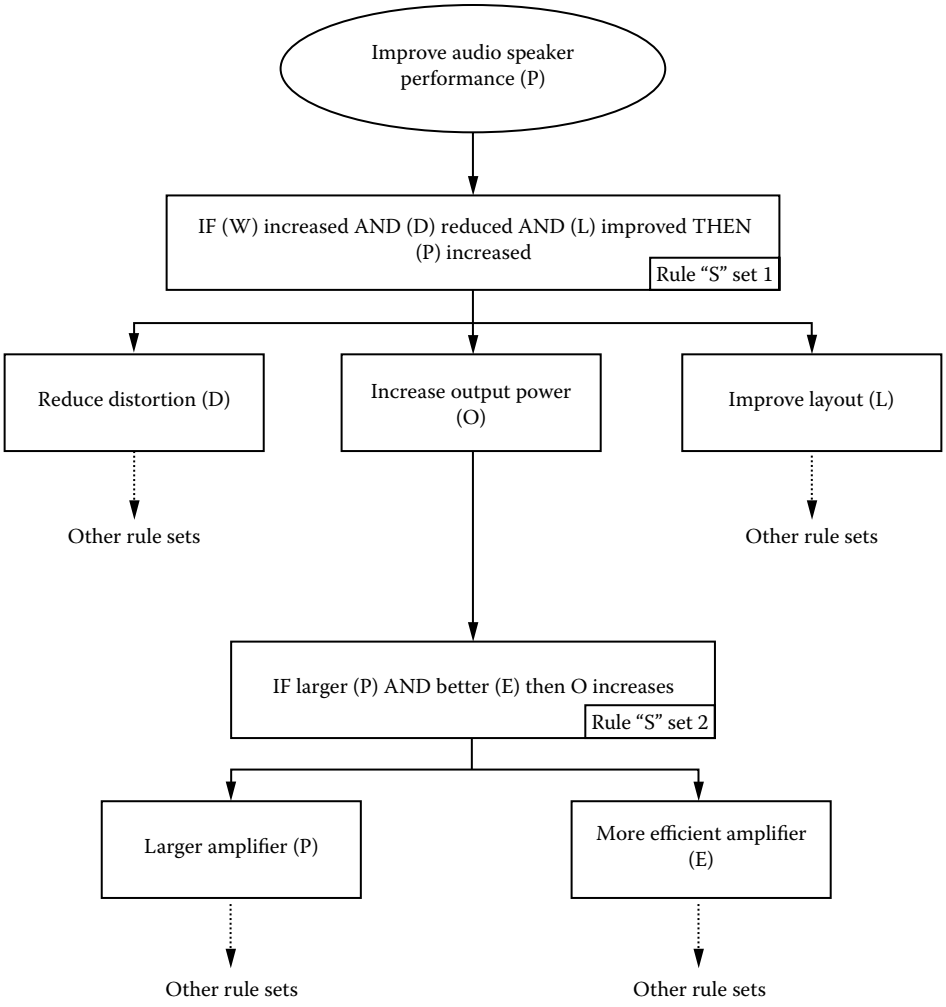


FIGURE 87.4 Part of a design knowledge tree for improving the performance of an audio speaker system.

Figure 87.4 is a tree giving some of the design options for improving the performance of an audio speaker system. This tree has been built by applying a designer's knowledge of speaker system design. The heuristic rule set for the top branch is stated as

IF	Speaker output (W) increases
AND	Distortion (D) is reduced
AND	Positioning (L) improved
THEN	Audio output (O) is improved

No part of the rule set could be first realized using formal mathematical or algorithmic thinking. Intuition, leap, common sense, and other terms that describe the human intelligence process must be applied to commence a solution.

At some stage a rule may become describable by an algorithm, when that occurs, a formal mathematical expression can be used to embed relationships. However, this is often not the case and so methods have been developed in computers to process heuristics.

The full set of AI techniques were originally all rolled into what became known as KBSs but this term is so overused that it has lost specific meaning.

Among the first AI processing methods were special computing ways to process the logic of a set of rules. These became known as ESs. In the example earlier, the rule tree is very sparse; a practical system for decision making is likely to have from 100 to several thousand rules.

The rules are considered by an inference engine (a software program) that is able to carry out Boolean logical operations of AND, OR, etc., to yield the outcome appropriate to the set of rules relevant to the problem at hand.

Trees can be traversed from the top down (downward chaining) or from the bottom up (upward chaining) and modern ES software applications carry out these operations with great sophistication.

KBS methods generally suffer from the feature that they seem to give answers all too easily, for they use only a few of the many rules available to come to a solution in a given situation. To help users feel more confident in their application, features are often offered that include plotting of the chaining used to get the solution or stating the rule set used.

Rule-based software applications are sold as empty shells. The user fills the shells with the rule set to make the application specific. These applications are now commonly used. They are relatively easy to use without the need for a competent computer programmer.

Rules are a primitive way to express knowledge. A better form of representation is the frame. This has the ability to hold more knowledge than a single rule and is a small database about a limited area of the system of interest. Frames are like objects in object-oriented programming. Advanced ES shells operate with frames.

ES shells are now very capable entities for decision making. They are a significant tool in the instrument signal processor's toolbox. Space restricts more explanation but enough has been stated here to allow further development of AI methods.

Some basic characteristics about rule- and frame-based processing are as follows (these apply variously):

- The software program is often self-explanatory as the rules can be read as (almost) normal language.
- They need considerable computer power as the rule number increases.
- They are relatively slow to yield a solution but are best used for cases where slow outcomes are applicable.
- They need special software.
- They are not that well known, so their application may be slow to find favor.

If the problem becomes complex, EAs soon run out of usefulness. The computer search becomes too slow because the number of rules needed rapidly rises with problem complexity. The approach used today for large problem-solving systems is to build an ES for each facet of the problem. These small-problem AI units are called agents. A set of agents is then combined using a conceptual software-based blackboard that calls on the agents to investigate a problem put to the system.

The chaining operation basically only carries out simple Boolean algebra operations using system parameters represented by a description called a rule. The system has no understanding of the wording of the rule. Thus, it is only processing as through the tree branching is either one way or the other. In its simplest form, it contains no concept of making that branching decision with a graded concept of which way to go.

Real life is full of unclear logical operations. The outcome of a decision will be clearly this way or that, but just where the change arises is problematic. The changeover point is unclear because the variables of the rule are fuzzy. It is desirable to process rules with regard of their likelihood of relevance depending on the state and selection of other rules. As is explained in the following section, the fuzzy logic method of carrying out a mapping is another way that allows the rule to have more variability than the two-state exactness of binary logic.

87.6.3 Fuzzy Logic

This explanation is not intended to be a rigorous mathematical examination of fuzzy sets and fuzzy logic but rather explain, through example, the application of fuzzy techniques in measurement systems, specifically in respect to mapping models. For more detail, see Mauris et al. [3].

The simple example used here is a measurement system (Figure 87.5) that maps two input sensors (temperature and humidity) into one output value (comfort index). Clearly, there is no formally accepted definition of comfort index as it is a subjective assessment. One of the advantages of fuzzy sets is that they are usually intended to model people's cognitive states.

In the mid-1960s, Professor Lofti Zadeh recognized the deficiencies of Boolean logic, in that its TRUE/FALSE nature did not deal well with the shades of gray that exist in real-life situations.

Boolean logic uses classical set theory where an element is either viewed as entirely true or completely false $A = \{0, 1\}$. These are often referred to as a crisp sets (Figure 87.6). In a crisp set the transition between sets is instantaneous, i.e., 36.9 °C is considered warm whereas 37.1 °C is considered hot. Hence, small changes in the input values can result in significant changes in the model output. Clearly, the real world is not like this.

Fuzzy logic uses a multivalued set where degrees of membership are represented by a number between 0 and 1, $A = [0, 1]$ μ_A:$U \rightarrow [0, 1]$, where μ_A is the membership function. With fuzzy logic, the transition between sets is gradual and small changes in input values result in a more graceful change in the model output (Figure 87.7).

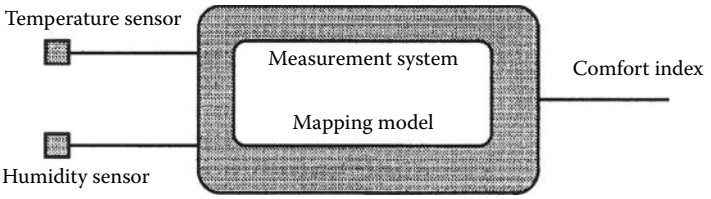


FIGURE 87.5 A simple comfort index measurement system uses temperature and humidity variables.

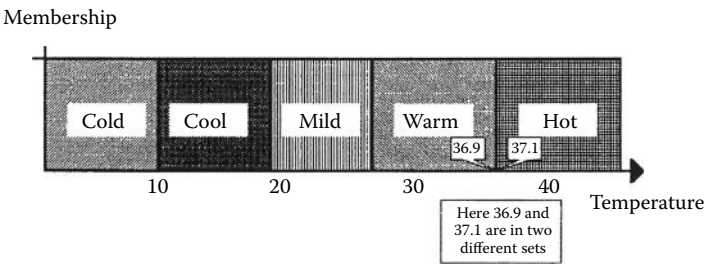


FIGURE 87.6 Conventional crisp set for the measurement system of Figure 87.5.

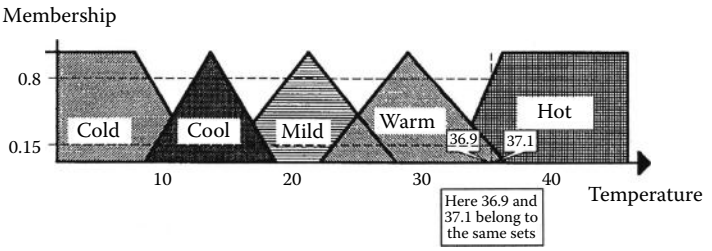


FIGURE 87.7 Fuzzy set representation temperature regimes in the comfort controller example.

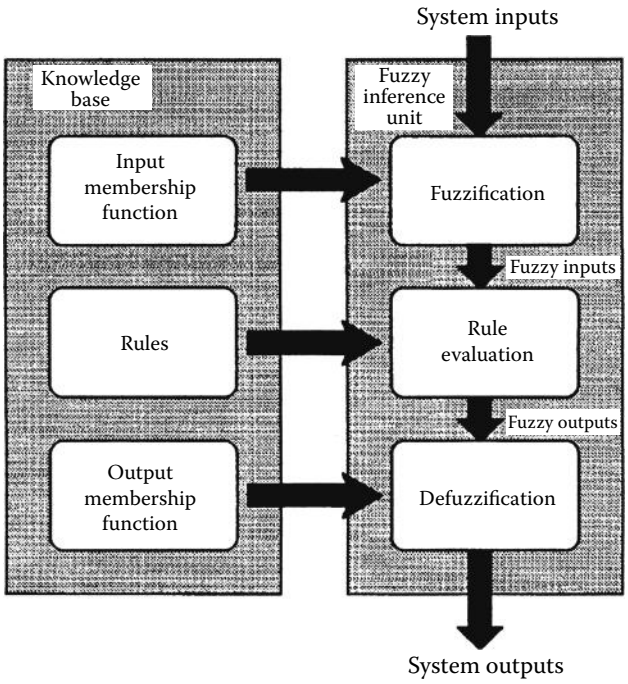


FIGURE 87.8 A fuzzy inference system combines rules in a way that allows them to be fuzzy in nature, that is, not crisp.

87.6.4 Fuzzy Expert Systems

A fuzzy ES (Figure 87.8) combines fuzzy membership functions and rules, in place of the often all-too-crisp Boolean logic, to reason about data. The fuzzy ES is usually composed of three processing sections:

Step 1: Fuzzification
Step 2: Rule evaluation
Step 3: Defuzzification

87.6.4.1 Step 1: Fuzzification

In fuzzification, crisp inputs from input sensors are converted into fuzzy inputs using the membership functions in the knowledge base. A fuzzy input value is generated for each linguistic label of each input. For example, in Figure 87.7, for an input temperature of 37°C the fuzzy input is COLD (0.0), COOL (0.0), MILD (0.0), WARM (0.15), HOT (0.80). A similar set of fuzzy inputs are generated for the humidity sensor (Figure 87.9).

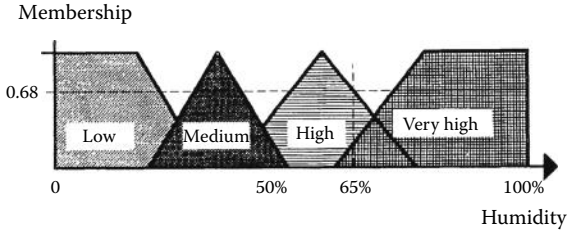


FIGURE 87.9 Humidity membership function for the controller example.

TABLE 87.2 Fuzzy Associative Matrix Links Inputs and Outputs

Humidity/ Temperature	Cold	Cool	Mild	Warm	Hot
Low	Uncomfortable	Uncomfortable	Uncomfortable	Uncomfortable	Uncomfortable
Medium	Uncomfortable	Acceptable	Comfortable	Acceptable	Uncomfortable
High	Uncomfortable	Acceptable	Acceptable	Acceptable	Uncomfortable
Very high	Uncomfortable	Uncomfortable	Uncomfortable	Uncomfortable	Uncomfortable

87.6.4.2 Step 2: Rule Evaluation

Rules provide a link between fuzzy inputs and fuzzy outputs. Rules are usually expressed in the form of IF … AND/OR … THEN … statements.

For example,

"IF" the TEMPERATURE is HOT "AND" the HUMIDITY is HIGH "THEN" it is UNCOMFORTABLE.

"IF" the TEMPERATURE is MILD "AND" the HUMIDITY is MEDIUM "THEN" it is COMFORTABLE.

"IF" the TEMPERATURE is WARM "AND" the HUMIDITY is LOW "THEN" it is UNCOMFORTABLE.

Rules can also be expressed in the form of a table or matrix, called the fuzzy associative matrix. This matrix (Table 87.2) provides a complete description of the system performance for all combinations of inputs.

The function of the rule evaluation step is to evaluate the relative strengths or truth of each of the rules in order to determine which rules dominate. In this example, the rules contain AND relationships and therefore the overall rule strength must be the minimum (MIN) value of the two strengths of the input values.

For example, at a temperature of 37 °C and a humidity of 65% the rule strengths of the three example rules are

"IF" the TEMPERATURE is HOT (0.8) "AND" HUMIDITY is HIGH (0.68) "THEN" it is UNCOMFORTABLE (Rule Strength = MIN (0.8,0.68) = 0.68).

"IF" the TEMPERATURE is MILD (0.0) "AND" the HUMIDITY is MEDIUM (0.0) "THEN" it is COMFORTABLE (Rule Strength = MIN (0.0,0.0) = 0.0).

"IF" the TEMPERATURE is WARM "AND" the HUMIDITY is LOW "THEN" it is UNCOMFORTABLE (Rule Strength = MIN (0.15,0.0) = 0.0).

All rules must be evaluated (in this case, all 20 in the matrix) to determine each rule strength. If two rule strengths exist for one fuzzy output label, then the maximum (MAX) rule strength is used because this represents the rule which is most true; that is, in the previous example the fuzzy output for UNCOMFORTABLE is the MAX(0.68,0.0) = 0.68.

87.6.4.3 Step 3: Defuzzification

Now that rule strengths exist for all the output fuzzy labels, a crisp output value can be determined from the fuzzy output values. The most common method used to defuzzify the fuzzy output value is the center of gravity (COG) method. The fuzzy rule strengths determined from rule evaluation are used to truncate the top of the output membership functions. Given this area curve, the COG or balance point can then be calculated. For example, in Figure 87.10, for fuzzy output values of

UNCOMFORTABLE = 0.68
ACCEPTABLE = 0.2
COMFORTABLE = 0

the COG or crisp "Comfort Index" evaluates to 25.

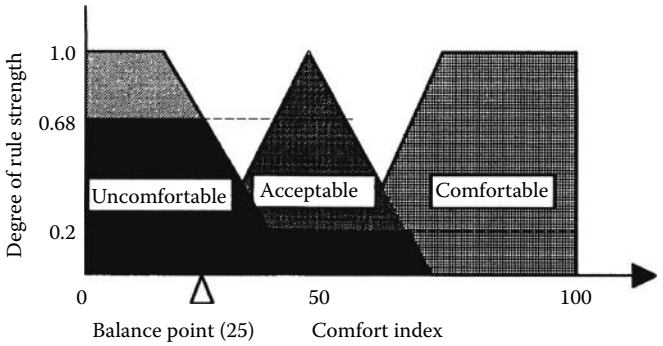


FIGURE 87.10 Output membership functions for controller. Each type of shaded area represents the three states of comfort.

Fuzzy logic signal processing is now a well-developed method. It is supported by copious teaching material including those on Internet and on CD ROMs provided by manufacturers of this form of special integrated circuit chip sets and setup software. Possibly, its best-known application has been in clothes washing machines where the wash parameters are set to suit the load. Although there are still limitations, this method can be considered to be a mature procedure.

87.6.5 Genetic Algorithms

The methods discussed thus far all have required the user to advise the system about the parameters to use. That is, they need to be taught efficient mappings.

In sharp contrast to these types of systems, there also exist other AI methods that possess the ability to learn, by themselves, what are the more optimal mapping configurations. Two main techniques are used—GAs and artificial neural networks (ANNs). These self-learning processes both work well in certain situations and are now commonly used in signal processing. On the face of it, both seem to possess magical properties because they defy logical thought processes and a clear understanding of how they actually operate.

We begin with an overview of GAs. These make use of the basic principles by which natural selection found in living things, that is, from genetics, is able to improve gradually the fitness of species. That the genetic principles found in nature can be used in human-made applications is accredited to pioneering work of John Holland in the mid-1970s. Today, it is large field in algorithm optimization research (see Tang et al. [4]). A very simplistic description now follows to give some insight.

The concept starts with the selection of a set of features (Figure 87.11) (these can take a wide range of forms and are not just measurement parameters) that represent the essential features of a system of interest. As examples, the DNA molecule carries the code of the characteristics of living beings and a computer string can carry a coded message that represents the features of the behavior of some human-devised system.

Various types of events (crossover, mutation, inversion are commonly encountered methods) can slightly alter the code of any particular string. When this happens, the new string then represents another closely similar, but different system having new properties. Consider, next, that a number of slightly different code strings have been formed.

When a change takes place in the code of a string, it is assessed against the other strings using rules for a predecided fitness test. If improvement has occurred in the overall properties, then it is adopted as one of the full set. If not better, then it is discarded. In this way the set of strings, and thus the total system capability, gradually improves toward an optimal state.

The operational aspects of such systems are beyond description here. Suffice to say that the technique is well established—but highly specialized—and is undergoing massive international research effort in hope of alleviating certain limitations.

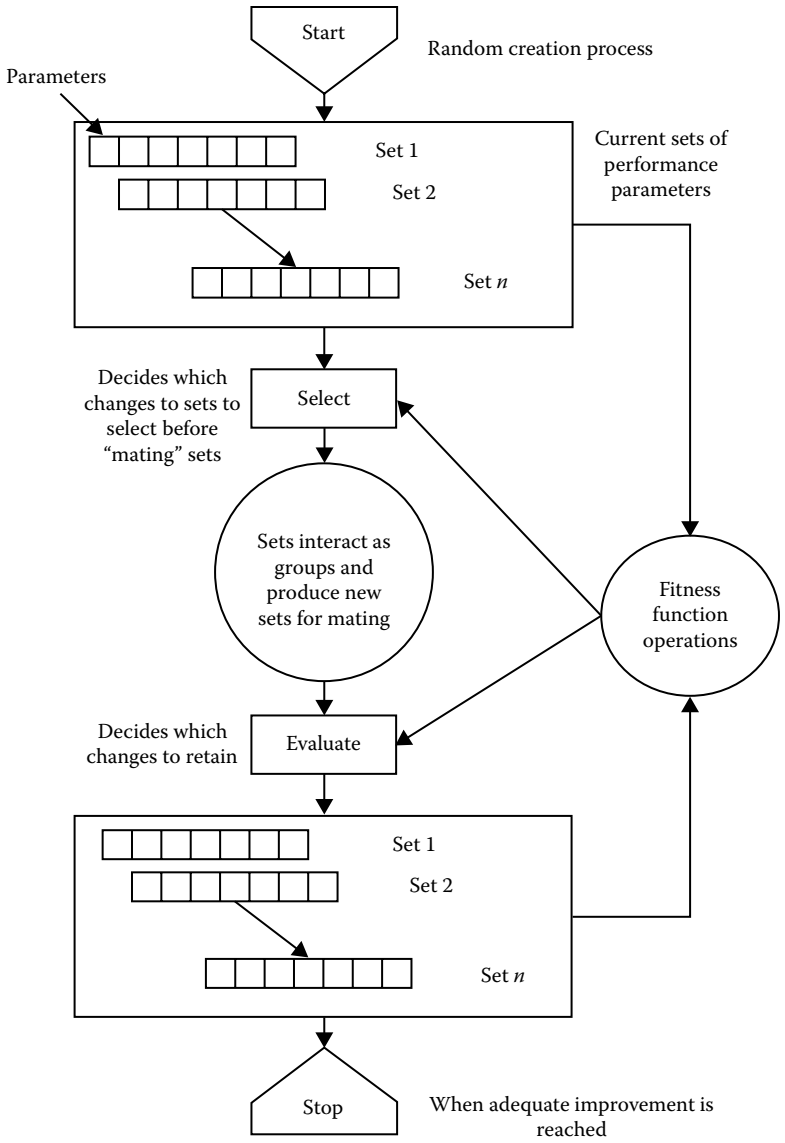


FIGURE 87.11 In the GA method, sets of code strings are first modified by some form of genetic operations. They are then intercompared using fitness functions to select a better code to use in the subsequent set of strings. This iterative process is continued until some event disrupts it.

The first limitation is that although it is generally agreed each adopted change for the better takes the overall system capability closer to the goal being sought, there is, as yet, no theory that can be applied to show the state of maximum optimization. GAs, therefore, always have doubt associated with their solutions as to how much more improvement might be possible.

The second limitation becomes obvious when the computational demands are considered in terms of the number of comparison operations needed to be run in the improvement process. This number can be truly huge, especially as the range of options rises with increase in string length. This kind of operation usually needs very large and fast computing power. In cases where there is plenty of time to determine an improvement, this is not a major limitation. An example of effective use would be finding how best to operate a wide range of functions in a complex system when the task is then fixed; this method

was used to set up more optimal use of power supplies in a space vehicle. In cases where the variables in a code string are fast changing, the method may not be applicable.

When running this form of computer program, it is essential to have a measure of the speed—the dynamic behavior—at which this highly iterative process is moving toward the optimization goal: the user can then decide if it is a viable method for obtaining improvement.

With this background, it is probably obvious why GAs are sometimes used to set up the membership functions of fuzzy logic systems. As explained earlier, the membership functions each are part of a set of individual functions that form the mapping for a multisensor system. The choice of the membership function is largely heuristic and thus may not be the best function to use. By setting up several sets of functions, it is then possible to apply GA computing methods iteratively to select a better set to use. Such systems have been used where the time of currency of a function set is longer than the time needed to improve the functions. Obviously, use of GAs increases system setup and operational complexity, but once implemented may yield worthwhile gains.

87.6.6 Artificial Neural Networks

The AI basket of potentially useful methods in signal processing is full of surprises. Attention is now directed to another method, which can learn, after an initial training period, how to better operate a given mapping process.

Neurocomputing or use of ANNs is another AI technique well suited to make a mapping processor in some circumstances. (These are sometimes called NNs but should not be used for human-devised systems as they can lead to confusion with life sciences research on living neural networks.) ANNs are particularly useful when mapping complex multidimensional information to a simpler representation. Because of their ability to deal with nonlinear relationships, they find application in areas where traditional statistical techniques are of limited use. Some application areas include pattern recognition and classification, categorization, function approximation, and control.

87.6.7 Biological Neuron

The ANN has been inspired by the biological structure of the brain, and it is an attempt to mimic processes within the biological nervous system. The neuron is an information-processing cell in the brain (Figure 87.12). It consists of

1. A body or *soma*
2. Input branches or *dendrites*
3. Output branch or *axon*

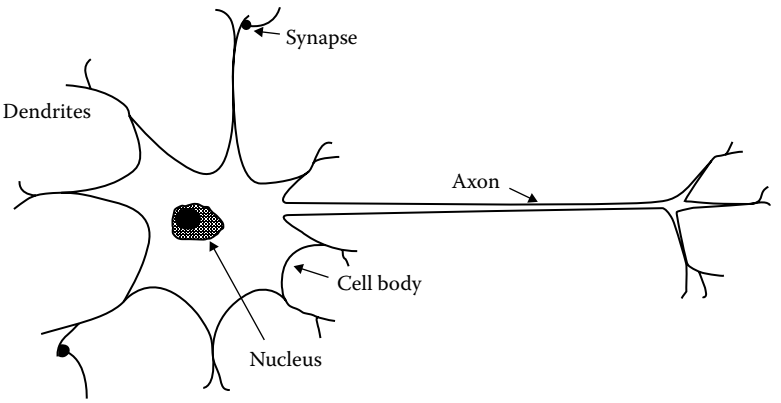


FIGURE 87.12 A biological neuron.

The neuron receives signals through its dendrites, and then transmits signals from its cell body along the axon. The neuron will generate an output (or fire) when the aggregation of the inputs reaches a threshold level. At the terminals of the axon are *synapses*. The synapse couples the signals from the axon of one neuron to the dendrites of another neuron. The strength with which these signals are transferred from the axon to the dendrite is controlled by the synapse and can be altered; hence, the synapses can learn from experience.

Desirable characteristics of neural systems include

- Massive parallelism
- Learning ability
- Ability to generalize
- Adaptability
- Fault tolerance

It is the attempt to construct machines that exhibit these characteristics that has led to the ANN methods of signal processing.

87.6.8 Artificial Neural Networks

In 1943, McCulloch and Pitts proposed the first model of an artificial neuron. This has formed the basis for the generally accepted form of synthetic neuron model or processing element (PE) (see Figure 87.13).

The output of the PE y is given by

$$y = g\left[\left(\sum\nolimits_i w_i \cdot x_i\right) - b\right]$$

where

x_i is the PE inputs with weights (synaptic strengths) w_i
b is the PE bias
g is the activation or transfer function

Many types of function have been proposed but the most popular is the sigmoid function, defined by

$$g(h) = \frac{1}{(1 + e^{(-\beta h)})}$$

where β is the slope parameter.

By themselves the PEs are very simple; however, when the individual elements are joined into large interconnected networks, complex relationships can be represented. Although a number of network architectures [5] exist, one of the most commonly discussed architectures found in the literature is the

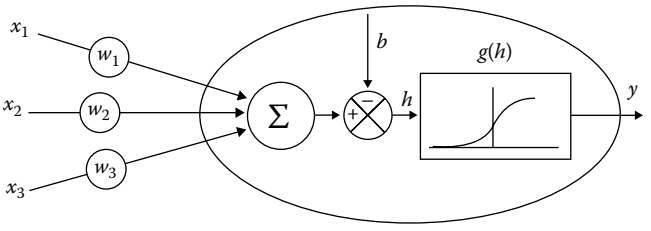


FIGURE 87.13 This neuron model is commonly used.

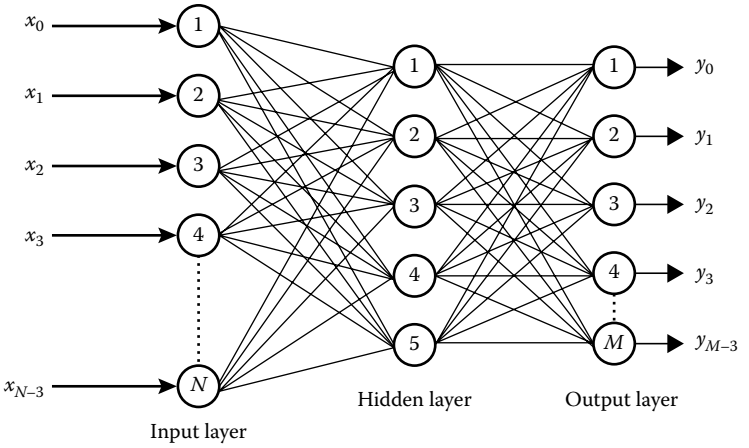


FIGURE 87.14 A typical three-layer neural network as is commonly used to create an optimal mapping from sensors to outputs.

feed-forward multilayered perceptron. Figure 87.14 provides a simple illustration of how a multilayered ANN maps a multidimensional input vector $x_0, \ldots, x_{N-1}$ in an input space to a vector $y_0, \ldots, y_{M-1}$ in an output space.

87.6.9 Learning

A fundamental characteristic of the ANN, once it has been set up as an operational tool in software form, is that it does not need to be programmed for the application. ANNs appear to learn rules from a representative set of examples, rather than having rules programmed in by an expert. The knowledge acquired by the system that controls how the system maps input to output is held within the connection weights of the network.

The focus of extensive ongoing research is the search for optimal training techniques. These techniques tend to fall into two broad categories—supervised and unsupervised learning.

In supervised learning, a representative set of inputs is presented to the network which then modifies its internal weights in order to achieve a desired output. With unsupervised learning, the input data only are presented to the network, following which the network organizes itself through self-modification of its internal weights so that it responds differently to each input stimulus.

It is beyond the scope of this chapter to review all the current learning techniques for ANNs. This is well documented in the literature [6]. One popular training algorithm—backpropagation—will be discussed as an example.

87.6.10 Backpropagation

Backpropagation is an example of a supervised learning paradigm commonly used with multilayer perceptron network architectures. Backpropagation follows the error-correction principle and uses the error signal {d (desired output) − y (actual output)} to modify the connection weights to reduce this error. The backpropagation algorithm is implemented as follows:

1. Initialize network weights to small random values.
2. Present an input vector $x_0, \ldots, x_{N-1}$ and the corresponding desired output vector $d_0, \ldots, d_{M-1}$
3. Calculate the actual output vector $y_0, \ldots, y_{M-1}$ by propagating the input through the network.

4. Use a recursive algorithm starting at the output layer and adjust the weights backward by

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j x_i'$$

where
$w_{ij}(t)$ is the weight from an input to node j at time t
x_i' is either the output of node i or an input
η is a gain term ($0.0 < \eta < 1.0$)
δ_j is an error term for node j

For the output layer $l = L$, the error is calculated:

$$\delta_j^L = g'\left(h_j^L\right)[d_j - y_j]$$

where
h_j^L is the net input to the jth unit in the L layer
g' is the derivative of the activation function g

For hidden layers $l = (L - 1), \ldots, 1$, the error is calculated:

$$\delta_j^L = g'\left(h_j^l\right)\sum_l w_{ij}^{l+1}\delta_j^{l+1}$$

where
h_j^l is the net input to the jth unit in the lth layer
g' is the derivative of the activation function g

5. Return to Step 2 and repeat for the next pattern until the error reaches a predefined minimum level.

Unfortunately, no method exists that allows the ANN to create or learn information that is not contained in the training data; that is, the ANN can only reproduce based on experience. Under certain conditions, however, the network can generalize, that is, approximate, output values for data not contained in the training set.

The neural network can be considered as a universal approximator. During supervised learning, the output eventually approximates a target value based on training data. While this is a useful function, the ability to provide output data for test cases not in the training data is more desirable. Loosely, generalization can be viewed in terms of interpolation and extrapolation based on training data. If a test case is closely surrounded by training data, then (as with interpolation) the output accuracy is generally reliable. If, however, the test case is outside of, and not sufficiently close to, training data, then (as with extrapolation) the accuracy is notoriously unreliable. Therefore, if the training cases are a sufficiently large sample of the total population of possible input data so that each test case is close to a training case, then the network will adequately generalize.

While multilayer feed-forward networks are finding increasing application in a wide range of products, many design issues such as determining an optimal number of layers, units, and training set for good generalization are research topics. Current theory provides loose guidelines and many of these

design issues are resolved by trial and error. Another disadvantage of ANNs is the high demand that many training algorithms can put on computing resources because of their recursive nature.

The ANN, then, provides another alternative for development of suitable mapping models for measurement systems. They can be used to describe complex nonlinear relationships using a network of very simple processing elements. The attraction of the ANN lies in its ability to learn. As long as there exists a sufficiently representative sample of input-to-output data available, the mathematical relationship of the mapping function need not be known. It is effectively taught to the network during a learning process.

Again, there exist important limitations. The worst is the time it might take to adjust the system nodes and weights to a nearly final state. This can often require considerably more time than the time-varying properties of the inputs allow. In many potential applications, they take too long to learn and are not effective. Again computational speed and power are governing factors.

Despite their shortcomings in some applications, ANNs are now a commonly used procedure to set up sensor mapping systems. Examples are banknote image detection and the increased sensitivity of the detection of aluminum in water.

87.7 Problems in Calibration of AI Processing Methods

Calibration of a measurement system is the result of using an agreed upon, often legally binding, process by which it is proven to possess a declared level of accuracy in its measurement outcome. In conventional instrument terms, this implies the system can be set up and compared with a measurement method of superior performance to give its error of accuracy plus its variance from the agreed value determined with a level of uncertainty. This kind of instrument system is then accepted to have a known behavior that could be explained by the laws of physics as causal and unique in performance. The prime example, the physical standard apparatus for a parameter, can and is defined such that it will always give very closely the same outcome, even if built in different laboratories. It will have predictable behavior. At the heart of this acceptability is that it can be modeled in terms of an algorithm. All parts in it follow formal laws and have the same outcomes from implementation to implementation.

Most of this notion has to be put aside because, as explained earlier, AI-based instrument signal processors are built on a raft of transformations that convert subjective situations into objective ones or they carry out unexplained processes. It is, therefore, not hard to see that calibration is a major issue with this type of processor.

Processes make use of heuristics to at least start them going. The designer, when invoking any of the decision-making methods, will almost certainly not select the same rules, processes, and parameters that another designer will choose. There is a lack of consistency in AI processors. The outcomes are fuzzy, not crisp as in instrumentation that complies with a physical law. They act like humans do in that they provide a range of solutions to the same problem.

At first sight, this seems to imply that we should ignore the AI possibilities for they cannot be calibrated according to long-standing metrological practices. However, their performance is often very worthy and will be workable where algorithmic methods are not. This calibration constraint must be considered in terms of human thinking, not so much in terms of physics and mathematical models.

At present, AI processing methods have been well proved in many fields, these tending to be fields in which performance does not need calibration with regard to the standards regime. Examples are the use of fuzzy logic in clothes washing machines to improve the wash by varying the wash cycle parameters, in neural network methods to aid the recognition of banknotes, and in rule-based controllers in industrial process plant controls. GAs have been used to schedule power supply and usage in space shuttles—said to be an impossible task by any other known means. These all are typified as being needs where much experience can be devoted to ensuring they work satisfactorily.

They should normally not be the critical processing element in safety-critical situations. They can certainly be used in routine management situations where they can outperform algorithmic methods, but there they should be backed up with conventional alarms. They are often used in off-line plant control

where the human and alarms are still the final arbiter. This, however, seems to be only a cautious step in our slow acceptance of new ideas.

The process of calibration here is more akin to that of conducting evaluation and validation. Does the system give the range of outcomes expected in given circumstances? Are the outcomes better than those without the processor? Could it be done as well or better by algorithmic-based processing? Is the speed it gives worth the problems it may bring? Problems in their testing, and thus calibration, are discussed by Sizemore [7]. The issues that need to be considered in the calibration of conventional instrumentation [8] are relevant to the calibration of AI-based processing but need much more care in their execution.

Such questions require consideration of the very same elements of decision theory upon which they are based to test them. They have been set up to think like humans so it is expected they will have to be calibrated and evaluated like humans—that is not at all easy.

At present, the calibration and validation of AI systems are not standardized well enough. This impedes acceptance, but standardization will improve as world knowledge of this relatively new method of processing develops to greater maturity inside instrumentation.

There will be opposition to the use of AI methods, but the performance gain they bring will ensure they are used. The forward-thinking academic measurement community is showing signs of addressing AI signal processing—but it will take time.

References

1. Kaufmann, A., *The Science of Decision Making*, Weidenfeld & Nicholson, London, U.K., 1968.
2. Baker, D., Kohler, D. C., Fleckenstein, W. O., Roden, C. E., and Sabia, R., Eds., *Physical Design of Electronic Systems*, Vol. 4, Prentice Hall, Englewood Cliffs, NJ, 1972.
3. Mauris, G., Benoit, E., and Foulloy, L., Fuzzy sensors for the fusion of information, in *Proceedings of 13th IMEKO Conference*, Turin, Italy, pp. 1009–1014, 1994.
4. Tang, K. S., Man, K. F., Kwong, S., and He, Q., Genetic algorithms and their applications, *IEEE Signal Process. Mag.*, 13(6), 22–37, 1996.
5. Jain, A. K. and Mao, J., Artificial neural networks: A tutorial, *IEEE Comput.*, 31–44, 1996.
6. Venmuri, *Artificial Neural Networks: Theoretical Concepts*, IEEE Computer Society Press, Los Angeles, CA, 1988.
7. Sizemore, N. L., Test techniques for knowledge-based systems, *ITEA J.*, XI(2), 34–43, 1990.
8. Nicholas, J., Relationship of legal issues to measurement, in *Handbook of Measurement and Control*, Vol. 3, Sydenham, P. H. and Thorn, R., Eds., John Wiley & Sons, Chichester, U.K., 1992.

88

Analog-to-Digital Converters

E.B. Loewenstein
National Instruments

88.1 Introduction

Almost every modern instrumentation system includes some form of digitizer, or *analog-to-digital converter* (ADC). An ADC converts real-world signals (usually voltages) into digital numbers so that a computer or digital processor can (1) acquire signals automatically, (2) store and retrieve information about the signals, (3) process and analyze the information, and (4) display measurement results. A digitizing system can do these jobs with greater speed, reliability, repeatability, accuracy, and resolution than a purely analog system normally can.

The two main functions of an ADC are *sampling and quantization*. These two processes convert analog signals from the time and voltage continuums (respectively) into digital numbers having discrete amplitudes, at discrete times. To represent changing signals at every instant in time or at every possible voltage would take an infinite amount of storage. So for every system there is an appropriate *sampling rate* and degree of quantization (*resolution*) so that the system retains as much information as it needs about the input signals while keeping track of manageable amounts of data. Ultimately, the purpose of sampling and quantization is to reduce as much as possible the amount of information about a signal that a system must store in order to reconstruct or analyze it meaningfully.

88.2 Sampling

To prevent having to digitize an infinite amount of information, an analog signal must first be sampled. Sampling is the process of picking one value of a signal to represent the signal for some interval of time. Normally, digitizers take samples uniformly in time, e.g., every microsecond. It is not necessary to sample uniformly, but doing so has some interesting and convenient mathematical properties, which we will see later.

Sampling is done by a circuit called a *sample-and-hold* (S/H), which, at a sampling instant, transfers the input signal to the output and holds it steady, even though the input signal may still be changing. An S/H usually consists of a signal buffer followed by an electronic switch connected to a capacitor. At a sampling instant, the switch briefly connects the buffer to the capacitor, allowing the capacitor to charge to the input voltage. When the switch is disconnected, the capacitor retains its charge and thus keeps the sampled input voltage steady while the ADC that follows does its job. Quite often, sampling is actually done by a circuit called a *track-and-hold* (T/H), which differs from an S/H only slightly. Whereas the S/H holds the analog signal until the next sampling instant, the T/H holds the analog signal still only until the ADC has finished its conversion cycle. After the ADC is through, the T/H reconnects the buffer to the capacitor and follows the input signal until the next sampling instant. The result is more accurate sampling, because the buffer has more time to charge the capacitor and "catch up" with (track) the input signal, which has changed since the last sampling instant. Nearly every modern ADC chip has a built-in S/H or T/H, and virtually all data acquisition systems include them.

Of course, sampling necessarily throws away some information, so the art of sampling is in choosing the right sample rate so that enough of the input signal is preserved. The major pitfall of *undersampling* (sampling too slowly) is *aliasing*, which happens whenever the input signal has energy at frequencies greater than one-half the sample rate. In Figure 88.1a, a signal (the fast sine wave) is sampled at a rate *Fs*, shown by the hash marks at the bottom of the graph. The sine wave has a frequency of 0.8*Fs*, which is higher than one-half the sample rates (0.5*Fs*). Notice that sampling the lighter sine wave of 0.2*Fs* produces the same set of samples. The resulting sampled data are ambiguous in that we cannot tell from the data what the frequency of the incoming sine wave actually is. In fact, even though the data set appears to represent a sine wave of 0.2*Fs*, the actual signal could be any sine wave having a frequency of $(n)Fs \pm 0.2Fs$, where n is any integer, starting with 0. So the original signal could be 0.2*Fs*, 0.8*Fs*, 1.2*Fs*, 1.8*Fs*, 2.2*Fs*, etc. (or even more than one of those). We say that 0.2*Fs* is the *alias* of a signal that may actually be at another frequency entirely. During interpretation of sampled data, it is customary to treat signals as though they occurred in the baseband (0–0.5*Fs*), whether or not that is the case. In general, in a system sampling at *Fs*, a signal at a frequency *F* will alias into the baseband at

$$Fa = \text{abs}[(n)Fs - F] \tag{88.1}$$

where

abs denotes absolute value
$n \geq 0$
$(n)Fs$ is the closest integer multiple of *Fs* to *F*

Everyone has seen a demonstration of aliasing at the movies, in the form of "wagon wheeling." As the stagecoach or wagon takes off, the wheels begin to turn, slowly at first, then faster. As the wagon speeds up, the spokes suddenly appear to be turning backward, even though the wagon is moving forward. Sometimes, the spokes appear to be standing still. The reason for this is that a motion picture camera shooting film at 24 frames/s is a sampling system operating at 24 samples/s. The turning wagon wheel is a periodic signal that the camera undersamples. When the wheel begins turning just fast enough that one spoke travels at least half the distance to the next spoke in $\frac{1}{24}$th of a second, the spokes begin to

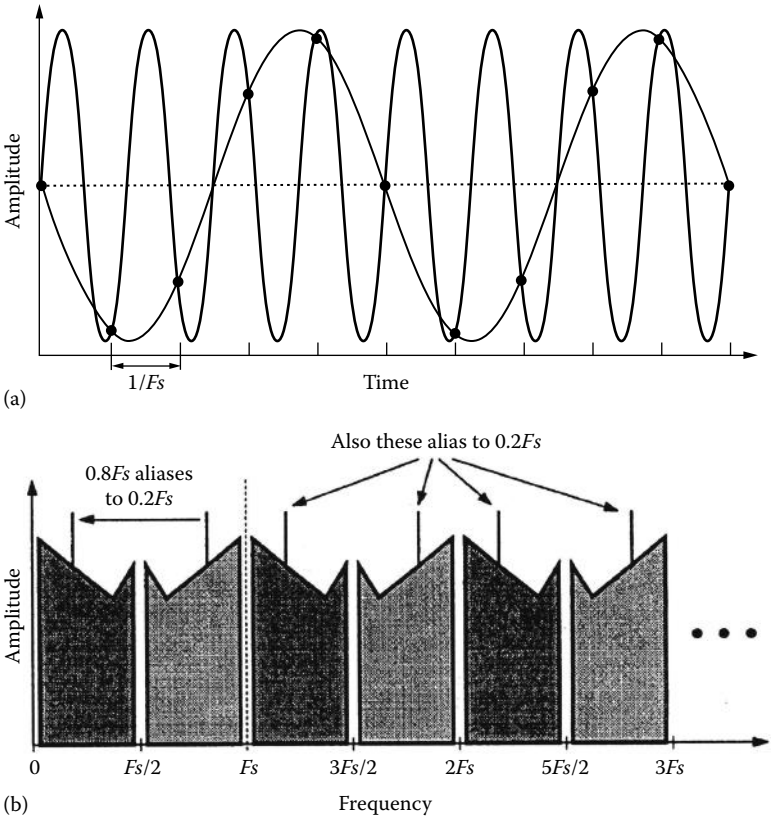


FIGURE 88.1 A demonstration of aliasing. An ADC sampling at rate *Fs* cannot distinguish between a 0.8*Fs* sine wave and a 0.2*Fs* sine wave: (a) a time-domain illustration and (b) a frequency-domain illustration. Theoretically, a sampler aliases an infinite number of 0.5*Fs* wide frequency bands into the baseband (0–0.5*Fs*). Practically, finite analog bandwidth eventually limits how far out in frequency aliases can come from.

appear to move backward, and the system is aliasing. When the wheel is turning so that a spoke moves exactly the distance between two spokes in $\frac{1}{24}$th of a second, the spokes appear to be standing still, since they all look the same to the camera.

It follows from Equation 88.1 that if we put into a sampler a signal with no energy at frequencies greater than one-half the sample rate (0.5*Fs*), then aliasing will not occur. This is the essence of the Shannon sampling theorem [1], which states that, with mathematical interpolation, the complete input waveform can be recovered *exactly* from the sampled data, at all times at and in between the sampling instants, as long as the sample rate is at least twice as high as the highest frequency content in the signal. Sometimes we refer to 0.5*Fs* as the *Nyquist frequency*, because Nyquist was concerned with the maximum bandwidth of signals [2]. Similarly, twice the highest frequency content of a signal (i.e., the minimum nonaliasing sample rate) is sometimes called the *Nyquist rate*. Sample rates are specified in samples/s, or S/s, and it is also common to specify rates in kS/s, MS/s, and even GS/s.

It is not always necessary to worry about aliasing. When an instrument is measuring slow-moving dc signals or is gathering data for statistical analysis, for instance, getting frequencies right is not important. In those cases, we choose the sample rate so that we can take enough data in a reasonable amount of time. On the other hand, if the instrument is a spectrum analyzer, where frequency does matter, or an oscilloscope, where fine time detail is needed, aliasing certainly is an issue. When aliased signals from beyond the

frequency band of interest can interfere with measurement, an instrument needs to have an *antialias filter* before the S/H. An antialias filter is a low-pass filter with a gain of 1 throughout most of the frequency band of interest. As frequency increases, it begins to attenuate the signal; by the Nyquist frequency, it must have enough attenuation to prevent higher-frequency signals from reaching the S/H with enough amplitude to disturb measurements. An efficient antialias filter must attenuate rapidly with frequency in order to make most of the baseband usable. Popular analog filters with rapid cutoff include elliptic and Chebyshev filters, which use zeros to achieve fast cutoff, and Butterworth filters (sixth order and above), which do not attenuate as aggressively, but have very flat passband response. A good book about filters is Ref. [3].

Some ADCs do not need a S/H or T/H at all. If the ADC is converting a slow-moving or dc signal and precise timing is not needed, the input may be stable enough during conversion that it is as good as sampled. There are also *integrating ADCs* (discussed later), which average the input signal over a period of time rather than sampling it. However, internally they actually sample the average.

88.3 Quantization

What sampling accomplishes in the time domain, quantization does in the amplitude domain. The process of digitization is not complete until the sampled signal, which is still in analog form, is reduced to digital information. An ADC quantizes a sampled signal by picking one integer value from a predetermined, finite list of integer values to represent each analog sample. Each integer value in the list represents a fraction of the total analog input range. Normally, an ADC chooses the value closest to the actual sample from a list of uniformly spaced values. This rule gives the *transfer function* of analog input-to-digital output a uniform "staircase" characteristic. Figure 88.2 represents a 3 bit quantizer, which maps a continuum of analog input values to only eight (2^3) possible output values. Each step in the staircase has (ideally) the same width along the *x*-axis, which we call *code width* and define as 1 *least significant bit* (LSB). In this case, 1 LSB is equal to 1 V. Each digital code corresponds to one of eight 1 LSB intervals making up the analog input range, which is 8 LSB (and also 8 V in this case).

Of course, we would like our measurement system to have a transfer function that is a straight line and has no steps at all. The bottom graph in Figure 88.2 is the ideal transfer function (a straight diagonal line) subtracted from the staircase function, or the *quantization error*. In an ideal ADC, the quantization error is bounded by ±½ LSB, and over the input range, the average error is 0 LSB and the standard deviation of error is $1/\sqrt{12}$ LSB. As the bottom graph shows, the quantization error at any point is a deterministic function of the input signal.

88.4 ADC Specifications

88.4.1 Range and Resolution

The *input range* of an ADC is the span of voltages over which a conversion is valid. The end points at the bottom and the top of the range are called *–full scale* and *+full scale*, respectively. When –full scale is 0 V the range is called *unipolar*, and when –full scale is a negative voltage of the same magnitude as +full scale the range is said to be *bipolar*. When the input voltage exceeds the input range, the conversion data are certain to be wrong, and most ADCs report the code at the end point of the range closest to the input voltage. This condition is called an *overrange*.

The *resolution* of an ADC is the smallest change in voltage the ADC can detect, which is inherently 1 LSB. It is customary to refer to the resolution of an ADC by the number of binary bits or decimal digits it produces; for example, "12 bits" means that the ADC can resolve one part in 2^{12} (=4096). In the case of a digital voltmeter that reads decimal digits, we refer to the number of digits that it resolves. A "6-digit" voltmeter on a 1 V scale measures from –0.999999 to +0.999999 V in 0.000001 V steps; it resolves one part in 2,000,000. It is also common to refer to a voltmeter measures from –1.999999 to +1.999999 as a "6½ digit" voltmeter. Table 88.1 compares the resolutions of common word lengths for ADCs.

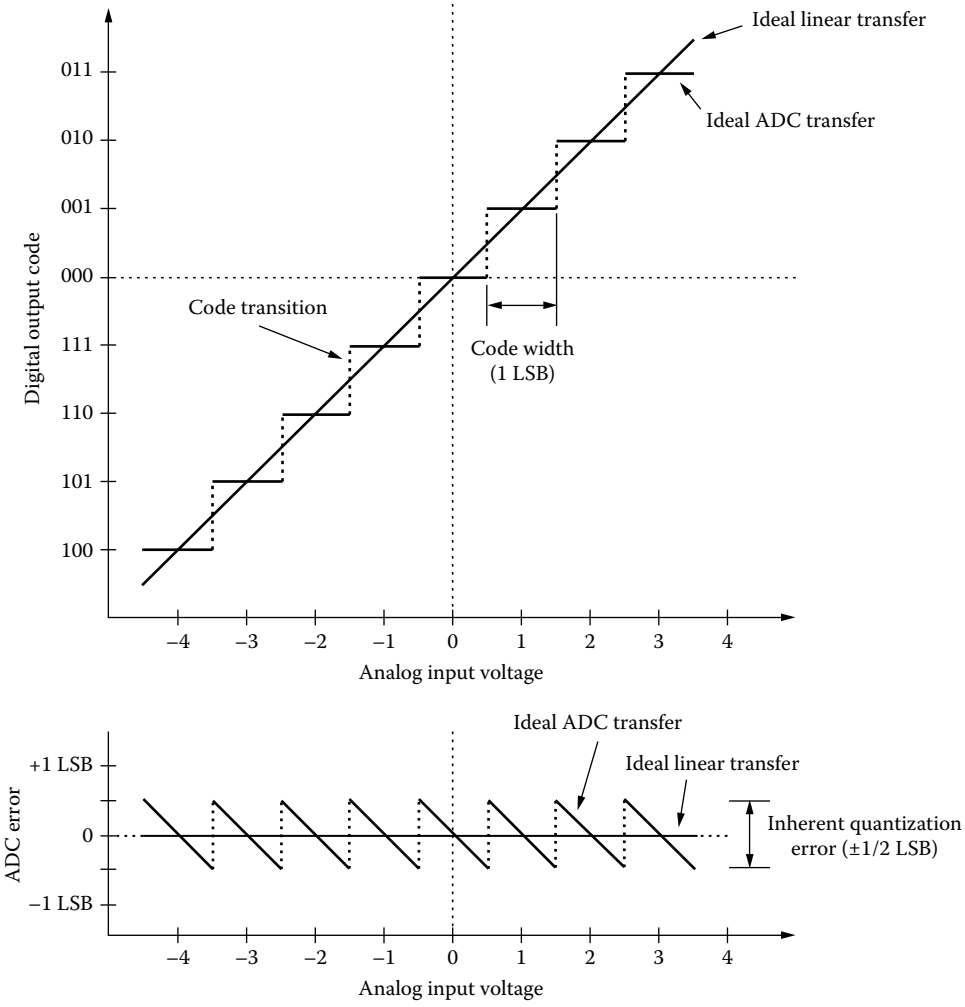


FIGURE 88.2 The ideal 3 bit quantizer has eight possible digital outputs. The analog input-to-digital output transfer function is a uniform staircase with steps whose width and height are 1 LSB exactly. The bottom graph shows the ideal transfer function (a straight line) subtracted from the staircase transfer function.

88.4.2 Coding Conventions

There are several different formats for ADC output data. An ADC using *binary* coding produces all 0s (e.g., 000 for the 3 bit converter) at –full scale and all 1s (e.g., 111) at +full scale. If the range is bipolar, so that –full scale is a negative voltage, binary coding is sometimes called *offset binary*, since the code 0 does not refer to 0 V. To make digital 0 correspond to 0 V, bipolar ADCs use *two's complement* coding, which is identical to offset binary coding except that the *most significant bit* (MSB) is inverted, so that 100 … 00 corresponds to –full scale, 000 … 00 corresponds to 0 V (*midscale*), and 011 … 11 corresponds to +full scale. All of the figures in this chapter depicting 3 bit ADC transfer functions use two's complement coding.

Decimal-digit ADCs, such as those used in digital voltmeters, use a coding scheme call *binary-coded decimal* (BCD). BCD data consist of a string of 4 bit groups of binary digits. Each 4 bit group represents a decimal digit, where 0000 is 0, 0001 is 1, and so on, up to 1001 for 9. The other six combinations (1010 through 1111) are invalid, or can be used for special information, such as the sign of the conversion.

TABLE 88.1 Comparison of Theoretical Resolutions of ADCs

Bits	Digits	Voltmeter "Digits"	Steps in FSR	Step Size (ppm)	Theoretical Dynamic Range (dB)
30	8.730		1,073,741,824	0.001	182.379
28.575	8.301	8½	400,000,000	0.003	173.802
28	8.128		268,435,456	0.004	170.338
27.575	8	8	200,000,000	0.005	167.782
26	7.526		67,108,864	0.015	158.297
25.253	7.301	7½	40,000,000	0.025	153.802
24.253	7	7	20,000,000	0.05	147.782
*24	6.924		16,777,216	0.060	146.255
22	6.322		4,194,304	0.238	134.214
21.932	6.301	6½	4,000,000	0.25	133.802
20.932	6	6	2,000,000	0.5	127.782
*20	5.720		1,048,576	0.954	122.173
18.610	5.301	5½	400,000	2.5	113.802
18	5.118		262,144	3.815	110.132
17.610	5	5	200,000	5	107.782
*16	4.515		65,536	15.259	98.091
15.288	4.301	4½	40,000	25	93.802
14.288	4	4	20,000	50	87.782
14	3.913		16,384	61.035	86.049
*12	3.311		4,096	244.141	74.008
11.966	3.301	3½	4,000	250	73.802
10.966	3	3	2,000	500	67.782
10	2.709		1,024	976.563	61.967
8.644	2.301	2½	400	2500	53.802
*8	2.107		256	3906.25	49.926
7.644	2	2	200	5000	47.782
6	1.505		64	15625	37.885

Note: "Bits" refers to binary word length, and "digits" refers to decimal word length. * denotes popular binary word lengths. FSR is full-scale range, and theoretical dynamic range is computed from the formula $1.7609 + 6.0206n$, where n is the number of bits (see discussion on dynamic range).

88.4.3 Linear Errors

Linear errors are the largest and most common errors in an ADC and are easily corrected by simple calibrations or by additions with and multiplications by correction constants. Linear errors do not distort the transfer function; they only change somewhat the input range over which the ADC operates.

Figure 88.3 shows the transfer function of an ideal 3 bit ADC with some *offset error*. The straight line joining the centers of the code transitions is raised, or offset, by 0.6 LSB, and the bottom graph shows the resulting error. Figure 88.4 shows an ideal 3 bit ADC with a +25% *gain error*. The slope of the line through the code transitions is 1.25 times the ideal slope of 1.00. If the slope of the line was 0.75 instead, the gain error would be –25%. The bottom graph shows the error resulting from excessive gain. Offset errors can be compensated for simply by adding a correcting voltage in the analog circuitry or by adding a constant to the digital data. Gain errors can be corrected by analog circuitry like potentiometers or voltage-controlled amplifiers or by multiplying the digital data by a correction constant.

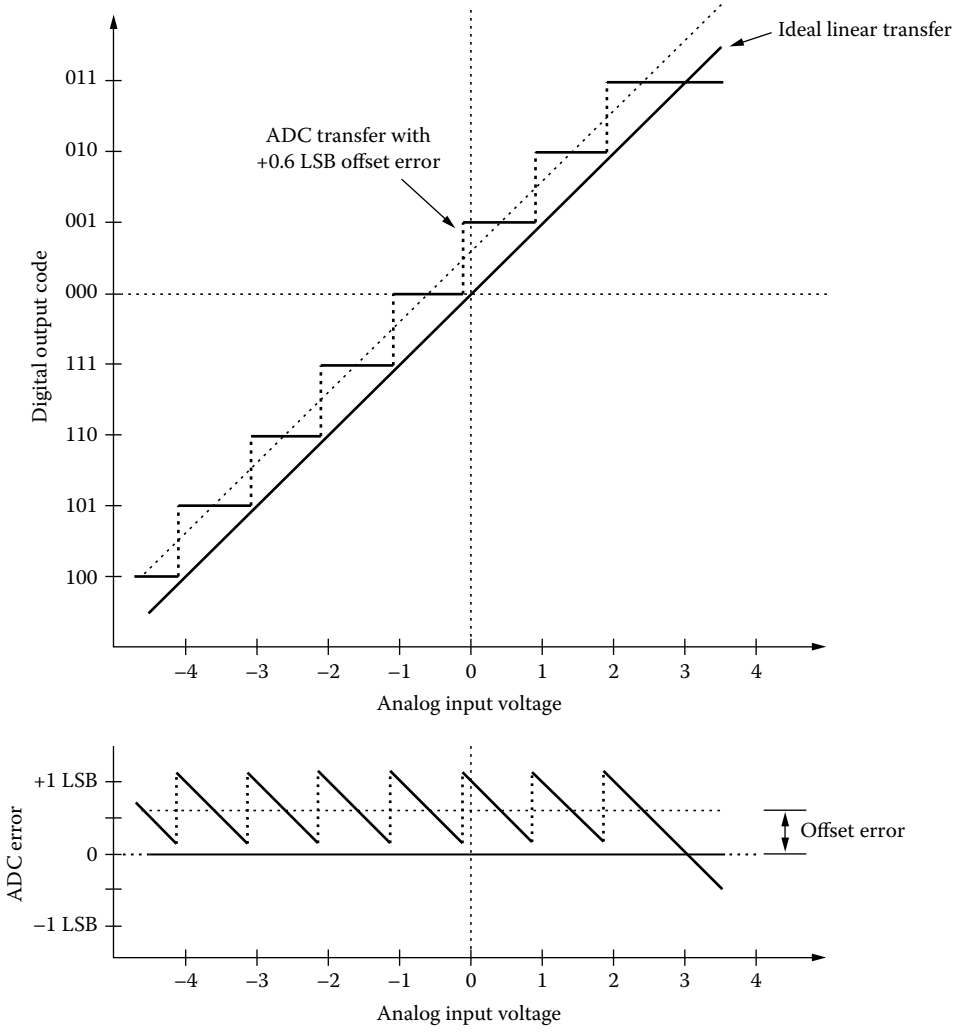


FIGURE 88.3 An ideal 3 bit quantizer, only with +0.6 LSB of offset error.

88.4.4 Nonlinear Errors

Nonlinear errors are much harder to compensate for in either the digital or analog domain, and are best minimized by choosing well-designed, well-specified ADCs. Nonlinearities are characterized in two ways: differential nonlinearity and integral nonlinearity.

DNL measures the irregularity in the code step widths by comparing their widths to the ideal value of 1 LSB. Figure 88.5 illustrates the 3 bit ADC with some irregular code widths. Most of the codes have the proper width of 1 LSB and thus contribute no DNL, but one narrow code has a width of 0.6 LSB, producing a DNL of –0.4 LSB, and one wide code has a width of 1.8 LSB, producing a DNL of +0.8 LSB at that code. This converter would be consistent with a DNL specification of ±0.9 LSB, for example, which guarantees that all code widths are between 0.1 and 1.9 LSB.

It is possible for a code not to appear at all in the transfer function. This happens when the code has a width of 0 LSB, in which case we call it a *missing code*. Its DNL is –1 LSB. If an ADC has a single missing code, the step size at that point in the transfer function is doubled, effectively reducing the local resolution of the ADC by a factor of two. For this reason it is important for an ADC specification to declare that

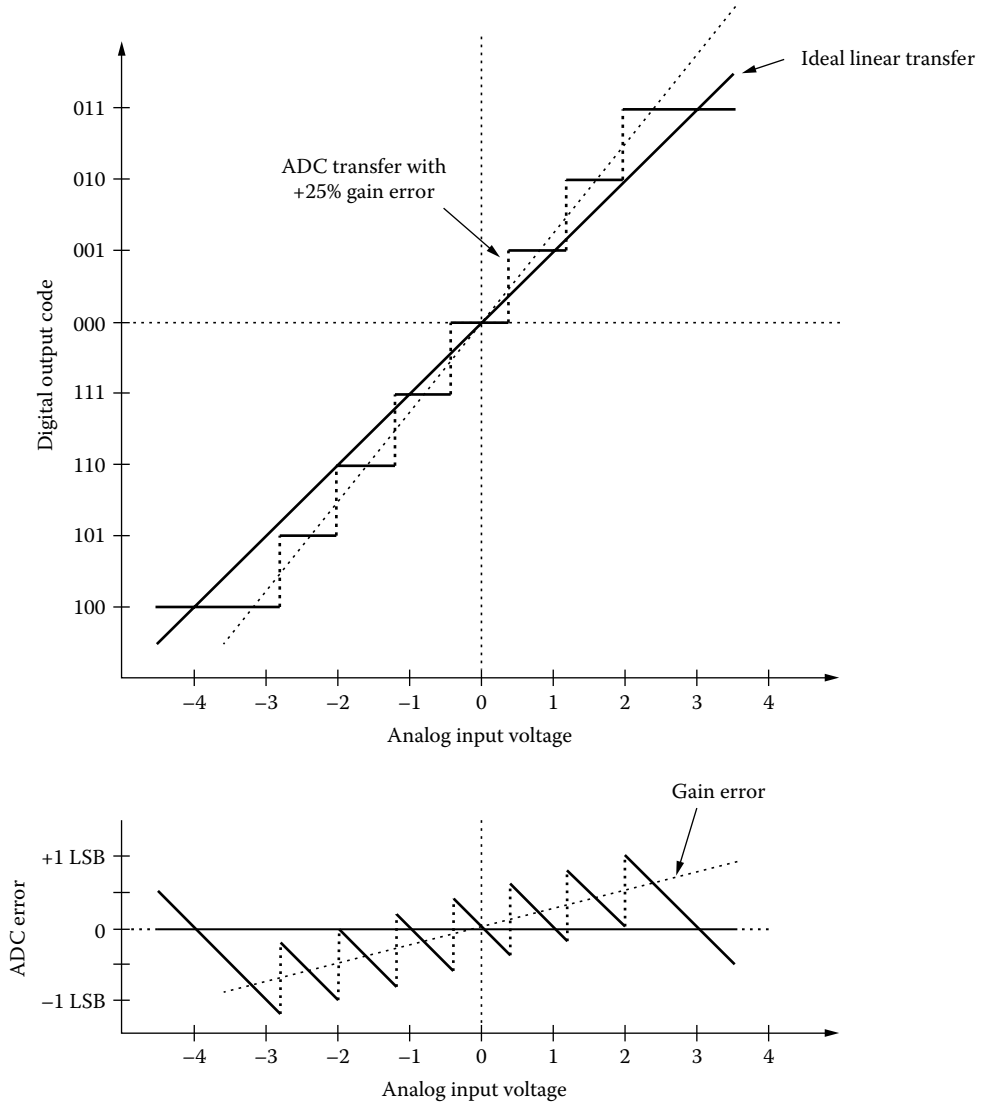


FIGURE 88.4 An ideal 3 bit quantizer, only with a gain of 1.25 instead of 1.00. This represents a +25% gain error.

the ADC has *no missing codes*, guaranteeing that every code has a width greater than 0 LSB. Even if an ADC has missing codes, no code can have a width <0 LSB, so the DNL can never be worse than –1 LSB.

INL measures the deviation of the code transitions from the ideal straight line, providing that the linear errors (offset and gain) have been removed. Figure 88.6 depicts an ADC with an INL error of +0.7 LSB. The offset and gain errors have been calibrated at the end points of the transfer function.

Relative accuracy (*RA*) is a measure of nonlinearity related to INL, but is more useful. It indicates not only how far away from ideal the code transitions are, but also how far any part of the transfer function, including quantization "staircase" error, deviates from ideal (assuming offset and gain errors have been calibrated at the end points). In a noiseless ADC, the worst-case RA always exceeds the worst-cast INL by ±0.5 LSB, as demonstrated in Figure 88.6. In an ADC that has a little inherent noise or has noise (called *dither*) added at the input, the RA actually improves because the addition of noise to the quantizer tends to smooth the *averaged* transfer function. Figure 88.7 shows the average of the digital output data as a function of the input voltage when 0.1 LSB rms of Gaussian random noise is intentionally added to the input.

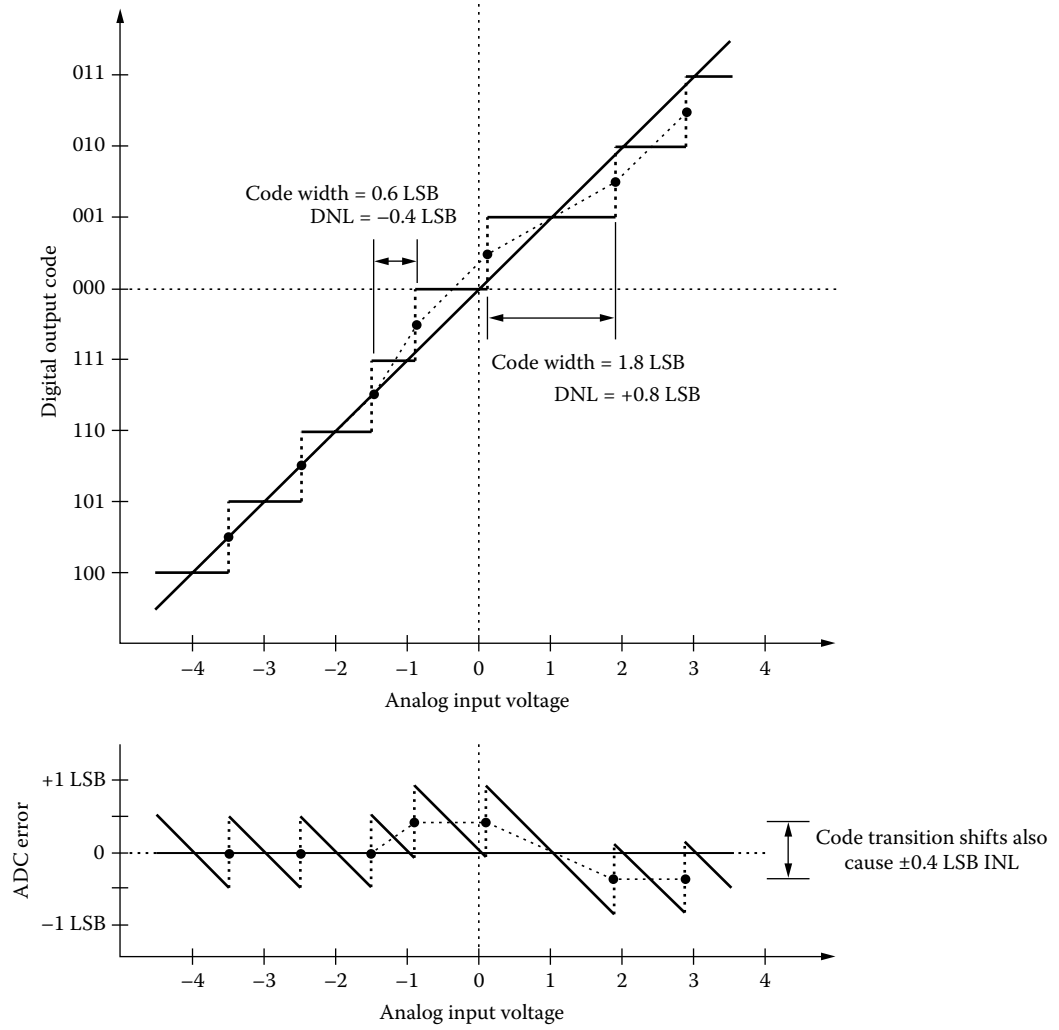


FIGURE 88.5 A 3 bit quantizer with substantial DNL errors. The bottom graph illustrates the resulting INL errors.

The RA improves to ±0.3 LSB from ±0.5 LSB in the noiseless case. If about 0.5 LSB rms of Gaussian noise is added, the quantization staircase becomes nearly straight. This improvement in linearity comes at the expense of the random error in each individual conversion caused by the noise. Adding more noise to the ADC does not improve the average quantization error much more, but it does tend to smooth out local nonlinearities in the averaged transfer function. For a good discussion of noise and dither, see Ref. [4].

88.4.5 Aperture Errors

Aperture errors have to do with the timing of analog-to-digital conversions, particularly of the S/H. *Aperture delay* characterizes the amount of time that lapses from when an ADC (S/H) receives a convert pulse to when the sample is held as a result of the pulse. Although aperture delay (sometimes called *aperture time*) is usually specified as a few nanoseconds for an ADC or S/H by itself, this delay is usually much more than negated by the group delay in any amplifiers that precede the S/H, so that the convert pulse arrives at the S/H quite some time before the analog signal does. For instance, a typical 1 MHz bandwidth amplifier has 160 ns of delay; if the ADC or S/H it was connected to have an aperture delay of 10 ns, the effective aperture delay for the system would be –150 ns.

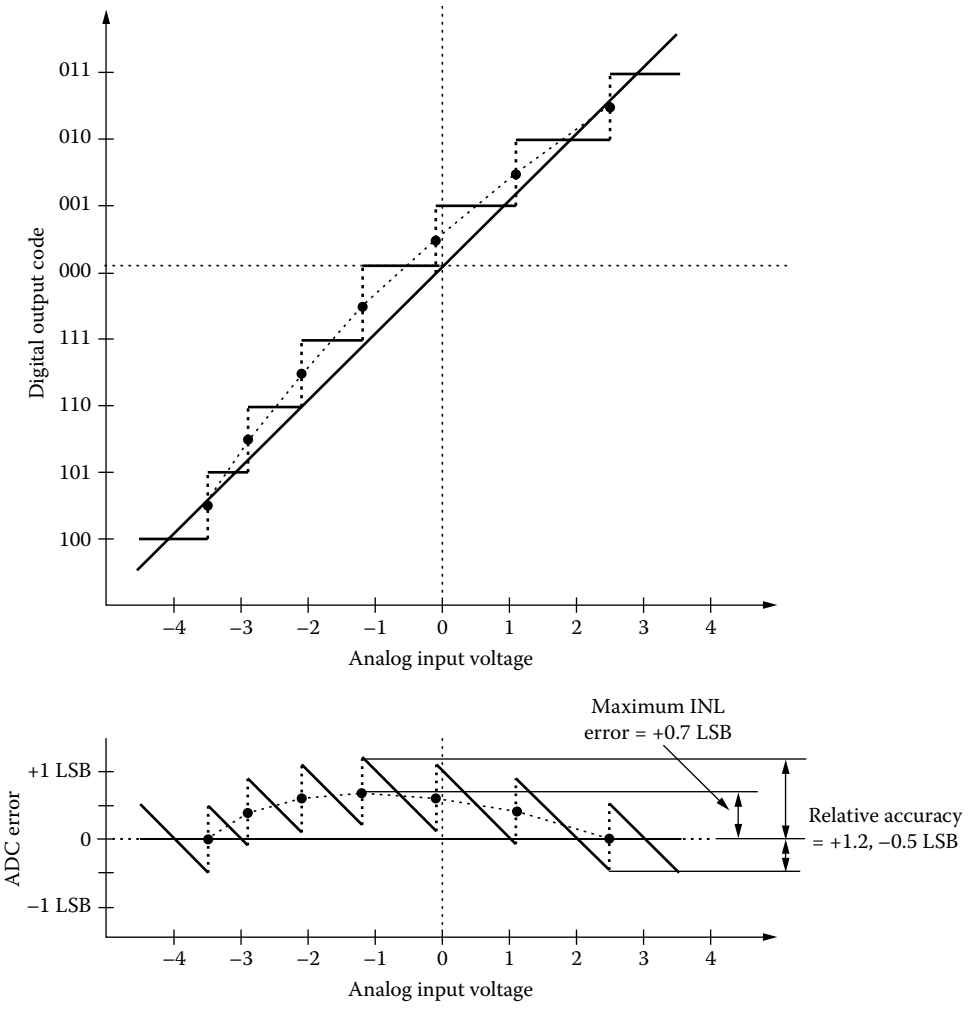


FIGURE 88.6 A 3 bit quantizer with substantial INL errors. Here, the DNL error is still significant, but, for example, a 12 bit converter with 0.7 LSB of INL from a smooth error "bow" like the one above could have negligible DNL because it would have so many more steps over which to accumulate error.

Jitter (or *aperture jitter*) characterizes the irregularity in times at which samples are taken. If the nominal period between samples in an ADC is 1 μs, the actual time may vary from 1 μs by as much as a few hundred picoseconds or even as much as a nanosecond from cycle to cycle. Contributions to these variations can come from the crystal clock source (if included under the jitter specification), digital clock circuitry, or the S/H. Jitter is usually specified in picoseconds peak-to-peak or picoseconds rms.

Jitter interferes with measurements (particularly spectral analysis) by effectively frequency modulating the input signal by the jitter profile. A jittery ADC sampling a pure sine wave would scatter energy from the sine wave all throughout the spectrum, perhaps covering up useful spectral information. In a typical ADC, however, most of the interference from jitter tends to occur at frequencies very close to the main signal.

88.4.6 Noise

Noise, whether inherent in an ADC or introduced intentionally (see dither earlier), limits the resolution of an ADC by adding an interfering waveform to the input signal as the data are converted. Noise comes

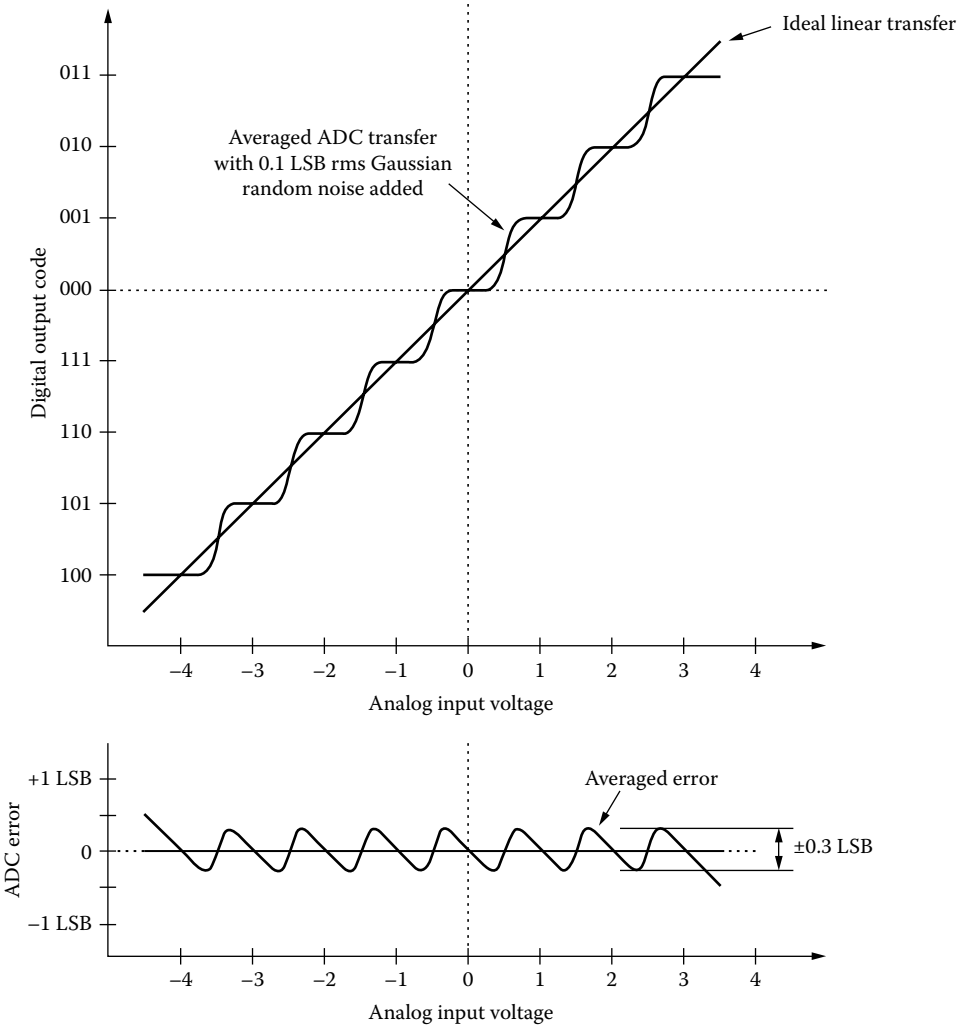


FIGURE 88.7 An ideal 3 bit quantizer with 0.1 LSB rms Gaussian random noise (dither) added at the input. The relative accuracy has improved to ±0.3 LSB rms from the ±0.5 LSB expected from a noiseless quantizer. With the application of 0.5 LSB rms Gaussian noise, the transfer function becomes almost perfectly straight. Larger amounts of dither produce essentially no improvement in linearity.

from many places. The most common kind of noise is *thermal noise*, which is caused by the random nature of electric conduction in resistors and transistors. Thermal noise is worse at higher temperatures and higher resistances. Most other ADC noise is coupled electromagnetically from nearby circuitry, such as clock or logic circuits, or from routing of other input signals. Noise is usually specified in volts rms or peak-to-peak, or LSBs rms or peak-to-peak.

Quantization error (see earlier) can sometimes be thought of as *quantization noise*. Although quantization error is perfectly predictable with respect to the input signal, when a signal is fairly "busy" (i.e., busy enough that consecutive conversions do not tend to repeat data) the quantization error becomes chaotic, and it can be thought of it as another source of random noise, whose statistical distribution is uniform from −0.5 LSB to +0.5 LSB and whose standard deviation is $1/\sqrt{12}$ LSB. This is sometimes the dominant source of noise in spectral analysis applications.

Once noise gets into an ADC, there are ways to process out the noise if it is independent of the signal. Acquisitions of dc signals can be quieted by collecting a number of points and averaging the collection. If the noise is *white random noise*, which has equal energy density at all frequencies, averaging can reduce the amount of noise by the square root of the number of samples averaged. The noise interfering with a repetitive waveform can be quieted by acquiring many waveforms using a level trigger and averaging the collection to produce an average waveform. Most digital oscilloscopes have waveform averaging. Quantization noise, as described earlier, cannot be averaged out unless other random noise is present.

The noise specifications for an ADC are for quiet, low-impedance signals at the input, such as a dead short. To preserve the noise performance of the ADCs, the user must carefully connect signals to the input with tidy cabling that keeps away from sources of electromagnetic noise. For more information on noise sources and treatment and prevention of noise, see Refs. [5,6].

88.4.7 Dynamic Range

The *dynamic range* (DR) of an ADC is the ratio of the largest to the smallest signals the converter can represent. The largest signal is usually taken to be a full-scale sine wave, and the smallest signal is usually taken to be the background noise level of the ADC. It can be expressed simply as a ratio, but it is more common to express it in decibels (dB):

$$\mathrm{DR} = 20\log\left(\frac{\mathrm{S}}{\mathrm{N}}\right) \tag{88.2}$$

where

DR is dynamic range in dB
S is the rms amplitude of the largest signal
N is the rms amplitude of the smallest signal, the noise

The noise must include the quantization noise of the ADC, which for a perfect, noiseless converter is $1/\sqrt{12}$ LSB rms. For an n bit converter, a full-scale sine wave has a peak amplitude of 2^{n-1} LSB, which corresponds to an rms amplitude of $2^{n-1}\sqrt{2}$ LSB, or $2^{n-1.5}$ LSB rms. Hence, a perfect ADC had a DR of

$$\begin{aligned}\mathrm{DR} &= 20\log\left(2^{n-1.5}\times\sqrt{12}\right)\\ &= 20\log(2^n) + 20\log\left(2^{-1.5}\times\sqrt{12}\right)\\ &= (n)[20\log(2)] + 20\log\left(\sqrt{1.5}\right)\\ &\approx 6.0206n + 1.7609.\end{aligned} \tag{88.3}$$

Equation 88.3 can be used to determine the *effective number of bits* (ENOB) of an imperfect ADC. ENOB may take only noise into account, or it may include noise and harmonic distortion products of the input signal. It is computed as

$$\mathrm{ENOB} = \frac{(\mathrm{DR} - 1.7609)}{6.0206} \tag{88.4}$$

For example, a 16 bit ADC with a DR of 92 dB has an ENOB of 14.988 bits.

88.5 Types of ADCs

The fundamental building block of analog-to-digital conversion is the *comparator*. Every type of ADC has at least one comparator in it, and some ADCs have many. The comparator itself is a 1 bit ADC; it has two analog voltage inputs and (usually) one digital output. If the voltage at the positive input is greater than the voltage at the negative input, the output of the comparator is a digital 1. If the voltage at the positive input is less than the voltage at the negative input, the output is a digital 0 (see Figure 88.8).

Another piece that all ADCs have in common is a linearity reference. This is what a comparator in an ADC compares the input signal with in the process of conversion. It directly determines the DNL and INL of the ADC. Examples of linearity references include capacitors (in integrating ADCs) and DACs (found in successive-approximation ADCs).

The third piece that every ADC has is a voltage reference. The reference(s) determine the full-scale input range of the ADC and are usually part of or closely associated with the linearity reference.

88.5.1 Flash

Flash converters are the fastest ADCs, achieving speeds near 1 GS/s and resolutions of 10 bits and below. The flash converter with n bits of resolution has $2^n - 1$ high-speed comparators operating in parallel (see Figure 88.9). A string of resistors between two voltage references supplies a set of uniformly spaced voltages that span the input range, one for each comparator. The input voltage is compared with all of these voltages simultaneously, and the comparator outputs are 1 for all voltages below the input voltage and 0 for all the voltages above the input voltage. The resulting collection of digital outputs from the comparators is called a "thermometer code," because the transition between all 1s and all 0s floats up and down with the input voltage. Fast logic converts the thermometer codes to normal n bit binary numbers.

Because of their simplicity, they are fast, but flash converters are limited to resolutions of 10 bits and below because the number of comparators and resistors goes up exponentially with resolution. Because the string resistor values typically vary only a few percent from each other in practical devices, the differential linearity of the flash ADC is quite good. But the same resistor variations can accumulate error across the input range and cause INL of a few LSBs.

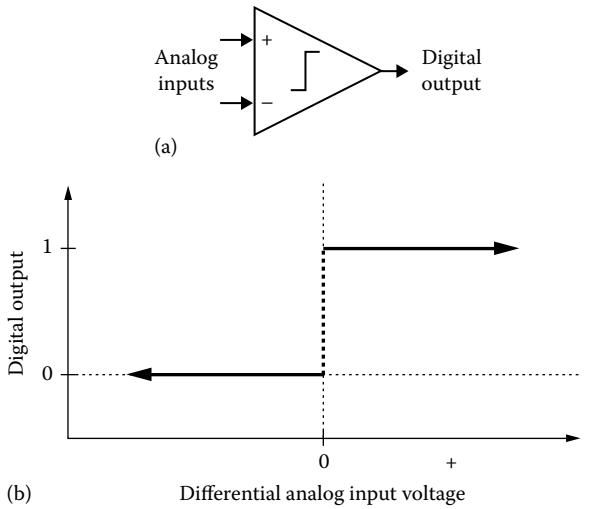


FIGURE 88.8 The comparator is the essential building block of all ADCs. (a) Comparator symbol and (b) comparator input/output transfer function.

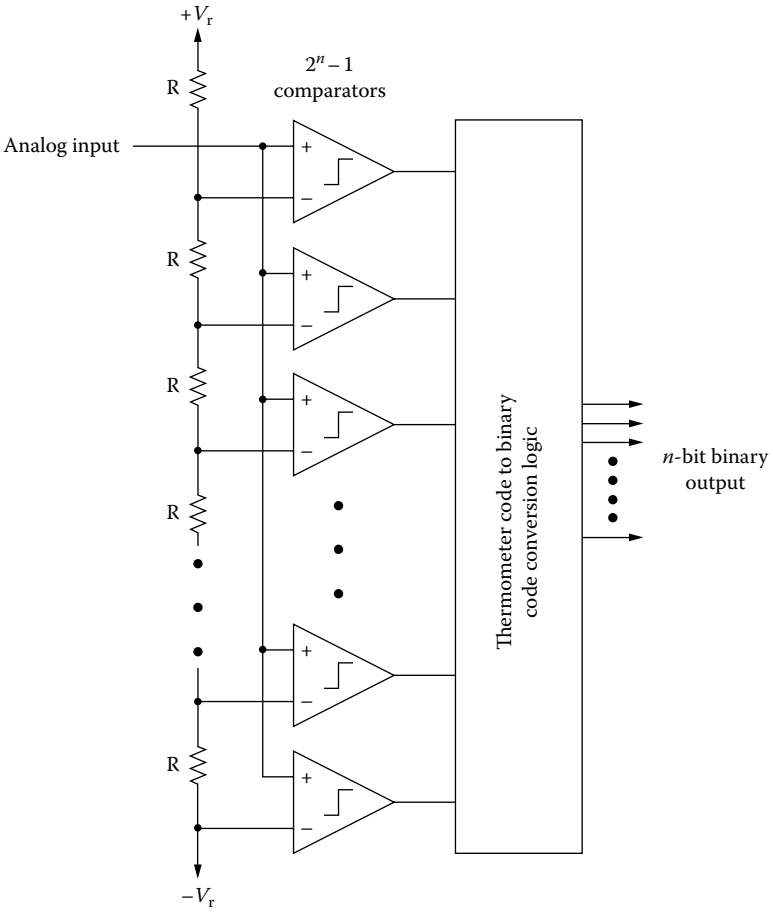


FIGURE 88.9 A flash converter has $2^n - 1$ comparators operating in parallel. It relies on the uniformity of the resistors for linearity.

88.5.2 Successive-Approximation Register

Successive-approximation register (SAR) ADCs are the most common ADCs, having resolutions of 8–16 bits and speeds of 1 MS/s and below. They are generally low in cost, and they typically have very good integral linearity. The *n* bit SAR ADC contains a high-speed *n* bit DAC and comparator in a feedback loop (see Figure 88.10). The SAR sequences the DAC through a series of *n* "guesses," which are compared with the input voltage (Figure 88.11). As the conversion progresses, the register builds the *n* bit binary

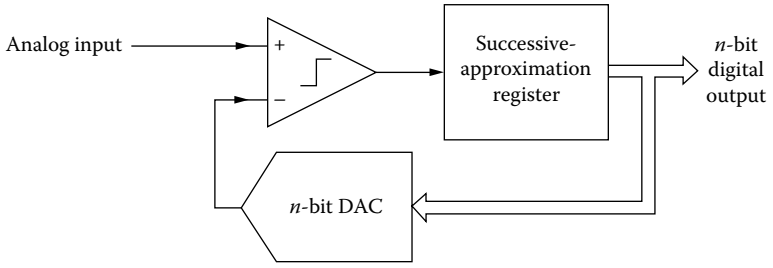


FIGURE 88.10 A successive-approximation converter has only one comparator and relies on an internal, precision DAC for linearity.

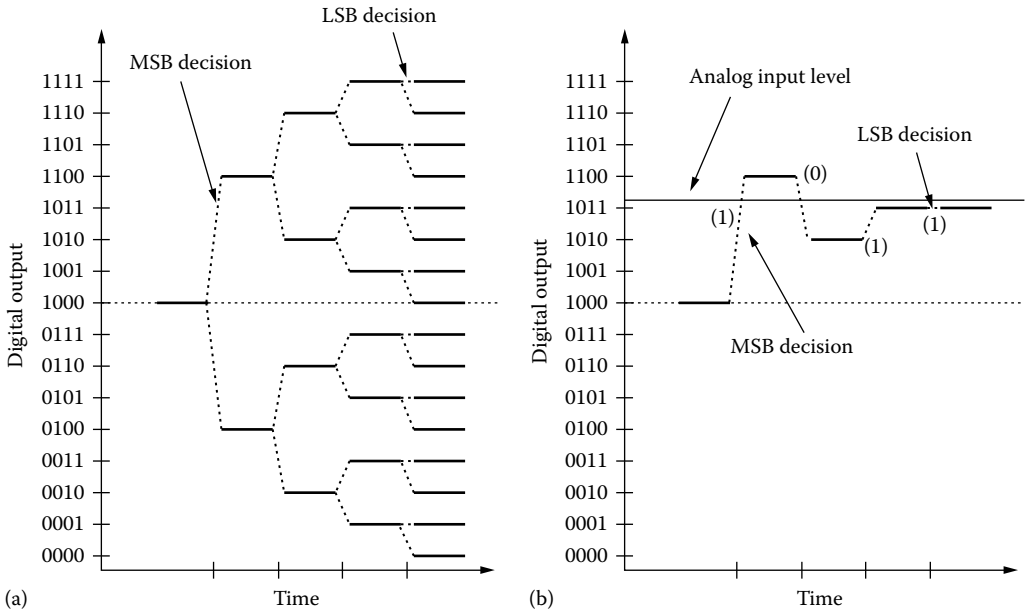


FIGURE 88.11 (a) Decision tree shows all the possible digital "guesses" of a 4 bit successive-approximation converter over time and (b) decision tree for conversion of 4 bit code 1011.

conversion result out of the comparator outputs. By the end of the sequence, the register has converged to the closest DAC value to the input voltage.

The speed of an SAR ADC is limited by the speed of the DAC inside the feedback loop. The DAC must settle n times to within $1/2^{-n}$ of full scale within the conversion time of the ADC. Current SAR technology achieves 12 bit resolution at 1 MS/s and 16 bit resolution at 200 kS/s. Faster conversion at these resolutions requires multistage architectures.

88.5.3 Multistage

To achieve higher sample rates than SAR ADCs at resolutions of 10–16 bits, *multistage* ADCs (sometimes called *subranging* or *multipass* ADCs) use the iterative approach of SAR ADCs but reduce the number of iterations in a conversion. Instead of using just a comparator, the multistage ADC uses low-resolution flash converters (4–8 bits) as building blocks. Figure 88.12 illustrates an example of a 12 bit

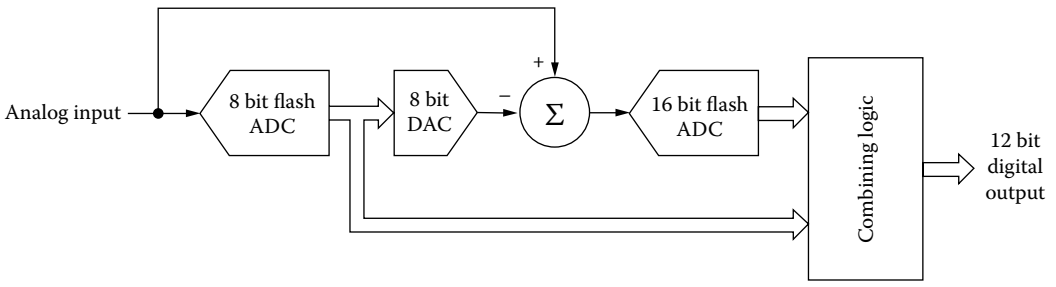


FIGURE 88.12 An example of a 12 bit multistage ADC built out of two flash ADCs and a fast DAC. The 8 bit flash ADC takes a first "guess" at the input signal and the 6 bit flash ADC converts the error in the guess, called the "residue." The 12 bit DAC actually needs to have only 8 bits, but it must be accurate to 12 bits. If the 8 bit flash ADC were perfect, the second flash ADC would only need 4 bits. But since the first flash actually may have some error, the second flash has 2 bits of "overlap."

two-stage ADC built out of two flash ADCs and a fast DAC. The 6 bit flash ADC converts the residual error of the 8 bit flash ADC. The two digital outputs are combined to produce a 12 bit conversion result.

If each flash ADC has a T/H at its input, then each stage can be converting the residual error from the previous stage while the previous stage is converting the next sample. The whole converter then can effectively operate at the sample rate of the slowest stage. Without the extra T/Hs, a new conversion cannot start until the residues have propagated through all the stages. This variation of the multistage ADC is called a *pipelined* ADC.

88.5.4 Integrating

Integrating converters are used for low-speed, high-resolution applications such as voltmeters. They are conceptually simple, consisting of an integrating amplifier, a comparator, a digital counter, and a very stable capacitor for accumulating charge (Figure 88.13). The most common integrating ADC in use is the dual-slope ADC, whose action is illustrated in Figure 88.14. Initially, the capacitor is discharged and so has no voltage across it. At time 0, the input to the integrator is switched to the analog input and the capacitor is allowed to charge for an amount of time, $T1$, which is always the same. Its rate of charging and thus its voltage at $T1$ are proportional to the input voltage. At time $T1$, the input switch flips over to the voltage reference, which has a negative value so that the capacitor will begin to discharge at a rate proportional to the reference. The counter measures how long it takes to discharge the capacitor completely. If the capacitor is of high quality, the ratio of the discharge time

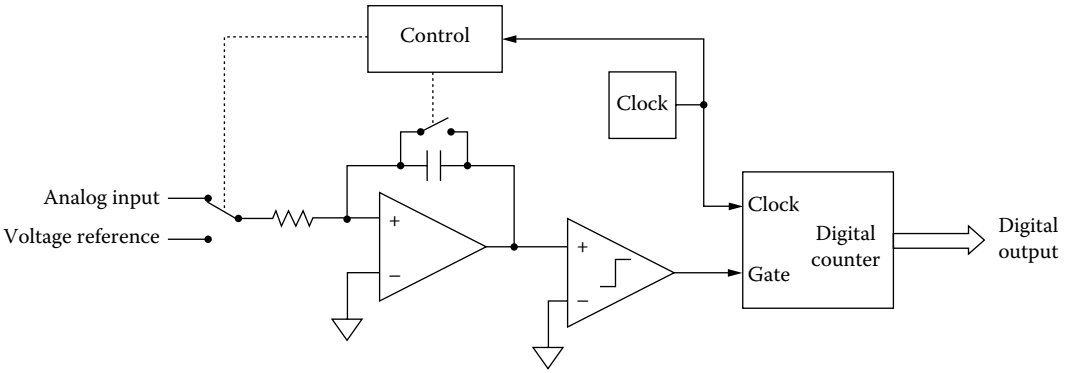


FIGURE 88.13 A dual-slope integrating converter uses a comparator to determine when the capacitor has fully discharged and relies on the capacitor for linearity.

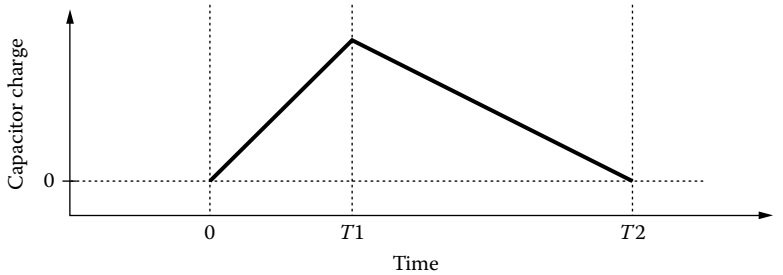


FIGURE 88.14 Charge on the integrating capacitor vs. time. At time 0, the input is switched to analog input and the switch across the capacitor opens. The capacitor integrates charge until fixed time $T1$. The input is then switched to the voltage reference to discharge the capacitor, and the counter begins counting a known clock. The comparator turns off the counter when the capacitor charge reaches 0 again, at time $T2$. The resulting count is proportional to the average input voltage over the time interval 0 to $T1$.

to the charge time is proportional to the ratio of the input voltage to the voltage reference, and so the counter output represents the analog input voltage.

An elaboration of the dual-slope ADC is the *multislope* integrating ADC. It achieves even higher resolution than the dual-slope ADC by discharging the capacitor at several progressively slower rates. At each rate, the counter is able to resolve finer increments of accumulated charge.

An important distinction between integrating converters and other ADCs is the way they sample the input voltage. Integrating converters do not sample the voltage itself; they *average* the voltage over the integration period and *then* they sample the average that is accumulated on the capacitor. This tends to reject noise that conventional sampling cannot, especially periodic noises. Most integrating ADCs operate with an integration period that is a multiple of one ac line period ($\frac{1}{60}$ or $\frac{1}{50}$ s) so that any potential interference from stray electric or magnetic fields caused by the power system is canceled.

Integrating converters are gradually being replaced in the marketplace with low-speed, high-resolution sigma–delta (SD) converters, which see. SD converters are generally more flexible than integrating ADCs, and they are easier to use because they do not require an external charging capacitor. The resolution and speed of the two types are comparable, although integrating converters still have the highest linearity.

88.5.5 Sigma–Delta ADCs

The SD ADC is quickly becoming one of the most popular types of ADC. SD ADCs typically have resolutions of 16–24 bits and sample rates of 100 kS/s down to 10 S/s. Because of their high resolution at 48 kS/s, they are the most common type of converters in modern digital audio equipment. SD ADCs defy intuition by quantizing initially with *very* low resolution (often 1 bit) at very high rates, typically 64× to 128× the eventual sample rate (called *oversampling*). The high-rate, low-resolution quantizer operates inside a feedback loop with an analog low-pass filter and a DAC to force the large amount of otherwise unavoidable quantization error (noise) to frequencies higher than the band of interest. The resulting spectral redistribution of the quantization noise is called *noise shaping*, illustrated in Figure 88.15. The low-resolution digital output of the ADC loop is fed into a digital filter that increases the resolution from the resolution of the ADC loop to the output resolution, reduces the data rate from

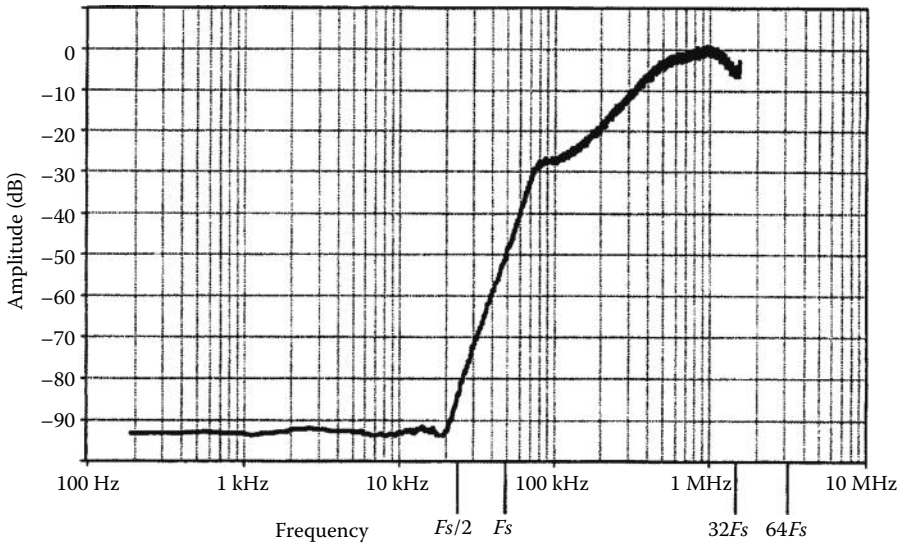


FIGURE 88.15 Spectrum of a 64-time oversampling SD ADC before the digital decimation filter. The modulator loop samples at 3.072 MS/s and the data come out of the filter at 48 kS/s. The filter cuts off sharply at *Fs*/2, or 24 kHz, leaving only the small amount of noise left below 24 kHz.

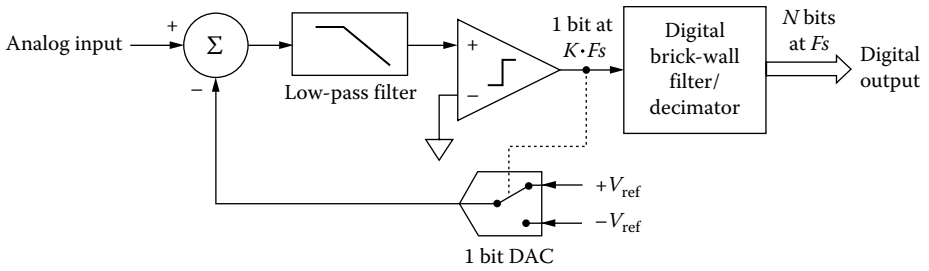


FIGURE 88.16 A SD modulating ADC uses a comparator simply as a 1 bit quantizer. The linearity of a SD ADC is theoretically perfect because the 1 bit DAC can only assume two values, and thus is linear by definition. Modern SD ADCs are made with switched-capacitor circuits which operate at *KFs*, where *Fs* is the output data sample rate and *K* is the oversampling ratio.

the rate of the ADC loop to the output sample rate, and applies a low-pass digital filter, leaving only the signals in the frequency band of interest and a little inherent electronic noise.

Figure 88.16 shows how a 1 bit SD ADC works. The comparator *is* the ADC, and its output is processed digitally, so that no further analog errors can accumulate. The comparator is in a feedback loop with a low-pass filter (typically third to fifth order) and a 1 bit DAC. The 1 bit DAC can take on only one of two values, +full scale and –full scale, so it is perfectly linear. The low-pass filter causes the loop gain to be high at low frequencies (the signal band of interest) and low at high frequencies. Since the error in a feedback loop is low when the gain is high and high when the gain is low, the errors dominate at high frequencies and are low in the band of interest. The result is a 1 bit output whose duty cycle is proportional to the input signal. Together, the elements of the feedback loop are called a *SD modulator*.

Figure 88.17 illustrates the operation of a simple discrete-time (switched-capacitor) SD ADC. In this first-order example, the low-pass filter is just an integrator. The loop tries to force the input to the comparator back to the baseline, and the figure shows how the duty cycle of the resulting digital output reflects input signal. The digital data here have undesirable patterns which tend to repeat, called *limit cycles*. They can appear in the band of interest and interfere with the signal. Higher-order loop filters (third and above) make the bit activity so chaotic that it has no substantial limit cycles.

The chief advantage of a SD converter is that it has a built-in antialias filter, and a darn good one at that. Most SD parts have a *finite-impulse response* (FIR) digital filter, which has an extremely flat frequency response in the passband and an extremely sharp cutoff, properties impossible to implement in analog filters. The ADC still needs an antialias filter to reject signals above one-half the oversampling rate. But this filter is simple to build, since it has to be flat only up to one-half the output sampling rate and has many octaves (all the way to near the oversampling rate) to fall off. The combination of the two filters provides watertight protection from aliases, often 96 dB of attenuation over the entire spectrum.

An important improvement of the 1 bit SD is the multibit SD, wherein the comparator is replaced by a flash converter with as much as 4 bits of resolution. This improves the ENOB of the whole converter by several bits.

88.5.6 Voltage-to-Frequency Converters

Voltage-to-frequency converters (VFCs) are versatile, low-cost circuits that convert analog voltages to periodic waveforms whose frequency is proportional to the analog input voltage. A VFC is conceptually similar to an integrating converter (see earlier) except that the digital counter is missing and is replaced with a short-pulse generator that quickly discharges the capacitor. The voltage reference is not connected intermittently to the input; instead, it appears all the time at the minus input of the comparator

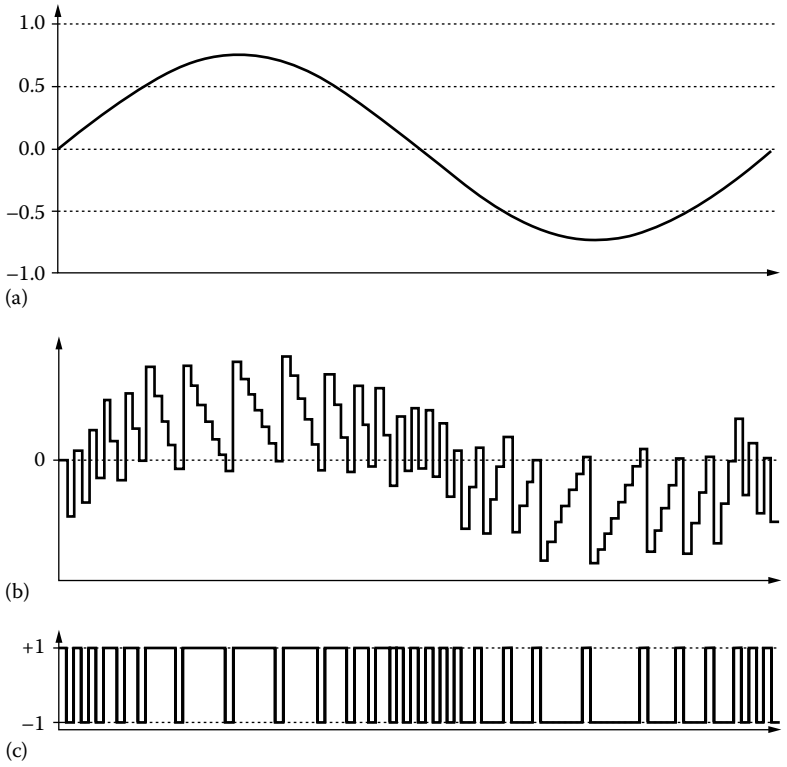


FIGURE 88.17 Behavior of a discrete-time (switched-capacitor) first-order SD modulator, where the low-pass filter is simply an integrator. In each graph, the *x*-axis represents time, and the *y*-axis represents signal level. (a) The input waveform, (b) input to the comparator, and (c) the 1 bit digital comparator output. The duty cycle of this waveform corresponds to the input waveform. The digital filter and decimator recover the original waveform from this 1 bit.

instead of ground. The capacitor charges at a rate proportional to the input voltage, until the voltage is equal to the voltage reference. Then the comparator trips the pulse generator, which quickly discharges the capacitor, and the cycle begins again. The periodic pulse at the comparator output can be used as the digital output.

The advantage of the VFC over conventional ADCs is that the 1 bit output can be transmitted digitally, through isolation transformers, through fiber-optic cable, or through any other isolating, nonisolating, long-distance, or short-distance transmission medium. All that is needed at the receiving end to complete the analog-to-digital conversion is a digital counter, which does not need to be synchronized to the VFC itself. Sometimes, the digital conversion is not needed; a VFC can be used with an isolating transformer and a *frequency-to-voltage converter* (FVC) to create an isolation amplifier. For a good discussion of VFCs, see Ref. [7].

88.6 Instrumentation and Components

88.6.1 Integrated Circuits

Table 88.2 lists several popular high-quality ADCs in integrated circuit form. By no means exhaustive, the list is sampling of a few of the most popular or best-performing chips of each type of ADC. Table 88.3 contains addresses, phone numbers, and Internet sites for the manufacturers in Table 88.2.

TABLE 88.2 ADC-Integrated Circuits

Part	Type	Sample Rate	Resolution, Bits	Manufacturer
ADC 160	Integrating	1 S/s	24	Thaler
AD7714	Sigma–delta	2.62 S/s	24	Analog Devices
MAX132	Integrating	6 S/s	19	MAXIM
CS5508	Sigma–delta	20 S/s	20	Crystal
HI7190	Sigma–delta	10 S/s	24	Harris
AD 1879	Sigma–delta	50 kS/s	18	Analog Devices
CS5390	Sigma–delta	50 kS/s	20	Crystal
ADS7809	SAR	100 kS/s	16	Burr-Brown
CS5101A	SAR	100 kS/s	16	Crystal
AD7893	SAR	117 kS/s	12	Analog Devices
AD976	SAR	200 kS/s	16	Analog Devices
AD7722	Sigma–delta	200 kS/s	16	Analog Devices
LTC1278	SAR	500 kS/s	12	Linear Technology
AD1385	Multistage	500 kS/s	16	Analog Devices
ADS7819	SAR	800 kS/s	12	Burr-Brown
AD9220	Multistage	10 MS/s	12	Analog Devices
AD775	Multistage	20 MS/s	8	Analog Devices
AD9050	Multistage	40 MS/s	10	Analog Devices
AD9066	Flash	60 MS/s	6	Analog Devices
HI1276	Flash	500 MS/s	8	Harris

TABLE 88.3 Companies That Manufacture ADC-Integrated Circuits

Analog Devices, Inc.
One Technology Way
P.O. Box 9106
Norwood, MA 02062-9106
(617) 329-4700
http://www.analog.com

Burr-Brown Corporation
P.O. Box 11400
Tucson, AZ 85734-1400
(520) 746-1111
http://www.burr-brown.com

Crystal Semiconductor Corporation
P.O. Box 17847
Austin, TX 78760
(512) 445-7222
http://www.cirrus.com/prodtech/crystal.html

Harris Corp Semiconductor
Products Division
P.O. Box 883
Melbourne, FL 37902
(407) 729-4984
http://www.semi.harris.com

Linear Technology Corporation
1630 McCarthy Blvd.
Milpitas, CA 95035-7417
(408) 432-1900
http://www.linear-tech.com

Maxim Integrated Products, Inc.
120 San Gabriel Drive
Sunnyvale, CA 94086
(408) 737-7600
http://www.maxim-ic.com

Thaler Corporation
2015 N. Forbes Boulevard
Tucson, AZ 85745
(520) 882-4000
http://www.thaler.com

88.6.2 Instrumentation

Plug-in data acquisition cards are becoming increasingly popular as personal computer prices come down and processor performance goes up. These cards typically contain one or more ADCs (with S/H), instrumentation amplifiers with gain and differential input, and multiplexers to switch to different inputs. Some have DACs on-board, and some have digital data and timing functions as well. Once considered low performance and hard to use, data acquisition cards have improved dramatically, equaling and in some cases exceeding capabilities of stand-alone instruments. Most come

TABLE 88.4 Plug-DATA Acquisition Boards

Part	Type	Sample Rate	Resolution, Bits	Manufacturer
AT-A2150	Sigma–delta	51.2 kS/s	16	National Instruments
AT-MIO-16XE-50	SAR	20 kS/s	16	National Instruments
AT-MIO-16E-10	SAR	100 kS/s	12	National Instruments
CIO-DAS1600/12	SAR	160 kS/s	12	ComputerBoards, Inc.
AT-MIO-16XE-10	SAR	100 kS/s	16	National Instruments
CIO-DAS1600/16	SAR	100 kS/s	16	ComputerBoards, Inc.
DT-3001	SAR	330 kS/s	12	Data Translation, Inc.
DAS-1800AO	SAR	333 kS/s	12	Keithley Metrabyte
AT-MIO-16E-1	SAR	1 MS/s	12	National Instruments
FAST16-1	Multistage	1 MS/s	16	Analogic

TABLE 88.5 Companies That Manufacture Plug-In Data Acquisition Boards

Analogic Corporation
360 Audubon Road
Wakefield, MA 01880
(508) 977-3000
http://www.analogic.com/

National Instruments Corporation
6504 Bridge Point Parkway
Austin, TX 78730
(512) 794-0100
http://www.natinst.com

Data Translation, Inc.
100 Locke Drive
Marlboro, MA 01752-1192
(508) 481-3700
http://www.datx.com

SenDEC, API Technologies Corp.
72 Perington Parrkway
Fairport, NY 14450
Tel: (585) 425-3390
http://www.sendec-cem.com

Keithley Metrabyte
440 Myles Standish Blvd.
Taunton MA 02780
(508) 880-3000
http://www.metrabyte.com

with drivers that interface to user-friendly software packages for creating easy-to-use yet custom-built computer instrumentation. Table 88.4 lists a few popular plug-in data acquisition boards and Table 88.5 lists how their manufacturers may be contacted.

References

1. C. E. Shannon, Communication in the presence of noise, *Proc. IRE*, 37(1): 10–21, 1949.
2. H. Nyquist, Certain topics in telegraph transmission theory, *AIEE Trans.*, 47: 617–644, April 1928.
3. A. B. Williams and F. J. Taylor, *Electronic Filter Design Handbook: LC, Active, and Digital Filters*, 2nd edn., New York: McGraw-Hill, 1988.
4. S. P. Lipshitz, R. A. Wannamaker, and J. Vanderkooy, Quantization and dither: A theoretical survey, *J. Audio Eng. Soc.*, 40: 355–375, 1992.
5. H. W. Ott, *Noise Reduction Techniques in Electronic Systems*, 2nd edn., New York: John Wiley & Sons, 1988.
6. R. Morrison, *Grounding and Shielding Techniques in Instrumentation*, 3rd edn., New York: John Wiley & Sons, 1986.
7. J. Williams, Designs for high performance voltage-to-frequency converters, Application Note 14, Linear Technology Corporation, March 1986.

Further Information

D. Dallet and J. M. da Silva, *Dynamic Characterization of Analogue-to-Digital Converters*, Springer, Kluwer Academic Publishers, Dordrecht, the Netherlands, 2005.

Y. Geerts, M. Steyaert, and W. M. C. Sansen, *Design of Multi-Bit Delta-Sigma A/D Converters*, Kluwer Academic Publishers, Dordrecht, the Netherlands, 2002.

M. J. M. Pelgrom, *Analog-to-Digital Conversion*, 2nd edn., Springer, New York, 2013.

R. van de Plassche, *CMOS Integrated Analog-to-Digital and Digital-to Analog Converter*, 2nd edn., Kluwer Academic Publishers, Dordrecht, the Netherlands, 2003.

89

Computers

A.M. MacLeod
University of Abertay Dundee

P.F. Martin
University of Abertay Dundee

W.A. Gillespie
University of Abertay Dundee

89.1 Introduction

Computers are an essential feature of most instrumentation systems because of their ability to supervise the collection of data and allow information to be processed, stored, and displayed. Many modern instruments are capable of providing a remote user with access to measurement information via standard computer networks.

89.2 Computer-Based Instrumentation Systems

The main features of a computer-based instrumentation system are shown in Figure 89.1. The actual implementation of such systems will depend on the application. Many commercially produced instruments such as spectrophotometers or digital storage oscilloscopes are themselves integrated computer-based measurement systems. These "stand-alone" instruments may be fitted with interfaces such as IEEE-488, RS-232, or USB to allow them to be controlled from personal computers (PCs) and also to support the transfer of data to the PC for further processing and display. Alternatively, the instrumentation system may be based around a PC/workstation or an industrial bus system such as a VME or Multibus, allowing the user the ability to customize the computer to suit the application by selecting

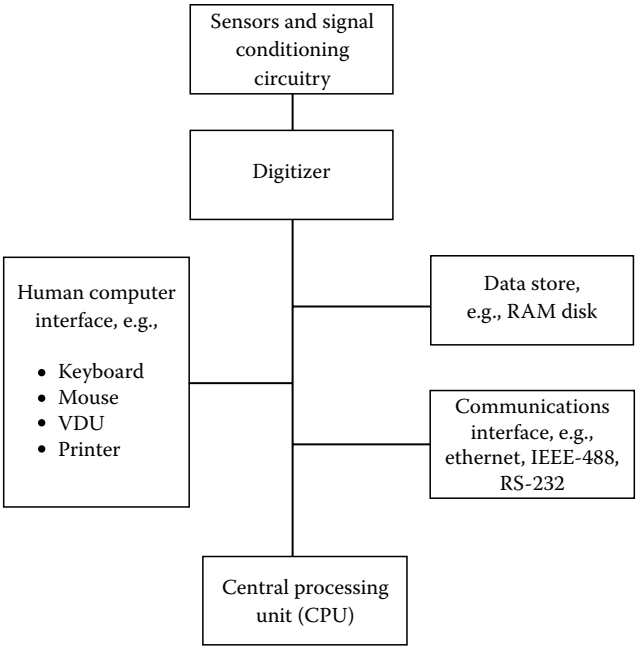


FIGURE 89.1 Elements of a computer-based instrumentation system.

an appropriate set of add-on cards. Recently, the introduction of laptop and notebook PCs fitted with PCMCIA interfaces with input/output capability has provided opportunities for the development of highly portable instrumentation systems.

89.2.1 Single-Board Computer

The simplest form of a computer is based around the single-board computer (SBC) which contains a microprocessor, memory, and interfaces for communicating with other electronic systems. The earliest form of PCs simply comprised an SBC together with a keyboard, display, disk drives, and a power supply unit. Today, the SBC still offers a solution for the addition of limited intelligence to instrumentation systems as well as forming an element of most computer systems, e.g., as the motherboard of a PC.

An overview of a simple SBC is shown in Figure 89.2. The microprocessor, which contains the central processor unit, is responsible for executing the computer program and for controlling the flow of data around the SBC. The random access memory (RAM) acts as a temporary (i.e., volatile) storage area for both program code and data. On the motherboard of a PC, the read-only memory (ROM) is largely used for storing the low-level code used to access the input/output hardware, e.g., the BIOS (basic input–output system). The operating system and applications software are loaded into RAM from the disk unit. In small, dedicated systems such as an oscilloscope the ROM may be used to store the program code to avoid the need for a disk drive. For small production runs or development work, erasable programmable ROM (EPROM) is used as an alternative to ROM, allowing the code to be upgraded without the high cost of commissioning a new ROM chip from the manufacturers.

Data are transferred around the SBC on its data bus, which will be typically 8, 16, or 32 bits wide (corresponding to the number of bits that can be transmitted at the same time). SBCs with 8 bit data buses are less complex and consequently of lower cost than 32 bit systems and may well be the optimum solution for those process control applications which require minimal processing of 8 bit data, e.g., from temperature and position sensors. However, the 32 bit bus width of most modern PCs and workstations is essential to ensure the fast operation of Windows-based applications software.

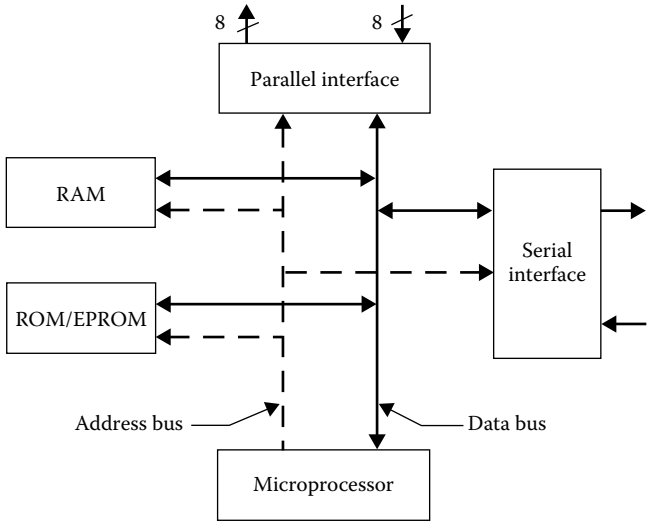


FIGURE 89.2 An overview of a single-board computer.

The address bus is used to identify the destination of the data on the data bus. Data transfer is usually between the microprocessor and the memory or interfaces. However, some SBCs support direct memory access (DMA) which allows data to be transferred directly between interfaces and memory without the need the information to pass through the processor. DMA is inherently faster than program-driven data transfer and is used for moving large blocks of data, e.g., loading programs from disk or the transfer of digitized video images.

SBCs are fitted with interfaces to allow them to communicate with other circuits. Interfaces carry out two main functions. First, they serve to ensure that the signals on the internal buses of the SBC are not affected by the connection of peripheral devices. Second, they ensure that signals can pass into and out of the computer and that appropriate voltage levels and current loading conditions are met. A well-designed interface should also provide adequate electrical protection for the SBC from transients introduced via the external connection. Parallel interfaces allow data to be transferred, usually 8 or 16 bits at a time, and contain registers that act as temporary data stores. Serial interfaces must carry out the conversion of data from the parallel format of the internal SBC data bus to and from the serial format used by the interface standard (e.g., RS-232, RS-422, or USB). Some SBCs may contain additional interfaces to support communication with a VDU, local area network (LAN), or a disk drive. Interfaces can range in complexity from a single parallel interface chip to a LAN controller which may require a significant fraction of the SBC board area.

89.2.2 Computer Bus Architectures

All but the simplest computer systems contain several circuit board cards which plug into a printed circuit board backplane. The cards will include at least one SBC and a number of "add-on" cards providing functions such as interfaces to peripherals (e.g., a LAN, graphics display, disk unit) or additional memory. The actual structure of bus architectures is quite variable but the simple model shown in Figure 89.3 contains the essential features. A group of tracks will carry the data and address information with a second group of tracks being used to control the flow of data and to ensure its reliable transfer. Other tracks are reserved for the signals which provide arbitration between SBC cards to ensure that only one such card has control of the bus at any given moment. There will also be tracks which provide utility functions such as clock signals and the transmission of event or interrupt signals.

The main advantage of bus-based systems is that they help one build a complex computerized system using standard cards. By conforming to agreed standard buses (see Table 89.1 for typical examples), the

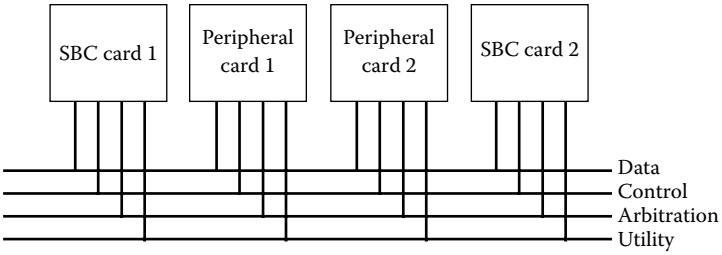


FIGURE 89.3 A simplified model of a computer bus.

TABLE 89.1 Typical Microprocessor Bus Standards

Bus Standard	STE	G96	VME	Multibus II
Data width (bits)	8	8/16	8/16/32	8/16/32
Max address (bytes)	1 M	32 M	4 G	4 G
Synchronous/asynchronous	Async	Async	Async	Sync
Connectors	64 pin	96 pin	96 pin	96 pin + 32 pin

system integrator can minimize problems of incompatibility and keep system costs to a minimum. The complexity of many bus systems is such that a significant fraction of the area of each card is dedicated to providing the logic required to interface to the bus, adding appreciably to the cost. In addition, data transfer between cards, even using DMA, is relatively slow compared with transfers between chips on the same SBC. There is, therefore, a trade-off between integrating functions on a single SBC and a lack of flexibility in customizing the system. An alternative to bus-based systems is to utilize processors such as transputers which are designed to be connected together directly into multiprocessor structures.

Most modern computerized systems adopt some form of distributed intelligence strategy in order to free the central processor unit to run the main application. Graphics controller cards generally employ a dedicated graphics processor able to handle the creation of the basic features such as line drawing or scaling the size of objects. LAN controller cards require intelligence to handle the network protocols and buffer the data flow to and from the network. Disk controller cards relieve the central processor of the need to handle the processes of reading and writing to the disk.

The use of a bus system with more than one SBC allows the designer to dedicate an SBC to the process of data acquisition and temporary data storage in order to ensure that the main SBC can concentrate on data processing and communication with the user. Such a strategy enables the data acquisition SBC to maintain a real-time response to making measurements while allowing the user access to processed data. In most systems, one of the SBCs acts as a system controller, which has ultimate control over the use of the bus and will normally contain the bus arbitration hardware. Only one SBC at a time can obtain the authority of the controller to act as a bus master and initiate data communication. The other cards will be configured as slaves allowing them to respond to requests for information from the master. In some systems, e.g., Multibus II, bus arbitration is distributed throughout the intelligent cards.

89.2.3 Industrial Computers

Many instrumentation systems utilize standard PCs or workstations either fitted with add-on cards or linked to intelligent instruments via standard protocols such as IEEE-488 or RS-232. In many industrial environments, there is a need to provide protection against hazards such as dust, damage by impact or vibration, and unauthorized access. The systems may have to fit more stringent electromagnetic compatibility (EMC) requirements. Industrial or ruggedized versions of desktop computers are available to meet this market. Cardframes designed to accommodate the standard industrial bus systems such as VME and Multibus can be readily customized to meet the demands of the industrial environment and

ruggedized versions of PCs are also available. In designing industrial computers, care must be paid to the specification of adequate power supplies, both to provide sufficient current for the add-on cards and also to provide protection against fluctuations in the mains supply. It may also be necessary in safety critical applications to use an uninterruptable power supply that will guarantee the operation of the system during short failures in the mains supply.

89.2.4 Software

All but the simplest computer systems require an operating system to support the operation of applications software. The operating system will allocate areas of memory for use by the applications programs and provide mechanisms to access system resources such as printers or displays. Some operating systems, such as MS DOS, are single-tasking systems, meaning that they will only support the operation of one program or task at a time. Instrumentation systems which simultaneously take measurements, process data, and allow the user to access information generally require a multitasking operating system, such as OS-9, UNIX, or Windows 95. In such instrumentation systems the applications software may well comprise a single program, but it may generate software processes or tasks each with their own local data and code. The multitasking operating system will allow the actual execution of these tasks to be scheduled and also provide mechanisms for communication of information between tasks and for the synchronization of tasks.

Instrumentation systems are real-time environments, i.e., their operation requires that tasks be carried out within specified time intervals; for example, the capture of data must occur at specified moments. Operating systems adopt a range of strategies for determining the scheduling of software tasks. Round-robin scheduling, for example, provides each task with execution time on a rota basis, with the amount of time being determined by the priority of the task. Operating systems which support preemptive scheduling allow high-priority tasks to become active if a specified event such as trigger signal occurs. The speed at which an operating system can switch from one task to another (i.e., context switch) is an important metric for real-time systems.

The extent to which an operating system can respond to a large number of events is extremely limited, as low-priority tasks will have a poor response time. If an instrumentation system needs to make a large number of measurements and also supports other activities, such as information display, the solution is generally to use a distributed system with dedicated SBCs operating as microcontrollers to capture the data. These microcontrollers will not usually require an operating system and will each operate a single program which will be either downloaded from the master SBC or located in ROM.

Commercial packages (e.g., see Table 89.2) are readily available to support PC/workstation-based instrumentation systems. A typical system will provide a Windows-type environment to control measurements and store, process, and display data and increasingly make use of the virtual instrument concept which allows the user to configure instrumentation systems from an on-screen menu. Most packages will allow stand-alone instruments to be controlled using interface cards supporting standards such as RS-232 or USB. Some packages also support the use of add-on cards for analog-to-digital conversion or a range of control functions such as channel switching and waveform generation; however, their main restriction is the limited range of hardware configurations supported by the software supplier. Each stand-alone instrument and add-on card requires a piece of code called a device driver so that the operating system can access the hardware resources of the card and instrument. Therefore, the development of device drivers requires an intimate knowledge of both the hardware and the operating system.

89.2.5 System Development

The software component of any computerized instrumentation system can form a significant percentage of its total cost. This is especially true of on–off systems which require software to be developed for a specific application, where the cost of the implementation of the software and its maintenance can

TABLE 89.2 Typical Data Acquisition and Display Software Packages for PC-Based Instrumentation Systems

Product	Platforms	Manufacturer
DTV	PC (Windows 3.1 or 95)	Data Translation Inc. 101 Locke Drive Marlboro, MA 01752-1192 Tel: (USA and Canada) (800) 525-8528 (worldwide) + 1 508-481-3700
Labtech Notebook/ Notebook Pro	PC (DOS, Windows 3.1 or 95)	Laboratory Technologies Corporation 400 Research Drive Wilmington, MA 01887 Tel: (USA and Canada) 800-8799-5228 (worldwide) +1 508-658-9972
Lab VIEW	PC, Mac, Sun, Power PC	National Instruments 6504 Bridge Point Parkway Austin, TX 78730-5039 Tel: +1 512-794-0100
Lab Windows	PC (DOS)	National Instruments 6504 Bridge Point Parkway Austin, TX 78730-5039 Tel: +1 512-794-0100
Lab Windows/ CVI	PC (Windows 3.1 or 95) Sun	National Instruments 6504 Bridge Point Parkway Austin, TX 78730-5039 Tel: +1 512-794-0100
ORIGIN	PC (Windows 3.1 or 95)	MicroCal Software Inc. One Roundhouse Plaza Northampton, MA 01060 Tel: +1 413-586-2013

considerably exceed the hardware costs. In such circumstances, the ability to reuse existing code and the access to powerful development systems and debugging tools are of crucial importance. Some debug software only provides support for stepping through the operation of code written at the assembly code level. When developing device drivers, for example, access to a source-level debugger which can link the execution of code written in a high-level language to the contents of processor registers is useful.

Many SBCs that are intended for operation as stand-alone microcontrollers, instead of in bus-based systems, are often supported by a high-level language such as C or FORTH and a development environment. A common practice is to utilize a PC to develop the code which can be downloaded via a serial link to the microcontroller. Some debugging support for the execution of the code on the microcontroller is also provided. The development of powerful compact Intel 486 or Pentium microcontroller cards running operating systems such as MS DOS can greatly reduce software development time because of the ready access to PC-based software. The implementation of systems based on dedicated microcontrollers may require the use of a logic analyzer to view data on system buses.

89.3 Computer Buses

89.3.1 VMEbus (IEEE P1014)

The VMEbus [1–3] utilizes a backplane fitted with two 96-pin connectors (called P1 and P2). The P1 connector provides access to all the bus signals with the exception of bits 16–31 of the data bus and bits 24–31 of the address bus, which are provided on the P2 connector. The location of cards in a VME system is important, as slot 1 must contain the system controller and the priority of the SBCs is determined by their proximity to the system controller SBC. As only 32 of the pins on the P2 connector are defined,

the remaining 64 tracks on the P2 backplane may be specified by the user. VME cards may be single height (fitted with a P1 connector only) or double height (fitted with both P1 and P2 connectors). Single-height boards are 100 mm high and 160 mm deep and fit 3U height cardframes. Double-height boards are 233.35 mm high and 160 mm deep and fit 6U height cardframes.

89.3.2 VMEbus Signals

The VMEbus can be described as having four distinct groups of signals.

89.3.2.1 Data Transfer Bus

The Data Transfer Bus (DTB) provides 32 bit data transfer with 32 bit address information using a non-multiplexed asynchronous approach. Transfer is supported by a simple handshake mechanism with the bus master initiating the transfer providing an address strobe (AS) signal to tell the destination card to read the address information. The master also controls two data strobe (DS1 and DS2) lines which in a write operation indicate the presence of valid data on the data bus and in a read operation that the master is ready to receive data from the slave. The slave card uses the data acknowledge signal (DTACK) in a read operation to indicate that it has placed valid data on the bus and in a write operation to confirm that it has read the data. There is clearly a danger that the bus will hang up if the slave does not respond to a request for a data transfer, so in most VME systems a watchdog timer is provided to produce a signal on the bus error (BERR) line if a DTACK signal is not detected within the time-out period.

Revision C of the VME bus specification will support 8, 16, and 32 bit data transfers, whereas revision D supports 64 bit transfers by making use of the address bus to transfer the additional 32 bits. Data may also be transferred in blocks of up to 256 bytes without the need to retransmit the address information. Most SBCs have both master and slave interfaces allowing their memory to be accessed from other SBCs in the system. The VME bus also supports three types of addressing mode, short (16 bits), standard (24 bits), and extended (32 bits), allowing the memory decoding circuitry of cards to be minimized for small systems or for input/output cards. Six address modifier lines (AM0–AM5) are used to indicate the addressing mode and also to give information about the mode of transfer (e.g., block transfer and whether the transfer is in privileged, supervisor, or nonprivileged mode).

89.3.2.2 Arbitration Bus

The arbitration bus provides a mechanism for deciding which master is allowed to gain control of the bus. The arbitration bus provides four bus request lines (BR0–BR3) and four bus grant lines. The SBC in slot 1 of the VME bus acts as a bus controller which provides arbitration using a number of scheduling algorithms. In a simple priority-based algorithm each of the bus request lines is assigned a priority, whereas in a round-robin system access to the bus cycles around the bus request lines. Each of the four bus grant lines is in fact allocated two pins on the P1 connector (see Figure 89.4). The bus grant signal therefore passes through each card (called daisy chaining); thus in Figure 89.4, the signal enters via BG1IN and exits via BG1OUT. This scheme allows a card to intercept the bus grant signal, giving cards nearer the system controller a higher priority. A consequence of daisy chaining is that any unused card slots must be fitted with jumpers between the bus grant pins.

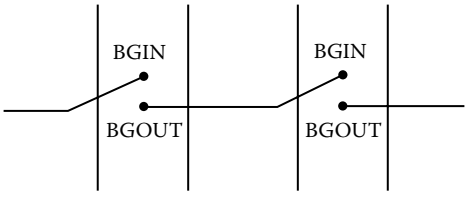


FIGURE 89.4 VMEbus arbitration daisy chain.

89.3.2.3 Priority Interrupt Bus

The priority interrupt bus provides seven interrupt request lines (IRQ1 to IRQ7) with IRQ7 having the highest priority. The interrupt requests may be handled by several SBCs, provided that all interrupts of the same level are handled by the same master. The interrupt acknowledge line is daisy chained through the cards in a similar way to the bus grant lines (the pins are called IACKIN and IACKOUT). Again, any unused card slots must be fitted with jumpers between these pins.

89.3.2.4 Utility Bus

The utility bus provides the power supply tracks, the system reset, a 16 MHz clock, and two system failure lines, SYSFAIL, which allows a card to indicate that it has suffered a failure, and ACFAIL, which is generated by the power supply monitor. Both failure lines are handled by the system controller SBC.

Associated with the VMEbus is VXI (VMEbus extensions for instrumentation). This blends the IEEE 1014 VMEbus standard with the IEEE-488 instrumentation bus using the uncommitted pins on the P2 connector to form the IEEE P1155 standard.

89.3.3 Multibus II (IEEE 1296)

Multibus II [3] is a synchronous bus system which is implemented on a single 96-track backplane. Cards conform to the Eurocard format of 220 × 233.65 mm; however, the bus requires only a single 96-pin DIN 41612-type connector. A second connector may be used to support a local bus or merely provide mechanical stability for a double-height card.

Unlike VME, the Multibus II standard has the following features:

1. A synchronous bus, i.e., data transfer is linked to the bus clock rather than the strobe/acknowledge control lines of the asynchronous VME system
2. A 32 bit multiplexed address and data bus
3. A message-passing strategy as an alternative to interrupt request lines
4. Distributed bus arbitration rather than a central system controller (each card contains a message-passing controller [MPC] chip which participates in implementing the message-passing algorithm)
5. Only intelligent cards (i.e., fitted with microprocessors) that can interface to the main bus; simple input/output (I/O) cards should make use of a local bus (e.g., the iLBX bus) using the second DIN connector

89.3.4 Multibus Signals

Multibus II may be described as having five distinct groups of signals.

89.3.4.1 Central Control

The central services module (CSM) in slot 0 is responsible for the generation of this group of signals. The CSM may be implemented on a dedicated card or incorporated onto a SBC. The CSM produces the 10 MHz bus clock as well as a range of reset signals to support both a cold and warm start as well as recovery from power supply failure.

89.3.4.2 Address/Data

Multibus II supports a 32 bit multiplexed address/data bus (AD0–AD31) with four parity lines (PAR0–PAR3) providing parity information for each data byte. As in all synchronous systems, data are only sampled on an edge of the system clock, a feature that enhances noise immunity.

89.3.4.3 System Control

There are 10 system control lines (SC0–SC9) which provide basic handshake information to assist data transfer, e.g., requester is ready or replier is not ready. In addition, these lines are used to convey information such as data width, data error indication, and whether the bus is in the request or reply phase.

89.3.4.4 Exception Signals

Two exception signals, bus error (BUSERR) and time-out (TIMOUT), are provided. Any card detecting a data integrity problem must report this to all other cards using BUSERR. The CSM generates a TIMOUT when it detects a data communications hang up on the bus.

89.3.4.5 Arbitration Group

The arbitration signals, which determine which card gains control of the bus, consist of a single bus request line (BREQ) and six Bus Arbitration lines (ARB0–ARB5). To request exclusive use of the bus, a card asserts the BREQ line and provides it arbitration ID on ARB0–ARB4. It also uses ARB5 to indicate whether the request has high priority or whether "fairness mode" is acceptable. In this latter mode, the card will not make another request until after all other requesters have used the bus. Each card has the same bus arbitration logic within its MPC chip. If several cards make high-priority requests, the order of access is determined by the numerical value of the arbitration ID.

89.3.5 Message Passing on Multibus II

Multibus II uses message passing to implement block data transfers and interrupt requests. Each MPC chip contains a first-in first-out (FIFO) buffer, which ensures that the full bandwidth of the bus can be utilized by storing data immediately before or after transfer. In "solicited" transfers, the MPCs cooperate by warning each other that a message is to be sent. These messages may be sent as 32 byte packets in a manner that is analogous to the block transfer mechanism in VME. "Unsolicited" packets not expected by the receiving MPC are used to set up a block transfer or to act as the equivalent of interrupt request signals.

89.3.6 System Configuration

Multibus II employs a software configuration approach in which information such as the base memory address and arbitration ID are sent down the bus rather than by the use of jumpers or dip switches. Some information such as card type and serial number are coded on the cards themselves.

89.3.7 Other System Buses for Small-Scale Systems

The G64/G96 and STE standards [3] are examples of buses well suited to small industrial real-time instrumentation and control systems because of their relatively simple and hence lower-cost bus interfaces, compared with VME and Multibus. Both buses support DMA, have multiprocessor bus arbitration, and use single-height Eurocard (100 × 160 mm) cards. Prototyping cards fitted with bus interfaces are readily available and may be used to develop custom-designed I/O cards. Power is supplied on the buses at +5 and ±12 V. In addition, the real-time multitasking operating system OS-9 has been ported onto SBCs which operate with these buses. Development systems that support the use of MS-DOS are also available, thus providing access to a wide range of PC-based software especially for graphics applications.

The G64 bus, which was defined by the Swiss company Gespac in 1979, specifies 64 bus lines which are mapped to rows of a DIN41612 connector. The bus has a 16 bit nonmultiplexed data bus and a 16 bit address bus; 32 bit transfers may be achieved by multiplexing the upper 16 bits of the data bus with the address bus. The G96 standard adds a further 32 bus lines by making use of the third row of a DIN41612 connector to extend the address bus width to 24 bits and provide additional interrupt request lines. The STE bus provides an unmultiplexed 8 bit data and 20 bit address bus using 64-pin DIN41612 connectors.

89.4 PC Buses

There are three main bus standards for PCs—industry standard architecture (ISA), extended ISA (EISA), and the microchannel architecture (MCA). In addition, the Personal Computer Memory Card International Association (PCMCIA) architecture has been developed primarily for use in laptop and notebook computers.

ISA and EISA are pin-compatible and are both synchronous buses with a clock rate of 8 MHz regardless of the clock rate of the main processor, whereas the MCA bus is an asynchronous bus. Slow slave add-on cards can utilize the ISA and EISA buses by using an appropriate number of wait states. The MCA architecture will not be covered in this chapter. Further details of these bus architectures are given in Refs. [4–6].

89.4.1 ISA Bus

The original IBM PC and its clones used the standard PC bus which supported 8 data bus and 20 address bus lines and employed a 62-pin printed circuit card edge connector. When IBM introduced the PC-AT, a second 36-pin connector was added to the motherboard backplane slots to provide a 16 bit data bus and increase the number of address lines to 24. This new bus subsequently became known as the ISA bus and is capable of supporting both cards designed for the original PC bus, and cards with two connectors providing full 16 bit data access. The bus master line allows a card to take over control of the bus; however, this is generally only suited to long-term takeovers. The ISA bus supports 8 and 16 bit DMA transfers allowing efficient transfer of blocks of data, e.g., while loading software from disk into RAM. The ISA bus supports the I/O addressing mode of the Intel 80 × 86 range of processors with an I/O address space of 768 locations in addition to the 256 locations reserved for the motherboard.

89.4.2 EISA Bus

The EISA bus provides full 32 bit data and 32 bit address bus lines. ISA cards can fit into EISA bus connectors; however, cards designed for the EISA standard have bilevel edge connectors, providing double the number of contacts as an ISA card. The presence of a notch in the EISA cards allows them to be inserted farther into the motherboard socket than the ISA cards and thus mates with the additional contacts. The EISA standard increases the maximum ISA data width in DMA transfer to 32 bits and also provides a much-enhanced bus master. The following features of EISA are worthy of note.

89.4.3 Bus Arbitration

All EISA systems have a central arbitration control (CAC) device on the motherboard. The CAC uses a multilevel rotating priority arbitration scheme. The top-priority level rotates around three customers, the DMA controller, the dynamic memory refresh controller, and, alternately, either the main CPU or one of the bus master cards. A device that does not make a request is skipped over in the rotation process. The bus masters take it in turns to gain access to the top-priority level. Whereas ISA supported a single bus request line, the EISA standard provides a dedicated pair of request lines (MREQO to 14) and acknowledges lines (MAKO to 14) for each bus master. (*Note*: Although this allows 15 bus masters to be used, in many systems the chip set which implements the CAC supports bus masters in a limited number of the EISA sockets.) The CAC supports preemption, i.e., it allows a device making a request to capture control from another device if it is next in turn. A bus master card must release the bus within 64 bus clock cycles, whereas a DMA controller has 32 clock cycles to surrender the bus.

89.4.4 Input/Output

The EISA bus provides 768 additional I/O locations for each slot in addition to the 768 ISA I/O locations that may be accessed from any slot. EISA cards contain nonvolatile memory to store configuration information (see later), and a minimum of 340 bytes of the I/O address space of each slot is reserved for this purpose.

89.4.5 System Configuration

The EISA system provides support for automatic system configuration to replace the use of jumpers and dip switches to specify parameters such as the base memory address, interrupt request line number, or DMA channel number used by each add-on card. Information on the product type and manufacturer are stored on each EISA add-on card is read by the CPU during system start-up, making it possible to identify the slots that are fitted with full EISA cards. Manufacturers of both ISA and EISA cards should supply a configuration file containing information on the programmable options that are available. When a system configuration program is run, an optimum configuration of the boards will be determined, and the configuration information written into the I/O space of each EISA card. For ISA cards, the user can be informed of the required jumper settings.

89.4.6 PCMCIA

The PCMCIA architecture was developed by the PCMCIA and the Japan Electronics Industry Development Association for removable add-on cards for laptop and notebook computers. Each PCMCIA socket has its own host bus adapter (HBA), which acts as an interface to the main computer bus. Cards may be plugged into PCMCIA sockets either before or after the computer has been powered up. There are three types of PC card all with the same planar dimensions (54.00 × 85.6 mm) but with differing thicknesses, namely, 3.3 mm for type I, 5.0 mm for type II, and 10.5 mm for type III. Cards and sockets are keyed to prevent them from being inserted the wrong way round. The PCMCIA standard supports cards that operate at several possible voltage levels, i.e., 5.0 V cards, 3.3 V cards, or dual voltage 5.0 V/3.3 V cards.

89.4.7 Configuration

PC cards contain configuration information called the card information structure (CIS) which is stored in nonvolatile memory. Cards may be configured either on system power-up or on insertion of the card into the socket (i.e., plug and play). Configuration is carried out using a form of device driver called an enabler. The enabler may make use of two additional software services called card services and socket services. Socket services, which may be contained in ROM on the PC or loaded from disk during power-up, provide function calls to allow the HBA to be configured to cooperates' computers. Card services, which may be an extension to the operating system or an installable device driver, act as a server for the enabler, which performs as a client. Card services provide a range of functions such as accessing the CIS of the card, requesting system resources required by the card, and telling the enabler that a card has been inserted or removed from the socket. Enablers may be classified as dedicated to one particular card or generic, i.e., designed for a range of cards. Note that early PCMCIA cards were not designed for use with card services.

89.4.8 PCMCIA Socket Interface

The PCMCIA socket comprises a 68-pin connector with 26 address lines providing access to 64 MB of memory space and a 16 bit data bus. Two V_{cc} pins and four ground pins are supplied. The maximum current that can be supplied to the card is 1.0 A with a maximum of 0.5 A from each of the two power supply pins. Release 2.x sockets apply 5.0 V to the V_{cc} pins on power-up and reduce this to 3.3 V if the card has dual-voltage capability. Cards that operate at 3.3 V only are keyed so they cannot fit into this type of socket and only into low-voltage sockets. The supply voltage provided by low-voltage sockets depends on

the logic state of its voltage sense inputs. A PCMCIA socket can be configured either as a *memory-only* socket or as a *memory or I/O socket*. Initially, the socket acts as a memory-only socket but is converted by the enabler to a memory or I/O socket if the card is required to support I/O functions. In this mode, the card can generate an interrupt request via a single IRQ pin and support both 8 and 16 bit I/O data transfers. DMA may be supported but not by Release 2.x systems.

89.4.9 PC/104

The enormous popularity of PC architecture resulted in its use in embedded systems. A need then arose for a more compact implementation of the ISA bus that could accommodate the reduced space and power requirements of embedded applications. The PC/104 specification (1992) was adopted as the base

TABLE 89.3 Manufacturers/Suppliers of Bus-Based Systems and Industrial PCs

System	Manufacturer/Supplier	System	Manufacturer/Supplier
VME	Wordsworth Technology Ltd. STE 6 Enterprise Way Edenbridge, Kent TN8 6HF UK Tel: +44 (0) 1732 866988	STE	Arcom Control Systems Units 8-10 Clifton Road Cambridge CB1 4WH, UK Tel: +44 (0) 1223 411200
	PEP Modular Computers, Inc. 750 Holiday Drive, Building 9 Pittsburg, PA 15220 Tel: (412) 921-3322	G64	Gespac SA 18 Chemin des Aulx 1228 Geneva, Switzerland Tel: +41 (22) 794 34 00
	BVM Ltd., Hobb Lane Hedge End Southampton, SO30 0GH, UK Tel: +44 (0) 1489 780144		Altek Microcomponents Ltd. Lifetrend House Heyes Lane Alderley Edge, Cheshire SK9 7LW Tel: +44 (0) 1625 584804
	Motorola, Inc. Computer Group 2900 S Diablo Way Tempe, AZ 85282 Tel: (800) 759-1107	Industrial PC	Blue Chip Technology Ltd. Chowley Oak, Tattenhall Chester, Cheshire CH3 9EX, UK Tel: +44 (0) 1829 772 000
VXI	National Instruments 6504 Bridge Point Parkway Austin, TX 78730-5039 Tel: (512) 794-0100		Capax Industrial PC Systems Ltd. Airport House, Purley Way Croydon, Surrey, CR0 0XZ, UK Tel: +44 (0) 181 667 9000
Multibus	Tadpole Technology, Inc. 2001 Gateway Place, Suite 550 West, San Jose, CA 95110 Tel: (408) 441-7920	PC/104	Computer Boards, Inc. 125 High Street Mansfield, MA 02048 Tel: (508) 261-1123
	Intel Corp. 3065 Bowers Avenue Santa Clara, CA 95051		Diamond Point International (Europe) Ltd. Unit 9, North Point Business Estate, Enterprise Close Rochester, Kent ME2 4LY, UK Tel: +44 (0) 1634 718100
	Syntel Microsystems Queens Mill Road Huddersfield, HD1 3PG, UK Tel: +44 (0) 1484 535101		

for an IEEE draft standard called the P996.1 Standard for Compact Embedded PC Modules. The key features of the PC/104 are

- Size reduced to 90 × 96 mm (3.550 × 3.775 in.)
- Self-stacking bus allowing modules to be "piggy-backed" and eliminating backplanes or cardframes
- Rugged 64-contact and 40-contact male and female headers replacing the PC edge connectors (64 + 40 = 104, hence PC/104)
- Lower power consumption (<2 W per module)

PC/104 CPU modules range from a basic 9.6 MHz, 8088 compatible XT with one serial port and a keyboard connector to a 100 MHz, 80486DX4 with four serial ports, parallel port, IDE disk controller, display controller, Ethernet adapter, keyboard port, and up to 64 MB of on-board RAM. Pentium-based systems are also available. Systems can be customized from a wide range of modules, including data acquisition boards, solid-state disk modules, and LAN support, all in the same 3.6 by 3.8 in. stackable format.

The wide availability of software development tools for the PC, the large number of software developers familiar with the PC environment, and the ease of transferring software developed on a conventional PC to the PC/104 make this an increasingly popular format.

Table 89.3 lists some manufacturers and suppliers of bus-based systems and industrial computers.

89.5 Peripherals

Computer peripherals fall conveniently into two categories. The first category may be considered to be internal to the computer system and comprises cards plugged directly into a computer bus slot. The second category comprises instruments external to the computer but controlled by it. These external instruments are usually themselves "intelligent," being controlled by their own CPU, and are linked to the main computer by a serial (RS-232 or USB) line or the IEEE-488 bus (GPIB). Such instruments may normally be operated in a stand-alone mode in response to their front panel controls without the requirement of an external computer.

89.5.1 Internal Cards

Internal cards are available to perform a wide range of functions. Table 89.4 provides a brief list of representative types. Note that both the serial interface and the IEEE-488 adapter required for the control

TABLE 89.4 Typical Internal Peripherals

Card Type	Facilities Offered
Display adapter	Provides video and graphics facilities
Serial communications adapter	Serial communication to a similar device using the RS232, RS422, or RS485 standard
IEEE-488 adapter	Communication with intelligent instruments using the IEEE-488 bus (GPIB)
Digital input/output (I/O)	Individual I/O lines, normally grouped as an 8 bit byte, which may be used to provide logic high and low signals (output) or sense logic high or low signals (input); inputs and outputs may be optically isolated; outputs may drive relays
Counter/timer	Hardware counter timer allowing external pulses to be counted, digital waveforms generated or pulse widths measured; counter/timers are often available as an additional facility on digital I/O cards
Analog input	Converts an analog input voltage to an integer value which may be read by the computer; resolution typically 8, 10, 12, or 16 bits; input voltage range may be fixed or may be user selectable; conversion times vary from several seconds to tens of nanoseconds
Analog output	Produces an analog output voltage proportional to a digital input; resolution typically be 8, 10, 12, or 16 bits

of external "intelligent" peripherals will be fitted as internal cards. Almost all internal peripherals need to be configured before use to set up such parameters as the base address, the interrupts, and/or DMA channels used. This may involve setting jumpers or switches on the card or may be accomplished under software control using a configuration file supplied by the manufacturer. To operate the cards, bytes are written or read from appropriate addresses on the cards by the controlling SBC. The mechanism for doing this is discussed further in the section concerned with software for data acquisition. Further detail on some of the card types is given later.

89.5.2 Display Adapter

While graphics support is normally available as standard on a PC-based system, this is not the case with other bus-based systems such as VME or Multibus. These systems are provided with a serial port that may be connected to a terminal to provide a text-only dialog with the operating system running on the SBC. In such circumstances the choice of display adapter will be determined by the user requirements, taking into account the support for the device provided by any software packages that are to be used. If the user intends to write custom graphics software, it is essential to ensure that a graphics library is available from the vendor providing as a minimum line drawing, arc drawing, and block color fill facilities.

89.5.3 IEEE-488 Adapter

This device provides support for communications across the IEEE-488 bus or GPIB (general-purpose interface bus). The bus itself comprises eight data lines, five interface management lines, and three handshake lines. Transfers are parallel, synchronous, and at rates up to 1.5 MB/s. The IEEE-488 bus and its applications are discussed further in the section on external instruments.

89.5.4 Serial Communications Adapter

This device provides support for serial communications using the RS-232, RS-422, or RS-485 standards. It is used to provide a text-based terminal for systems based on Multibus or VME and to communicate with "intelligent" instruments such as position controllers, multimeters, or storage oscilloscopes fitted with similar interfaces.

Serial adapters convert parallel data to and from a bit stream in which each data byte (5, 7, or 8 bits) is framed by a start bit, an optional parity bit, and one or more stop bits. The bit stream may be sent via a circuit consisting of only two wires at rates of up to 115,200 bits per second (commonly referred to as 115,200 baud). Common bit rates are 300, 600, 1200, 2400, 4800, 9600, 19,200, 28,800, 38,400, 57,600, 115,200 baud. Both the transmitting and receiving adapters must be configured, normally by software, for the same baud rate, number of data bits, stop bits, and parity. Some form of flow control to prevent a receiver from being overloaded with incoming data is essential. This is accomplished either by separate handshake lines (e.g., those denoted by RTS and CTS in the standard) or by software where the receiver sends a special control byte (XOFF) back to the transmitter telling it to stop sending until it receives a second control byte (XON) to reenable it. For historical reasons, the RS-232 standard does not define a bidirectional handshake procedure, and manufacturers have been forced to implement their own schemes which are not always compatible with each other.

Serial communication between devices may be

Full duplex, where either device may transmit or receive data at any time
Half duplex, where both devices are capable of transmission or reception but only one may transmit at any instant

Simplex, where one device is a transmitter, one is a receiver, and data can only flow in a single direction

RS-232, developed by the Electronics Industries Association (EIA), is the oldest standard, originally developed in the early 1960s, to allow mainframe computers to communicate with terminals via modems

and telephone lines. This is the origin of the names of some of the connections (e.g., RI ring indicator, DCD data carrier detect), which have no relevance in the applications considered here. A related problem is that the standard expects that the devices being connected are data terminal equipment (DTE), at one end of the link, and data communication equipment (DCE) at the other. Computers and terminals are DTE, while modems are DCE. The most commonly used revision of the standard, RS-232-C (revisions D and E also exist), was made in 1969 and is still widely used. Serial communication is made using voltage levels in the region of ±12 V, over distances up to 15 m (50 ft) at speeds up to 20,000 baud.

RS-422, also developed by the EIA, is an enhancement of the RS-232 standard. Differential transmitters and receivers are employed which allow one transmitter to drive up to 10 receivers, using a twisted-pair connection for each circuit, at bit rates up to 10 Mbaud at distances up to 12 m (40 ft) or 100 kbaud at distances up to 1200 m (4000 ft). The RTS and CTS lines (defined in the standard) are used for flow control, while the RXD and TXD lines are used to transmit and receive data. Thus, a two-twisted-pair cable is required for duplex connection without hardware handshaking. A four-twisted-pair cable is required if hardware handshaking is used.

RS-485 is based on RS-422 and allows up to 32 driver/receiver pairs to be connected to a common data bus (two-twisted pairs). Clearly, only one device can be allowed to transmit at any one time. The RTS circuit is used to disable the other transmitters connected to the bus if a device is required to transmit data. Handshaking is performed using software.

The serial interfaces on instruments are usually configured as DTE devices. We are faced with the problem of connecting one DTE device (the computer serial interface) to another (the instrument), which is not what the RS-232 standard was designed for. Furthermore, since the standard does not define a bidirectional handshake to control data flow, several incompatible handshaking schemes exist. A comprehensive survey of these is presented in Ref. [7]. A common solution to the DTE to DTE connection problem is the so-called *null modem*, which nothing more than a specially wired cable. Figure 89.5 shows two such connection schemes. One requires the software handshaking procedure and the other implements a bidirectional hardware handshake. The reader should refer to Ref. [7] for details of other schemes and for the definitions of the mnemonics used to label the connections.

Some common problems encountered in practice are

1. The received data are garbage. This is almost always due to the baud rate, parity, and number of stop bits not being the same at both ends of the link.
2. Data initially correct, but parts in the middle are missing. This is probably a handshake problem. The transmitting device is sending data faster than the receiver can process it.
3. No communications at all. Probably a handshake problem where the transmitter does not sense that the receiver is ready.

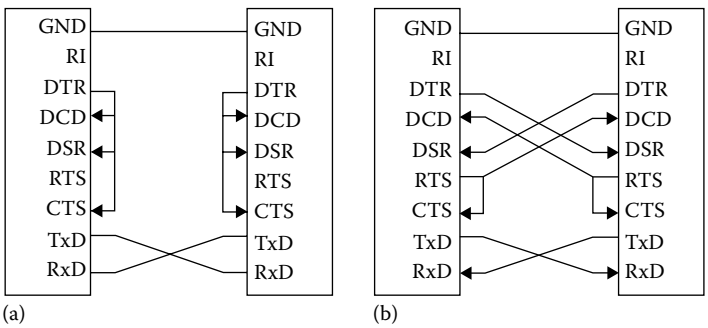


FIGURE 89.5 Two null modems for connecting DTE to DTE. In (a), all handshaking must be in software. The DTR line "fools" the serial interface that it is connected to the handshake lines of another device. Scheme (b) implements a hardware handshake. The DTR–DSR connection shows each device that the other is present. The RTS lines, connected to the DCD of the other device, which it can monitor, are used to control the flow of data in either direction.

89.5.5 Digital Input and Output

These cards provide I/O lines, normally in groups of eight, which may be used to sense or generate digital signals for devices outside the computer. A group of eight input lines is referred to as an (8 bit) input port and a group of eight output lines as an (8 bit) output port. Input and output levels vary from card to card and it is best to consult the appropriate data sheet. Typically, voltages between 2.5 and 5.0 V are considered as high logic levels, whereas voltages between 0.0 and 0.5 are considered as low logic levels. These levels are sometimes referred to loosely as transistor–transistor logic (TTL) levels. Note that the actual logic levels used by the various TTL families differ from these slightly. The "high" and "low" ranges may be slightly different for input and output lines. Output lines often have limited current sourcing and sinking abilities compared with TTL, and it is therefore often necessary to buffer them. It is important that voltages exceeding the maximum rated values do not appear on inputs or outputs (e.g., attempting to switch an inductive load might produce a dangerously high transient voltage at an output); otherwise, the device may be damaged. Where this is likely to be a problem, inputs and outputs should be suitably buffered or even optically isolated, which provides protection up to a few kilovolts. Outputs may also drive appropriately connected relays. I/O cards with these facilities on board are readily available.

As a minimum, I/O card may be expected to support a control register, two I/O ports each with an associated data register, and some handshake lines. Handshake lines may sometimes be used to generate interrupts on the controlling SBC. A byte written to the control register is used to configure the I/O ports, i.e., to determine if they are to behave as input or output ports as well as to select the function, if any, of the handshake lines. It may not be possible to select the direction of optically isolated or buffered ports. Writing a byte to an output port causes a pattern of high and low voltages to appear on the lines reflecting the pattern of zeros and ones in the binary representation of the byte written. Similarly, when a byte is read from an input port, the number read is specified in binary representation by the pattern of high and low logic levels on the input lines. The following example illustrates this.

The Intel 8255 Programmable Peripheral Interface (PPI) is commonly used in digital I/O cards for the PC. Data for this device are readily available [8]. The 8255 provides three ports, denoted by A, B, and C, and a control register. Ports A and B may be designated as an 8 bit input port or an 8 bit output port. Port C may be considered as two independent 4 bit ports, which may be chosen independently as input or output. Port C may also provide handshake functions. There are three modes in which the chip may operate. The simplest case, mode 0, provides basic input and output without automatic handshaking, is used in our example. Note that the 8 bits of an I/O port are conventionally labeled bits 0–7. This is because bit 0 is weighted 2^0, bit 1 weighted 2^1, etc., in the binary representation of the number read from or written to the port.

To configure the PPI to operate in mode 0 with port A as an input port and port B as an output port, the bit pattern 10011001 must be written to the control register. This number is equivalent to 99 in hexadecimal or 153 in decimal. Suppose now that switches connected to port A hold bits 2 and 7 high and the remaining bits low. The resulting binary pattern will be 10000100, which is equivalent to hexadecimal 84 or decimal 132. When port A is read, the number 132 (decimal) will therefore be obtained. To hold the lines connected to bits 2, 3, and 5 of port B high while leaving the remaining lines low, we see that the binary pattern 00101100 must appear at port B. 00101100 binary is equivalent to 2C hexadecimal or 44 decimal. We therefore write the decimal number 44 to port B.

89.5.6 Counter/Timers

These devices typically provide software programmable event counting, pulse, and frequency measurement. As output devices, they may generate a single pulse (one shot) when a programmable number of input pulses have been counted and produce square waves of arbitrary frequency and complex duty cycles. Frequencies generated are normally based on an on-board crystal clock to provide independence

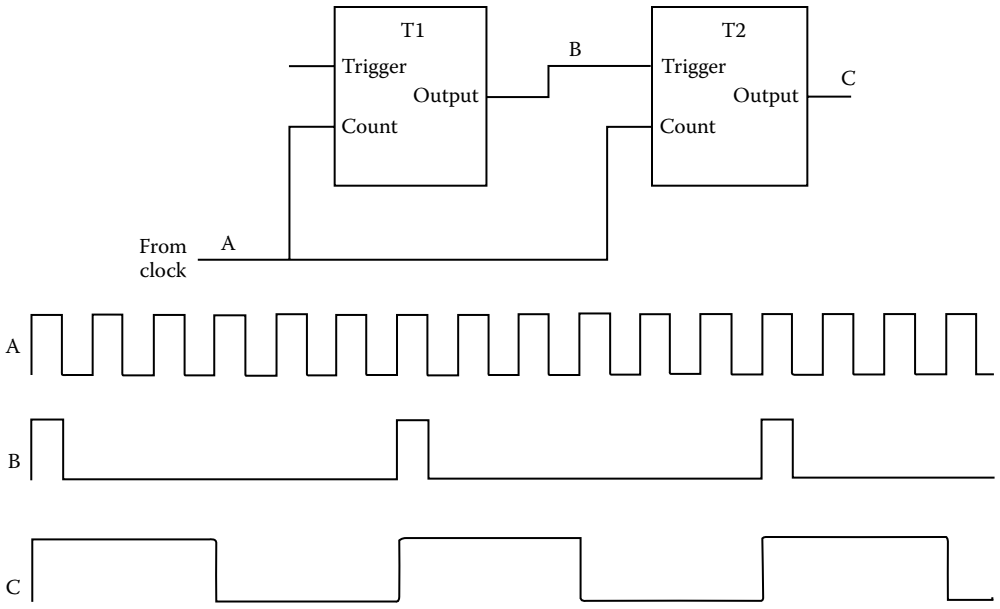


FIGURE 89.6 Using two timers to produce pulse-width modulation under software control. Both timers receive input pulses at a constant frequency from an external clock, as shown at A. Timer T1 operates in continuous mode. The trigger has no effect in this mode. The output of T1 is a single positive-going pulse when it has counted the specified (by software) number of input pulses, as shown at B. Timer T2 operates in one-shot mode. Each time it receives a trigger pulse, its output goes high for a specified (again by software) number of counts, as shown at C. In this way, the frequency of the output at C is controlled by the count specified for T1 and the width of the positive-going part of C is controlled by the count specified for T2.

from the internal clock speed of the computer. Counter/timer cards commonly support at least three independent 16 bit counters.

Common applications include

1. Alarms. The counter is in one-shot mode and generates a single pulse on time-out. This is connected to interrupt the computer and alert the user in the middle of the currently executing task.
2. Watchdog timer. This is used to detect problems, particularly in systems which are intended to operate without operator intervention. It is similar to the Alarm described earlier except that the interrupt is used to reset the computer. In normal operation, this will never occur as all software tasks executing are designed to update the counter constantly so that it never reaches its terminal count. Only if a problem develops, e.g., a software "crash," will the counter time-out and the system be reset.
3. The generation of complex waveforms, e.g., for pulse-width modulation. This application uses two counters in cascade, one (T1) to provide regular pulses at the carrier frequency triggering another (T2), in one-shot mode, to provide the variable duty cycle as shown in Figure 89.6.

89.5.7 Analog Input

An analog input card uses an analog-to-digital converter (ADC) that accepts an input voltage and supplies an integer proportional to that voltage to the computer. Many cards now are produced with on-board signal conditioning circuits that provide for variable gain either by means of switches or under program control. Cards with specialized signal conditioning circuits for common applications such as thermocouple linearization or interfacing to strain gages and other bridge sensors are available. Signal conditioning to protect cards destined to be used in hostile electrical environments is also available.

TABLE 89.5 Common ADC Parameters

Parameter	Description
Resolution	The smallest change in input detectable in the digital output; resolutions are commonly expressed in the number of significant bits in the digital output; hence, 8 bit resolution means 1 part in 256; 10 bit resolution, 1 part in 1024, and 12 bit resolution, 1 part in 4096
Linearity	The extent to which the output deviates from a linear relationship with the input; good devices will be linear to ±1 least-significant bit, i.e., the output value is guaranteed to be within ±1 of an exactly linear conversion
Range	The maximum (and minimum) input voltages; inputs may be unipolar, e.g., 0–5 V or bipolar ±5 V; voltages are specified relative to ground unless the inputs are differential, e.g., those designed for bridge sensors
Conversion speed	The time taken to convert an input voltage into a digital output, typically 1 s–1 μs; may also be quoted as a sample rate
Linearity	The extent to which the conversion is linear, e.g., a linearity of ±1 least-significant bit means that the output value is within ±1 of the ideal linear conversion
Input impedance	The impedance between the input terminal and ground or between differential inputs

Cheaper cards may provide fixed gain and require additional signal conditioning circuits to be provided external to the computer. Many cards also feature multiplexed inputs where one of the several inputs may be selected under program control to be fed to the ADC. Some important parameters to consider in selecting an analog input board are given in Table 89.5.

At rates above a few tens of kilohertz, interrupt-driven data capture is essential to maintain speed. Faster data rates require on-board memory to avoid degrading the performance of the controlling SBC. In this case, DMA may be used to transfer data to main memory and increase performance further.

A timer function is often incorporated to allow samples to be taken at regular intervals independently of what the controlling SBC is doing. An interrupt is generated when the conversion is complete, and an interrupt service routine is then activated to read the result of the conversion into memory. The writing and installation of interrupt service routines is not a trivial task and is best left to those with an intimate knowledge of the operating system running on the SBC. Fortunately, most manufacturers supply software (device drivers) for this purpose. The simpler analog input boards may be driven by writing values directly to registers on them in a similar manner to the example given for digital I/O cards. Manufactures now commonly provide a software library which may be called from a variety of high-level languages to allow the user to access the card in a more intuitive way. This is discussed further in the section on software.

89.5.8 Analog Output

Analog output cards are available as 8, 10, or 12 bit devices. Frequently, a card will support several channels of analog output with provision for delaying the updating of channels so that all can be updated simultaneously. Output voltages may be unipolar or bipolar and current outputs (4–20 mA) are also available. Signal conditioning (buffering) is necessary to drive loads drawing currents of more than a few milliamps. Special care should be taken with inductive loads, e.g., motors to avoid damage to the device by transient voltages. Specially designed position control modules incorporating suitably buffered analog and digital I/O are available for this purpose.

89.5.9 External Peripherals

These are usually "intelligent" devices which can operate via their front panel controls without another computer but are additionally capable of being controlled by a computer. Many common laboratory instruments are available with such facilities, including power supplies, signal generators, storage oscilloscopes, voltmeters, spectrophotometers, and position controllers. A computer can coordinate the

actions of several such instruments to gather then manipulate and display data in a way which enhances the power of the instrumentation system. An almost trivial example is the use of a computer to control a signal generator and a voltmeter in order to generate the frequency response of an amplifier automatically. Such a system has an obvious role in automatic testing rigs.

Two common methods are used to control such devices: a serial link or the IEEE-488 bus. In both cases, the devices are controlled by sending messages consisting of sequences of ASCII characters. Usually the sequences are chosen to have an obvious meaning, as in the example that follows, but this is not always the case—particularly with older devices where user friendliness was often sacrificed as a result of limited memory and processing power. The message sequence

"FREQ10kHz"
"SINE"
"1.0VOLTRMS"

might be used to set a signal generator to produce a 10 kHz sinusoidal signal at 1 V rms. There is little standardization in the form of device messages used although the IEEE standard 488.2 goes some way in this direction. Responses from instruments are sent in the same way, i.e., as ASCII characters so that a voltmeter might respond to a command to make a measurement with the data

"AC2.01mV"

to indicate that it was on an ac voltage range and measured 2.1 mV. Again, there is little standardization in the format of responses. Large blocks of data may be sent in a binary format where possible (8 bit serial links or IEEE-488 bus) to minimize the amount of data to be transferred.

89.5.10 Serial Devices

Serial control of devices is accomplished using links conforming to one of the serial standards (RS-232, RS-422, or RS-485) described elsewhere in this chapter. This is a relatively simple method of control and has the advantage that much of the preliminary testing and debugging of a system can be done using a terminal or a terminal emulator program such as the public domain KERMIT available from Kermit Distribution (Columbia University Academic Information Systems, 612 West 115th Street, New York, NY 10025, Phone: 212-854-3703). The writing of custom software that accesses the serial interface of the controlling computer is relatively easy under common operating systems including DOS, Windows, Windows 95, UNIX, and OS-9 using languages such as C, Pascal, or BASIC. It is increasingly common, particularly for DOS and Windows applications, for manufactures to provide software support for their devices.

Disadvantages of serial transfers are the relative slowness when large amounts of data are transferred, the lack of standardization in device messages, and the limited control facilities available. Advantages are the ease of testing, the simplicity of the controlling software, the relative simplicity of the interconnection scheme, and—for remote instrumentation systems—the fact that with the use of modems data can be transferred over large distances using standard telephone lines or even a radio link.

89.5.11 IEEE-488 Devices

The IEEE standard 488 was developed in the 1970s and rapidly became an industry standard for the interconnection and control of test equipment. This standard was modified slightly in 1987 (IEEE standard 488.1) to allow for the considerable enhancements of IEEE standard 488.2 which was introduced at the same time [9]. The original IEEE standard 488 specifies the electrical characteristics of the bus, the mechanical characteristics of its connectors, and a set of messages to be passed between interfaces.

It does not attempt to provide any syntax or structure for communicating these messages, to specify commonly used commands, or to establish a standard for device-specific messages. These issues are addressed in IEEE standard 488.2.

The bus itself supports synchronous parallel transfers of data using three groups of lines, a bidirectional 8 bit data bus, five interface management lines, and three handshake lines, over distances of up to 20 m and at data transfer rates of up to 1 MB/s.

Devices on the bus are classed as *talkers*, *listeners*, or *controllers*. In general, the computer system is the bus controller which can also talk (send data) or listen (receive data). Most devices are both talkers and listeners, for example, a digital voltmeter will be a listener when receiving instructions to set the voltage range prior to making a measurement but will be a talker when returning the result of the measurement to the controller, which is itself acting as a listener. Each device on the bus must be assigned a unique address which is a number between 0 and 30. This may be done from the front panel of the device or, less conveniently, by setting switches elsewhere on the device.

It is a difficult, time-consuming, and error-prone process to write software to drive an IEEE-488 card. Purchasers of new IEEE-488 interfaces are strongly advised to obtain a device driver from the manufacturer. Such device drivers are now readily available and integrate the card into the filing system of the operating system running on the controlling computer. This allows the interface to be accessed in a natural way from high-level languages running on the controller.

89.6 Software for Instrumentation Systems

The difficulty involved in writing software for instrumentation systems depends largely on the support available from the manufacturers of the subsystems, on the operating system (if any), and on the development tools available. On one extreme, one may be working in a virtual instrument environment, such as that provided by the National Instrument LabVIEW, where software development is entirely graphical and, for small projects at least, is readily undertaken by users with little or no prior experience. On the other extreme, one is faced with the problem of developing software, which is at the very least interrupt-driven and probably multitasking, for a target SBC with no resident operating system; this requires considerable expertise in software design and development together with the availability of development tools such as cross compilers and source-level debuggers.

89.6.1 Virtual Instruments

Figure 89.7 shows a layer model of the software for a generalized instrumentation system. The application layer handles the data acquisition, analysis, and presentation. The instrument drivers provide a mechanism for communicating with the instruments in a standard way without requiring the user to know about the often cryptic data strings which need to be sent. For example, all digital multimeters will need the facility to choose a specific input voltage range. The range coding, resolution, etc. that have to be sent to the multimeter to achieve this will vary from instrument to instrument; however, the instrument driver allows the software writer programming in the application layer to call a procedure such as SetVoltageRange (Voltage Value) and this procedure call is the same for all multimeters. Although some manufacturers use the term slightly differently, the instrument driver is in effect the virtual instrument. Writing *instrument drivers* is a time-consuming but not too difficult task. Instrument drivers for

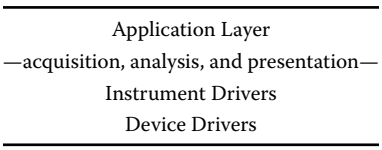


FIGURE 89.7 A layer model for instrumentation systems.

proprietary instrumentation software design packages are readily available from instrument manufacturers. Device drivers integrate the controlling interface (e.g., IEEE-488, RS-232, or internal card) into the operating system of the computer. Writing a *device driver* requires a detailed knowledge of the device hardware and of the computer operating system. This is a difficult task and new interfaces should be purchased with a device driver appropriate for the operating system wherever possible.

A number of development environments which are based on the virtual instrument concept are now available. These free the user from the problems of writing conventional software to control instruments and handle the data produced. Instruments appear to the software developer as "front panels" drawn on the computer screen, complete with familiar buttons, knobs, and displays. Data flow is handled by linking instruments in a block diagram using a mouse in an environment that resembles an ordinary drawing package. The software developer is working only in the application layer.

While the graphical environment allows simple systems to be developed rapidly, experienced programmers may find it restrictive. There are software development systems available that give the programmer access libraries containing instrument drivers, data analysis routines, graphics functions, and data visualization facilities in commonly used high-level languages such as C, Pascal, BASIC, and FORTRAN. Table 89.2 lists representative software packages.

89.6.2 Working Directly with Peripherals

It may occasionally be the case that the cost of software support for a virtual instrument development environment is not justified for a small application. Software must then be written to interface directly with the peripheral. The earlier section on digital input and output explained in principal what was necessary to program a simple interface chip. We now continue this example and show using the language C how this might be achieved.

The method of accessing the registers of peripheral cards depends on the microprocessor involved and may not even be a standard feature of the language being used. Where the peripheral forms part of the same address space as the computer memory, such as in the Motorola 680XX series, pointers can be used to read and write values in the registers. The Intel 80 × 86 series of processors often place peripherals in a separate address space which may not be accessed by pointers. In this case, an extension to the language is required. Borland's Turbo C and C++ provide functions to read and write I/O mapped devices:

unsigned char inportb(int portid)

void outportb(int portid, unsigned char value)

These are used in the code fragment which implements the software for our earlier example. We assume that the 8255 PPI has base address 0 × 1b0 and that the program copies the value read from the input port (port A) directly to the output port (port B), until bit 0 of the input becomes zero:

```
/* define the addresses of the registers for the PPI */
#define BASE 0×1b0
#define PABASE
#define PBBASE+1
#define CONTROLBASE+3
unsigned char xin;/* declare an 8-bit variable */
outportb(CONTROL, 153)/* set port A as input, port B as output */
do
{
          xin = inportb(PA);/* read the input */
          outportb(PB,xin);/* write the output */
}
while(xin & 1);/* loop if bit 0 is still 1 */
```

This fragment also illustrates that high-level languages generally only support input and output of bytes. Masking techniques must be used to access individual bits.

89.6.3 Choice of Operating System

The software development support discussed so far is typically available under DOS, Windows 3.1, Windows 95, Windows NT, and UNIX. When a system is multitasking, it has to meet stringent real-time constraints; however, none of these operating systems is particularly appropriate. DOS does not support multitasking, and the others are not optimized for real-time systems which require speedy context switching and rapid interrupt response. The most fundamental requirement of real-time applications is the ability of the system to respond to external events with very short, bounded, and predictable delays. Table 89.6 lists some important real-time operating systems and kernels.

Real-time operating systems tend not to have the mature and powerful software development support available for conventional operating systems. It is not possible simply to develop the software on a familiar operating system and then transfer the working programs to the target system. Much of the debugging, testing, and system integration will have to be done on the target itself to access the hardware. A common solution is a development system in which a conventional workstation is linked to the target system. Software is developed on the workstation using familiar tools, e.g., a Windows-based editor, support for version control, and a powerful filing system. At any time, code can be cross compiled (i.e., compiled for the processor on the target system) and downloaded to the target. The workstation may then monitor the execution of the software running on the target processor. Features that allow the user to single step through the source code, seen in a workstation window, while viewing the status of key variables in another, are available. It is also possible to set breakpoints and allow the processes to run until one is encountered. Manufacturers of real-time operating systems are often able to provide development support of this type for their product for a variety of workstations.

TABLE 89.6 Some Important Real-Time Operating Systems and Kernels

System	OS/Kernel	Manufacturer
OS-9	OS	Microware Systems Corporation 1900 NW 114th Street Des Moines, IA 50325
LynxOS	OS	Lynx Real-Time Systems, Inc. 16870 Lark Avenue Los Gatos, CA 95030-2315
VxWorks	OS	Wind River Systems 1010 Atlantic Avenue Alameda, CA 94501
VRTX/OS	OS	Microtec Research 2350 Mission College Blvd. Santa Clara, CA 95054
VRTX	Kernel	Microtec Research 2350 Mission College Blvd. Santa Clara, CA 95054
iRMX	Kernel	Intel Corp. 3065 Bowers Avenue Santa Clara, CA 95051

References

1. Anonymous, *Cypress VMEbus Handbook*, Cypress Semiconductor Corp, 1996, Digitalized on June 23, 2011.
2. S. Heath, *VMEbus: A Practical Companion*, Boston, MA: Butterworth Heinemann, 1993.
3. J. Di Giacomo, Ed., *Digital Bus Handbook*, New York: McGraw-Hill, 1990.
4. T. Shanley and D. Anderson, *ISA System Architecture*, 3rd ed., Reading, MA: Addison-Wesley, 1995.
5. T. Shanley and D. Anderson, *EISA System Architecture*, 2nd edn., Reading, MA: Addison-Wesley, 1995.
6. T. Shanley and D. Anderson, *PCMCIA System Architecture*, 2nd edn., Reading, MA: Addison-Wesley, 1995.
7. M.D. Seyer, *RS232 Made Easy*, Englewood Cliffs, NJ: Prentice Hall, 1984.
8. S. Walsh, J. Linton, R. Grace, S. Marshall, and K. Knutti, Microsystems, micromachines, technology and applications, in *MEMS and MOEMs Technology and Applications*, P. Rai-Chouhury, Ed., Bellingham, WA: SPIE Publications, pp. 479–517, 2000.
9. A.J. Caristi, *IEEE-488: General Purpose Instrumentation Bus Manual*, New York: Academic Press, 1989.

90

Grounding and Shielding

Daryl Gerke
Kimmel Gerke Associates, Ltd.

William Kimmel
Kimmel Gerke Associates, Ltd.

90.1 Introduction

Electromagnetic compatibility (EMC) is crucial to successful operation of industrial systems. Due to the increased electronic content of most industrial controls, electromagnetic interference (EMI) problems have increased dramatically in recent years. Two keys to EMC success are grounding and shielding. This section will briefly discuss how to implement these two crucial EMC strategies. It will also provide a general introduction to EMI problems in today's industrial electronic systems. The primary emphasis will be on practical insights and ideas gained in dealing with numerous industrial control problems.

90.2 Understanding EMI Problems

Here are three general observations on dealing with EMI problems in industrial electronics:

1. *The industrial environment is harsh*. The primary EMI threats are power disturbances, radio-frequency interference (RFI), and electrostatic discharge (ESD). In addition, analog sensor circuits are often plagued with 50/60 Hz "ground loop" problems. Industrial electronics need more EMC care than most commercial electronics, and even more than many military systems.
2. *Electronics often play a secondary role in electronics systems*. Unlike a computer system, where electronics is the core technology, industrial electronics are often used to support another technology, such as chemical, mechanical, or process function. This leads to EMC challenges when integrating the electronics to nonelectronic technologies.
3. *EMC rules and regulations are finally catching up to industrial electronics*. For many years, industrial electronics were exempt from mandatory EMC rules, so unless there was an actual problem, EMC was often ignored. With the EMC directives of the European Union (EU) now in force, industrial electronics are no longer exempt.

90.2.1 Three Types of EMC Problems

There are three aspects of the EMC problem: *emissions*, *susceptibility* (also known as immunity), and *self-compatibility*. *Emissions* originate within the equipment, and may upset other nearby equipment. On the other hand, external energy may upset equipment, leading to *susceptibility* (or a lack of immunity). Finally, energy internal to the system may interfere with the other internal circuits, resulting in a self-compatibility problem.

Problems with both emissions and susceptibility have led to EMC regulations. Two of the best known are the Federal Communications Commission (FCC) regulations for emissions in the United States, and the EU regulations for both emissions and immunity in Europe. Industrial controls have always been exempt from the FCC regulations, but they are not exempt from the EU regulations which became mandatory in January 1996.

90.2.2 Four Major EMC Threats

Most industrial EMC problems fall into one of four key areas: *emissions*, *power disturbances*, *RFI*, and *electrostatic discharge*. In the past, industrial systems were usually only concerned with power disturbances. Today, all four threats must be considered.

90.2.2.1 Emissions

Emissions refer to electric energy originating within equipment that can interfere with other equipment. The prime concern of this threat is jamming nearby television receivers, which is the basis for the now mandatory EU emissions regulations. Emissions problems between industrial electronic systems, however, are rare. While it is possible to interfere with any other nearby equipment, most industrial electronics generate only minute amounts of conducted and radiated interference, well below upset thresholds for digital or analog circuits.

Emissions are best addressed at the equipment design stage. Strategies include printed circuit board design techniques, high-frequency filtering on power and signal interfaces, shielded cables, and enclosure shielding. Fixes in the field are usually limited to shielded cables or enclosures, add-on filters, and ferrite clamps on cables.

90.2.2.2 Power Disturbances

Power disturbances can take many forms, from short transients to long sags, surges, or complete power outages. The three most serious power threats to industrial controls are transients, voltage sags, and power outages. Stray 50/60 Hz currents can cause also problems with sensitive analog circuits, particularly due to ground loops (to be discussed later). Other power disturbances, like frequency or waveform variations, often have little effect on electronic systems.

Power disturbances are very common in industrial environments. As a result, most industrial systems are pretty robust against this threat, at least at low frequencies. High-frequency threats, such as fast transients or RF on the power lines, can still cause problems. The EU tests simulate these threats with the electrical fast transient (EFT) and injected RF tests.

Most power disturbances are caused by nearby equipment, rather than external sources. (One critical exception is lightning, which can result in some nasty voltage and current surges.) Power disturbances solutions include grounding, power filters, transient protectors, and in extreme cases, uninterruptible power systems (UPS).

90.2.2.3 Radio-Frequency Interference

RFI deals with threats in the RF range. RFI is quite common in industrial environments, and will likely get worse with the proliferation of handheld radios and cellular telephones. It is expected that wireless local area networks (LANs) will also provide some interesting EMI challenges. There have been cases where handheld radios were banned from use due to repeated EMI problems with industrial electronics.

It turns out that the nearby handheld radio is a much bigger threat than a large commercial broadcast station several kilometers away. A key metric is electric field intensity, measured in "volts/meter." This is a function of both transmitter power, and distance from the antenna, and can be quickly predicted by the formula:

$$E\,(\text{V/m}) = 5.5\frac{\sqrt{PA}}{d} \tag{90.1}$$

where

P is the transmitter power in watts
A is the antenna gain
d is the distance from the antenna in meters

For example, the electric field from a 1 W radio with a zero gain antenna at 1 m is about 5 V/m, while the electric field from a 10,000 W broadcast station at 1 km is about 0.5 V/m. Since unprotected equipment can fail in the 0.1–1 V/m range, problems can and do occur. The EU "heavy industrial" limits of 10 V/m are clearly aimed at protecting against the nearby handheld radio.

Solutions to RFI problems include high-frequency filtering on power and signal cables, shielded cables, and shielded enclosures. Analog circuits are particularly vulnerable to RFI, so they often need extra protection. Do not overlook banning radio transmitters in the immediate vicinity. Often, maintaining a 3–10 m distance is enough to solve the problem.

90.2.2.4 Electrostatic Discharge

ESD refers to the sudden discharge that can occur after a gradual buildup of electric charge. ESD is most commonly associated with humans (touching controls or keyboards), but ESD can also be caused by internal arcing due to the movement of paper, plastic, etc. Internal ESD problems are increasing in industrial systems.

Although the static buildup can take a long time (seconds or even minutes), the discharge is almost instantaneous (nanoseconds or less). Furthermore, it is the sudden current, not the voltage that is the culprit. The effect is a bit like having a dam burst—the ESD current is like water running down a mountain, destroying anything in its path. Fortunately, the current surge does not last too long, so the energy levels are not high. They are high enough, however, to damage or upset electronic devices.

The extremely fast discharge results in high frequencies well into the UHF range. At 1 ns, the transient bandwidth is over 300 MHz. As a result, it does not take a "direct ESD hit" to cause a problem. ESD upsets 5–10 m away are not uncommon, due to the intense electromagnetic fields associated with an ESD event. These problems are particularly insidious, since the ESD event may be occurring on a different piece of equipment.

Solutions to ESD problems include transient protection, high-frequency filtering, cable shielding, and enclosure shielding. Grounding is a very important factor in ESD protection, but it must be designed for high frequencies. Since many times ESD causes "reset" problems, extra attention to microprocessor reset circuits is beneficial.

90.2.3 Sources, Paths, and Receptors

A common EMI problem is gathering and organizing data. This is particularly important when troubleshooting EMI in the field. The "source–path–receptor" model is popular. Simply stated, three elements are necessary for any EMI problem:

1. There must be a source of energy.
2. There must be a receptor that is upset by that energy.
3. There must be a coupling path between the source and receptor.

Any interference problem can be broken down into
- Source of interference
- Receptor of interference
- Path coupling the source to the receptor

Sources	Paths	Receptors

• Microprocessors
• Video drivers
• ESD
• Transmitters
• RF generators
• Power disturbances
• Lightning
	• Radiated
	  • EM fields
	  • Cross talk
	  Capacitive
	  Inductive
	• Conducted
	  • Signal
	  • Power
	  • Ground
		• Digital
		  • Microprocessors
		  • Reset
		  • Other logic
		• Low-level analog
		• Receivers

FIGURE 90.1 The source–path–receptor model for assessing EMI problems. All three elements must be necessary for an EMI problem to occur.

All three elements must exist at the same time, and if any one is missing, there is no EMI problem. Sometimes one can identify all three, and, other times, one can only guess. While this may seem simple, it is a useful tool to organize EMI information.

Figure 90.1 illustrates this model, giving typical sources, paths, and receptors. Several possible sources have been discussed: emissions from digital circuits, ESD, RFI from communications transmitters, and power disturbances (including lightning). Several different receptors have also been suggested: communications receivers, analog electronics, and digital electronics. Note the two types of paths: radiated and conducted. In both cases, the object is to block unwanted energy from reaching a receptor, which is done with shielding (for the radiated path) and filtering (for the conducted path).

90.3 Grounding

Grounding is probably the most important, yet least understood, aspect of EMI control. Every circuit is connected to some sort of "ground," so every circuit is affected by EMI grounding issues.

90.3.1 What Is a Ground?

A major problem with the subject of grounding is the ambiguity of the term. Our favorite definition is one popular in the EMC community, which says that a *ground is simply a return path for current flow*. These currents can be intended or unintended. The unintended currents are often referred to as "sneak grounds," and can cause many kinds of EMI problems. Finally, a physical connection is not even necessary at higher frequencies, where parasitic capacitance or inductance may form part of a ground path.

90.3.2 Different Types of Grounds

Grounds are used for many reasons, including power, safety, lightning, EMI, and ESD. Although they may share common functions, they may vary widely when it comes to frequencies and current amplitudes. Recognizing these key differences is key to understanding grounding issues.

Table 90.1 shows some frequency and amplitude requirements of several different types of grounds. Note that power and safety grounds must handle high currents, but only at low frequencies. Grounds for EMI and ESD, on the other hand, must often handle high frequencies at relatively low current levels. Lightning grounds must handle extremely high currents, but at moderate frequencies.

TABLE 90.1 A Ground May Work Over Wide Frequency Ranges

Type	Frequency	Typical Current Levels	Typical Duration
Power	50/60 Hz	10–1,000 A	Seconds or minutes
Lightning	300 kHz	100,000 A	Tens of milliseconds
ESD	300 MHz	10–50 A	Tens of nanoseconds
EMI	dc–Daylight	μA–A	Nanoseconds to years

The frequency of transient events is calculated using the formula $f = 1/(\pi t_r)$, where f is the equivalent frequency and t_r is the transient rise/fall time. This relationship can be derived using Fourier analysis. For example, ESD has an equivalent frequency of about 300 MHz based on a typical 1 ns rise time, and lightning has an equivalent frequency of about 300 kHz based on a 1 μs rise time.

Note that out of all these types of grounds only one actually needs an Earth connection—lightning. Other grounds may be connected to Earth by convention or for other safety reasons. For example, power neutrals are connected to Earth in many parts of the world to help provide lightning protection. On the other hand, in many other parts of the world, the power systems do not have Earth connections. When dealing with power grounding, the local safety codes will determine the proper Earth grounding methods.

90.3.3 Ground Impedances

A good ground must have a low enough impedance to minimize voltage drop in the ground system, and must provide the preferred path for current flow. The key to success is maintaining that low impedance over the entire frequency range of interest. We cannot overemphasize this point. Most EMI grounding problems are due to using the wrong approach for a given range.

The impedance of a ground conductor consists of both resistance and inductance ($Z = R + j\omega L$). For frequencies from dc through about 10 kHz, the resistance is the major factor, so heavy-gage wires are often used for low-frequency ground conductors. As the frequency increases, however, the inductance becomes the limiting factor for impedance. As a rule of thumb, the inductance for round wires is in the range of 10 nH/cm.

Table 90.2 gives the resistance, inductance, and inductive reactance for a typical wire sizes used in instrumentation power and signal circuits. It is apparent that at power and audio frequencies (dc to 10 kHz), resistance is the dominant factor in ground impedance. Thus, at low frequencies, look for ways to reduce resistance, typically by using larger wires. At frequencies above the audio range (>10 kHz), inductance becomes the dominant factor in ground impedance. Thus, at higher frequencies, look for

TABLE 90.2 Impedance Parameters for 10 cm Length Wires

Gage	Ω/m	μH/m	Z at 10 kHz	Z at 1 MHz	Z at 100 MHz
10	0.0033	1.01	0.006	0.63	63
12	0.0052	1.05	0.007	0.66	66
14	0.0083	1.10	0.007	0.69	69
16	0.0132	1.15	0.007	0.72	72
18	0.0209	1.19	0.007	0.75	75
20	0.0333	1.24	0.008	0.78	78
22	0.0530	1.29	0.009	0.81	81
24	0.0842	1.33	0.010	0.84	84
26	0.1339	1.38	0.012	0.87	87
28	0.1688	1.40	0.019	0.88	88
30	0.2129	1.43	0.022	0.90	90

TABLE 90.3 Impedance Values for Ground Plane Impedance

	Thickness		
Frequency	0.1 mm	1 mm	10 mm
60 Hz	172 μΩ	17.2 μΩ	1.83 μΩ
1 kHz	172	17.5	11.6
10 kHz	172	33.5	36.9
100 kHz	175	116	116
1 MHz	335	369	369
10 MHz	1.16 mΩ	1.16 mΩ	1.16 mΩ
100 MHz	3.69	3.69	3.69
1000 MHz	11.6	11.6	11.6

ways to reduce the inductance of the ground path. This is accomplished by using ground planes, grids, and straps to lower the inductance.

Table 90.3 gives the impedances for solid ground planes at various frequencies. In this case, the impedances are in "ohms per square," which is a measure of impedance across a diagonal surface. By comparing this with Table 90.2, one can see that at high frequencies (such as 100 MHz) the ground plane impedance may be several orders of magnitude below the impedance of a wire. Furthermore, at high frequencies the thickness is not a factor, since the impedance is limited by the skin effect.

90.3.4 Ground Topologies

Now that we have looked at ground impedance vs. frequency, we are ready to look at ground topologies vs. frequency. The impedance limitations yield two different grounding approaches, dependent on frequency. For low-frequency problems (dc to 10 kHz), single-point grounds are preferred, while at high frequencies (above 10 kHz), multipoint grounds with planes or grids become the preferred approach.

This dichotomy often causes confusion with industrial controls, but this can be minimized by determining the frequency of the EMI threat and then selecting the appropriate grounding approach. In many cases, both approaches may be necessary at the same time, leading to "hybrid" grounds, which use capacitors and inductors to alter the ground topology with frequency.

90.3.4.1 Single-Point Grounds

At low frequencies, one can usually steer current via wires. Since the inductance is low, the limiting factor is the wire resistance itself. Furthermore, capacitive coupling from the ground wires to adjacent wires or surfaces is small, so virtually all the current follows the wiring path.

Figure 90.2 shows a typical industrial grounding scheme. Note what happens if the system is grounded at both ends. Any common noise current in the common ground path is now coupled into the circuit via the "common ground impedance." This results in the severe "ground loop," which will be discussed shortly. A single-point ground eliminates the ground loop, since there is no common impedance across which a common-mode voltage can be generated. Thus, single-point grounding is a very practical way to limit "ground noise" problems with the threat of low-frequency ground currents. This is very typical of 50/60 Hz currents getting into sensitive analog instrumentation circuits.

90.3.4.2 Multipoint Grounds

Unfortunately, as the frequency increases, the inductive reactance of the wires increases. At the same time, parasitic capacitive reactance to adjacent wires or surfaces decreases, and soon it is no longer possible to maintain a true single-point ground, even if the system is wired that way. The only option left

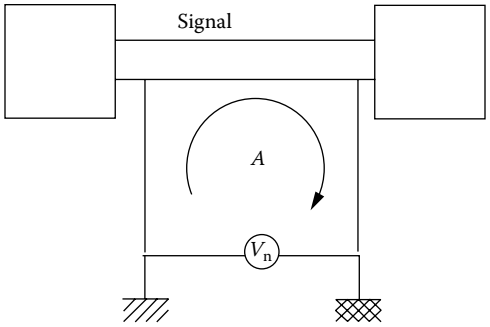


FIGURE 90.2 Typical industrial grounding situations, which also illustrate a ground loop.

is to lower the ground path impedance, and that is accomplished with planes or grids. Furthermore, a single-point connection to a grid or plane is usually not adequate because of transmission line effects, so multipoint grounds (combined with planes/grids) become the preferred approach above 10 kHz.

Ground grids have been used for years in computer facilities, and are seeing increasing use in industrial facilities. The recommended spacing for grids is no more than $\frac{1}{20}$ of a wavelength at the highest frequency of concern. Computer room grids are often spaced about 0.7 m (about 2 ft), which meets this criteria from dc to about 25 MHz. This is very beneficial in addressing ground noise due to lightning and other power transients, which are usually in the 1 MHz range and below. But a 0.7 m grid does not help with VHF/UHF radio problems or ESD. In those cases, solid surfaces may be necessary.

90.3.4.3 Ground Loops

Ground loops are a serious problem for sensitive analog circuits facing low-frequency threats. At high frequencies, ground loops generally do not pose serious threats if proper high-frequency precautions are taken when designing the ground system.

A ground loop exists whenever multiple ground paths exist. Unwanted currents can take unwanted paths, resulting in unwanted noise voltages at unwanted places. The problem is particularly acute with sensitive analog systems, where even a few microvolts can jam intended signals. A classic example is 60 Hz ground currents causing hum in an audio system.

Figure 90.2 shows a typical ground loop problem. Note that there must be the three conditions of any EMI problem: a source, a path, and a victim. In this case, the source can be circulating power currents, the path the common ground impedance, and the victim is often the sensitive analog circuit. With many systems problems, one cannot do anything about either the source or victim, so the solution is with the ground path. As we have already seen, single-point grounding is effective at low frequencies, and ground planes/grids are effective at higher frequencies.

If one cannot change the ground paths, one can still attack the ground loop by "breaking" it in other places. For example, transformers or optical isolators (or even fiber optics) can be used in cable connections, which will block common-mode noise currents while passing intended differential mode signals. Balanced input/output (I/O) circuits can be used to "cancel" the noise through common-mode rejection. All of these are most effective at 50/60 Hz, and become less effective at higher frequencies due to parasitic capacitance.

90.3.5 Grounding Guidelines

By now it should be apparent that there is no magic solution for grounding. Rather, different methods and approaches are necessary for different circuits and operating conditions. Two key parameters are the threat frequency (low vs. high), and the circuit operating levels. Here are some guidelines, but keep in mind that even these may need to be modified for a particular situation.

90.3.5.1 Analog Circuits

Since most analog circuits operate at low frequencies and are subject to low-frequency threats, single-point grounds are preferred. Typical threats are 50/60 Hz power return currents, stray switching power supply currents, and perhaps digital circuit return currents (if separate analog and digital power and grounds are not provided). Low-level analog circuits are the most vulnerable, since the signal levels are small.

Keep in mind that high-frequency threats (such as a VHF radio) to low-frequency circuits may require high-frequency grounding solutions, such as multipoint grounds. Often, this can be accomplished by using small high-frequency capacitors (1000 pF typical) which appear as a short at 100 MHz, yet still appear as a high impedance at 50/60 Hz.

90.3.5.2 Digital Circuits

Most digital circuits today operate at relatively high frequencies, so multipoint grounds and ground planes and grids are preferred. The connections between the circuits and their grounds need to be short, fat, and direct to minimize inductance.

Digital circuits, particularly I/O circuits, are vulnerable to external high-frequency threats like RF and ESD. They are also a key source of high-frequency emissions and internal problems like cross talk. For digital circuits, multilayer boards with internal ground planes are preferred. These ground planes typically are connected to a metallic enclosure through multiple low-inductance connections.

Pay particular attention to where digital and analog circuits meet. A single-point connection is usually preferred to minimize ground loops, but installing a small resistor (1–10 Ω typical) or inductor (1–100 μH typical) at that point is often helpful in providing additional isolation. One may need to experiment with this to determine the optimum solutions.

90.3.5.3 Power Safety Grounding

Entire books have been written about this subject, and rightly so; this is an extremely important safety issue. The key concern here is human safety and prevention of electric shock. In most parts of the world, exposed metal on line-powered equipment must be bonded to a safety grounding conductor. Furthermore, the electric wiring codes (such as the National Electrical Code in the United States) give very specific guidelines on how power grounding must be accomplished.

These guidelines must be followed when wiring any industrial control system, and must never be compromised by "isolated" power grounds or other similar foolishness. *Finally, if there is ever a conflict between EMI and safety grounding, the safety issues must always prevail!*

90.4 Shielding

Many systems today require at least some *shielding* for proper operation or to meet radiated emission or immunity requirements. Many engineers consider shielding purely a mechanical issue, but nothing could be farther from the truth. EMI shielding needs both an electrical and a mechanical understanding of key issues to assure success.

Two of these key issues are selecting the right material and maintaining the shielding integrity over the desired frequency range. While most people worry more about the selection, shield integrity is usually much more important. We will soon see that even very thin metallic coatings can be effective shields, yet even very small holes or penetrations can completely destroy a shield. Like grounding, shielding cannot be left to chance, and must be properly designed and implemented.

90.4.1 How Shielding Works

EMI shielding involves two independent mechanisms: *reflection* and *absorption*. In reflection, an electromagnetic wave bounces off the surface, just like light off a mirror. In absorption, the electromagnetic wave penetrates the material and is absorbed as it passes through, much like heat loss through an insulating wall.

Shielding effectiveness is usually expressed as follows:

$$\text{SE(dB)} = R(\text{dB}) + A(\text{dB}) \tag{90.2}$$

where

SE is the total shielding effectiveness in dB

R and A are the reflection and absorption losses expressed in dB

Reflection is the primary mechanism for high-frequency shielding (emissions, RFI, ESD), while absorption is the key mechanism for low-frequency magnetic field shielding. The actual formulas for calculating reflection and absorption losses are a bit complex, and beyond the scope of this chapter, but several sources are included in the further information section.

90.4.2 Three Types of Fields

It is customary when dealing with shielding to use three types of "fields" to explain shielding. These three fields account for differences in shielding performance due to differences in frequency and circuit impedance levels. They also explain why the same shield can behave differently for different energy sources. These are plane waves, magnetic fields, and electric fields. Figure 90.3 shows typical shielding curves for copper, with references to each type of field.

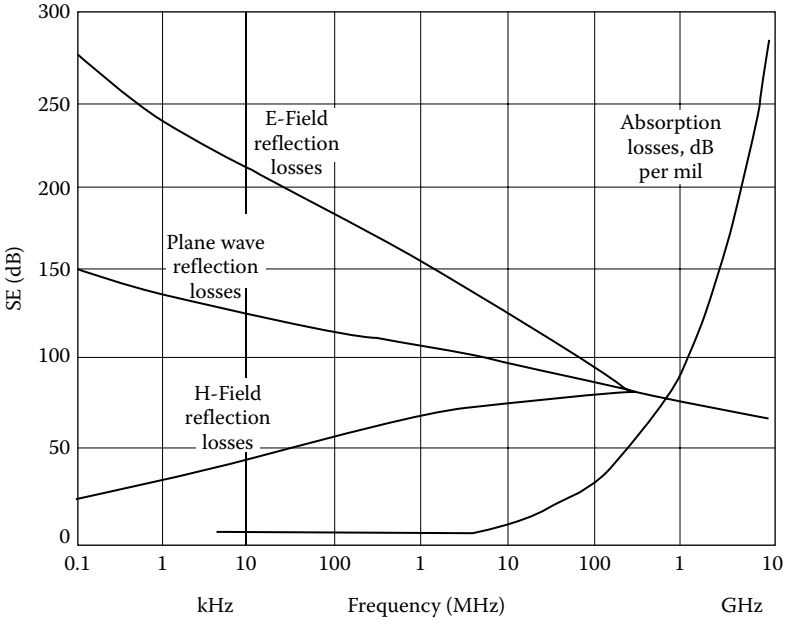


FIGURE 90.3 Typical shielding effectiveness curves for copper. Note two mechanisms (reflection and absorption) and three types of fields (electric, magnetic, and plane wave). Shielding for aluminum is almost the same as for copper.

90.4.2.1 Plane Wave Fields

If one is located greater than about one-sixth wavelength from a point source, the wave impedance (ratio of electric field intensity to magnetic field intensity) is a constant 377 Ω in free space. This field is known as the "far field" or "radiation field," since real energy predominates here and propagates as a "plane wave." Since reflection losses are due to a mismatch between the wave impedance (377 Ω) and a metallic shield surface impedance (typically milliohms or less), shielding effectiveness is usually very high for plane wave sources.

At frequencies 30 MHz and above, once one is more than about 1 m away, one is in the plane wave region. Thus, even very thin shields work well for emissions, ESD, and RFI problems, with reflection as the prime shielding mechanism.

90.4.2.2 Electric and Magnetic Fields

If one is located less than about one-sixth wavelength from the source, then the wave impedance is dependent on the circuit impedance. This region is known as the "near field," since reactive energy predominates here. This region is further divided into "electric" and "magnetic" fields, both dependent on source circuit impedance. For high-impedance sources (electric fields), the reflection losses are still high, but for low-impedance sources (magnetic fields), the impedance can be quite low. In the latter case, the reflection losses can become minimal.

For power line frequencies, the near field almost always predominates. As a result, materials like aluminum or copper have no reflection losses and are virtually transparent to power line magnetic fields. (As a rule, remember that aluminum foil is transparent to 60 Hz magnetic fields.) To solve this problem, permeable materials are needed to boost the electric thickness for a given physical thickness. Steel or high-permeability mu-metals are normally used to absorb (not reflect) the magnetic fields. Even so, it can still be very difficult to shield for low-frequency magnetic fields.

90.4.3 Why Shielding Fails

While material selection is important, other factors must also be considered. For low-frequency/low-impedance threats (power supply or power line magnetic fields), steel or other high-permeability materials are needed. For high-frequency threats, however, even very thin materials like conductive paints provide high levels of shielding. Two problems at high frequencies, however, are shield openings and shield penetrations. Lack of attention to these areas can result in a loss of virtually all shielding effectiveness at high frequencies. Figure 90.4 illustrates these two high-frequency failure modes.

Intuition suggests that any opening in a shield can leak, much like an open window. The surprise is that for electromagnetic leakage it is not the area that is critical, but the longest dimension. For example, a 100 × 1 mm opening is about 10 times leakier than a 10 × 10 mm square hole. And that slot may not even be obvious. It could be a painted seam or a poorly fitting panel or door.

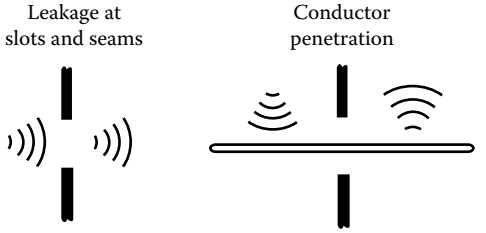


FIGURE 90.4 Two shielding failure modes, due to slots/seams and due to penetration of conductors. In both cases, the critical dimension is $\frac{1}{20}$ wavelength for the highest frequency of concern.

Slots act like small antennas. Because they are antennas, the longer they are, the better they radiate. While a half wavelength is very efficient, as a rule of thumb, we like to limit slots to $\frac{1}{20}$ wavelength or less at the highest frequency of concern. For 100 MHz, this is 15 cm (about 6 in.); at 300 MHz (ESD frequencies), this drops to 5 cm (about 2 in.), and at 1000 MHz, it is only 1.5 cm ($\frac{2}{3}$ in.). And even these dimensions may be too large, as they only assure 20 dB (10-fold) of shielding through the slot. Clearly, even small slots and other openings mean big shielding problems at high frequencies.

The other way to destroy a high-frequency shield is to pass unterminated metal through a shield. Hole dimensions do not matter here, and even a pinhole with an insulated wire passing through can carry large amounts across the shielding barrier. The dimension that does matter is how far the penetration extends on either side of the shield. Once again, the critical distance is $\frac{1}{20}$ wavelength or more.

90.4.4 Shielding Guidelines

Now that we have looked at how shielding works (and fails), let us look at how to design good electromagnetic shields. Most of our focus will be on RF shielding in the 30–1000 MHz range, necessary for emissions, ESD, and RFI.

90.4.4.1 Material Selection

We have already seen that for low-frequency magnetic interference problems, ferrous material like steel or mu-metals are necessary. Most instrumentation problems are either high impedance or high frequency in nature, so most of the time, thin conductive materials will work fine. For high frequencies, however, attention must be given to slots and penetrations.

Many enclosures today are made of plastic. For high-frequency shielding, conductive coatings also work quite well. Popular surface treatments include conductive paints, vacuum deposition, electrodeless plating, and even metal fabrics. Conductive plastics are also available, but they generally do not perform nearly as well as surface treatments for high frequencies.

90.4.4.2 Gasketing and Screening

Large openings, such as ventilation ports or display areas, can be sealed with screening material. Seams or slots can be filled with conductive gaskets. In both cases, the secret is to provide complete and continuous metal-to-metal contact at all junctions. For high-frequency shielding, the connections must be almost watertight. Anything less is asking for problems.

For screening material, the smaller is the openings the higher is the shielding. Window screen spacing is almost as effective as solid materials from dc to 1000 MHz, and even 5 mm (about $\frac{1}{4}$ in.) openings are often acceptable at 1000 MHz. In any case, do not exceed $\frac{1}{20}$ wavelength at the highest frequency of concern.

90.4.4.3 Cable Terminations and Filters

Poor termination of shielded cables can cause big problems at high frequencies. If a shielded cable is not terminated directly at the shielding barrier, a lot of energy leaks, degrading both the cable and the enclosure shield. Pigtail connections, popular for terminating low-frequency cable shields, are particularly bothersome at high frequencies. In fact, this is a leading cause of EMI failures for RFI, emissions, and ESD. As a rule of thumb, pigtail connections should not be used on cable shields at frequencies above about 1 MHz.

Unshielded cables can also cause problems at high frequencies. In those cases, high-frequency filtering is needed directly at the interface to assure that the shield is not degraded at high frequencies. Common solutions are EMI filters on power and signal lines, or ferrite beads on the lines or cables. These must be installed as close to the shield penetration as possible. The best situation is to mount the filter directly in the shield itself, although this is not always necessary for moderate problems.

90.4.4.4 Internal Shields

Finally, do not overlook using internal shield on critical circuits. Radio and television designers have been doing this for years, using selective shields on oscillators, power amplifiers, and the like. A classic example of this approach is the TV tuner, the most sensitive part of a television receiver. It is an inexpensive, yet highly effective shielding strategy.

Defining Terms

Conducted: Energy or interference propagated by a conductor, e.g., power, grounding, or signal interface wiring.

EFT: Electrical fast transient; a high-frequency burst of energy on power wiring.

EMC: Electromagnetic compatibility; the condition wherein electric and electronic equipment operate successfully in close proximity.

EMI: Electromagnetic interference; unwanted electric energy that may impair the function of electronic equipment.

Emissions: Electric energy emanating from an electronic source.

ESD: Electrostatic discharge; the rapid discharge that often follows a buildup of static charge.

EU: European Union; formerly called the European Community.

FCC: Federal Communications Commission (US government).

Ground: A return path for current.

Radiated: Energy or interference is propagated by electromagnetic radiation through space.

RFI: Radio-frequency interference; an older term for EMI, now normally used to describe interference caused by a nearby radio transmitter.

Shield: A metallic enclosure used to reduce electric or magnetic field levels.

Susceptibility: Vulnerability of electronic equipment to external sources of interference; often used interchangeably with immunity.

Further Information

Gerke, D.D. and W.D. Kimmel, *EDN's Designer's Guide to Electromagnetic Compatibility*, Originally published by Cahners, now available directly from Kimmel Gerke Associates, Ltd., Basic introduction on EMC issues written by the authors. It is very good for non-EMC engineers, updated 2005.

Kimmel, W.D. and D.D. Gerke, *Electromagnetic Compatibility in Medical Equipment*, New York: IEEE Press and Interpharm Press, 1995. Detailed introduction to EMC issues in medical instrumentation, written by the authors. Applicable to most industrial control instrumentation.

Kimmel, W.D. and D.D. Gerke, Internet World Wide Web Home Page [online]. Available at http://www.emiguru.com and contains a wide range of useful EMC information.

Ott, H.W., *Electromagnetic Compatibility Engineering*, New York: Wiley, 2009. Good introduction to EMC issues, particularly system and analog issues.

Paul, C.R., *Introduction to Electromagnetic Compatibility*, 2nd edn., New York: Wiley, 2006. Written as a college text and it is quite analytical with very good coverage of EMC issues.

Morrison, R., *Grounding and Shielding Techniques in Instrumentation*, New York: John Wiley & Sons, 1986. A classic on instrumentation, with emphasis on low-level signal issues.

91

Amplifiers and Signal Conditioners

Ramón Pallás-Areny
Universitat Politècnica de Catalunya

91.1 Introduction

Signals from sensors do not usually have suitable characteristics for display, recording, transmission, or further processing. For example, they may lack the amplitude, power, level, or bandwidth required, or they may carry superimposed interference that masks the desired information. Further, implementing other functions such as amplitude demodulation, linearization, and temperature compensation before digitizing signals can sometimes be cost effective.

Signal conditioners, including amplifiers, adapt sensor signals to the requirements of the receiver (circuit or equipment) to which they are to be connected. The functions to be performed by the signal conditioner derive from the nature and properties of both the signal and the receiver. Commonly, the receiver requires a single-ended, low-frequency ("dc") voltage with low-output impedance and amplitude range close to its power supply voltage(s). A typical receiver here is an analog-to-digital converter (ADC). Sampling ADCs are preceded by a sample-and-hold amplifier that keeps constant the voltage during its digitization. Integrating ADCs obtain the average of the input voltage during the integration time.

Signals from sensors can be analog or digital. Digital signals come from position encoders, electromechanical or magnetic switches, or oscillator-based sensors connected to frequency counters or time-to-digital converters. The amplitude for digital signals must be compatible with logic levels for the digital receiver, and their edges must be fast enough to prevent any false triggering. Slow edges can be accelerated by a Schmitt trigger. Small digital voltages can be amplified and large voltages can be attenuated by *logic-level translators*. Alternatively, large voltages can be attenuated by a voltage divider.

Analog sensors are either self-generating or modulating. *Self-generating sensors* yield a voltage (thermocouples, photovoltaic, and electrochemical sensors) or current (piezo- and pyroelectric sensors) whose bandwidth equals that of the measurand. *Modulating sensors* yield a variation in resistance, capacitance, self-inductance or mutual inductance, or other electrical quantities. Modulating sensors need to be powered by an excitation voltage or current in order to provide an output voltage or current. Some sensors based on electromagnetic effects such as the Hall effect or based on semiconductor junctions (*p-n* junction temperature sensors, photodiodes) need to be biased to yield an output voltage or current. Impedance variation–based sensors are normally placed in voltage dividers, or in either Wheatstone bridges (resistive sensors) or ac bridges (resistive and reactance-variation sensors). The bandwidth for signals from modulating sensors equals that of the measurand in dc-excited or biased sensors, and is twice that of the measurand in ac-excited sensors (sidebands about the carrier excitation frequency). Capacitive and inductive sensors require an ac excitation, whose frequency must be at least 10 times higher than the maximal frequency variation of the measurand in order to simplify the filtering of undesired harmonics in the demodulator. Pallas-Areny and Webster [1] give the equivalent circuit for different sensors and analyze their electronic interface.

Current signals can be converted into voltage signals by inserting a series resistor into the circuit, or using other current-to-voltage converters described in Section 91.6. Alternatively, current input ADCs, such as the DDC series from Texas Instruments, can directly digitize low-level currents. Henceforth, we will refer to voltage signals to analyze transformations to be performed by signal conditioners.

91.2 Dynamic Range

The *dynamic range* for a measurand is the quotient between the measurement range and the desired resolution. The dynamic range for the input or output of any signal-processing stage can be defined as the quotient between the corresponding voltage range and the resolution, often determined by noise. Any stage for processing the signal from a sensor must have a dynamic range equal to or larger than that of the measurand. For example, to measure a temperature from 0 °C to 100 °C with 0.1 °C resolution, we need a dynamic range of at least (100 − 0)/0.1 = 1000 (60 dB). Hence, a 10 bit ADC should be appropriate to digitize the signal because $2^{10} = 1024$ (>1000). Let us assume we have a 10 bit ADC whose input range is 0–10 V; its resolution will be 10 V/1024 = 9.8 mV. If the sensor sensitivity is 10 mV/°C and we connect it to the ADC, the 9.8 mV resolution for the ADC will result in a 9.8 mV/(10 mV/°C) = 0.98 °C resolution! In spite of having the suitable dynamic range, we do not achieve the desired resolution in temperature because the output range of our sensor (0–1 V) does not match the input range for the ADC (0–10 V).

The basic function of voltage amplifiers is to amplify the input signal so that the output voltage extends across the input range of the subsequent stage. Further, the voltage levels of the amplified and further processed signal must match the input voltage levels of the ADC; otherwise, the voltage level must be shifted, even if the voltage ranges match. In the earlier example, an amplifier with a gain of 10 V would match the sensor output range to the ADC input range of 10 V and the 0 V sensor output at 0°C to the 0 V ADC input. In addition, the signal source should not be disturbed when connecting the amplifier, the output of the amplifier should depend only on the input signal, and the output noise should be smaller than the equivalent input noise of the stage that follows the amplifier. To fulfill these requirements, the amplifier must be selected according to the characteristics of the input signal. Amplifiers improve the sensitivity and resolution of the system with respect to that of the ADC, but because they amplify both the signal and the noise that comes with the signal, they do not improve the dynamic range

of the signal. Filters that reduce noise without modifying the information-carrying signal improve the dynamic range. Whenever a sampling ADC is used, an *antialiasing filter* must be used to reduce noise amplitude above the *Nyquist frequency* (half the sampling frequency) to less than the voltage resolution of the ADC. Integrating ADCs do not need antialiasing filters as the converter itself behaves as a low-pass filter, rather than behaving as a multiplier as sampling ADCs do.

Alternatively, instead of using an amplifier in front of the ADC, an ADC can be selected whose resolution matches the desired resolution for the measurand. In the example given earlier, 0.1°C resolution implies 1 mV change at the sensor output and the ADC should have N bits such that $10\ V/2^N = 1$ mV. That is, $N = 14$. The output voltage levels of the sensor must be within the input voltage levels of the ADC. Then, the four most significant bits of the ADC will not change and the 10 least significant bits will change when the temperature changes from 0 °C to 100 °C. In general, we can either use an N bit ADC preceded by an amplifier with gain G, or use an ADC with N bits and use algorithm for a gain G and no analog amplifier.

91.3 Signal Classification

Signals can be classified according to their amplitude, the relationship between their source terminals and ground (signal ground—also termed circuit common—chassis ground, earth [safety] ground), their bandwidth, and the value of their output impedance. Signals lower than around 100 mV are considered to be low level and need either amplification or an ADC whose voltage resolution is equal or better than the desired resolution for the signal. Signals larger than 100 mV may also need amplification depending on the input range of the receiver.

91.3.1 Single-Ended and Differential Signals

A *single-ended signal* source has one of its two output terminals at 0 V (signal ground). For example, Figure 91.1a shows a voltage divider whose terminal L remains at the power supply reference voltage (circuit common or signal ground terminal) regardless of the sensor resistance, as shown in Figure 91.1b. If terminal L is at (earth) ground potential (grounded power supply in Figure 91.1a), then the signal is single-ended and grounded. If terminal L is completely isolated from ground (for example, if the power supply is a battery), then the signal is single-ended and floating. If terminal L is at some voltage with respect to earth ground, then the signal is single-ended and driven-off ground. The voltage at terminal H will be the sum of the signal plus the off-ground voltage. Therefore, the off-ground voltage is common to H and L; hence, it is called the *common-mode voltage* and carries no information. For example, a thermocouple yields a voltage that depends on its temperature; a thermocouple bonded to a power transistor yields a signal whose amplitude depends on the temperature of the transistor case, riding on a common-mode voltage equal to the case voltage. Common-mode voltages can also arise because of a limited isolation of signal ground with respect to earth ground, that is, a finite impedance between circuit common and earth ground.

A *differential signal* source has two output terminals whose voltages change simultaneously by the same magnitude but in opposite directions with respect to the common (signal ground) terminal, hence the alternative name *push–pull signal*. The Wheatstone bridge in Figure 91.1c provides a differential signal. Its equivalent circuit (Figure 91.1d) shows that there is a differential voltage ($v_d = v_H - v_L$) proportional to x and a common-mode voltage ($v_c = (v_H + v_L)/2 = V_r/2$) that does not carry any information about x. Further, the two output impedances are balanced. We thus have a balanced differential signal with a superimposed common-mode voltage. Were the output impedances different, the signal would be unbalanced. If the bridge power supply is grounded, then the differential signal will be grounded; otherwise, it will be floating. When the differential signal is very small as compared with the common-mode voltage, in order to simplify circuit analysis it is common to use the equivalent circuit in Figure 91.1e, where v_H and v_L are not longer symmetrical with respect to v_c. Other measurement signals, such as impedances

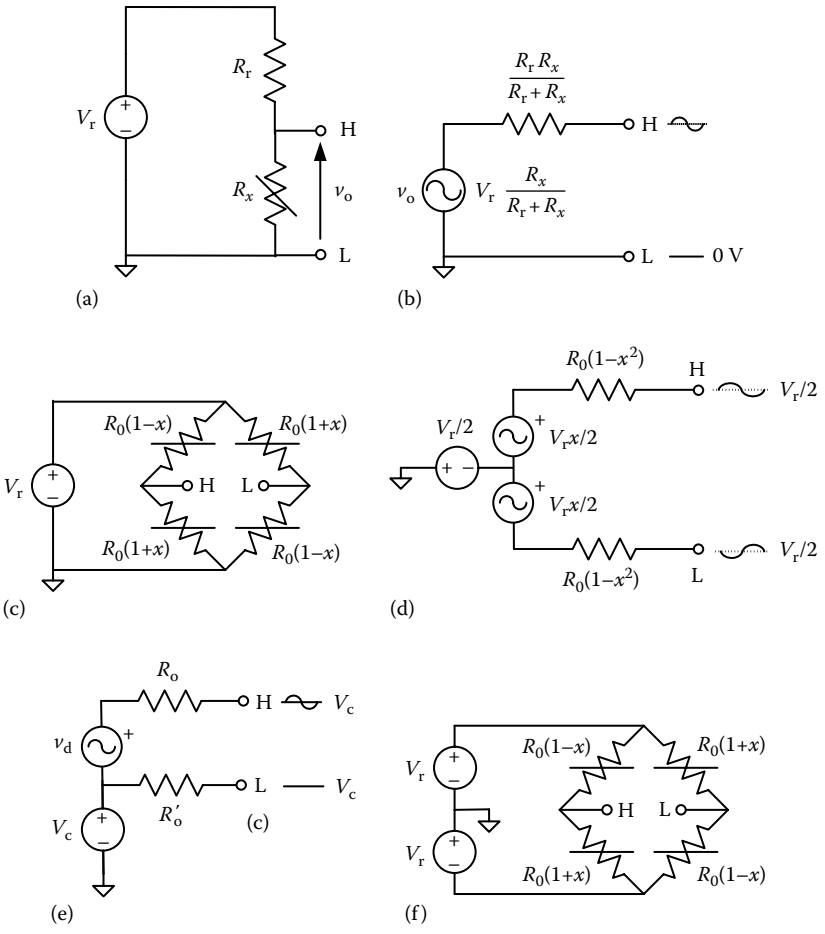


FIGURE 91.1 Classes of signals according to their source terminals. A voltage divider (a) provides a single-ended signal (b) where terminal L is at a constant voltage. A Wheatstone bridge with four sensors (c) provides a balanced differential signal which is the difference between two voltages v_H and v_L having the same amplitude but opposite signs and riding on a common-mode voltage v_c and the equivalent circuit of the Wheatstone bridge is shown in (d). For differential signals much smaller than the common-mode voltage, the equivalent circuit in (e) is used. A Wheatstone bridge with four sensors supplied by a split power source (f) yields a differential voltage with zero common-mode voltage. In all cases, if circuit common (signal ground) is (earth) grounded, the signal (single-ended or differential) will be grounded; if the reference point is floating, the signal will also be floating.

measured with the four-wire (or four-lead) method, can also be modeled by the equivalent circuit in Figure 91.1e, where v_c can be an ac voltage; because in this case v_H and v_L do not change in a push–pull fashion, the signal is termed *pseudodifferential*. Some differential signals (grounded or floating) have a zero common-mode voltage, for example, the output of a full-sensor Wheatstone bridge supplied by a split voltage (Figure 91.1f). Note that the output signal in voltage dividers and Wheatstone bridges depends on the excitation voltage V_r (or current); therefore, that excitation, the same as bias voltages or currents applied to sensors, must be stable with time and temperature. If the same V_r is used as voltage reference in the ADC, its actual value and stability are irrelevant; therefore, the power supply voltage can be used instead, provided it is not too noisy. Reference [5] analyzes several circuits able to obtain a stable voltage or current from a voltage reference IC.

Signal conditioning must ensure the compatibility between sensor signals and receivers (amplifiers), which will depend on the relationship between input terminals and ground. For example, a differential

and grounded signal is incompatible with an amplifier having only two input terminals, one of which is grounded. Hence, amplifiers must also be described according to their input topology.

91.3.2 Narrowband and Broadband Signals

A *narrowband signal* has a very small frequency range relative to its central frequency, defined as the geometric mean of its minimal and maximal frequency components ($f_c^2 = f_{min} \times f_{max}$), that is, $f_{max} - f_{min} < f_c$. Narrowband signals can be dc, or static, resulting in very low frequencies, such as those from a thermocouple or a weighing scale, or ac, such as those from an ac-driven modulating sensor, in which case the exciting frequency (carrier) becomes the central frequency.

Broadband signals, such as those from sound and vibration sensors, have a large frequency range relative to their central frequency $f_{max} - f_{min} > f_c$. Therefore, the value of the central frequency is crucial to determine whether a signal is narrowband or broadband; a signal ranging from 1 Hz to 10 kHz is a broadband instrumentation signal, but two 10 kHz sidebands around 1 MHz are considered to be a narrowband signal. Signal conditioning of ac narrowband signals is easier because the conditioner performance only needs to be guaranteed within a narrow band around the carrier frequency.

91.3.3 Low- and High-Output Impedance Signals

The output impedance of signals determines the requirements of the input impedance of the signal conditioner. Figure 91.2a shows a voltage signal connected to a device whose input impedance is Z_d. The voltage detected will be

$$v_d = v_s \frac{Z_d}{Z_d + Z_s} \tag{91.1}$$

Therefore, the voltage detected will equal the signal voltage only when $Z_d \gg Z_s$; otherwise $v_d \neq v_s$ and there will be a *loading effect*, also called *voltage-divider effect*. Furthermore, it may happen that a low Z_d disturbs the sensor (the signal source), changing the value of v_s and rendering the measurement useless or, worse still, damaging the sensor. Qualitatively, Z_d/Z_s should be large enough for the ADC not to detect that Z_d is finite, that is, the difference between v_s and v_d should be less than the resolution of the ADC (1 LSB). This requires $|Z_d + Z_s| > 2^N|Z_s|$.

At low frequencies, it is relatively easy to achieve large input impedances even for high-output impedance signals, such as those from piezoelectric sensors. At high frequencies, however, stray input capacitances make it more difficult. For narrowband signals this is not a problem, unless v_s were disturbed, because the value for Z_s and Z_d will be almost constant and any attenuation because of a loading effect can be compensated for later. However, if the impedance seen by broadband signals is frequency-dependent, then each frequency signal undergoes different attenuations which are impossible to compensate for.

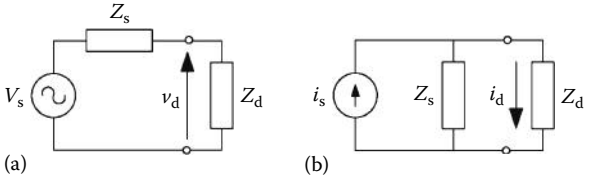


FIGURE 91.2 Equivalent circuit for a voltage signal connected to a voltage detector (a) and for a current signal connected to a current detector (b). We require $Z_d \gg Z_s$ in (a) to limit the voltage-loading effect, and $Z_d \ll Z_s$ in b) to limit the current shunting effect.

It is better to model signals with very high-output impedance as current sources (Figure 91.2b). The current through the detector will be

$$i_d = i_s \frac{Z_s}{Z_d + Z_s} \tag{91.2}$$

In order for $i_d = i_s$, it is required that $Z_d \ll Z_s$, which is easier to achieve than $Z_d \gg Z_s$. If Z_d is not low enough, then there is a *shunting effect*. Qualitatively, Z_d/Z_s should be small enough for the ADC not to detect that Z_d is nonzero.

91.4 General Amplifier Parameters

A *voltage amplifier* produces an output voltage which is a proportional reproduction of the voltage difference at its input terminals, regardless of any common-mode voltage and without loading the voltage source. Figure 91.3a shows the equivalent circuit for a general (differential) amplifier. If one input terminal is connected to one output terminal as in Figure 91.3b, the amplifier is single-ended; if this common terminal is (earth) grounded, the amplifier is single-ended and grounded; if the common terminal is isolated from ground, the amplifier is single-ended and floating. In any case, the output power comes from the power supply, and the input signal only controls the shape of the output signal, whose amplitude is determined by the *amplifier gain*, defined as

$$G(f) = \frac{V_o(f)}{V_d(f)} \tag{91.3}$$

The ideal amplifier would have any required gain for all signal frequencies. A practical amplifier has a *gain-bandwidth trade-off*, which implies that the larger the gain, the lower is the frequency where the gain starts to roll off; even for unity gain, at high frequency the gain decreases because of parasitic capacitances. In order to reduce noise and reject interference, it is common to add reactive components to reduce the gain for out-of-band frequencies further. If the gain decreases by $n \times 10$ when the frequency increases by 10, we say that the gain (downward) slope is $20n$ dB/decade. The corner (or –3 dB) frequency f_0 for the amplifier is that for which the voltage gain is 70% of that in the bandpass. (*Note*: 20 log 0.7 = –3 dB.)

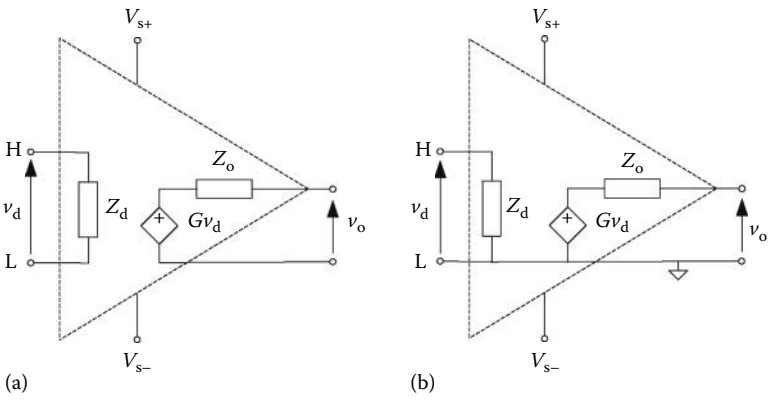


FIGURE 91.3 General amplifier with differential (a) or single-ended (b) input. The input voltage controls the amplitude of the output voltage, whose power comes from the power supply.

The *gain error* at f_0 is then 30%, which may be too large for some measurement applications. If a maximal error ε is accepted at a given frequency f_ε, then the corner frequency for the amplifier should be

$$f_0 = \frac{f_\varepsilon(1-\varepsilon)}{\sqrt{2\varepsilon-\varepsilon^2}} \approx \frac{f_\varepsilon}{\sqrt{2\varepsilon}} \tag{91.4}$$

For example, $\varepsilon = 0.01$ requires $f_0 = 7f_\varepsilon$, $\varepsilon = 0.001$ requires $f_0 = 22.4f_\varepsilon$. A broadband signal with frequency components larger than f_ε would undergo amplitude distortion. A narrowband signal centered on a frequency larger than f_ε would be amplified by a gain lower than expected, but if the actual gain is measured, the gain error can later be corrected.

Whenever the gain decreases, the output signal is delayed with respect to the output. In the aforementioned amplifier, an input sine wave of frequency f_0 will result in an output sine wave delayed by 45° (and with relative attenuation 30% as compared with a sine wave of frequency $f \ll f_0$). Complex waveforms that include frequency components close to f_0 would undergo shape (or phase) distortion. In order for a waveform to be faithfully reproduced at the output, the phase delay should be either zero or proportional to the frequency (linear phase shift). This last requirement is difficult to meet. Hence, for broadband signals it is common to design amplifiers whose bandwidth is larger than the maximal input frequency, even though this means that more amplifier noise will add to the amplified signal. (Power spectral density for amplifier noise has a white component plus an $1/f$ component at frequencies below 1 kHz for common models.) Narrowband signals undergo a delay which can be measured and corrected.

An ideal voltage amplifier would have infinite *input impedance*. Then no input current would flow when connecting the signal (Figure 91.2a), and no energy would be taken from the signal source, which would remain undisturbed. A practical voltage amplifier, however, will have a finite, yet large, input impedance at low frequencies, decreasing at larger frequencies because of stray input capacitances. If sensors are connected to conditioners by coaxial cables with grounded shields, then the capacitance to ground can be very large (from 70 to 100 pF/m, depending on the diameter of the cable). This capacitance can be reduced by using driven shields (or guards). If twisted pairs are used instead, the capacitance between wires is only about 5–8 pF/m, but there is an increased risk of capacitive interference; this interference, however, can be rejected by a balanced differential signal-to-amplifier connection.

Signal conditioners wired to remote sensors or working in environments where voltage or current surges are common, such as in automotive vehicles, or when there is risk of an electrostatic discharge (ESD), must be protected by *limiting* both *voltage* and *input currents*. Current can be limited by inserting a power resistor (e.g., 100 Ω–1 kΩ, 1 W), a polymeric positive temperature coefficient (PPTC) resistor, or a fuse between each signal source lead and conditioner input terminal; PPTC devices are resettable which is an advantage with respect to fuses that must be replaced when blown out. Current limiters increase the voltage-loading effect because they are connected in series to the internal impedance of the signal source.

Input voltages can be limited by connecting reverse-biased diodes, zeners, metal oxide varistors (MOV), gas-discharge devices, or other surge-suppression nonlinear devices, from each input line to dc power supply lines or to ground, depending on the particular protecting device. Voltage limiters reduce the conditioner input impedance because they are connected in parallel to it. Some devices, such as PolyZen® from Tyco Electronics, incorporate a Zener diode for voltage clamping and a PPTC resistor that responds to either diode heating or overcurrent events by transitioning from a low to high resistance state. Fault detection circuits intended to detect broken wires also add resistors from the conditioner input terminal to the supply voltages, which also limits their input resistance.

The ideal amplifier would also have zero *output impedance*. This would imply no loading effect because of possible finite input impedance for the following stage, low-output noise, and unlimited output power. Practical amplifiers can indeed have a low-output impedance and low noise, but their output power is very limited. Common signal amplifiers provide at best about 40 mA output current and sometimes only 10 mA, or even 1 mA in low-power amplifiers. The power gain, however, is quite noticeable, as input

currents can be in the picoampere range (10^{-12} A) and input voltages in the millivolt range (10^{-3} V); a 10 V, 10 mA output would mean a power gain of 10^{14}! Yet the output power available is very small (100 mW). Power amplifiers are quite the opposite; they have a relatively small power gain but provide a high-power output. For both signal and power amplifiers, output power comes from the power supply, not from the input signal, and maximal and minimal input and output voltage levels are determined by power supply voltages. The headroom between power supply voltages and maximal and minimal input and output voltages can be from 2 to 3 V in devices built with bipolar technologies, to a few millivolts in the so-called *rail-to-rail input* (RRI), *rail-to-rail output* (RRO), or *rail-to-rail input and output* (RRIO) amplifiers.

Some sensor signals do not require amplification but only *impedance transformation*, for example, to match their output impedance to that of a low-impedance ADC. Amplifiers for impedance transformation (or matching) and $G = 1$ are called *voltage buffers*.

91.5 Instrumentation Amplifiers

For instrumentation signals, the so-called *instrumentation amplifier* (IA, *ins amp*) offers performance closest to the ideal amplifier, at a moderate cost (from about $1.00 up). Figure 91.4a shows the symbol for the IA and Figure 91.4b its input/output relationship; ideally this is a straight line with slope G and passing through the point (0,0), but actually it is an off-zero, seemingly straight line, whose slope is somewhat different from G. The output voltage is

$$v_o = v_a + (V_{os} + V_b + v_r + v_n)G + V_{ref} \tag{91.5}$$

where

- v_a depends on the input voltage v_d, the second term includes offset, drift, noise, and interference-rejection errors
- G is the designed gain, set by an external resistor or a single connection between IC pins
- V_{ref} is the reference voltage, commonly 0 V (but not necessarily, thus allowing output level shifting)

Equation 91.5 describes a worst-case situation where absolute values for error sources are added. In practice, error sources will seldom achieve their worse value simultaneously. Nevertheless, if systematic effects are not calibrated out, the accuracy will seldom suit a 10 bit ADC. Calibration improves accuracy to about 14 bit for very low-frequency signals and 12 bit for ac signals.

Figure 91.5 shows a circuit model for *error analysis* when a practical IA is connected to a signal source (assumed to be differential for completeness). Impedance from each input terminal to ground (Z_c) and between input terminals (Z_d) are all finite. Furthermore, if the input terminals are both connected to (signal) ground, v_o is not zero and depends on G; this is modeled by V_{os}. If the input terminals are grounded through resistors, then v_o also depends on the value of these resistors; this is modeled by current sources

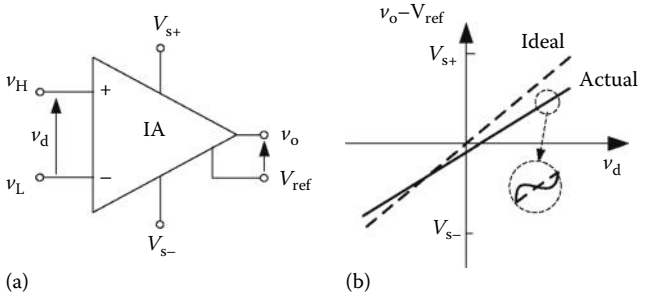


FIGURE 91.4 Instrumentation amplifier. (a) Symbol and (b) ideal and actual input/output relationship. The ideal response is a straight line through the point (0,0) and slope G, the gain.

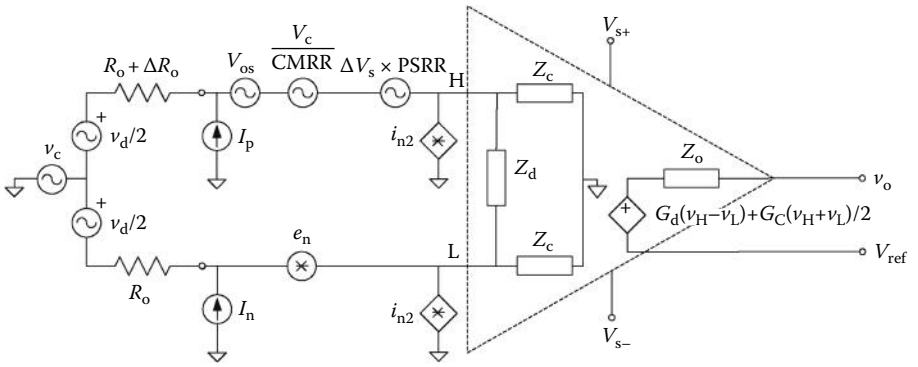


FIGURE 91.5 A circuit model for a practical instrumentation amplifier including major error sources.

I_p and I_n, which represent input bias or leakage currents. These currents need a return path, and therefore a third lead is required that connects the signal source to the amplifier, or to a common ground in the circuit. Neither V_{os} nor I_p nor I_n are constant; rather, they change with temperature and time: slow changes (<0.01 Hz) are called drifts and fast changes are described as noise (hence the noise sources e_n, i_{n1}, and i_{n2} in Figure 91.5). Common specifications for instrumentation amplifiers are defined in Ref. [2].

If a voltage v_c is simultaneously applied to both inputs, then v_o depends on v_c and its frequency. The *common-mode gain* is

$$G_c(f) = \frac{v_o(f)\big|_{v_d=0}}{v_c(f)} \tag{91.6}$$

v_c will be acceptable as long as $v_c \times G_c < 1$ LSB. In order to describe the output voltage due to v_c as an input error voltage, we must divide the corresponding v_o (when $v_d = 0$) by G (the normal or differential-mode gain, $G = G_d$). The *common-mode rejection ratio* (CMRR) is defined as

$$\text{CMRR} = \frac{G_d(f)}{G_c(f)} \tag{91.7}$$

and is usually expressed in decibels ($\{\text{CMRR}\}_{dB} = 20 \log \text{CMRR}$). Note that CMRR is not a gain but a quotient of gains, and it that it will depend on the frequency of the common-mode signal, not on that of the (differential-mode) signal. The input error voltage will be

$$\frac{v_o\big|_{v_c}}{G_d} = \frac{G_c v_c}{G_d} = \frac{v_c}{\text{CMRR}} \tag{91.8}$$

In the aforementioned analysis, we have assumed $Z_c \gg R_o$; otherwise, if there were any unbalance (such as ΔR_o for the source impedance in Figure 91.5), v_c at the voltage source would result in a differential-mode voltage at the amplifier input,

$$v_d\big|_{v_c} = v_c\left(\frac{R_o + \Delta R_o}{Z_c + R_o + \Delta R_o} - \frac{R_o}{Z_c + R_o}\right) = v_c \frac{z_c \Delta R_o}{(Z_c + R_o + \Delta R_o)(Z_c + R_o)} \approx v_c \frac{\Delta R_o}{Z_c} \tag{91.9}$$

which would be amplified by G_d. Then, the *effective CMRR* would be

$$\frac{1}{\text{CMRR}_e} = \frac{\Delta R_o}{Z_c} + \frac{1}{\text{CMRR}} \tag{91.10}$$

where CMRR is that of the IA alone, expressed as a fraction, not in decibels. A large $CMRR_e$ requires large common-mode input impedances Z_c relative to the signal source imbalance ΔR_o. Often, the CMRR for the amplifier is specified for $\Delta R_o = 1\ k\Omega$, which means that if the actual imbalance is smaller, the effective CMRR will be better than that specified. Stray capacitances from input terminals to ground will decrease Z_c, therefore reducing $CMRR_e$.

The ideal amplifier is unaffected by power supply fluctuations. The practical amplifier shows output fluctuations when supply voltages change. For slow changes ΔV_S, the equivalent input error can be expressed as a change in input offset voltages in terms of the *power supply rejection ratio* (PSRR):

$$\text{PSRR} = \frac{\Delta V_{os}}{\Delta V_S} \tag{91.11}$$

The terms in Equation 91.5 can be detailed as follows. Because of gain errors, we have

$$v_a = v_d\left(G + e_G + \frac{\Delta G}{\Delta T} \times \Delta T + e_{NLG}\right) \tag{91.12}$$

where

G is the differential gain designed (assumed to be constant in the bandpass for v_d)
e_G is its absolute error
$\Delta G/\Delta T$ is its thermal drift
ΔT is the difference between the actual temperature and that at which the gain G is specified
e_{NLG} is the nonlinearity gain error, which describes the extent to which the input/output relationship deviates from a straight line (insert in Figure 91.4b)

The actual (working) temperature T_J is calculated by adding to the current ambient temperature T_A the temperature rise produced by the power P_D dissipated in the device. This rise depends on the thermal resistance θ_{JA} for the case

$$T_J = T_A + P_D \times \theta_{JA} \tag{91.13}$$

where P_D can be calculated from the respective voltage and current supplies

$$P_D = |V_{S+}||I_{S+}| + |V_{S-}||I_{S-}| \tag{91.14}$$

The terms for the equivalent *input offset error* will be

$$V_{os} = V_{os}(T_a) + \frac{\Delta V_{os}}{\Delta T} \times (T_J - T_a) \tag{91.15}$$

$$V_b = (I_p - I_n)\ R_o + I_p \Delta R_o = I_{os} R_o + I_b \Delta R_o \tag{91.16}$$

where

T_a is the ambient temperature in data sheets
$I_{os} = I_p - I_n$ is the offset current
$I_b = (I_p + I_n)/2$

and all input currents must be calculated at the actual temperature,

$$I = I(T_a) + \frac{\Delta I}{\Delta T} \times (T_J - T_a) \tag{91.17}$$

Temperature must be corrected because offset voltage and current specifications for instrumentation amplifiers are measured shortly after turning the device on, hence before it warms up.

Error contributions from finite *interference rejection* are

$$v_r = \frac{v_c}{\text{CMRR}_e} + \Delta V_S \times \text{PSRR} \tag{91.18}$$

where

the CMRR_e must be that at the frequency for v_c

the PSRR must be that for the frequency of the ripple ΔV_S

It is assumed that both frequencies fall inside the bandpass for the signal of interest v_d.

The equivalent *input voltage noise* is

$$v_n = \sqrt{e_n^2 B_e + i_{n1}^2 R_o^2 B_{i1} + i_{n2}^2 R_o^2 B_{i2}} \tag{91.19}$$

where

e_n^2 is the voltage noise power spectral density of the IA

i_{n1}^2 and i_{n2}^2 are the current noise power spectral densities for each input of the IA (normally $i_{n1}^2 = i_{n2}^2$, and to simplify they are assumed to be uncorrelated, even though they are not)

B_e, B_{i1}, and B_{i2} are the respective noise equivalent bandwidths of each noise source

In Figure 91.5, the transfer function for each noise source is the same as that of the signal v_d, but this is not always the case, for example, when signals are ac-coupled to the IA using a high-pass filter. If the signal bandwidth is determined as $f_H - f_L$ by sharp filters, then

$$B_e = f_H - f_L + f_{ce} \ln \frac{f_H}{f_L} \tag{91.20}$$

$$B_{i1} = B_{i2} = f_H - f_L + f_{ci} \ln \frac{f_H}{f_L} \tag{91.21}$$

where f_{ce} and f_{ci} are, respectively, the frequencies where the value of voltage and current noise power spectral densities is twice their value at high frequency, also known as corner or –3 dB frequencies.

Another noise specification method for instrumentation amplifiers states the peak-to-peak noise at a given low-frequency band (f_A to f_B), usually 0.1–10 Hz, and the noise spectral density at a frequency at which it is already constant, normally 1 or 10 kHz. In these cases, if the contribution from noise currents is negligible, the equivalent input voltage noise can be calculated from

$$v_n = \sqrt{v_{nL}^2 + v_{nH}^2} \tag{91.22}$$

where v_{nL} and v_{nH} are, respectively, the voltage noise in the low-frequency and high-frequency bands expressed in the same units (peak-to-peak or rms voltages). To convert rms voltages into

peak-to-peak values, multiply by 6.6. If the signal bandwidth is from f_L to f_H, and $f_L = f_A$ and $f_H > f_B$, then Equation 91.22 can be written

$$v_n = \sqrt{v_{nL}^2 + (6.6e_n)^2(f_H - f_B)} \tag{91.23}$$

where

v_{nL} is the specified peak-to-peak value
e_n is the rms voltage noise as specified in data books

Equation 91.23 results in a peak-to-peak calculated noise that is somewhat lower than the real noise, because noise spectral density is not constant from f_B up to f_H. However, it is a simple approach providing an useful estimate.

For signal sources with high-output resistors, thermal and excess noise from resistors must be included. For first- and second-order filters, noise bandwidth is slightly larger than signal bandwidth. Motchenbacher and Connelly [3] show how to calculate noise bandwidth, resistor noise, and noise transfer functions when different from signal transfer functions.

Low-noise design always seeks the minimal bandwidth required for the signal. When amplifying low-frequency signals, if a large capacitor C_i is connected across the input terminals in Figure 91.5, then noise and interference having a frequency larger than $f_0 = 1/(2\pi 2R_oC_i)$ ($f_0 \ll f_s$) will be attenuated. As for low-frequency noise, minimal signal bandwidth does not necessarily ensure minimal noise. Hence, ac coupling, for example to cancel out offset, must consider that the noise bandwidth for input noise currents is not the same as the signal bandwidth.

Another possible source of error for any IA, not included in Equation 91.5, is the *slew rate limit* of its output stage. Because of the limited current available inside the IA, the voltage at the output terminal cannot change faster than a specified value SR. Then, if the maximal amplitude A of an output sine wave of frequency f exceeds

$$A = \frac{SR}{2\pi f} \tag{91.24}$$

the waveform will be distorted near zero crossings.

Table 91.1 lists some basic specifications for IC instrumentation amplifiers whose gain G can be set by an external resistor or a single connection.

91.5.1 Instrumentation Amplifiers Built from Discrete Parts

Instrumentation amplifiers can be built from discrete parts by using operational amplifiers (op amps) and a few resistors. By so doing, the amplifier bandwidth can be better tailored to the signal bandwidth, and some shortcomings of IC instrumentation amplifiers can be overcome, such as limited high-frequency gain or slew rate. An *op amp* is basically a differential voltage amplifier whose gain A_d is very large (from 10^5 to 10^7) at dc and rolls off (20 dB/decade) from frequencies of about 1 to 100 Hz, becoming 1 at frequencies from 1 to 10 MHz for common models (Figure 91.6a), and whose input impedances are so high (up to $10^{12}\ \Omega||1$ pF) that input currents are almost negligible. Op amps can also be modeled by the circuit in Figure 91.5, and their symbol is that in Figure 91.4a, either deleting IA or replacing it by OA. However, because of their large gain, op amps cannot be used directly as amplifiers; a mere 1 mV dc input voltage would saturate any op-amp output. Furthermore, op-amp gain changes from unit to unit, even for the same model, and for a given unit it changes with time, temperature, and supply voltages.

TABLE 91.1 Basic Specifications for Some Instrumentation Amplifiers

	AD624A	AMP02F	INA110KP	LT1101AC	Units
Gain range	1–1000	1–1000	1–500	10,100	V/V
Gain error, e_G					
$G = 1$	±0.05	0.05	±0.02	n.a.	%
$G = 10$	n.s.	0.40	±0.05	±0.04	%
$G = 100$	±0.25	0.50	±0.10	±0.04	%
$G = 1000$	±1.0	0.70	n.a.	n.a.	%
Gain nonlinearity error e_{NLG}[a]					
$G = 1$	±0.005	0.006	±0.005	n.a.	%
$G = 10$	n.s.	0.006	±0.005	±0.0008	%
$G = 100$	±0.005	0.006	±0.01	±0.0008	
$G = 1000$	±0.005	0.006	n.a.	n.a.	%
Gain drift $\Delta G/\Delta T$					
$G = 1$	5	50	±10	n.a.	μV/V/°C
$G = 10$	n.s.	50	±10	5	μV/V/°C
$G = 100$	10	50	±20	5	μV/V/°C
$G = 1000$	25	50	n.a.	n.a.	μV/V/°C
V_{os}	200 + 5/G	200	±(1000 + 5000/G)	160	μV
$\Delta v_{os}/\Delta T$	2 + 50/G	4	±(2 + 50/G)	2	μV/°C
I_B	±50	20	0.05	10	nA
$\Delta I_B/\Delta T$	±50 typ	250 typ	[b]	30	pA/°C
I_{os}	±35	10	0.025	0.90	nA
$\Delta I_{os}/\Delta T$	±20 typ	15 typ	n.s.	7.0	pA/°C
Z_d	1 || 10 typ	10 typ	5000 || 6 typ	12	GΩ
Z_c	1 || 10 typ	16.5 typ	2000 || 1 typ	7	GΩ
CMRR at dc					
$G = 1$	70 min	80 min	70 min	n.a.	dB
$G = 10$	n.s.	100 min	87 min	82	dB
$G = 100$	100 min	115 min	100 min	98	dB
$G = 1000$	110 min	115 min	n.a.	n.a.	dB
PSRR at dc					
$G = 1$	70 min	80 min	[c]	n.a.	dB
$G = 10$	n.s.	100 min	[c]	100	dB
$G = 100$	95 min	115 min	[c]	100	dB
$G = 1000$	100 min	115 min	n.a.	n.a.	dB
Bandwidth (−3 dB) (typ)					
$G = 1$	1000	1200	2500	n.a.	kHz
$G = 10$	n.s.	300	2500	37	kHz
$G = 100$	150	200	470	3.5	kHz
$G = 1000$	25	200	n.a.	n.a.	kHz
Slew rate (typ)	5.0	6	17	0.1	V/μs
Settling time to 0.01%					
$G = 1$	15 typ	10 typ	12.5	n.a.	μs
$G = 10$	15 typ	10 typ	7.5	n.a.	μs
$G = 100$	15 typ	10 typ	7.5	n.a.	μs
$G = 1000$	75 typ	10 typ	n.a.	n.a.	μs

(*continued*)

TABLE 91.1 (continued) Basic Specifications for Some Instrumentation Amplifiers

	AD624A	AMP02F	INA110KP	LT1101AC	Units
e_n (typ)					
$G = 1$	4	120	66	n.a.	nV/√Hz
$G = 10$	4	18	12	43	nV/√Hz
$G = 100$	4	10	10	43	nV/√Hz
$G = 1000$	4	9	n.a.	n.a.	nV/√Hz
v_n 0.1–10 Hz (typ)					
$G = 1$	10	10	1	0.9	μVp-p
$G = 10$	n.s.	1.2	1	0.9	μVp-p
$G = 100$	0.3	0.5	1	0.9	μVp-p
$G = 1000$	0.2	0.4	1	0.9	μVp-p
i_n 0.1–10 Hz (typ)	60	n.s.	n.s.	2.3	pAp-p
i_n (typ)	n.s.	400	1.8	20	fA/√Hz

Note: All parameter values are maximum, unless otherwise stated (typ, typical; min, minimum; n.a., not applicable; n.s., not specified). Measurement conditions are similar; consult manufacturers' data books for further detail.

[a] For the INA110, the gain nonlinearity error is specified as percentage of the full-scale output.

[b] Input current drift for the INA110KP approximately doubles for every 10 °C increase, from 25 °C (10 pA-typ) to 125 °C (10 nA-typ).

[c] The PSRR for the INA110 is specified as an input offset ±(10 + 180/G) μV/V maximum.

Nevertheless, external negative feedback around op amps yields a variety of analog circuits with excellent performance, thus rendering op amps very flexible. They are also cheaper than instrumentation amplifiers as they cost below $0.20 for packages with four units with modest performance. But when the cost for external feedback components and their connections and overall reliability are also considered, the optimal solution depends on the situation.

Figure 91.6b shows an amplifier built from an op amp with external feedback. If input currents are neglected, the current through R_2 will flow through R_1 and we have

$$v_d = v_i - v_o \frac{R_1}{R_1 + R_2} \tag{91.25}$$

$$v_o = A_d v_d \tag{91.26}$$

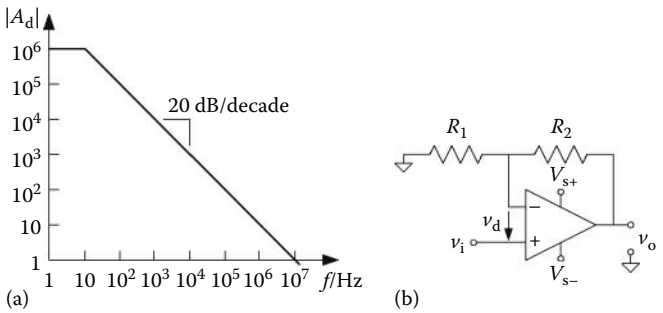


FIGURE 91.6 (a) Open-loop gain for an op amp and (b) voltage amplifier based on an op amp with external feedback.

Therefore,

$$\frac{v_o}{v_i} = \frac{A_d\left(1+\dfrac{R_2}{R_1}\right)}{A_d+1+\dfrac{R_2}{R_1}} = \frac{G_i}{1+\dfrac{G_i}{A_d}} \tag{91.27}$$

where $G_i = 1 + R_2/R_1$ is the ideal gain for the amplifier. If G_i/A_d is small enough (G_i small, A_d large), the gain does not depend on A_d but only on external components. At high frequencies, however, A_d becomes smaller and, from Equation 91.7, $v_o < G_j v_i$ so that the bandwidth for the amplifier will reduce for large gains. Franco [4] analyzes different op-amp circuits useful for signal conditioning.

Figure 91.7 shows an IA built from three op amps in two stages. The input stage is a fully differential amplifier and the output stage is a *difference amplifier* converting a differential voltage into a single-ended output voltage. Difference amplifiers (op amp and matched resistors) are available in IC form from several manufactures (Analog Devices, National Semiconductor, Texas Instruments). The gain equation for the complete IA is

$$G = \left(1+2\frac{R_2}{R_1}\right)\frac{R_4}{R_3} \tag{91.28}$$

Pallas-Areny and Webster [5] have analyzed matching conditions in order to achieve a high CMRR. Resistors R_2 do not need to be matched but the input op amps must. Resistors R_3 and R_4 need to be closely matched, hence the interest on an IC difference amplifier. If R_1 can be selected from a resistor network, for example by analog switches, we have a *programmable-gain amplifier* (PGA); *digitally controlled potentiometers*, available from several manufactures, include such an array of resistors and analog switches.

The *three-op-amp IA* has a symmetrical structure making it easy to design and test. Instrumentation amplifiers based on an IC difference amplifier do not need any user trim for high CMRR. The circuit in Figure 91.8 is an IA that lacks these advantages but uses only two op amps. Its gain equation is

$$G = 1+\frac{R_2}{R_1}+\frac{2R_2}{R_G} \tag{91.29}$$

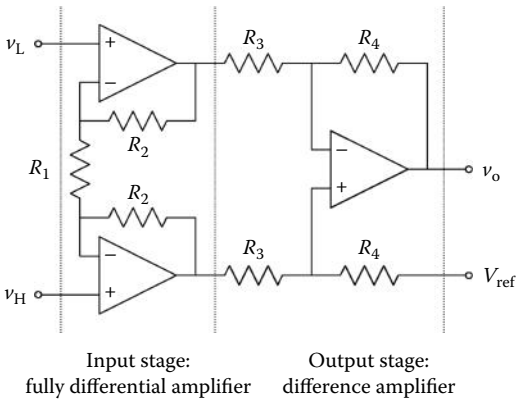


FIGURE 91.7 Instrumentation amplifier built from three op amps. The two input op amps, and R_3 and R_4 must be matched. IC difference amplifiers include an op amp and matched R_3 and R_4. This amplifier can be replaced by a LT1167 amplifier.

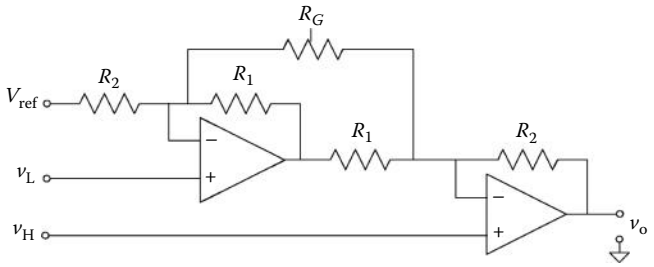


FIGURE 91.8 Instrumentation amplifier built from two op amps. R_1 and R_2 must be matched.

R_1 and R_2 must be matched and R_G should be comparable to R_2. Noise in instrumentation amplifiers built from discrete op amps may exceed that from integrated IAs because noise from separate op amps will be uncorrelated.

91.5.2 Composite Instrumentation Amplifiers

Instrumentation amplifiers have a gain limited to about 10,000, at low frequencies, and a bandwidth around 4 MHz for a gain of 10. Moreover, their inputs must be either dc-coupled or, if ac-coupled with input series capacitors, there must be a path for bias currents; if that path is a resistor from each input to ground, then the common-mode input impedance Z_c decreases and interference from the common-mode voltage increases.

Bandwidth can be increased by cascade connection of two or more amplifiers. However, if the additional gain is provided by a single-ended amplifier after the IA, then the overall CMRR is that of the IA, which is small at high frequencies. The circuit in Figure 91.9a is a broadband IA with large CMRR because the CMRR for the second stage is multiplied by the differential gain for the first stage, which can be very high if implemented by broadband op amps. The overall gain is

$$G = G_1 G_2 = \left(1 + \frac{2R_b}{R_a}\right) G_{IA} \tag{91.30}$$

The input stage can also be used as fully differential amplifier for differential-input ADCs.

An IA can be *ac-coupled* by feeding back its dc output to the reference terminal as shown in Figure 91.9b. The high-pass corner frequency is $f_0 = 1/(2\pi R_0 C_0)$. However, the dc level of the input signal will be amplified by the gain of the IA's first stage, whose output can saturate. Alternatively, a passive fully

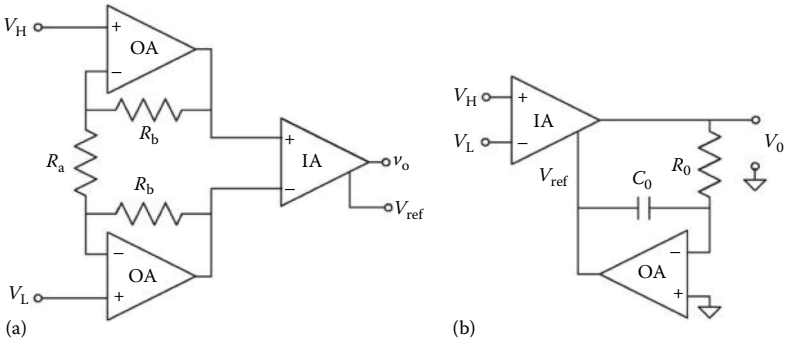


FIGURE 91.9 Composite instrumentation amplifiers. (a) Broadband IA with large CMRR built by preceding an IA by a fully differential amplifier and (b) the ac-coupled IA; the maximal dc level of the input voltage is limited by the gain of the input stage.

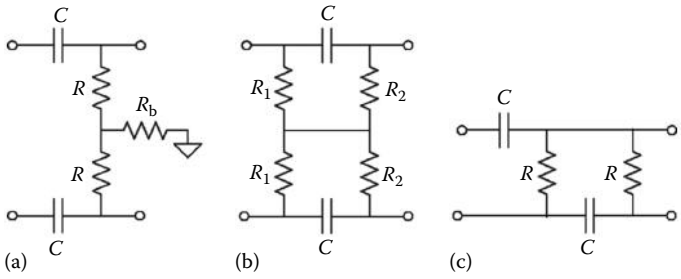


FIGURE 91.10 Fully differential ac-coupling networks, with first-order response (a) and (b), and second-order response (c).

differential high-pass filter can be placed in front of the IA, either in Figure 91.9b or in Figure 91.9a, which has a higher input impedance than that in Figure 91.9b. Figure 91.10 shows three alternative ac-coupling networks that do not need any component matching to achieve a large CMRR. Networks (a) and (b) are first-order filters, whereas network (c) is a second-order filter. R_b in circuit (a) and resistors R_1 in network (b) provide a means to bias the IA correctly; R_b reduces the CMRR, and R_1 limits the input impedance. Networks (b) and (c) yield the best CMRR because they have no connections to signal ground. Network (b) has the longer transient response to differential signals (damping ratio $\zeta = 1.5$), whereas the transient response of network (a) to common-mode input signals depends on component tolerance. Reference [6] compares the performance of each network when the limited input common-mode input impedance and input bias currents of the ensuing differential amplifier are considered.

91.6 Single-Ended Signal Conditioners

Single-ended floating signals can be connected to amplifiers with single-ended grounded input. Single-ended grounded signals can be connected to *single-ended amplifiers* provided the difference in ground potentials from signal to amplifier is not too large. Otherwise, a differential amplifier must be connected between signal terminals; thus, the voltage drop between (signal and amplifier) ground potentials becomes a common-mode voltage to the amplifier.

91.6.1 Inverting and Noninverting Voltage Amplifiers

Figure 91.11a shows a simple single-ended inverting amplifier based on an IA. However, op amps are normally better suited than IAs for single-ended amplifiers and signal conditioners that must perform functions additional or alternative to voltage amplification, such as filtering or integration. Figure 91.11b shows an op amp–based *inverting amplifier* whose gain is $G = -R_2/R_1$, and whose input impedance is R_1,

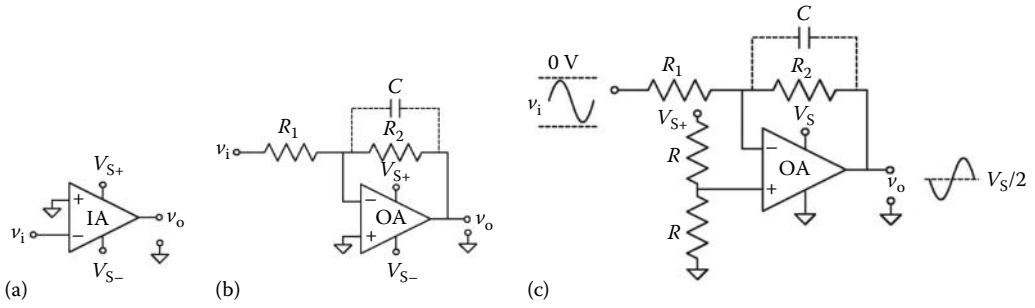


FIGURE 91.11 Single-ended inverting voltage amplifiers based on (a) an IA, (b) an op amp with split power supply, and (c) an op amp with a single power supply.

much smaller than that of the inverting amplifier built from an IA (Figure 91.11a). The capacitor on the dashed line (10 pF or larger) prevents gain peaking and oscillation. Signals with frequency larger than $f_c = 1/(2\pi R_2 C)$ are attenuated at a rate of 20 dB/decade; attenuation at f_c is 3 dB. If a capacitor C replaces R_2, input voltages are integrated and inverted. If C replaces R_1, instead, input voltages are differentiated and inverted. If the op amp is supplied by a single power source V_S instead of a split power source (V_{S+}, V_{S-}), the noninverting input must be biased to $V_S/2$, as shown in Figure 91.11c; the output voltage v_o can swing from about 0 V to about V_S, and will be close to $V_S/2$ when $v_i = 0$ V. For $v_i > 0$ V, the output will saturate. The same can be said for the inverting amplifier in Figure 91.11a if the IA is supplied between V_S and 0 V.

The circuit in Figure 91.6b is an op amp–based *noninverting amplifier* whose gain is $G = 1 + R_2/R_1$ and whose input impedance is close to the common-mode input impedance of the op amp, hence very large. By adding capacitors, undesired low- or high-frequency signals can be attenuated. For example, the amplifier in Figure 91.12a has $G = 1$ for dc signals, very low-frequency signals relative to $f_1 = 1/(2\pi R_1 C_1)$ (offset and drift included), and very high-frequency signals relative to $f_2 = 1/(2\pi R_2 C_2)$. The asymptotic gain in the passband is $G = 1 + R_2/R_1$; if f_2 is not at least 10 times larger than f_1, the actual gain will deviate from G by >10%. If the op amp is supplied by a single power source V_S, negative input voltages will saturate the output, and input voltages whose average value is relatively far from $V_S/2$ can result in a distorted output. Alternatively, if a single power supply is available, ac signals can be amplified with the circuit in Figure 91.12b. A noninverting amplifier can also be built from an IA as shown in Figure 91.12c; however, low-pass filtering must be then performed in later stages. The circuit in Figure 91.12d calculates the average for n voltages and amplifies it by $1 + R_2/R_1$. The difference between two voltages can be obtained from the difference amplifier in Figure 91.7 (output stage).

Op amps (and IAs) must be carefully selected according to the application. For dc circuits, *autozero op amps* (and autozero IAs) and *chopper-stabilized op amps* offer the best performance, with offset

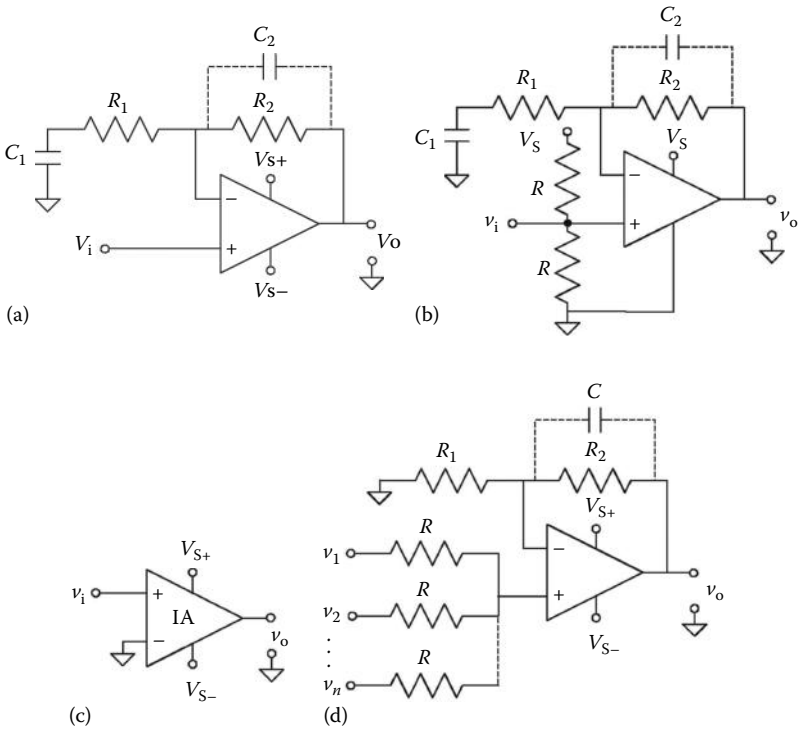


FIGURE 91.12 Single-ended noninverting amplifiers: ac amplifier based on (a) an op amp with split supply or (b) single supply, (c) dc amplifier based on an IA, and (d) voltage average amplifier.

voltages as low as 1–10 μV and temperature drifts below 0.1 μV/°C. For low-impedance signals, op amps with bipolar input transistors are better because of their lower offset voltage and voltage noise. For high-impedance signals, op amps with FET input transistors (J-FET or CMOS) offer lower input currents than bipolar amplifiers at room temperature, but they have larger drift and voltage noise, and their bias (input leakage) currents almost double each 10°C. Bipolar op amps offer wide power supply rails (+15 V/–15 V, +12 V/–12 V), whereas power supply rails for CMOS op amps seldom exceed +5 V/–5 V, yet these offer low-power consumption and there are low-supply voltage models, from 5 to 1.8 V. Manufacturers provide interactive selection guides in their websites that identify their most appropriate models for a set of user-defined values for some parameters.

91.6.2 Transimpedance Amplifiers

Current signals in instrumentation come from current-output sensors, such as photodiodes and some *p-n* junction–based temperature sensors (AD590 and the like), and also from 4 to 20 mA current loops used for signal transmission. Current can be sensed by inserting a resistor R in the current loop and measuring the drop in voltage across R using a single-ended or a differential voltage amplifier, depending on whether one of the resistor terminals is grounded or not. However, small currents may need $R = 100$ MΩ or even larger resistor values, which result in a measurement circuit with a large equivalent resistance, hence susceptible to capacitive interference as any small current coupled to the circuit will produce a large drop in voltage that can saturate it. A grounded shield enclosing the circuit reduces interference but at the cost of a large capacitance from the signal line to ground, which will reduce signal bandwidth. A transimpedance amplifier (Figure 91.13a) improves the bandwidth by placing the sensing resistor in the feedback loop of an op amp, whose input bias current must be negligible compared to the sensor current if this one includes low-frequency components. Stray capacitance from the sensor terminal to ground has a minor effect on bandwidth as the voltage at the inverting terminal is very close to signal ground; that stray capacitance, however, can induce oscillations, which can be prevented by shunting R with a capacitor C, in the range of picofarads. C will ultimately limit the signal bandwidth, but not that much as sensor and cable capacitance because these can be in the order of tens of nanofarads. In any case, the maximal "gain" I will be limited by the op-amp open-loop gain. Graeme [7] provides a thorough analysis of amplifiers for photodiodes.

91.6.3 Charge Amplifiers/Converters

Sensors whose equivalent circuit is a voltage source with an internal capacitive impedance are prone to interference the same as current-output sensors. That is the case for piezoelectric sensors and sensors based on a variation of capacitance. Measuring the voltage without incurring in an unacceptable voltage-loading effect requires an amplifier with very large input impedance. This is feasible at low

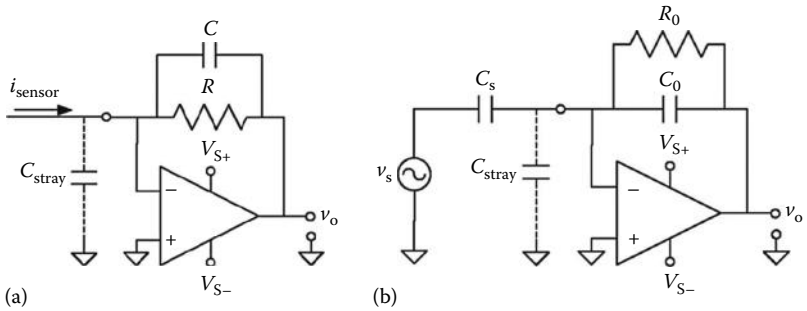


FIGURE 91.13 (a) Transimpedance amplifier to convert current to voltage and (b) charge amplifier/converter to convert charge to voltage.

frequencies, as a noninverting amplifier can provide it, but it is hard to achieve at medium to high frequencies because of stray capacitances to ground. The charge amplifier in Figure 91.13b is an alternative to voltage amplifiers: the sensor voltage yields a current through its internal capacitance C_s, and this current is converted into a voltage by flowing through C_0. The output voltage will be $v_o = -v_s C_s/C_0$, and its internal impedance will be very small as it corresponds to the op-amp output with negative feedback. Because there is no charge amplification but a charge transfer from C_s to C_0, this circuit is also called *charge converter*: it converts the charge across the sensor into a low-impedance voltage (v_o). R_0 provides a path for op-amp bias currents; it limits the (–3 dB) low-frequency response to $f_0 = 1/(2\pi R_0 C_0)$ and contributes to the output offset voltage ($I_b R_0$), albeit this offset can be canceled by high-pass coupling the output voltage to the next signal-processing stage.

91.6.4 Current Integrators

A current connected to the input of a charge amplifier will charge C_0 until the output voltage reaches the limits of the output signal range for the op amp, close to the power supply rails in the best case. The input–output relation if the integration time τ is short enough to not reach those limits will be

$$v_o(t) = -\frac{1}{C_0}\int_{t_1}^{t_1+\tau} i_i(t)dt \tag{91.31}$$

To obtain the response in the frequency domain, we assume $i_i(t) = I \sin\omega t$ to obtain

$$V_o(\omega) = -\frac{1}{C_0}\int_{t_1}^{t_1+\tau} I\sin\omega t\, dt = -\frac{I\omega}{C_0} 2\sin\left(\omega t_1 + \frac{\omega\tau}{2}\right)\sin\frac{\omega\tau}{2} \tag{91.32}$$

The absolute value of this voltage will be maximal when $\omega t_1 + \omega\tau/2 = \pi/2$. Then,

$$V_o(f) = \frac{I\tau}{C_0}\frac{\sin\pi f\tau}{\pi f\tau} = V_o(0)\frac{\sin\pi f\tau}{\pi f\tau} \tag{91.33}$$

where $V_o(0)$ is the dc amplitude response. Integration, it turns out, attenuates high-frequency signals, and those signals whose frequency is such that $f\tau$ is an integer, that is, multiple of the reciprocal of the integration time. The ability to discriminate signals of interest from those added to them is described by the *series* (or *normal*) *mode rejection ratio* (SMRR, NMRR), and is usually expressed in decibels. In the present case, if the signals of interest are those that remain constant during $\tau (f\tau \to 0)$,

$$\text{SMRR} = 20\log\frac{V_o(0)}{V_o(f_c)} = 20\log\frac{\pi f\tau}{\sin\pi f\tau} \tag{91.34}$$

which is represented in Figure 91.14a. Integrating ADCs have this same comblike response, which is used to reduce power-line interference.

Figure 91.14b shows the simplified equivalent circuit of an IC integrator; opening switch S1 stops the integration time and closing switch S2 resets the integration capacitor before starting to integrate the input current again. The ACF2101 and IVC102 are switched integrators that use this approach, the same as the DDC112, a dual input ADC that digitizes low-level currents by first converting them to voltage; current integration is continuous as each input uses two integrators: while one is being digitized, the other is integrating.

Table 91.2 lists some parameters for several different op amps. Some manufacturers provide selection guides which suggest the most appropriate model for a set of *r*-defined values for some parameters.

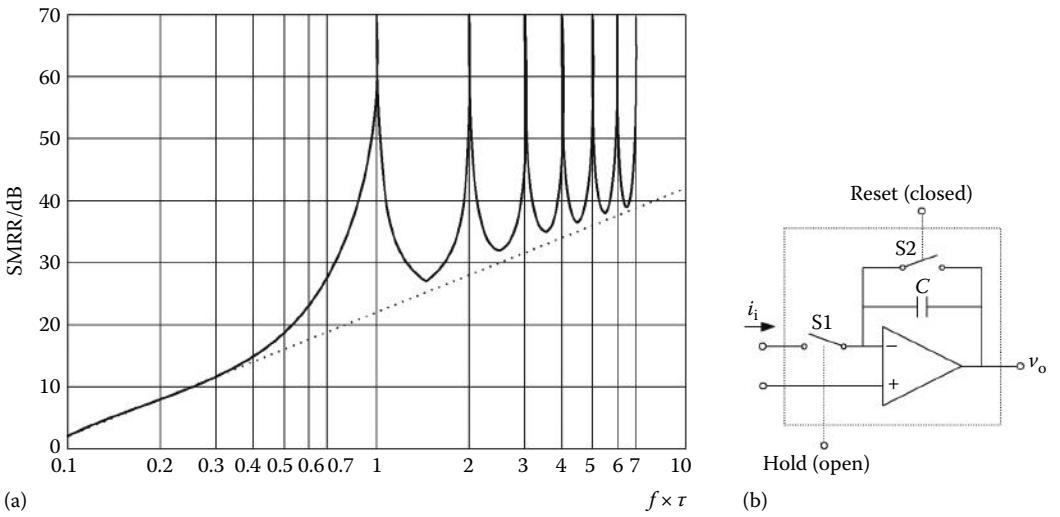


FIGURE 91.14 (a) Series-mode rejection ratio (SMRR) for an integrator: interference rejection is maximal for frequencies multiple of the reciprocal of the period of integration; (b) simplified schematic for a switching IC current integrator.

91.7 Carrier Amplifiers

A *carrier amplifier* is a conditioner for extremely narrowband ac signals from ac-driven sensors. A carrier amplifier is made of a sine wave oscillator, to excite the sensor (either directly as in LVDTs, or placed in a voltage divider or sensor bridge, as in sensors based in a variable capacitance, inductance, or resistance), an ac voltage amplifier for the sensor output, a synchronous demodulator, and a low-pass filter (Figure 91.15). The AD598, AD698, and NE5520/1 are carrier amplifiers in IC form intended for (but not limited to) LVDTs driven at a frequency from 2 to 20 kHz.

Carrier amplifiers make it possible to recover the amplitude and phase of the modulating signal after amplifying the output modulated waveform from the bridge or voltage divider. This is useful first because ac amplifiers are not affected by offset, drift, or low-frequency noise, and therefore the bridge or voltage-divider output can easily be amplified. Second, the *phase-sensitive demodulator* yields not only the amplitude but also the sign of the measurand. If the measurement range includes positive and negative values for the measurand, phase detection is essential.

A further advantage of carrier amplifiers is their extremely narrow frequency response, determined by the output low-pass filter. In the demodulator, the product of the modulated carrier of frequency f_c by the reference signal, also of frequency f_c, results in a baseband component and components at nf_c ($n \geq 2$). The output low-pass filter rejects components other than the baseband. If the corner frequency for this filter is f_0, then the passband for the system is $f_c \pm f_0$. Therefore, any interference of frequency f_i added to the modulated signal will be rejected if falling outside that passband. The SMRR when the output low-pass filter is Butterworth with order n is

$$\text{SMRR} = 20\log\frac{|v_o(f_c)|}{|v_o(f_i)|} = 20\log\sqrt{1+\left(\frac{f_c - f_i}{f_0}\right)^{2n}} \approx 20n\log\frac{|f_c - f_i|}{f_0} \tag{91.35}$$

A power-line interference superimposed on a 10 kHz carrier will undergo an 80 dB attenuation if the output low-pass filter is first order and has $f_0 = 1$ Hz. The same interference superimposed on the baseband signal would be attenuated by only 35 dB.

TABLE 91.2 Basic Specifications for Operational Amplifiers of Different Technologies

	V_{os}, μV	$(\Delta v_{os}/\Delta T)_{av}$, μV/°C	I_B, pA	$\Delta I_B/\Delta T$, pA/°C	I_{os}, pA	BW_{typ} ($G = 1$), MHz	e_n (1 kHz), nV$\sqrt{\text{Hz}}$	f_{ce}, Hz	$v_{n(p\text{-}p)}$, μV	i_n (1 kHz), fA/$\sqrt{\text{Hz}}$
Bipolar										
μA741	6,000	15	500,000	500[a]	200,000	1.5	20	200	—	550
LM358A	3,000	20	100,000	—	±30,000	1	—	—	—	—
LT1028	80	0.8	180,000	—	10,0000	75	0.9	3.5	0.035	1000
OP07	75	1.3	3,000	50[a]	2,800	0.6	9.6	10	0.35	170
OP27C	100	1.8	80,000	—	75,000	8	3.2	2.7	0.09	400
OP77A	25	0.3	2,000	25[a]	1,500	0.6	9.6	10	0.35	170
OP177A	10	0.1	1,500	25[a]	1,000	0.6	—	—	0.8	—
TLE2021C	600	2	70,000	80[a]	3,000	1.2	30	—	0.47	90
TLE2027C	100	1	90,000	—	90,000	13	2.5	—	0.05	400
FET input										
AD549K	250	5	0.1	[b]	0.03 typ	1	35	—	4	0.16
LF356A	2,000	5	50	[b]	10	4.5	12	—	—	10
OPA111B	250	1	1	[b]	0.75	2	7	200	1.2	0.4
OPA128J	1,000	20	0.3	[b]	65	1	27	—	4	0.22
TL071C	10,000	18	200	[b]	100	3	18	300	4	10
TLE2061C	3,000	6	4 typ	[b]	2 tip	2	60	20	1.2	1

CMOS										
ICL7611A	2	10 typ	50	[b]	30	0.044	100	800	—	10
LMC660C	6,000	1.3 typ	20	[b]	20	1.4	22	—	—	0.2
LMC6001A	350	10	0.025	[b]	0.005	1.3	22	—	—	0.13
TLC271CP	10,000	2 typ	0.7 typ	[c]	0.1 typ	2.2	25	100	—	n.s.
TLC2201C	500	0.1 typ	1 typ	[d]	0.5 typ	1.8	8	—	0.7	0.6
BiMOS										
CA3140	15,000	8	50	[b]	30	4.5	40	—	—	—
CMOS chopper										
LTC1052	5	0.05	30	[e]	30	1.2	—	—	1.5	0.6
LTC1150C	5	0.05	100	[f]	200	2.5	—	—	1.8	1.8
MAX430C	10	0.05	100	[g]	200	0.5	—	—	1.1	10
TLC2652AC	1	0.03	4 typ	[d]	2 typ	1.9	23	—	2.8	4
TLC2654C	20	0.3	50 typ	0.65	30 typ	1.9	13	—	1.5	4
TSC911A	15	0.15	70	—	20	1.5	—	—	11	—

Note: Specified values are maximal unless otherwise stated and those for noise, which are typical (typ, typical; av, average; nonspecified parameters are indicated by a dash).

[a] Values estimated from graphs.

[b] I_B doubles every 10 °C.

[c] I_B doubles every 7.25 °C.

[d] I_B is almost constant up to 85 °C.

[e] I_B is almost constant up to 75 °C.

[f] I_{B+} and I_{B-} show a different behavior with temperature.

[g] I_B doubles every 10 °C above about 65 °C.

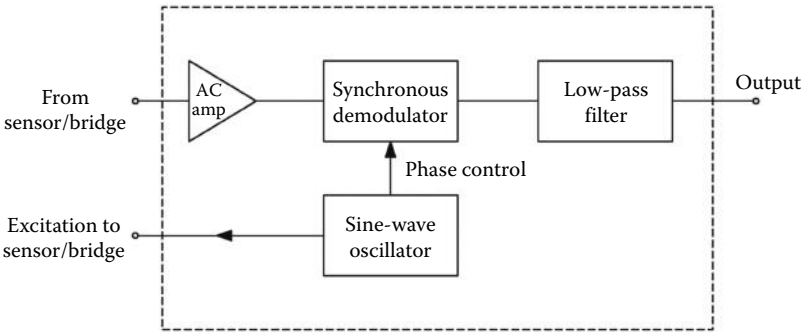


FIGURE 91.15 Elements for a carrier amplifier.

Carrier amplifiers can be built from a precision sine wave oscillator, and a demodulator (plus the output low-pass filter). The AD630 is an IC demodulator based on switched amplifiers. Alternatively, amplitude demodulation can be performed in the digital domain [8].

91.8 Lock-In Amplifiers

A *lock-in amplifier* is based on the same principle as a carrier amplifier, but instead of driving the sensor, here the carrier signal drives the experiment, so that the measurand is frequency translated. Lock-in amplifiers are manufactured as equipment intended for recovering signals immersed in high (asynchronous) noise. These amplifiers provide a range of driving frequencies and bandwidths for the output filter. Some models are vectorial because they make it possible to recover the in-phase and quadrature (90° out-of-phase) components of the incoming signal, by using two demodulators whose reference signals are delayed by 90°. Still other models use bandpass filters for the modulated signal and two demodulating stages. Meade [9] analyzes the fundamentals and specifications of analog lock-in amplifiers. Signal recovery (a division of Ametek, Inc.) offers, upon registration in its website, several application notes about analog and digital lock-in amplifiers.

91.9 Isolation Amplifiers

The maximal common-mode voltage withstood by common amplifiers is smaller than their supply voltage range and seldom exceeds 10 V. Exceptions are some difference amplifiers that integrate an input voltage divider and whose common-mode range is up to ±200 V. Signals with large off-ground voltages, for example, in voltage dividers or sensor bridges, or differences in ground potentials between signals and amplifiers exceeding the input common-mode range, result in permanent amplifier damage or destruction, and a safety risk, in spite of an exceptional CMRR: a 100 V common-mode 60 Hz voltage at the input of a common IA having a 120 dB CMRR at power-line frequency does not result in a $100\ V/10^6 = 100\ \mu V$ output, but a burned-out IA.

Figure 91.16a shows a signal source grounded at a point 1 far from the amplifier ground (point 2). Both ends of the signal loop have paths to ground, and the difference in voltage between grounds $v_i = v_2 - v_1$ forces a current into a signal line (and back to ground) that not only distorts the true measurement because of the drop in voltage across the nonnegligible lead impedance, but also can destroy the amplifier. Avoiding the *ground loop* by "lifting" one of the two ground connections would solve the problem but electrical safety codes mandate local earth ground connections. The solution is to break the ground loop to prevent voltage between grounds from forcing any large current through the signal circuit and at the same time to provide an information link between the signal source and the amplifier receiver. Figure 91.16b shows a solution: the signal source and the amplifier have separated (isolated) power supplies and the signal is coupled to the amplifier through a transformer acting as an isolation barrier for v_i;

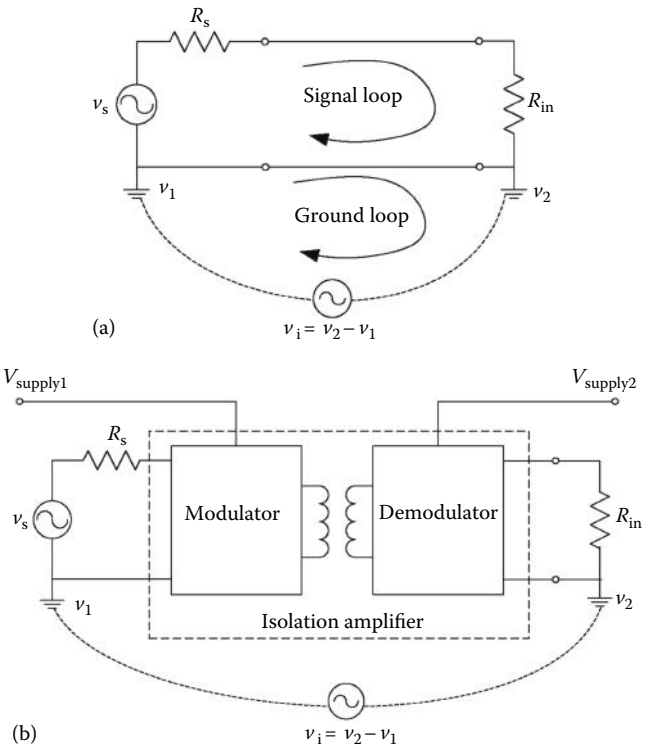


FIGURE 91.16 (a) A large difference in ground potentials damages amplifiers. (b) An isolation amplifier breaks the ground loop thus preventing large currents caused by the difference in ground potentials from flowing through the circuit.

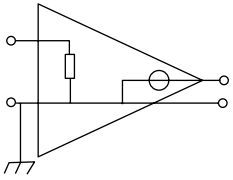
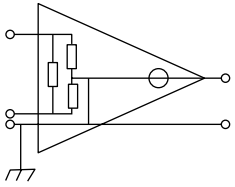
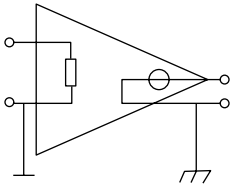
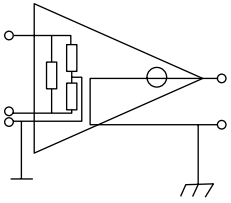
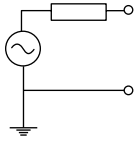
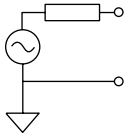
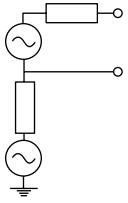
the only current forced by v_i will be that allowed by the stray capacitance between the transformer's primary and secondary windings. Other isolation barriers are optocouplers, series capacitors, and magnetoresistors. Those barriers impose a large series impedance (isolation impedance, Z_i) but do not usually have a good low-frequency response, hence the need to modulate and then demodulate the signal to be transferred through Z_i. The subsystem made of the modulator and demodulator, plus sometimes an input and an output amplifier and a dc–dc converter for the separate power supply, is called an *isolation amplifier*. The ability to reject the voltage difference across the barrier (isolation-mode voltage, v_i) is described by the *isolation mode rejection ratio* (IMRR), expressed in decibels:

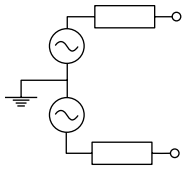
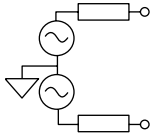
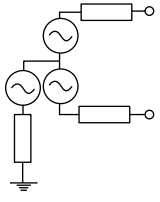
$$\text{IMRR} = 20 \log \frac{\text{Isolation-mode voltage}}{\text{Output voltage}} \tag{91.36}$$

Ground isolation also protects people and equipment from contact with high voltage because Z_i limits the maximal current. Some commercial isolation amplifiers are the AD20x series, and AD210/5 (Analog Devices), the ISO121/2/4 (Texas Instruments), and the HCPL-7850 and 5962-9755701 (Avago Technologies). Sometimes, it is more cost-effective to isolate digital rather than analog signals because digital isolators are far less expensive than isolation amplifiers. This means to postpone isolation after A-to-D conversion and implies a larger isolated power supply for the circuits previous to the isolation barrier.

Table 91.3 summarizes the compatibility between signal sources and amplifiers. When grounded, amplifiers and signals are assumed to be grounded at different physical points. Floating signals can be connected to any amplifier, provided other requirements are fulfilled; for example, if a single-ended

TABLE 91.3 Compatibility between Signal Sources and Conditioners

Conditioner → / Input signal source ↓				
	Incompatible unless grounds are very close	Compatible if CMRR is large	Compatible	Compatible
	Compatible	Compatible	Compatible	Compatible
	Incompatible unless grounds are very close	Compatible if CMRR is large	Compatible for large Z_i	Compatible

	Incompatible	Compatible	Compatible for large Z_i	Compatible
	Compatible	Compatible	Compatible	Compatible
	Incompatible	Compatible if CMRR is large	Compatible for large Z_i	Compatible

Note: When grounded, signals sources and amplifiers are assumed to be grounded at different points. Isolation impedance is assumed to be very high for floating signal sources but finite (Z_i) for conditioners. Single-ended signals are considered to be compatible with isolated differential conditioners because resistors can be included to bias the conditioner correctly.

floating signal is connected to a differential amplifier, a bias resistor must be connected from each amplifier input to ground because ground currents cannot flow to ground through the floating (isolated) signal ground. Conversely, isolation amplifiers can be connected to any signal, but a bias path must be provided if the signal source does not provide it.

91.10 Analog Linearization

Nonlinearity in instrumentation can result from the measurement principle, from the sensor, or from sensor conditioning. In pressure-drop flowmeters, for example, the drop in pressure measured is proportional to the square of the flow velocity; hence, flow velocity can be obtained by taking the square root of the pressure signal. Calculations are better performed using digital processors. Many sensors are linear only in a restricted measurand range; others are essentially nonlinear (NTC thermistors, LDR); still others are linear in some ranges but nonlinear in other ranges of interest (thermocouples). Linearization techniques for particular sensors are described in the respective chapters.

Nonlinearity attributable to sensor conditioning is common, for example, when resistive (linear) sensors are placed in voltage dividers or sensor bridges. The Wheatstone bridge in Figure 91.17a, for example, includes a (single) linear sensor but yields a nonlinear output voltage:

$$v_s = V_r\left(\frac{1+x}{2+x} - \frac{1}{2}\right) = \frac{V_r x}{2(2+x)} \tag{91.37}$$

The nonlinearity arises from the dependence of the current through the sensor on its resistance, because the bridge is supplied at a constant voltage. The circuit in Figure 91.17b provides a solution based on 1 op amp which forces a constant current V_r/R_0 through the sensor. The bridge output voltage is

$$v_s = \frac{V_r - v_a}{2} + v_a = \frac{V_r + v_a}{2} = V_r\frac{x}{2} \tag{91.38}$$

In addition, v_s is single-ended whereas in Figure 91.17a it was differential. The op amp must have a good dc performance and the output will need further amplification.

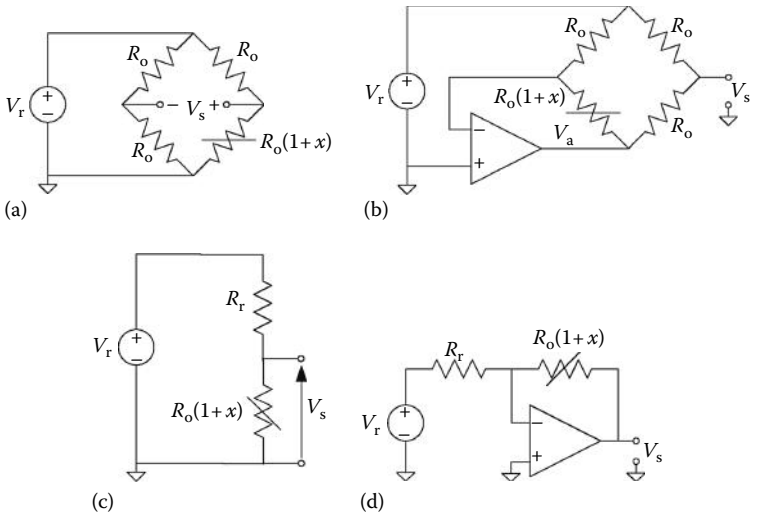


FIGURE 91.17 (a) A Wheatstone bridge with a single sensor yields a nonlinear output voltage; (b) by adding an op amp which forces a constant current through the sensor, the output voltage is linearized; (c) a voltage divider with a single linear sensor also yields a nonlinear output voltage; and (d) forcing a constant current through the sensor linearizes the output voltage.

The same applies to voltage dividers. Figure 91.17c shows a voltage divider with a single linear sensor, and whose output voltage is

$$v_s = V_r \frac{R_0(1+x)}{R_r + R_0(1+x)} = V_r \frac{1+x}{R_r/R_0 + 1 + x} \tag{91.39}$$

If $R_r/R_0 + 1 \gg x$, the output will be approximately linear, at the cost of a reduced sensitivity (which would require $R_r = R_0$). Forcing a constant current through the sensor, as shown in Figure 91.17d, would yield a linear output for any R_r and x.

If the sensor in Figure 91.17a and c was nonlinear, the nonlinearity of the Wheatstone bridge and voltage divider can be used in some cases to linearize the output voltage, to some extent. For example, if the sensor has a concave-upward calibration curve (such as NTC and PTC thermistors) and its values for x_1, x_2, and x_3, with $x_1 - x_2 = x_2 - x_3$, are R_{x1}, R_{x2}, and R_{x3}, in a voltage divider we only need to select

$$R_r = \frac{R_{x2}(R_{x1} + R_{x3}) - 2R_{x1}R_{x3}}{R_{x1} + R_{x3} - 2R_{x2}} \tag{91.40}$$

to obtain a linearized output voltage. In the sensor bridge in Figure 91.17a, we would replace R_0 in the same branch than the sensor by R_r. Additional circuits for bridge output linearization, as well as further signal conditioning and sensor lead compensation can be found in Ref. [1,10].

91.11 Time- and Frequency-Based Sensor Conditioners

Because ADCs digitize voltages rather than currents, most sensor signal conditioning is oriented toward obtaining voltages that are processed to suit ADC requirements relative to voltage range and levels, and input impedance. However, indirect ADCs such as dual- and triple-slope models internally convert voltage to time, and count time to obtain a digital output. These are low-speed ADCs that suit many low-frequency measurements. It follows that converting the measurand to time instead of voltage can save components and footprint in printed circuit boards (PCBs). Modulating sensors placed in a relaxation oscillator can readily yield an analog output with information encoded in a time parameter (oscillation period, pulse with, duty cycle) instead of an amplitude parameter (voltage or current). Some IC sensors for acceleration (ADXL202), temperature (TMP04, MAX6676), and light (TSL237, S9705) offer such an output, and are commonly designed digital sensors; actually, they must be considered *quasi-digital sensors* because their output is not a digital code but a two-level voltage that can yield a digital code by using a time counter. No ADC is required, hence no amplification, no level shifting, no demodulation, no ac-to-dc conversion, no antialiasing filtering.

Further on, microcontroller units (MCU) can easily implement by their I/O ports the switching actions involved in relaxation oscillators and can also count time. Figure 91.18 describes the working principle for an RC network; the MCU is assumed to have a Schmitt trigger (ST) input, otherwise an external STs should be used. First, the capacitor is charged by placing a high-output voltage (V_{OH}) in port 1 while port 2 is kept in a high-impedance state (Figure 91.18a). Afterward, port 1 (ST input) is placed at a high-impedance state, while port 2 is set to the low-output voltage (V_{OL}), which makes C to discharge through R_x (Figure 91.18b). The voltage across C during discharge will be

$$v_c(t) = V_0 + (V_{OH} - V_0)e^{-t/\tau} \tag{91.41}$$

where

V_0 is the initial voltage across C when it was charged

$\tau = R_xC$

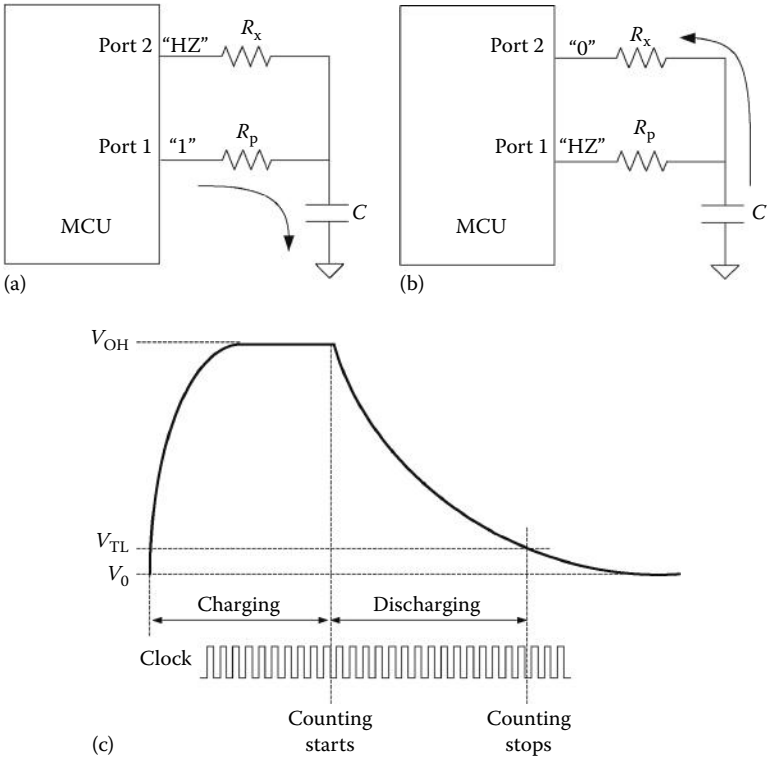


FIGURE 91.18 A microcontroller can (a) charge a capacitor and (b) then discharge it through a sensor R_x to obtain (c) a number of counts proportional to R_x, hence a digital output.

The discharging time until the low threshold (V_{TL}) is detected by the ST circuit in port 1 is

$$T_x = R_x C \ln \frac{V_{OH} - V_0}{V_{TL} - V_0} \tag{91.42}$$

hence proportional to R_x and C. If R_x is the sensor, to obtain an output that neither depends on C nor on V_0, V_{OH}, and V_{TL}, which are internal MCU parameters that depend on temperature and power supply, the charging–discharging procedure can be repeated through a known reference resistor R_r connected between C and an additional output port. The ratio between discharging times will be proportional to R_x/R_r. Better yet, a two-point calibration based on two known resistors can compensate for both multiplicative factors (such as those in Equation 91.42) and additive factors such as internal port resistance. Figure 91.19a shows the circuit to implement that calibration; R_{c1} and R_{c2} are known, stable resistors, and R_p limits the current surge when charging C. The sensor resistance can be obtained from

$$R_x = \frac{T_x - T_1}{T_2 - T_1}(R_{c2} - R_{c1}) + R_{c1} \tag{91.43}$$

This calibration is easier to implement than calibrations based on known voltages because stable resistors of any (standard) value are easier to obtain than stable voltages other than 0 V and voltages close to the full-scale range of ADCs.

The resolution and accuracy that can be achieved with this method is limited by several factors [11]. First of all, the resolution is not that of the time counter available because here there is no equivalent to

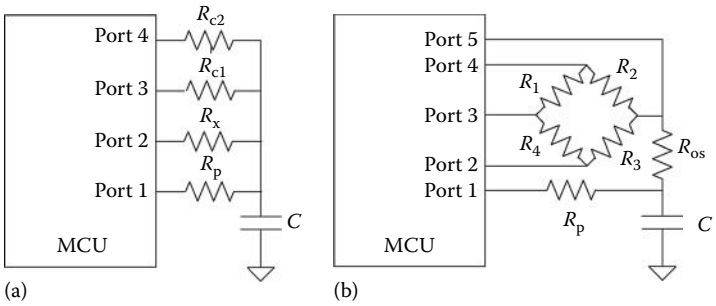


FIGURE 91.19 By successively charging a capacitor and discharging it through different resistors and performing basic calculations with the respective discharging times, additive and multiplicative effects can be corrected for (a) a single resistive sensor and (b) a Wheatstone bridge including one, two, or four sensors.

the "range matching" performed by amplifiers. Instead, if the time counts corresponding to the maximal and minimal sensor values are T_{max} and T_{min}, the resolution in bits is

$$M = \text{lb}\frac{T_{max} - T_{min}}{T_{clock}} = \text{lb}\left((x_{max} - x_{min})\frac{\tau}{T_{clock}} \right) \tag{91.44}$$

where in the last step the sensor has been assumed to be linear. The resolution improves with a higher clock frequency, the same as in any timing process, and also with a large τ, albeit at the cost of a longer measurement time, hence a reduced bandwidth for the measurand.

The accuracy is mostly determined by trigger noise that introduces uncertainty in the measured times, particularly in the time to stop counting. This noise can be inherent to the input voltage $v_c(t)$ (that across capacitor C), or be superimposed on the voltage threshold V_{TL}. Because V_{TL} is derived from the MCU supply voltage, to reduce this trigger noise the power supply voltage for the MCU must be bypassed by a large capacitor connected to ground, and the internal MCU activity, which involves switching, hence voltage transients, suspended during the discharging time. The equivalent number of bits (ENOB) is [11]

$$\text{ENOB} = \frac{R_{max} - R_{min}}{R_{max}} \frac{(V_{TL} - V_0)\ln\dfrac{V_{OH} - V_0}{V_{TL} - V_0}}{\sqrt{12}\sqrt{E_{ns}^2 + E_{ni}^2}} \tag{91.45}$$

where

E_{ns} is voltage noise in the capacitor decaying voltage
E_{ni} is the voltage noise in V_{TL}

Overall, a 10–12 bit ENOB can be obtained for single measurements of a 1000 Ω sensor that last 2–3 ms, which can be improved by about 2 bits by averaging 10 measurements, which obviously take 10 times longer.

This performance is useful enough to consider the same approach for sensor bridges. Figure 91.19b shows a sensor bridge connected to a MCU but not in the usual method for voltage measurements with two input and two output terminals. Instead, there is one terminal connected to C, through R_{os}, and three terminals each connected to ports 2, 3 and 4. First, C is charged through R_p and then it is discharged through port 2 by placing it at the low-output voltage, while ports 3, 4, and 5 are in high-impedance state. Then C is charged again and the discharging procedure repeated through ports 3, 4, and 5. Therefore, we will have three discharging times (T_2, T_3, T_4) related to sensor bridge resistances, and one discharging time T_5 that will depend on R_{os} plus the internal resistance of port 5. These times

can be combined to obtain a linear estimate x^* of the sensor fractional resistance change x, regardless of the number of sensors in the bridge [12]. For example, if there is a single sensor $R_3 = R_0(1 + x)$, we can use

$$x^* = \frac{2(T_2 - T_4)}{T_3 + T_4 - T_2 - T_5} \tag{91.46}$$

For a full-sensor bridge with $R_1 = R_3 = R_0(1 + x)$ and $R_2 = R_4 = R_0(1 - x)$, we can use

$$x^* = \frac{T_2 - T_4}{T_3 - T_5} \tag{91.47}$$

T_5 is used to reduce the effect of internal resistances of MCU ports.

Capacitive sensors can be measured in a similar way by interchanging resistors and capacitors in the circuits above, but stray capacitances to ground and between MCU pins can reduce the performance for low-capacitance sensors.

A resistive, capacitive, or inductive sensor that replaces a similar component in a harmonic oscillator (Wien bridge, Colpitts, Hartley) will make the oscillation frequency to depend on the measurand. Whenever that frequency is higher than that of the clock available to measure it, it is faster to measure frequency than period, for a given resolution. Counting events (oscillation cycles) is equivalent to analog integration, hence the frequency response is as shown in Figure 91.14a. The high SMRR at specific frequencies is of particular interest to reduce power-line interference. To extend this capability to voltage- or current-output sensors, there are IC voltage- (and current-) to-frequency converters that offer an output frequency of up to about 3 MHz for a 2.5 V full-scale input.

91.12 Special-Purpose and Programmable Signal Conditioners

Industrial sensors were a response to the need for process control, particularly to produce aircraft fuel during World War II. Harsh environmental conditions in process control forced the remote placement of signal conditioners in control rooms, so that sensors and conditioners were separate electronic units. Progress in electronics involved a better knowledge of silicon properties, including its electrical response to temperature, light, magnetic fields, and stress (pressure), and this lead to sensors with integrated electronics, such as photointerrupters. Sensors were increasingly used in buildings, machinery, cars, and consumer products, which meant large markets that paid for the cost of developing microelectromechanical sensors (MEMS) based on mechanical properties of silicon and for specific signal conditioning circuits, sometimes cofabricated with MEMS, able to withstand, for example, the special environmental conditions in automotive applications. Further, many functions in signal conditioning apply to all sensors, the main difference being in degree, that is, for example, how much offset or gain needs to be applied; this flexibility is offered by programmable circuits, whose lowest-cost units are microcontrollers.

Overall, in addition to general-purpose integrated circuits that can be used to build any signal conditioning circuit on a multichip System on Board (SoB), there are (a) sensors that integrate the signal conditioner and use external components for further signal processing; (b) sensors that integrate analog and digital signal processing; (c) analog signal conditioners aimed as specific sensors, such as thermocouples, LVDTs, inductive sensors, and ultrasound-based sensors; (d) ADCs with analog front ends for a particular group of sensors that share a common operating principle, e.g., resistive or capacitive sensors, or low-level dc or ac voltages coming from sensors; and (e) microcontrollers with additional nonvolatile memory and data-acquisition capabilities that can be almost directly connected to a variety of sensors. Some of these ICs are the core of electronic modules intended for industrial applications based on rail-mounted systems or modular units for rack-mounted instruments. Table 91.4 lists some signal conditioners in IC form intended for specific sensors or sensor groups and their respective functions.

TABLE 91.4 Special-Purpose and Programmable Integrated Circuit Signal Conditioners

Model	Function	Manufacturer
ACF2101	Low-noise switched integrator	Texas Instruments
IVC102	Integrating amplifier	Texas Instruments
AD536	True rms-to-dc converter	Analog Devices
AD636/7	True rms-to-dc converter	Analog Devices
AD736/7	True rms-to-dc converter	Analog Devices
LTC1966/7/8	True rms-to-dc converter	Linear Technology
AC1226	Thermocouple cold junction compensator	Analog Devices
AD594/5	Instrumentation amplifier with thermocouple cold junction compensation	Analog Devices
AD596/7	Thermocouple conditioner with temperature set-point controller	Analog Devices
AD8494/5	Instrumentation amplifier with thermocouple cold junction compensation	Analog Devices
AD8496/7	Thermocouple conditioner with set-point controller	Analog Devices
LT1025	Thermocouple cold junction compensator	Linear Technology
MAX6674/5	Thermocouple-to-digital converter with cold junction compensation	Maxim Semiconductor
AD598/698	LVDT signal conditioner	Analog Devices
AD630	Balanced modulator/demodulator	Analog Devices
MCP2036	Inductive sensor analog front end	Microchip Technology
NV1124	Variable reluctance sensor interface	On Semiconductor
AD7142/3/7/8	Capacitance-to-digital converter	Analog Devices
AD7150/1/2/3/6	Capacitance-to-digital converter for proximity sensing	Analog Devices
AD7745/6/7	Capacitance-to-digital converter (sigma–delta)	Analog Devices
C8051F70x/71x	Capacitance-to-digital converter	Silicon Labs
C80518xx	Capacitance-to-digital converter	Silicon Labs
C8051F99x	Capacitance-to-digital converter	Silicon Labs
LDS6000	Capacitance-to-digital converter (sigma–delta)	Integrated Device Technology (IDT)
MS3110	Universal Capacitive Readout™	Irvine Sensors
ZSC31210	Capacitance-to-digital converter	ZMDI
AD5933/4	Impedance converter	Analog Devices
ic-MSB, ic-MSB2	Sin/cos sensor signal conditioner	iC-Haus Integrated Circuits
UTI	Sensor-to-time signal converter	Smartec
AD8290	Sensor amplifier with current excitation	Analog Devices
AD8555/6	Programmable sensor signal amplifier	Analog Devices
ADS1014/5	12 bit ADC with PGA, voltage reference, and oscillator	Texas Instruments
AMC7820/3	Analog monitoring and control circuit, 12 bit	Texas Instruments
ic-TW3	Signal conditioner with analog output	iC-Haus Integrated Circuits
MAX1450/2/5/8	Signal conditioner for piezoresistive sensors	Maxim Integrated Products
MAX6603	Platinum RTD signal conditioner	Maxim Integrated Products
MLX90320/6	Resistive bridge sensor interface circuit	Melexis
PGA308	Auto-zero sensor amplifier with programmable gain and offset	Texas Instruments
PGA309	Voltage output programmable sensor conditioner	Texas Instruments
X96010	Sensor conditioner with dual lookup table memory and DAC	Intersil
ZSC31010	Sensor signal conditioner for resistive sensor bridges	ZMDI
ZSC31015	Sensor signal conditioner for resistive sensor bridges	ZMDI

(*continued*)

TABLE 91.4 (continued) Special-Purpose and Programmable Integrated Circuit Signal Conditioners

Model	Function	Manufacturer
ZSC31014	Amplifier and ADC for differential and half-bridge input signals	ZMDI
ZSC31050	Amplifier and ADC for sensor bridge signals	ZMDI
ZSC31150	Amplifier and ADC for sensor bridge signals	ZMDI
ZSSC3053	Amplifier and sensor-specific correction of mV-dc signals	ZMDI
ZSI21013	Resistive sensor interface	ZMDI
AD7705/6/7	16 bit sigma–delta ADC with PGA	Analog Devices
AD7708/18	Analog front end for low-frequency measurements	Analog Devices
AD7709/19	Analog front end for low-frequency measurements	Analog Devices
AD7710/1	24 bit signal conditioning ADC with current source	Analog Devices
AD7712/3/4	24 bit sigma–delta signal conditioning ADC	Analog Devices
AD7730	24 bit sigma–delta bridge transducer ADC for load cell applications	Analog Devices
AD7190/4	24 bit sigma–delta ADC with PGA	Analog Devices
AD7792/3	16/24 bit analog front end for measurement applications	Analog Devices
AD7796/8	16 bit sigma–delta ADC for bridge sensors	Analog Devices
AD7797/9	24 bit sigma–delta ADC for bridge sensors	Analog Devices
ADS1230	20 bit sigma–delta ADC for bridge sensors	Texas Instruments
ADS1232/4	24 bit sigma–delta ADC for bridge sensors	Texas Instruments
CS5521/23/25/31/33	16 bit sigma–delta ADC with integrated amplifier	Cirrus Logic
CS5526	20 bit sigma–delta ADC with integrated amplifier	Cirrus Logic
CS5522/24728/30/32/34	24 bit sigma–delta ADC with integrated amplifier	Cirrus Logic
MAX1462	16 bit ADC with analog front end and digital correction	Maxim Integrated Products
RPIC1.2	Resistive sensor interface	e2v Technologies
ADUC	Analog microcontrollers	Analog Devices
MAX1464	Sensor signal processor	Maxim Integrated Products
MAXQ7665	Microcontroller-based data-acquisition system	Maxim Integrated Products
MAXQ7667	Mixed-signal microcontroller for ultrasonic applications	Maxim Integrated Products
MAXQ7670	Mixed-signal microcontroller for differential sensors	Maxim Integrated Products
MLX90308/14	Programmable interface for bridge or differential sensors	Melexis
MSC1210	ADC with microcontroller and flash memory	Texas Instruments
AD693	4–20 mA sensor transmitter, loop-powered	Analog Devices
AD694	4–20 mA current transmitter, local power	Analog Devices
MAX1459	4–20 mA sensor signal conditioner, loop-powered	Maxim Integrated Products
MLX90323	4–20 mA current loop output for bridge or differential sensors	Melexis
XTR101/5/6/8	4–20 mA two-wire transmitter	Texas Instruments
XTR110/1/2/4/5/6/7	4–20 mA current loop transmitter	Texas Instruments
RCV420	4–20 mA current loop receiver	Texas Instruments
XTR300	Voltage/current-output driver	Texas Instruments
ADE7878	Polyphase energy metering IC	Analog Devices
TLE2425/6	Virtual ground	Texas Instruments

The decision whether to design a signal conditioner from parts or use a model from Table 91.4 is a matter of cost, reliability, and availability. As a practical matter, whereas many general purpose and some specific op amps are pin-to-pin compatible or at least functionally equivalent, most other ICs are not compatible neither there are any second sources for them. This can be a major risk as integrated circuits can be discontinued and no replacement offered.

Defining Terms

Carrier amplifier: Voltage amplifier for narrowband ac signals, that includes in addition a sine wave oscillator, a synchronous demodulator, and a low-pass filter.

Common-mode rejection ratio (CMRR): The gain for a differential voltage divided by the gain for a common-mode voltage in a differential amplifier. It is usually expressed in decibels.

Common-mode voltage: The average of the voltages at the input terminals of a differential amplifier.

Differential amplifier: Circuit or device that amplifies the difference in voltage between two terminals, none of which is grounded.

Dynamic range: The measurement range for a quantity divided by the desired resolution.

Instrumentation amplifier: Differential amplifier with large input impedance and low offset and gain errors.

Isolation amplifier: Voltage amplifier whose ground terminal for input voltages is independent from the ground terminal for the output voltage (i.e., there is a large—isolation—impedance between both ground terminals).

Isolation Mode Rejection Ratio (IMRR): The amplitude of the voltage across the isolation impedance divided by the amplitude of the output voltage that produces.

Series (Normal) Mode Rejection Ratio (SMRR, NMRR): The gain of a circuit to signals of a frequency of interest divided by the gain at other frequencies.

Signal conditioner: Circuit or device that adapts a sensor signal to an ensuing circuit, such as an ADC.

Voltage buffer: Voltage amplifier whose gain is 1, or close to 1, and whose input impedance is very large while its output impedance is very small.

References

1. R. Pallas-Areny and J. G. Webster, *Sensors and Signal Conditioning*, 2nd edn., New York: John Wiley & Sons, 2001.
2. C. Kitchin and L. Counts, *A Designer's Guide to Instrumentation Amplifiers*, 3rd edn., Norwood, MA: Analog Devices, 2006.
3. C. D. Motchenbacher and J.A. Connelly, *Low-Noise Electronic System Design*, New York: John Wiley & Sons, 1993.
4. S. Franco, *Design with Operational Amplifiers and Analog Integrated Circuits*, 3rd edn., New York: McGraw-Hill, 2001.
5. R. Pallas-Areny and J.G. Webster, *Analog Signal Processing*, New York: John Wiley & Sons, 1999.
6. O. Casas, E. Spinelli, and R. Pallas-Areny, Fully-differential AC-coupling networks: A comparative study, *IEEE Trans. Instrum. Meas.*, 58, 94–98, 2009.
7. J. Graeme, *Photodiode Amplifiers, Op Amp Solutions*, New York: McGraw-Hill, 1996.
8. R. Poley, Signal conditioning an LVDT using a TMS320F2812 DSP, Application Report SPRA946, Texas Instruments, 2003.
9. M. L. Meade, *Lock-in Amplifiers: Principles and Applications*, London, U.K.: Peter Peregrinus, 1984.
10. W. G. Jung (ed.), *Op Amp Applications Handbook*, Burlington, MA: Newnes, 2005.

11. F. Reverter and R. Pallas-Areny, Effective number of resolution bits in direct sensor-to-microcontroller interfaces, *Meas. Sci. Technol.*, 15, 2157–2162, 2004.
12. E. Sifuentes, O. Casas, F. Reverter, and R. Pallas-Areny, Direct interface circuit for resistive sensor bridges, *Sensors Actuators A*, 147, 210–215, 2008.

Informational Websites

www.dataforth.com (Dataforth Corp.)
www.digikey.com—Techzones, Sensor solutions—(Digi-Key Corp.)
www.mccdaq.com (Measurement Computing Corp.)
www.sensorland.com (Copidate Technical Publicity)
www.sensorsmag.com (Questex Media Group)
Manufacturers of IC signal conditioners
Analog Devices (www.analog.com)
Cirrus Logic (www.cirrus.com)
Intersil (www.intersil.com)
Linear Technology (www.linear.com)
Maxim Integrated Products (www.maxim-ic.com)
Microchip Technology (www.microchip.com)
National Semiconductor (www.national.com)
ON Semiconductor (www.onsemi.com)
Texas Instruments (www.ti.com)

92

Data Acquisition Systems

Edward McConnell
National Instruments

92.1 Introduction

The fundamental task of a *data acquisition system* is the measurement or generation of real-world physical signals. Before a physical signal can be measured by a computer-based system, a sensor or transducer is used to convert the physical signal into an electrical signal, such as voltage or current. Often only a plug-in data acquisition (DAQ) board is considered the data acquisition system; however, a board is only one of the components in the system. A complete DAQ system consists of sensors, signal conditioning, interface hardware, and software. Unlike stand-alone instruments, signals often cannot be directly connected to the DAQ board. The signals may need to be conditioned by some signal-conditioning accessory before they are converted to digital information by the plug-in DAQ board. Software controls the data acquisition system—acquiring the raw data, analyzing the data, and presenting the results. The components are shown in Figure 92.1.

92.2 Signals

Signals are physical events whose magnitude or time variation contains information. DAQ systems measure various aspects of a signal in order to monitor and control the physical events. Users of DAQ systems need to know the relation of the signal to the physical event and what information is available in the signal. Generally, information is conveyed by a signal through one or more of the following signal parameters: state, rate, level, shape, or frequency content. The physical characteristics of the measured signals and the related information help determine the design of a DAQ system.

All signals are, fundamentally, analog, time-varying signals. For the purpose of discussing the methods of signal measurement using a plug-in DAQ board, a given signal should be classified as one of five

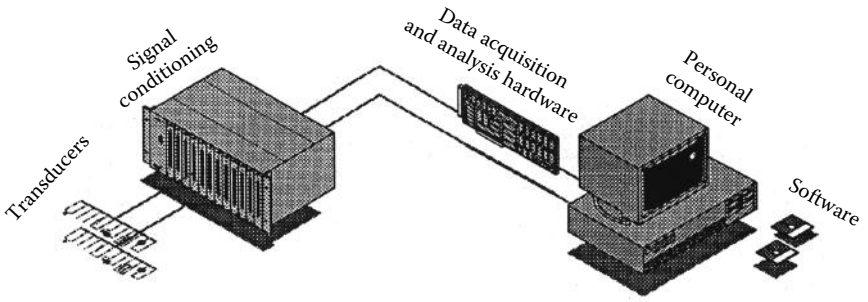


FIGURE 92.1 Components of a DAQ system.

signal types. Because the method of signal measurement is determined by the way the signal conveys the needed information, a classification based on this criterion is useful in understanding the fundamental building blocks of a DAQ system.

As shown in Figure 92.2, any signal can generally be classified as analog or digital. A digital, or binary, signal has only two possible discrete levels of interest—a high (on) level and a low (off) level. The two digital signal types are on–off signals and pulse train signals. An analog signal, on the other hand, contains information in the continuous variation of the signal with time. Analog signals are described in the time or frequency domains depending upon the information of interest. A dc-type signal is a low-frequency signal, and if the phase information of a signal is presented with the frequency information, then there is no difference between the time or frequency domain representations. The category to which a signal belongs depends on the characteristic of the signal to be measured. The five types of signals can be closely paralleled with the five basic types of signal information—state, rate, level, shape, and frequency content.

Basic understanding of the signal representing the physical event being measured and controlled assists in the selection of the appropriate DAQ system.

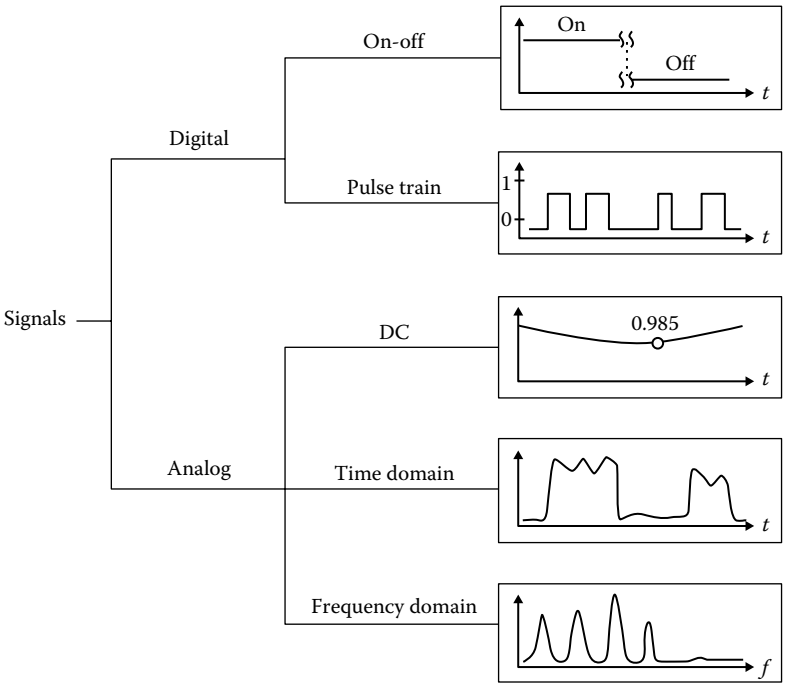


FIGURE 92.2 Classes of analog and digital signals.

92.3 Plug-In DAQ Boards

The fundamental component of a DAQ system is the plug-in DAQ board. These boards plug directly into a slot in a PC and are available with analog, digital, and timing inputs and outputs (I/O). The most versatile of the plug-in DAQ boards is the multifunction I/O board. As the name implies, this board typically contains various combinations of analog-to-digital converters (ADCs), digital-to-analog converters (DACs), digital I/O lines, and counters/timers. ADCs and DACs measure and generate analog voltage signals, respectively. The digital I/O lines sense and control digital signals. Counters/timers measure pulse rates, widths, delays, and generate timing signals. These many features make the multifunction DAQ board useful for a wide range of applications.

Multifunction boards are commonly used to measure analog signals. This is done by the ADC, which converts the analog voltage level into a digital number that the computer can interpret. The analog multiplexer (MUX), the instrumentation amplifier, the sample-and-hold (S/H) circuitry, and the ADC compose the analog input section of a multifunction board (see Figure 92.3).

Typically, multifunction DAQ boards have one ADC. Multiplexing is a common technique for measuring multiple channels (generally 16 single-ended or 8 differentials) with a single ADC. The analog MUX switches between channels and passes the signal to the instrumentation amplifier and the S/H circuitry. The MUX architecture is the most common approach taken with plug-in DAQ boards. While plug-in boards typically include up to only 16 single-ended or 8 differential inputs, the number of analog input channels can be further expanded with external MUX accessories.

Instrumentation amplifiers typically provide a differential input and selectable gain by jumpers or software. The differential input rejects small common-mode voltages. The gain is often software programmable. In addition, many DAQ boards also include the capability to change the amplifier gain while scanning channels at high rates. Therefore, one can easily monitor signals with different ranges of amplitudes. The output of the amplifier is sampled, or held at a constant voltage, by the S/H device at measurement time so that voltage does not change during digitization.

The ADC transforms the analog signal into a digital value which is ultimately sent to computer memory. There are several important parameters of A/D conversion. The fundamental parameter of an ADC is the number of bits. The number of bits of an A/D determines the range of values for the binary output of the ADC conversion. For example, many ADCs are of 12 bits, so a voltage within the input range of the ADC will produce a binary value that has one of $2^{12} = 4096$ different values. The more bits an ADC has, the higher the resolution of the measurement. The resolution determines the smallest amount of change that can be detected by the ADC. Resolution is expressed as the number of digits of a voltmeter or dynamic range in decibels, rather than with bits. Table 92.1 shows the relation among bits, number of digits, and dynamic range in decibels.

The resolution of the A/D conversion is also determined by the input range of the ADC and the gain. DAQ boards usually include an instrumentation amplifier that amplifies the analog signal by

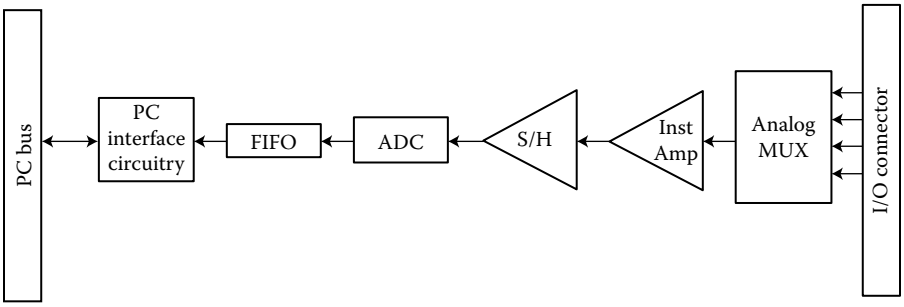


FIGURE 92.3 Analog input section of a plug-in DAQ board. FIFO, first-in first-out buffer; S/H, sample and hold; Inst Amp, instrumentation amplifier; MUX, analog multiplexer.

TABLE 92.1 Relation among Bits, Number of Digits, and Dynamic Range (dB)

Bits	Digits	dB
20	6.0	120
16	4.5	96
12	3.5	72
8	2.5	48

a gain factor prior to the conversion. This gain amplifies low-level signals so that more accurate measurements can be made.

Together, the input range of the ADC, the gain, and the number of bits of the board determine the minimum resolution of the measurement. For example, suppose a low-level ±30 mV signal is acquired using a 12 bit ADC that has a ±5 V input range. If the system includes an amplifier with a gain of 100, the resulting resolution of the measurement will be range/(gain × 2 bits) = resolution, or 10 V/(100 × 2^{12}) = 0.0244 mV.

Finally, an important parameter of digitization is the rate at which A/D conversions are made, referred to as the sampling rate. The A/D system must be able to sample the input signal fast enough to measure the important waveform attributes accurately. In order to meet this criterion, the ADC must be able to convert the analog signal to digital form quickly enough.

When scanning multiple channels with a multiplexing DAQ system, other factors can affect the throughput of the system. Specifically, the instrumentation amplifier must be able to settle to the needed accuracy before the A/D conversion occurs. With multiplexed signals, multiple signals are being switched into one instrumentation amplifier. Most amplifiers, especially when amplifying the signals with larger gains, will not be able to settle to the full accuracy of the ADC when scanning channels at high rates. To avoid this situation, consult the specified settling times of the DAQ board for the gains and sampling rates required by the application.

92.4 Types of ADCs

Different DAQ boards use different types of ADCs to digitize the signal. The most popular type of ADC on plug-in DAQ boards is the successive approximation ADC, because it offers high speed and high resolution at a modest cost.

Subranging (also called half-flash) ADCs offer very high-speed conversion with sampling speeds up to several million samples per second.

The state-of-the-art technology in ADCs is sigma–delta modulating ADCs. These ADCs sample at high rates, are able to achieve high resolution, and offer the best linearity of all ADCs.

Integrating and flash ADCs are mature technologies still used on DAQ boards today. Integrating ADCs are able to digitize with high resolution but must sacrifice sampling speed to obtain it. Flash ADCs are able to achieve the highest sampling rate (gigahertz) but are available only with low resolution. The different types of ADCs are summarized in Table 92.2.

92.5 Analog Input Architecture

With a typical DAQ board, the multiplexer switches among analog input channels. The analog signal on the channel selected by the multiplexer then passes to the programmable gain instrumentation amplifier (PGIA), which amplifies the signal. After the signal is amplified, the S/H keeps the analog signal constant so that the ADC can determine the digital representation of the analog signal. A good DAQ board will then place the digital signal in a first-in first-out (FIFO) buffer, so that no data will be lost if the sample cannot transfer immediately over the PC I/O channel to computer memory.

TABLE 92.2 Types of ADCs

Type of ADC	Advantages	Features
Successive approximation	High resolution	1.25 MS/s sampling rate
	High speed	12 bit resolution
	Easily multiplexed	200 kS/s sampling rate
		16 bit resolution
Subranging	Higher speed	1 MHz sampling rate
		12 bit resolution
Sigma–delta	High resolution	48 kHz sampling rate
	Excellent linearity	16 bit resolution
	Built-in antialiasing	
	State-of-the-art technology	
Integrated	High resolution	15 kHz sampling rate
	Good noise rejection	
	Mature technology	
Flash	Highest speed	125 MHz sampling rate
	Mature technology	

Having a FIFO becomes especially important when the board is run under operating systems that have large interrupt latencies, such as Microsoft Windows.

92.5.1 Basic Analog Specifications

Almost every DAQ board data sheet specifies the number of channels, the maximum sampling rate, the resolution, and the input range and gain.

The number of channels, which is determined by the multiplexer, is usually specified in two forms—differential and single-ended. Differential inputs are inputs that have different reference points for each channel, none of which is grounded by the board. Differential inputs are the best way to connect signals to the DAQ board because they provide the best noise immunity.

Single-ended inputs are inputs that are referenced to a common ground point. Because single-ended inputs are referenced to a common ground, they are not as good as differential inputs for rejecting noise. They do have a larger number of channels, however. Single-ended inputs are used when the input signals are high level (greater than 1 V), the leads from the signal source to the analog input hardware are short (less than 5 m), and all input signals share a common reference.

Some boards have pseudo-differential inputs which have all inputs referenced to the same common—like single-ended inputs—but the common is not referenced to ground. These boards have the benefit of a large number of input channels, like single-ended inputs, and the ability to remove some common-mode noise, especially if the common-mode noise is consistent across all channels. Differential inputs are still preferable to pseudo-differential, however, because differential is more immune to magnetic noise.

Sampling rate determines how fast the analog signal is converted to a digital signal. When measuring ac signals, sample must be at least twice faster than the highest frequency of the input signal. Even when measuring dc signals, the oversample and average the data increase the accuracy of the signal by reducing the effects of noise.

If the physical event consists of multiple dc-class signals, a DAQ board with interval scanning should be used. With interval scanning, all channels are scanned at one sample interval (usually the fastest rate of the board), with a second interval (usually slow) determining the time before repeating the scan. Interval scanning gives the effects of simultaneously sampling for slowly varying signals without requiring the additional cost of input circuitry for true simultaneous sampling.

Resolution is the number of bits that are used to represent the analog signal. The higher the resolution, the higher the number of divisions the input range is broken into, and therefore the smaller the possible detectable voltage. Unfortunately, some DAQ specifications are misleading when they specify the resolution associated with the DAQ board. Many DAQ board specifications state the resolution of the ADC without stating the linearities and noise, and therefore do not give the information needed to determine the resolution of the entire board. Resolution of the ADC, combined with the settling time, *integral nonlinearity* (INL), *differential nonlinearity* (DNL), and noise will give an understanding of the accuracy of the board.

Input range and gain determine the level of signal that should be connected to the board. Usually, the range and gain are specified separately, so the two must be combined to determine the actual signal input range as

$$\text{Signal input range} = \frac{\text{range}}{\text{gain}}$$

For example, a board using an input range of ±10 V with a gain of 2 will have a signal input range of ±5 V. The closer the signal input range is to the range of the signal, the more accurate the readings from the DAQ board will be. If the signals have different input ranges, use a DAQ board with the feature of different gains per channel.

92.6 Data Acquisition Software

The software is often the most critical component of the DAQ system. Users of DAQ systems usually program the hardware in one of two ways—through register programming or through high-level device drivers.

92.6.1 Board Register–Level Programming

The first option is not to use vendor-supplied software and program the DAQ board at the hardware level. DAQ boards are typically register-based; that is, they include a number of digital registers that control the operation of the board. The developer may use any standard programming language, such as C, C++, or Visual BASIC, to write series of binary codes to the DAQ board to control its operation. Although this method affords the highest level of flexibility, it is also the most difficult and time-consuming, especially for the inexperienced programmer. The programmer must know the details of programming all hardware, including the board, the PC interrupt controller, the DMA controller, and PC memory.

92.6.2 Driver Software

Driver software typically consists of a library of function calls usable from a standard programming language. These function calls provide a high-level interface to control the standard functions of the plug-in board. For example, a function called SCAN_OP may configure, initiate, and complete a multiple-channel scanning DAQ operation of a predetermined number of points. The function call would include parameters to indicate the channels to be scanned, the amplifier gains to be used, the sampling rate, and the total number of data points to be collected. The driver responds to this one function call by programming the plug-in board, the DMA controller, the interrupt controller, and CPU to scan the channels as requested.

92.6.3 What Is Digital Sampling?

Every DAQ system has the task of gathering information about analog signals. To do this, the system captures a series of instantaneous "snapshots" or samples of the signal at definite time intervals. Each sample contains information about the signal at a specific instant. Knowing the exact time of each conversion and the value of the sample, one can reconstruct, analyze, and display the digitized waveform.

92.6.4 Real-Time Sampling Techniques

In real-time sampling, the DAQ board digitizes consecutive samples along the signal (Figure 92.4). According to the *Nyquist sampling theorem*, the ADC must sample at least twice the rate of the maximum frequency component in that signal to prevent aliasing. *Aliasing* is a false lower-frequency component that appears in sampled data acquired at too low a sampling rate. The frequency at one-half the sampling frequency is referred to as the Nyquist frequency. Theoretically, it is possible to recover information about those signals with frequencies at or below the Nyquist frequency. Frequencies above the Nyquist frequency will alias to appear between dc and the Nyquist frequency.

For example, assume the sampling frequency, f_s, is 100 Hz. Also assume the input signal to be sampled contains the following frequencies—25, 70, 160, and 510 Hz. Figure 92.5 shows a spectral representation of the input signal.

The mathematics of sampling theory show us that a sampled signal is shifted in the frequency domain by an amount equal to integer multiples of the sampling frequency, f_s. Figure 92.6 shows the spectral content of the input signal after sampling. Frequencies below 50 Hz, the Nyquist frequency ($f_s/2$), appear correctly. However, frequencies above the Nyquist appear as aliases below the Nyquist frequency. For example, F1 appears correctly; however, F2, F3, and F4 have aliases at 30, 40, and 10 Hz, respectively.

The resulting frequency of aliased signals can be calculated with the following formula:

Apparent (Alias) Frequency = ABS (Closest Integer Multiple of Sampling Frequency – Input Frequency)

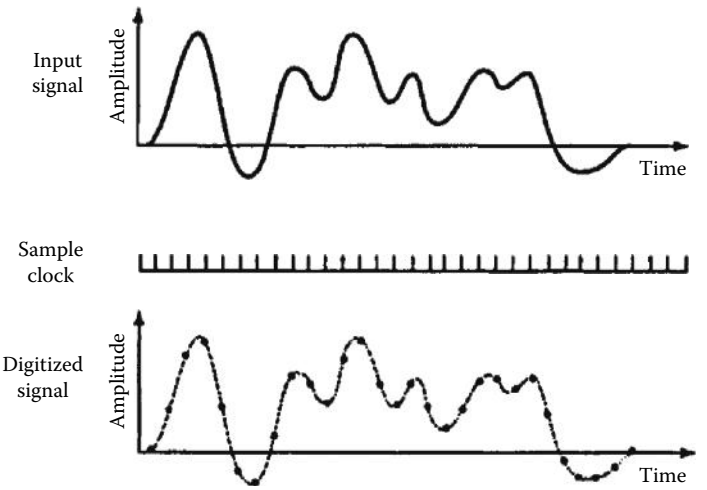


FIGURE 92.4 Consecutive discrete samples recreate the input signal.

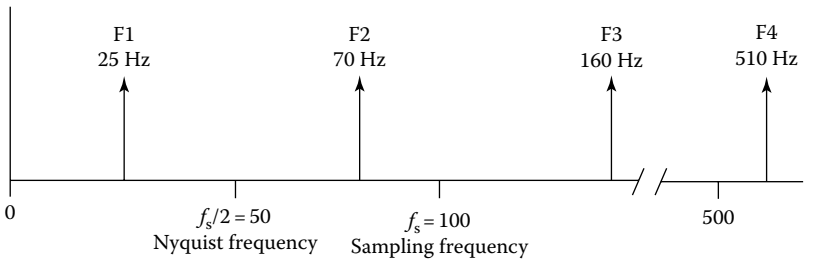


FIGURE 92.5 Spectral of signal with multiple frequencies.

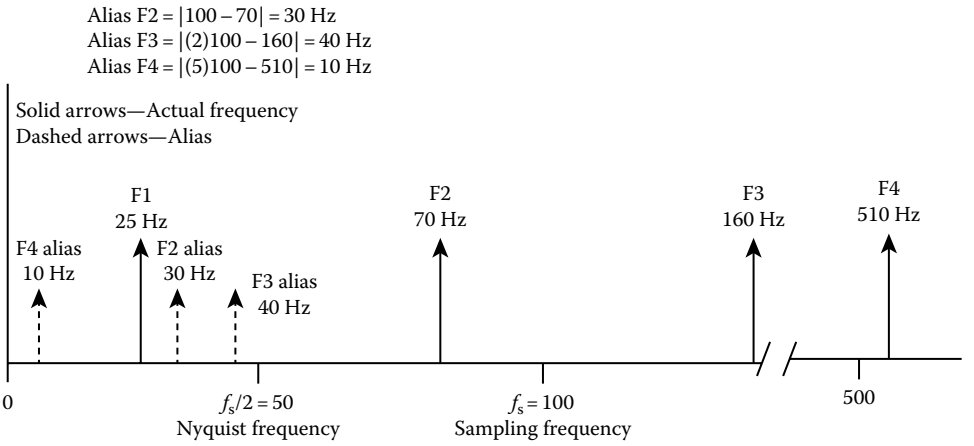


FIGURE 92.6 Spectral of signal with multiple frequencies after sampling at f_s = 100 Hz.

For the example of Figures 92.5 and 92.6:

$$\text{Alias F2} = |100 - 70| = 30 \text{ Hz}$$

$$\text{Alias F3} = |(2)100 - 160| = 40 \text{ Hz}$$

$$\text{Alias F4} = |(5)100 - 510| = 10 \text{ Hz}$$

92.6.5 Preventing Aliasing

Aliasing can be prevented by using filters on the front end of the DAQ system. These antialiasing filters are set to cut off any frequencies above the Nyquist frequency (half the sampling rate). The perfect filter would reject all frequencies above the Nyquist; however, because perfect filters exist only in textbooks, one must compromise between sampling rate and selecting filters. In many applications, one- or two-pole passive filters are satisfactory. The rule of thumb is to oversample (5–10 times) and use these antialiasing filters when frequency information is crucial.

Alternatively, active antialiasing filters with programmable cutoff frequencies and very sharp attenuation of frequencies above the cutoff can be used. Because these filters exhibit a very steep roll-off, the DAQ system can sample at two to three times the filter cutoff frequency. Figure 92.7 shows a transfer function of a high-quality antialiasing filter.

The computer uses digital values to recreate or to analyze the waveform. Because the signal could be anything between each sample, the DAQ board may be unaware of any changes in the signal between samples. There are several sampling methods optimized for the different classes of data; they include software polling, external sampling, continuous scanning, multirate scanning, simultaneous sampling, interval scanning, and seamless changing of the sample rate.

92.6.6 Software Polling

A software loop polls a timing signal and starts the A/D conversion via a software command when the edge of the timing signal is detected. The timing signal may originate from the internal clock of the computer or from a clock on the DAQ board. Software polling is useful in simple, low-speed applications, such as temperature measurements.

The software loop must be fast enough to detect the timing signal and trigger a conversion. Otherwise, a window of uncertainty, also known as jitter, will exist between two successive samples. Within the window of uncertainty, the input waveform could change enough to reduce the accuracy of the ADC drastically.

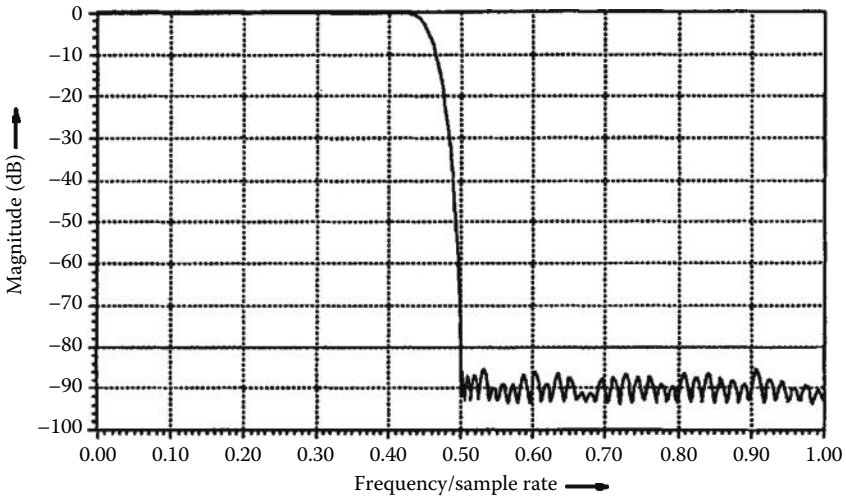


FIGURE 92.7 Magnitude portion of transfer function of an antialiasing filter.

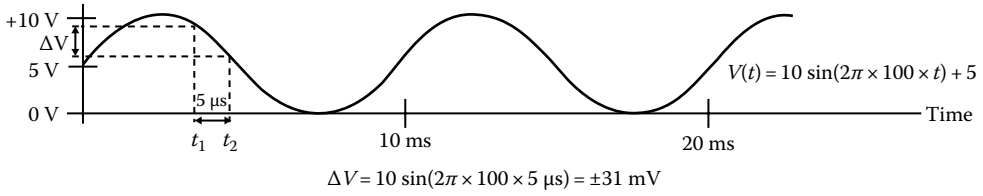


FIGURE 92.8 Jitter reduces the effective accuracy of the DAQ board.

Supposing that a 100 Hz, 10 V full-scale sine wave is digitized (Figure 92.8); if the polling loop takes 5 ms to detect the timing signal and to trigger a conversion, then the voltage of the input sine wave will change as much as 31 mV [$\Delta V = 10 \sin(2\pi \times 100 \times 5 \times 10^{-6})$]. For a 12 bit ADC operating over an input range of 10 V and a gain of 1, one least significant bit (LSB) of error represents 2.44 mV:

$$\left(\frac{\text{Input range}}{\text{gain} \times 2^n}\right) = \left(\frac{10\text{ V}}{1 \times 2^{12}}\right) = 2.44\text{ mV}$$

But because the voltage error due to jitter is 31 mV, the accuracy error is 13 LSB:

$$\left(\frac{31\text{ mV}}{2.44\text{ mV}}\right)$$

This represents uncertainty in the last 4 bits of a 12 bit ADC. Thus, the effective accuracy of the system is no longer 12 bits but rather 8 bits.

92.6.7 External Sampling

Some DAQ applications must perform a conversion based on another physical event that triggers the data conversion. The event could be a pulse from an optical encoder measuring the rotation of a cylinder. A sample would be taken every time the encoder generates a pulse corresponding to n degrees of rotation. External triggering is advantageous when trying to measure signals whose occurrence is relative to another physical phenomenon.

92.7 Scanning

92.7.1 Continuous Scanning

When a DAQ board acquires data, several components on the board convert the analog signal to a digital value. These components include the analog MUX, the instrumentation amplifier, the S/H circuitry, and the ADC. When acquiring data from several input channels, the analog MUX connects each signal to the ADC at a constant rate. This method, known as continuous scanning, is significantly less expensive than having a separate amplifier and ADC for each input channel.

Continuous scanning is advantageous because it eliminates jitter and is easy to implement. However, it is not possible to sample multiple channels simultaneously. Because the MUX switches between channels, a time skew occurs between any two successive channel samples. Continuous scanning is appropriate for applications where the time relationship between each sampled point is unimportant or where the skew is relatively negligible compared with the speed of the channel scan.

If samples from two signals are used to generate a third value, then continuous scanning can lead to significant errors if the time skew is large. In Figure 92.9, two channels are continuously sampled and added together to produce a third value. Because the two sine waves are 90° out of phase, the sum of the signals should always be zero. Because of the skew time between the samples, an erroneous sawtooth signal will result.

92.7.2 Multirate Scanning

Multirate scanning, a method of scanning multiple channels at different rates, is a special form of continuous scanning. Applications that digitize multiple signals with a variety of frequencies use multirate scanning to minimize the amount of buffer space needed to store the sampled signals. Channel-independent ADCs are used to implement hardware multirate scanning; however, this method is extremely expensive. Instead of multiple ADCs, only one ADC is used. A channel/gain configuration register stores the scan rate per channel and software divides down the scan clock based on the per-channel scan rate. Software-controlled multirate scanning works by sampling each input channel at a rate that is a fraction of the specified scan rate.

Suppose the system scans channels 0 through 3 at 10 kS/s, channel 4 at 5 kS/s, and channels 5 through 7 at 1 kS/s. A base scan rate of 10 kS/s should be used. Channels 0 through 3 are acquired at the base scan rate. Software and hardware divide the base scan rate by 2 to sample channel 4 at 5 kS/s, and by 10 to sample channels 5 through 7 at 1 kS/s.

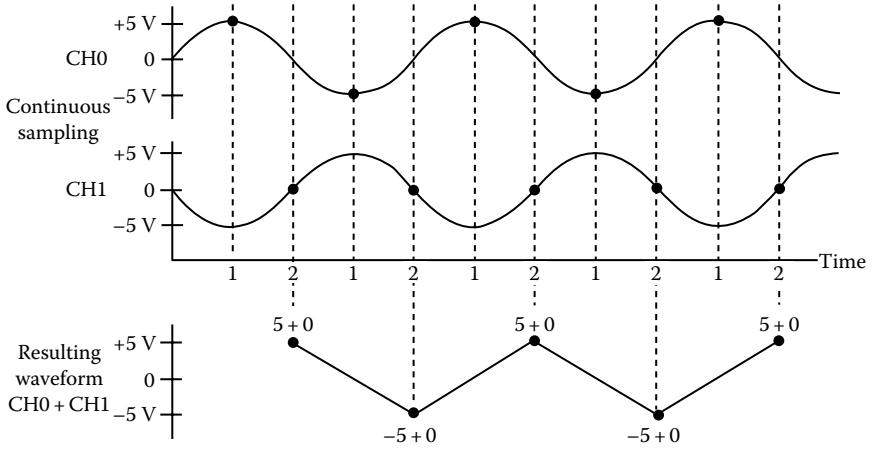


FIGURE 92.9 If the channel skew is large compared with the signal, then erroneous conclusions may result.

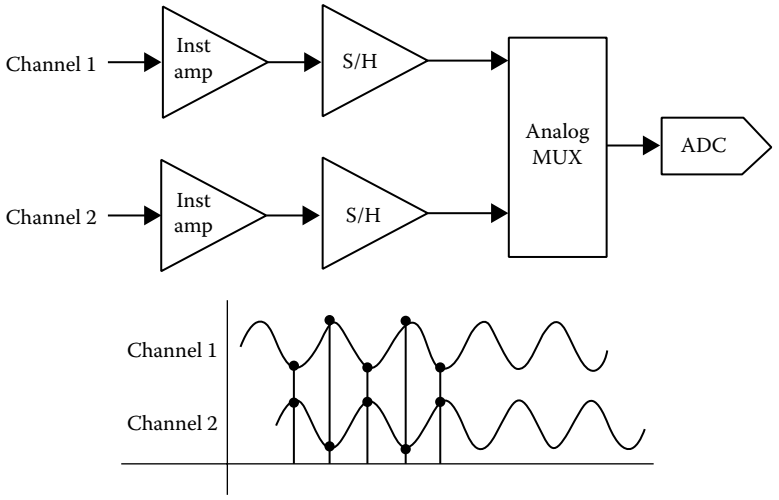


FIGURE 92.10 Block diagram of DAQ components used to sample multiple channels simultaneously.

92.7.3 Simultaneous Sampling

For applications where the time relationship between the input signals is important, such as phase analysis of ac signals, simultaneous sampling must be used. DAQ boards capable of simultaneous sampling typically use independent instrumentation amplifiers and S/H circuitry for each input channel, along with an analog MUX, which routes the input signals to the ADC for conversion (as shown in Figure 92.10).

To demonstrate the need for a simultaneous-sampling DAQ board, consider a system consisting of four 50 kHz input signals sampled at 200 kS/s. If the DAQ board uses continuous scanning, the skew between each channel is 5 μs (IS/200 kS/s) which represents a 270° [(15 μs/20 μs) × 360°] shift in phase between the first channel and fourth channel. Alternatively, with a simultaneous-sampling board with a maximum 5 ns interchannel time offset, the phase shift is only 0.09° [(5 μs/20 μs) × 360°]. This phenomenon is illustrated in Figure 92.11.

92.7.4 Interval Scanning

For low-frequency signals, interval scanning creates the effect of simultaneous sampling, yet maintains the cost benefits of a continuous-scanning system. This method scans the input channels at one rate and uses a second rate to control when the next scan begins. If the input channels are scanned at the fastest rate of the ADC, the effect of simultaneously sampling the channels is created. Interval scanning is

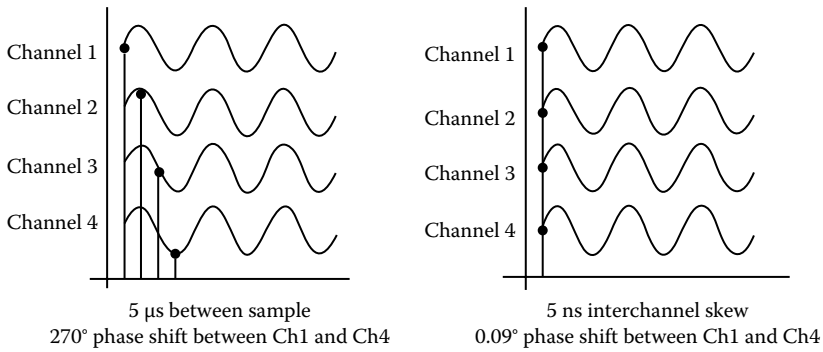


FIGURE 92.11 Comparison of continuous scanning and simultaneous sampling.

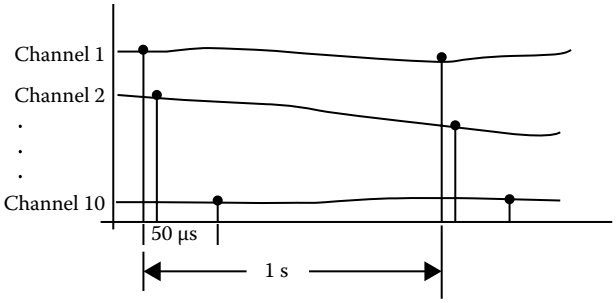


FIGURE 92.12 Interval scanning—all 10 channels are scanned within 45 μs; this is insignificant relative to the overall acquisition rate of 1 S/s.

appropriate for slow-moving signals, such as temperature and pressure. Interval scanning results in a jitter-free sample rate and minimal skew time between channel samples. For example, consider a DAQ system with 10 temperature signals. By using interval scanning, a DAQ board can be set up to scan all channels with an interchannel delay of 5 μs, and then repeat the scan every second. This method creates the effect of simultaneously sampling 10 channels at 1 S/s, as shown in Figure 92.12.

To illustrate the difference between continuous and interval scanning, consider an application that monitors the torque and RPMs of an automobile engine and computes the engine horsepower. Two signals, proportional to torque and RPM, are easily sampled by a DAQ board at a rate of 1000 S/s. The values are multiplied together to determine the horsepower as a function of time.

A continuously scanning DAQ board must sample at an aggregate rate of 2000 S/s. The time between which the torque signal is sampled and the RPM signal is sampled will always be 0.5 ms (1/2000). If either signals change within 0.5 ms, then the calculated horsepower is incorrect. But using interval scanning at a rate of 1000 S/s, the DAQ board samples the torque signal every 1 ms, and the RPM signal is sampled as quickly as possible after the torque is sampled. If a 5 μs interchannel delay exists between the torque and RPM samples, then the time skew is reduced by 99% [(0.5 ms − 5 μs)/0.5 ms], and the chance of an incorrect calculation is reduced.

92.8 Factors Influencing the Accuracy of Measurements

How does one determine if a plug-in DAQ will deliver the required measurement results? With a sophisticated measuring device like a plug-in DAQ board, significantly different accuracies can be obtained depending on the type of board used. For example, one can purchase DAQ products on the market today with 16 bit ADCs and get <12 bits of useful data, or one can purchase a product with a 16 bit ADC and actually get 16 bits of useful data. This difference in accuracies causes confusion in the PC industry where everyone is used to switching out PCs, video cards, printers, and so on, and experiencing similar results between equipment.

The most important thing to do is to scrutinize more specifications than the resolution of the ADC that is used on the DAQ board. For dc-class measurements, one should at least consider the settling time of the instrumentation amplifier, DNL, *relative accuracy*, INL, and noise. If the manufacturer of the board under consideration does not supply these specifications in the data sheets, ask the vendor to provide them or run tests to determine these specifications.

Defining Terms

Alias: A false lower-frequency component that appears in sampled data acquired at too low a sampling rate.

Asynchronous: (1) Hardware—a property of an event that occurs at an arbitrary time, without synchronization to a reference clock. (2) Software—a property of a function that begins an operation and returns prior to the completion or termination of the operation.

Conversion time: The time required, in an analog input or output system, from the moment a channel is interrogated (such as with a read instruction) to the moment that accurate data are available.

DAQ (data acquisition): (1) Collecting and measuring electric signals from sensors, transducers, and test probes or fixtures and inputting them to a computer for processing. (2) Collecting and measuring the same kind of electric signals with ADC and/or DIO boards plugged into a PC, and possibly generating control signals with DAC and/or DIO boards in the same PC.

DNL (differential nonlinearity): A measure in LSB of the worst-case deviation of code widths from their ideal value of 1 LSB.

INL (integral nonlinearity): A measure in LSB of the worst-case deviation from the ideal A/D or D/A transfer characteristic of the analog I/O circuitry.

Nyquist sampling theorem: A law of sampling theory stating that if a continuous bandwidth-limited signal contains no frequency components higher than half the frequency at which it is sampled, then the original signal can be recovered without distortion.

Relative accuracy: A measure in LSB of the accuracy of an ADC. It includes all nonlinearity and quantization errors. It does not include offset and gain errors of the circuitry feeding the ADC.

Further Information

House, R., Understanding important DA specifications, *Sensors*, 10(10), June 1993.

House, R., Understanding inaccuracies due to settling time, linearity, and noise, *National Instruments European User Symposium Proceedings*, Munich, Germany, November 10–11, 1994, pp. 11–12.

McConnell, E., New achievements in counter/timer data acquisition technology, *MessComp 1994 Proceedings*, September 13–15, 1994, pp. 492–498.

McConnell, E., PC-Based data acquisition users face numerous challenges, *ECN*, August 1994.

McConnell, E., Choosing a data-acquisition method, *Electronic Design*, 43(6), 147, 1995.

McConnell, E., Equivalent time sampling extends DA performance, *Sensors Data Acquisition*, Special Issue, June 13, 1995.

McConnell, E. and Jernigan, D., Data acquisition, in *The Electronics Handbook*, J.C. Whitaker (ed.), CRC Press, Boca Raton, FL, 1996, 1795–1822.

Park, J. and MacKay, S., *Practical Data Acquisition for Instrumentation and Control Systems*, Newnes, Elsevier, London, U.K., 2003.

Potter, D., Sensor to PC—Avoiding some common pitfalls, *Sensors Expo Proceedings*, Cleveland, OH, September 20, 1994.

Potter, D., Signal conditioners expand DAQ system capabilities, *I&CS*, 68, 25–33, August 1995.

Potter, D. and Razdan, A., Fundamentals of PC-based data acquisition, *Sensors*, 11(2), 12–20, February 1994.

Taylor, H. R., *Data Acquisition for Sensor Systems*, 2nd edn., Chapman & Hall, London, U.K., 2010.

X

Displays and Recorders

93

Plasma-Driven Flat Panel Displays

Robert T. McGrath
The Pennsylvania State University

Ramanapathy Veerasingam
The Pennsylvania State University

William C. Moffatt
Sandia National Laboratories

Robert B. Campbell
Sandia National Laboratories

93.1 An Introduction to Plasma-Driven Flat Panel Displays

93.1.1 Development History and Present Status

Plasma-driven flat panel display pixels were invented by Bitzer and Slottow at the University of Illinois in 1966 [1–3]. Figure 93.1 shows one of the inventors' early designs and demonstrates its simplicity. Parallel sets of thin conducting wires are deposited on two glass substrates which are then mounted with the conductor sets perpendicular to one another as shown in the Figure 93.1. A spacer, in this case a perforated glass dielectric, is used to maintain a gap separation of about 100 μm between the glass plates. The gap region then is filled with an inert gas, typically at a pressure of half an atmosphere. Individual pixels formed by the intersection of two conductor wires are aligned with the perforations. Pixels are illuminated by applying a voltage between two intersecting wires sufficient to initiate gas breakdown. Over the years, this basic pixel design has undergone a multitude of refinements and improvements, but the fundamental concept is still widely used.

Throughout the 1980s, plasma display products on the market were monochrome and operated with neon-based gases, directly producing within the discharge volume the red–orange (585–640 nm) visible photons that are characteristic of the quantum energy level structure of the neon atom. Dot matrix displays of the type shown in Figure 93.2 were widely used [3,4]. Early work by Owens—Illinois led to improvements in glass-sealing and spacer supports [5,6], and work by IBM led to improved understanding and control of the discharge [7–15]. These advances ultimately paved the way for manufacture of large-area, high-resolution monochrome displays. The largest area plasma display panels ever

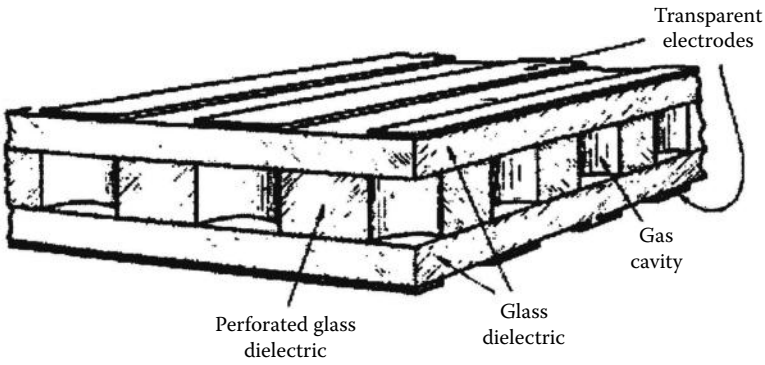


FIGURE 93.1 Structure of the ac plasma display invented at the University of Illinois. (From Bitzer, D.L. and Slottow, H.G., *AFIPS Conference Proceedings*, 29, 541, 1966. With permission.)

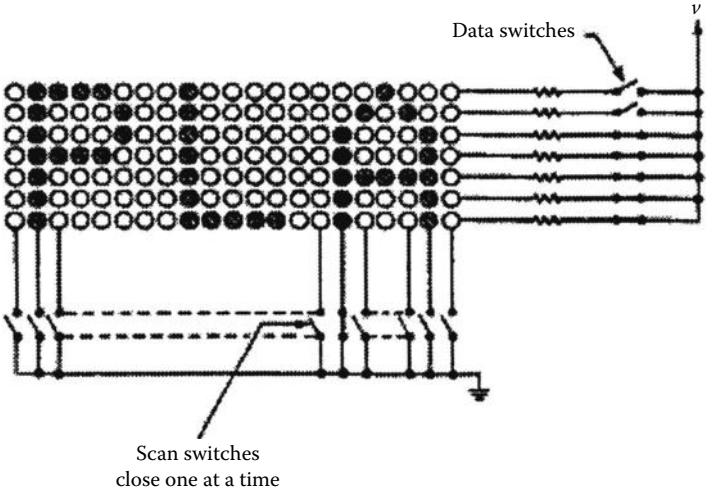


FIGURE 93.2 A simple dot matrix plasma display and data scanning switches. (From Weber, L.F., *Flat Panel Displays and CRTs*, Tannas, L.E., Jr., Ed., Van Nostrand Reinhold, New York, 1985. With permission.)

manufactured were produced by Photonics Imaging. These monochrome displays had a 1 m diagonal dimension and contained over 2 million pixels with a pixel pitch of 20 pixels/cm (50 lines/in.) [4].

Advances in lithography, patterning, and phosphors have enabled continued improvement of plasma display performance and resolution. Today, many of companies offer full-color plasma flat panel displays. Table 93.1 presents a summary list compiled by the National Research Institute of display panel specifications for some of the major companies investing in plasma flat panel manufacturing [16]. According to Stanford Research, Inc., sales for plasma display panels in 1995 totaled $230 million, but projected sales for 2002 are $4.1 billion [17]. NEC projects a more aggressive market growth reaching $2.0 billion by the year 2000 and $7.0 billion by 2002 [17]. The production capacities listed in Table 93.1 represent investments committed to manufacturing as of January 1997. As these production facilities come online, color plasma flat panel display production will grow to nearly 40,000 units per month by the end of 1998, and to over 100,000 units per month by early in 2000. In 1993, Fujitsu was the first to market a high-information-content full-color plasma flat panel display, a 21 in. diagonal, ac-driven system with a 640 × 480 pixel array [17,18]. Two examples of more recent market entries are shown in Figures 93.3 and 93.4. The first example is the NHK full color, 102 cm (40 in.) diagonal, high definition

TABLE 93.1 Plasma Flat Panel Display Specifications and Manufacturer's Business Plans

	Product Specification			Efficiency Specification				Plan			
Company	Inch	Aspect	Pixels	Luminescence (cd/m^2)	Contrast	lm/W	Power (W)	Factory	Capital Cost ($M)	Product Ability (unit/month)	Target Region
Fujitsu	42	16:9	852 × 480	300	70:1	0.7	350 (set) 300 (panel)	Miyazaki	20	10,000	Europe (Philips), Japan
NEC	33	4:3	640 × 480	200	150:1	1.2	270 (set) 190 (panel)	Tamagawa, Kagoshima	5	2,000	Japan
Pioneer	40	4:3	640 × 480	400	150:1	1.2	350 (set)	Kofu	5	10,000	Japan
Mitsubishi	40	4:3	640 × 480	350	200:1	0.8	350 (set) 300 (panel)	Kyoto	14.8	10,000	United States
MEC	42	16:9	852 × 480	450	150:1	10	300 (panel)	Kyoto	10	5,000	Japan, United States
Photonics	21	5:4	1280 × 1024	100	50:1	—	300 (panel)	Ohio	—	—	—
Hitachi	25	4:3	1024 × 768	150	50.1	—	250 (set)	Yokohama	3	1,000	—
NHK	40	16:9	1344 × 800	93	80:1	—	—	—	—	—	—

Source: Wakabayshi, H., paper presented at *Imaging 2001: The U.S. Display Consortium Business Conference*, San Jose, CA, January 28, 1997. With permission.

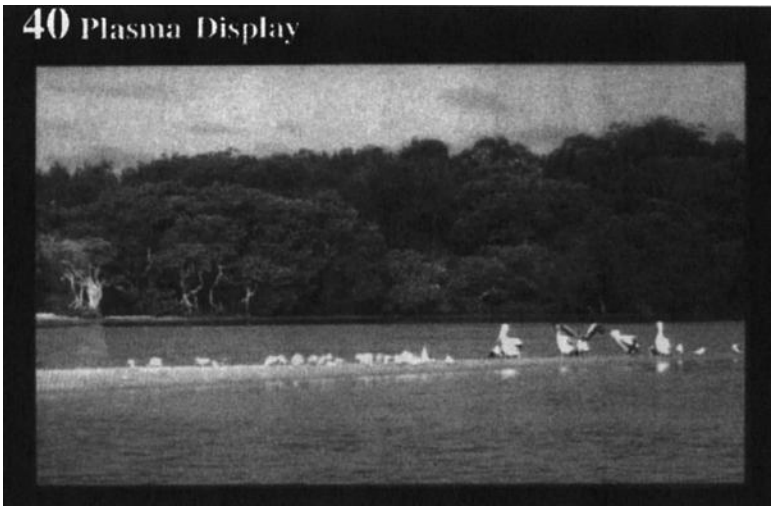


FIGURE 93.3 The 40 in. diagonal dc-driven plasma display from NHK. (From Mikoshiba, S., *Inf. Display*, 10(10), 21, 1994. With permission.)

FIGURE 93.4 The 40 in. diagonal ac-driven plasma display from Photonics Imaging. (From Friedman, P.S., *Inf. Display*, 11(10), 22, 1995. With permission.)

television (HDTV) [19–22]. The system comprises 1,075,000 full-color pixels (1344 × 800) with a pixel pitch of 0.65 mm in both horizontal and vertical directions (15.4 pixels/cm). This pulsed, dc-driven display has a peak luminance of 93 cd/m^2 with a contrast ratio of 80 to 1, and produces 256 gray levels. The display has an overall thickness of only 8 cm and weighs only 8 kg. The dimensions of the display panel itself are 87.5 × 52.0 cm with a width of only 6 mm. Shown in Figure 93.4 is the 76 cm (30 in.) diagonal,

full-color ac plasma display manufactured by Photonics Imaging [23,24]. The display contains an array of 1024 × 768 full-color pixels. At 16.8 pixels/cm (pixel pitch = 0.59 mm), this is the highest resolution full-color, plasma display manufactured to date. This unit has 64 gray levels per color channel and an average area (white) luminance greater than 103 cd/m^2 (30 fL).

93.1.2 DC and AC Plasma Pixels

As indicated earlier, plasma display pixels can be designed for either ac or dc operation. Figure 93.5 shows schematic diagrams for the simplest dc and ac pixel designs. In either case, sets of parallel conductor wires are deposited on glass substrates. In most cases, display costs are kept low by utilizing ordinary soda-lime float glass. The two glass plates are then mounted with a separation of about 100 μm and with the conductor wire sets perpendicular to one another. The gap region between the glass plates is filled with an inert gas, which discharges and illuminates the pixels when sufficient voltage is applied across two intersecting wires.

For dc pixels, shown in Figure 93.5a, the working gas is in direct contact with the electrodes. Electrons produced within the discharge volume flow rapidly to the anode, while ions produced flow more slowly toward the cathode. At 53.3 kPa (400 torr), a gap gap of 100 μm and an applied voltage of 200 V, the electron and ion transit times across the gap are roughly 0.2 and 20 ns, respectively. Once breakdown is initiated, the electrical resistance of the discharge is negligible. Consequently, dc operation requires that external resistors in series with each pixel be included in the circuit in order to limit the current amplitude. Often, dc pixels are operated in pulsed discharge mode with frequency modulation used to define the pixel brightness. For ac or dc pixels, a base firing frequency of 50 kHz

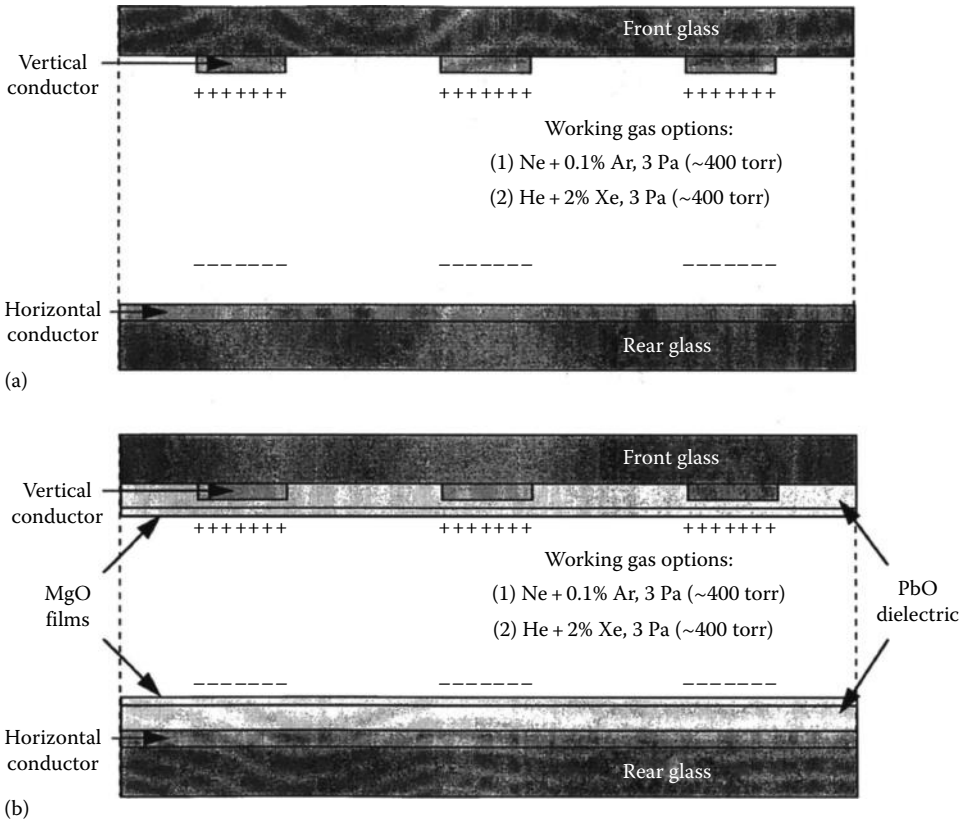


FIGURE 93.5 Schematic diagrams of (a) dc- and (b) ac-opposed electrode plasma pixels.

is typical. This frequency is too fast for the human eye to detect any *on–off* flicker, but allows sufficient flexibility for intensity and refresh control. In reviewing the literature on plasma displays, it is easy to confuse dc and ac pixels since dc pixels are often operated in pulsed mode and with electrode polarity reversal which distributes sputter damage over both electrode surfaces. The dc pixels are readily identified by conducting electrodes in direct contact with the discharge gas and the inclusion of a current-limiting resistor in the circuit for each pixel. While polarity reversal is optional for dc pixel operation, it is inherently required for ac pixel operation as discussed later. Drive electronics, current limiting, gray scale, and other aspects of both dc and ac pixel operation are discussed in greater detail in subsequent sections.

Figure 93.5b shows a schematic representation of an ac plasma pixel configuration. One can see that the differences between ac and dc pixel geometry are slight; however, the resulting operational differences are significant. In the ac pixel, the conductor wires are covered with a dielectric film. Typically, lead oxide (PbO), which has a dielectric constant of about 15, is deposited at a film thickness of about 25 μm. Most ac pixels are made with a thin-film (50–200 nm) magnesium oxide (MgO) dielectric coating covering the PbO and in contact with the working gas. This dual material dielectric film serves two principal functions: charge storage and secondary electron emission.

The exact voltage required for gas breakdown depends upon the gap width, the gas pressure, the gas composition, and MgO surface conditioning. For the pixel parameters shown in Figure 93.5b, an externally applied voltage of about 120–180 V is required to initiate a discharge. In the ac pixel, once the discharge is initiated, electrons and ions flow toward the anode and cathode, respectively, as in the dc pixel. However, in the ac case, charge carriers are unable to reach the conductor wires and instead collect as a surface charge on the dielectric coating. The electric field within the gas gap is always the sum of that produced by the externally applied voltage and that produced by the surface charge. During pixel firing, if the externally applied voltage is held constant for only a few microseconds, the net electric field within the gas gap very quickly decreases (~100–200 ns). The gap potential drop, produced by the surface charge, shields out the one produced by the externally applied voltage. Eventually, the gap electric field is insufficient to sustain the discharge and the pixel turns *off*. Thus, each ac pixel is inherently self-current-limiting and, unlike the dc pixel, requires no external resistance in series with it. At the start of the next ac half cycle, the externally applied voltage is reversed. When this occurs, the voltage across the gas gap is the sum of the external voltage and the voltage produced by the surface charge established during the previous discharge. If a sufficient surface charge is present, a new discharge pulse can be initiated by application of an external voltage, by which it would be insufficient to break down the gas. Within the new discharge, charge carriers flow quickly to reverse the polarity of the surface charge concentrations. Once again, the field within the gap is diminished and the discharge turns *off*. Storage of surface charge makes ac pixels easily controllable and provides them with their inherent memory properties. The presence or absence of surface charge determines whether or not a given pixel will discharge at the onset of the next ac half cycle of the externally applied voltage. The details of how these discharge dynamics are used to write, erase, and sustain each pixel are discussed in subsequent sections, along with drive mechanisms for gray scale and for pixel array refresh.

93.1.3 General Attributes of Plasma Displays

Plasma-driven flat panel displays offer a number of advantages over competing display technologies. The highly nonlinear electrical behavior of each pixel, with inherent memory properties, can be used to advantage in design of the drive electronics required to refresh and to update the pixel array of the display. The simplicity of the pixel design makes large-area manufacturing problems, such as alignment and film thickness uniformity, somewhat more manageable. Relative to color active matrix liquid crystal displays (AMLCDs) which use a thin-film transistor (TFT) to control each pixel, less-complicated manufacturing and less-complicated drive electronics give plasma flat

panel displays advantage for large-area applications. On the other hand, plasma displays require more robust drive electronics with voltages of 100–275 V. Plasma displays are also not well suited for portable applications since power consumption is high relative to other display technologies, but not restrictive for office or domestic use. The 76 cm (30 in.) diagonal color display manufactured by Photonics Imaging shown in Figure 93.4 has a peak power consumption of only 300 W [23]. At high power levels, plasma-driven flat panel displays are bright enough to be readable in sunlight. The displays are also easily adjusted to a low-ambient-light condition by discharge amplitude or frequency modulation.

Plasma flat panel displays are well suited for large-area (0.5–5 m) applications such as videoconferencing, large meeting room displays, outdoor displays, and simulators requiring large viewing areas. Thin, high-resolution, large-area, color plasma displays are also very attractive for desktop workstation or personal computer applications requiring high-resolution graphics. Note, too, that plasma flat panel displays have very large viewing angles, greater than 160° in many designs [22–24]. For displays using metal electrodes, one often finds that the best viewing angle is slightly off normal since the front electrode wire blocks out a portion of the pixel emission. This occurs both for monochrome pixels producing visible emissions within the discharge and for color plasma displays where the viewer sees visible red, green, and blue (RGB) emissions from vacuum ultraviolet (VUV) photon-stimulated phosphors. Some manufactures have investigated use of transparent electrodes, such as indium–tin oxide (ITO), but there is a trade-off with power consumption since the conductivity of ITO is less than that of metal electrodes [18]. In contemporary designs, the metal conductor width is thin (~20 μm) and its opacity does not present a major problem.

For color pixels, the discharge gas mixture is modified to produce emissions in the VUV. In all other respects, the operational principals of the plasma discharge by the pixel are identical for color and for monochrome displays. Ideally in color plasma displays, no visible emissions are produced within the discharge itself and VUV-photostimulated luminous phosphors are used to produce the required RGB visible light. The ac color pixel design concept shown in Figure 93.6 is that utilized by Fujitsu [18]. Long, straight barrier structures, each about 100 μm tall, are constructed parallel to and between each of the vertically oriented conductor wires on the rear glass plate. The sidewalls of these barriers are alternately coated with RGB-photostimulated phosphors. Note that the Fujitsu panel employs a three-electrode, ac-driven surface discharge pixel design which is slightly more complicated than the opposed electrode ac design shown in Figure 93.5b. This chapter will return to surface discharge configurations and other aspects of color pixel design and operation after reviewing fundamentals of the discharge physics and electrical behavior governing pixel operation.

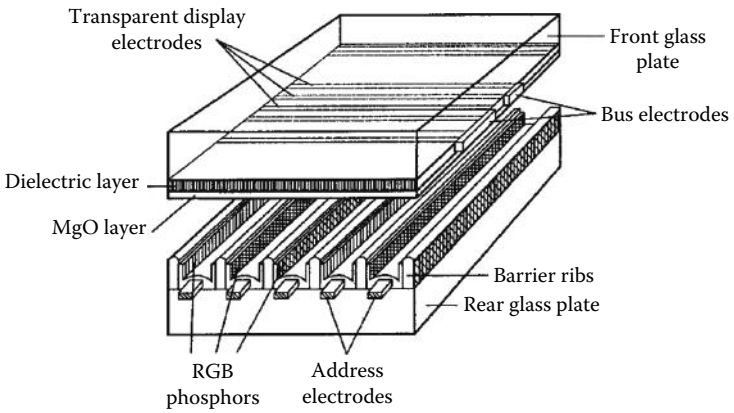


FIGURE 93.6 Structure of the ac color plasma display manufactured by Fujitsu. (From Mikoshiba, S., *Inf. Display*, 10(10), 21, 1994. With permission.)

93.2 Fundamentals of Plasma Pixel Operation

93.2.1 Atomic Physics Processes

Although simplistic in design, the plasma display pixel is a rich environment for study of basic atomic physics, electron collisional processes, photon production and transport, and plasma–surface interactions. The coupling of these processes for a neon–argon monochrome pixel discharge was nicely summarized in the diagram from Weber which is reproduced here as Figure 93.7 [4]. The reader interested in additional information on fundamental discharge physics is directed to one of the excellent textbooks in this field [25–27].

The physical processes governing of the behavior of the pixel discharge are closely coupled and form a closed-loop system. The discussion begins by assuming that a seed electron is resident within the gas gap and is subjected to an electric field which results from application of an externally applied voltage to the two conductors forming that pixel. Some of the gas and surface processes for production of the seed electrons will become evident as the discussion progresses. In order to ensure reliable discharge initiation, seed particles, which are either electrons or electron-producing photons or metastable atoms, are often provided by a controlled source which may be external to the pixel being fired. Some display panels include electrodes for production of seed particles at the edges of the panel outside the field of view or hidden behind opaque conductor wires. Other display panels use well-controlled temporal

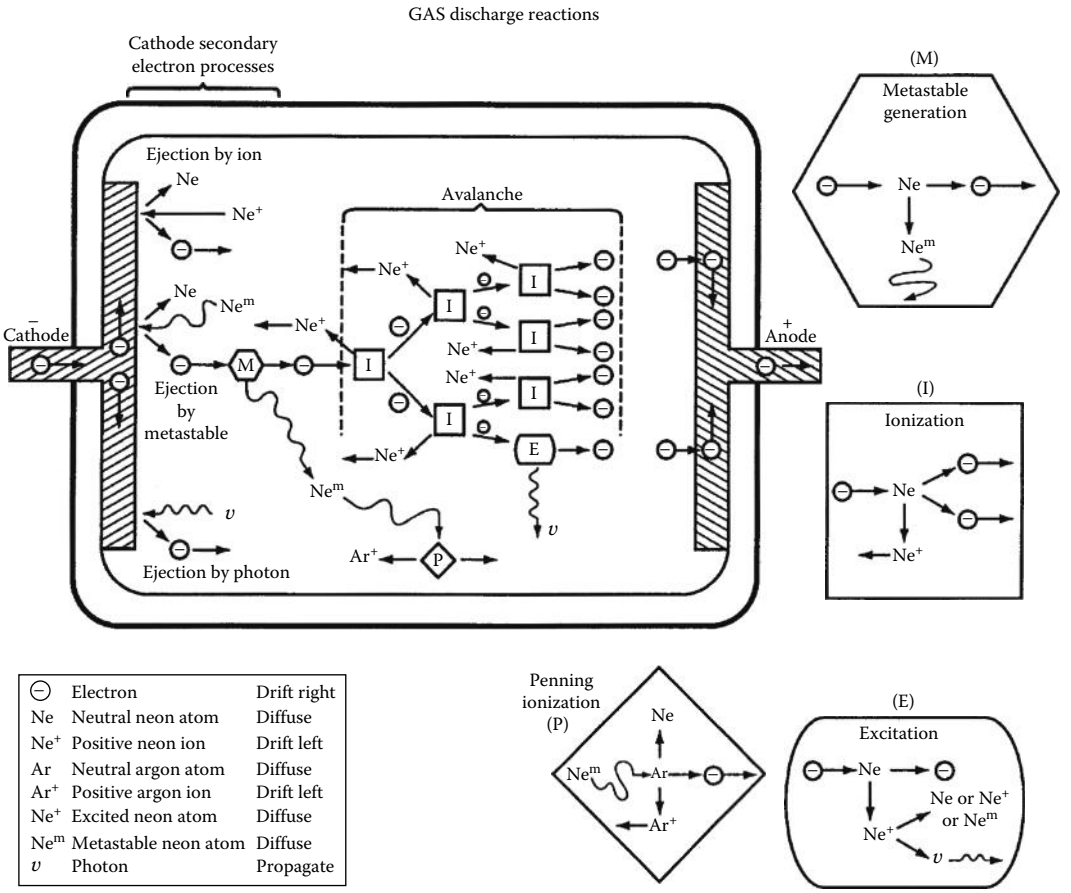


FIGURE 93.7 Collisional and surface interactions in a gas discharge. (From Weber, L.F., *Flat Panel Displays and CRTs*, Tannas, L.E., Jr., Ed., Van Nostrand Reinhold, New York, 1985. With permission.)

sequencing to ensure that the nearest-neighbor pixels provide seed particles for one another [4,19,28]. Pixel-addressing sequences are discussed further later in this chapter.

The transport of electrons or ions across the gas gap is a balance between field acceleration and collisional energy loss. In the example of Figure 93.7, the gas is mostly neon (98%–99.9%) and field-accelerated electrons will predominantly collide with Ne atoms. The quantum energy level diagram for excitation of the Ne atom is shown schematically in Figure 93.7 [29]. Note that the lowest-lying excited state is 16.6 eV above the ground state, while the ionization energy is 21.6 eV. This means that electrons with energies <16.6 eV can only experience elastic collisions with the Ne atoms. When an electron is field-accelerated to energy in excess of 16.6 eV, inelastic collisions which transfer energy from the incident electron to one of the outer-shell electrons in the Ne atom can take place. Incident electrons with kinetic energies in excess of 21.6 eV can drive ionization reactions:

$$\mathrm{Ne} + \mathrm{e}^- \rightarrow \mathrm{Ne}^+ + 2\mathrm{e}^- \tag{93.1}$$

Excitation and ionization collisions transfer energy from the electron population to the neutral atoms in the gas. At the same time, the electron population available to ionize the Ne further is increased with every ionizing event. The result is the discharge avalanche schematically shown in Figure 93.9, which manifests itself experimentally as a rapid increase in electric current flowing in the pixel gas gap. In dc panels, an external resistor of about $R = 500\ \mathrm{k}\Omega$ is placed in series with each pixel. The amplitude of the externally applied voltage provided by the driving electronics, V_a, is held constant and the total voltage across the gas gap, $V_g = V_a - IR$, decreases as the circuit current, I, increases. Very quickly, a steady-state dc current in the gas gap and in the circuit is established. Brightness and gray scale are controlled by frequency modulation of the pulsed dc pixel firing using a base frequency of about 50 kHz. In ac pixel discharges, electrons and ions are driven by the applied field to the dielectric-covered anode and cathode, respectively. The buildup of charge on the dielectric surfaces shields the gap region from the field produced by the externally applied voltage. Eventually, the electric field in the gap drops below a level sufficient to sustain the discharge and the pixel turns *off*.

For electron energies greater than 16.6 eV, collisions with Ne atoms can excite outer-shell electrons in the atom to one of the numerous excited energy states shown in Figure 93.7:

$$\mathrm{Ne} + \mathrm{e}^- \rightarrow \mathrm{Ne}^{\mathrm{ex}} + \mathrm{e}^- \tag{93.2a}$$

$$\mathrm{Ne}^{\mathrm{ex}} \rightarrow \mathrm{Ne}^* + h\nu \tag{93.2b}$$

Most of these excited states have short lifetimes ranging from fractions to tens of nanoseconds [30] and quickly decay to lower-lying atomic quantum states accompanied by the emission of a characteristic photon, indicated in Equation 93.2 by $h\nu$, the product of Planck's constant times the photon frequency. As can be seen in Figure 93.8, the characteristic red–orange Ne gas emissions result from electron transitions within the atom from higher-energy 2p quantum states to lower-lying 1s energy levels [30,31]. Two of the four 1s energy levels radiate to ground-emitting VUV photons with wavelengths of 74.4 and 73.6 nm. Due to quantum mechanical exclusion principles, electron decay from the other two 1s levels is more complex and depends upon fine details of the electronic wave function and upon very small perturbing interactions [31]. Consequently, decay lifetimes for these so-called metastable states are measured in seconds, which is very long relative to other dynamic physical processes governing pixel discharge behavior, such as charge or neutral particle transport. An Ne atom with an electron trapped in one of these metastable levels harbors 16.6 eV of latent energy. The metastable atom, Ne*, is unable to dissipate its stored energy in collisions with ground-state Ne atoms, yet readily liberates its energy whenever a lower-lying energy configuration can be accessed. The principal channels in this system to lower-energy configurations are Ne* collisions with Ar or Ne* incidence onto pixel interior surfaces.

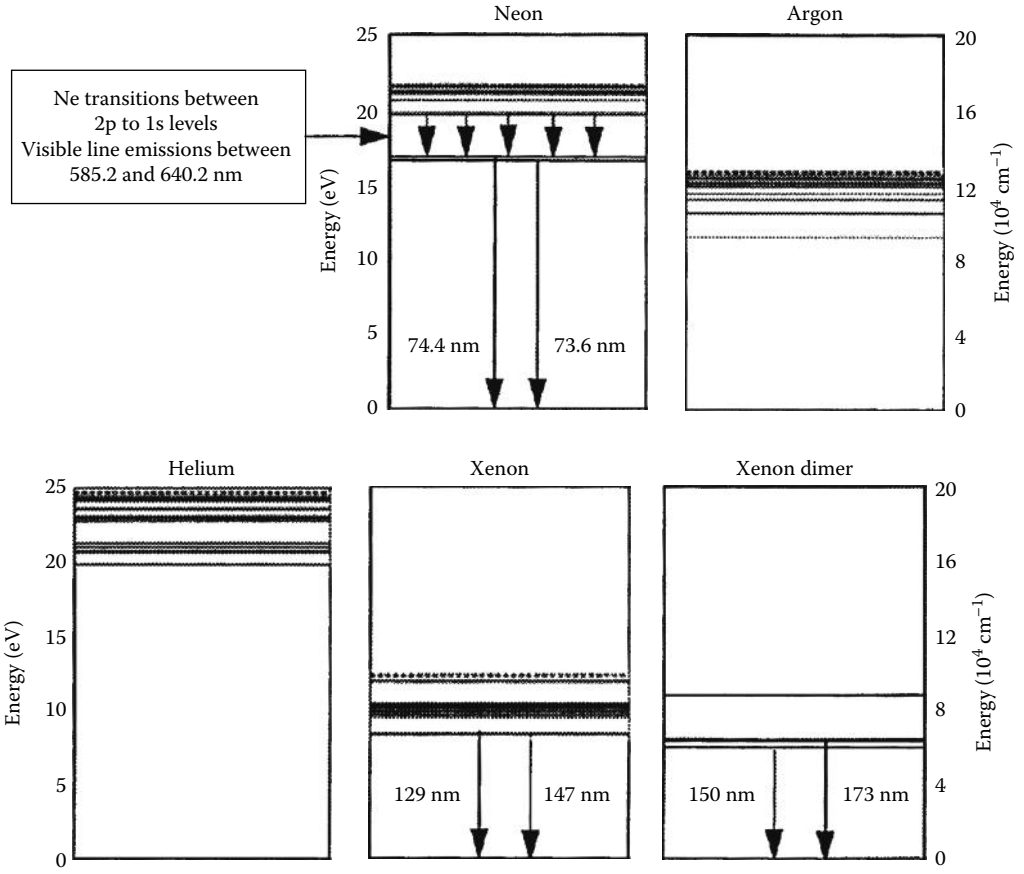


FIGURE 93.8 Quantum energy level diagrams for He, Ne, Ar, Xe, and the Xe^{2*} dimer.

Figure 93.8 shows simplified energy level diagrams for several inert gases. The relative positioning of the allowable energy levels provides insight into the energy exchange that occurs in collisional coupling. The ionization energy of the Ar atom is 15.8 eV and lies 0.8 eV below the metastable-state Ne*. Consequently, the Ne* has sufficient stored energy to ionize the Ar atom:

$$Ne^* + Ar \rightarrow Ne + Ar^+ + e^- \tag{93.3}$$

Ionizing reactions of this type are called Penning reactions, and gas mixtures that rely on metastable states of the majority gas constituent (Ne) for ionization of the minority gas constituent (Ar) are referred to as Penning gas mixtures [25,26,32]. Figure 93.9 shows the efficiency with which charge pairs are produced through ionization within Ne/Ar Penning gases containing various fractions of Ar. The curves show that for any given pressure, ion pair production per volt applied is optimal at low Ar gas fractions (0%–10%) except for very large values of E/P, greater than 75 V/m/Pa (100 V/cm/torr), where E is the electric field strength and P is the gas pressure. Penning gas mixtures have been studied for many years. Figure 93.9 shows the original data on Ne/Ar gas breakdown published by Kruithof and Penning in 1937 [32]. An extensive volume of literature has been published on inert gas Penning processes since then, and the interested reader is referred to the excellent texts which have recently been rereleased through the American Vacuum Society and MIT Press [25,26].

Plasma display pixels usually operate at pressures near 53.3 kPa (400 torr) in order to achieve sufficient photon production and brightness. Typical pixel fields are roughly 100 MV/m. Consequently,

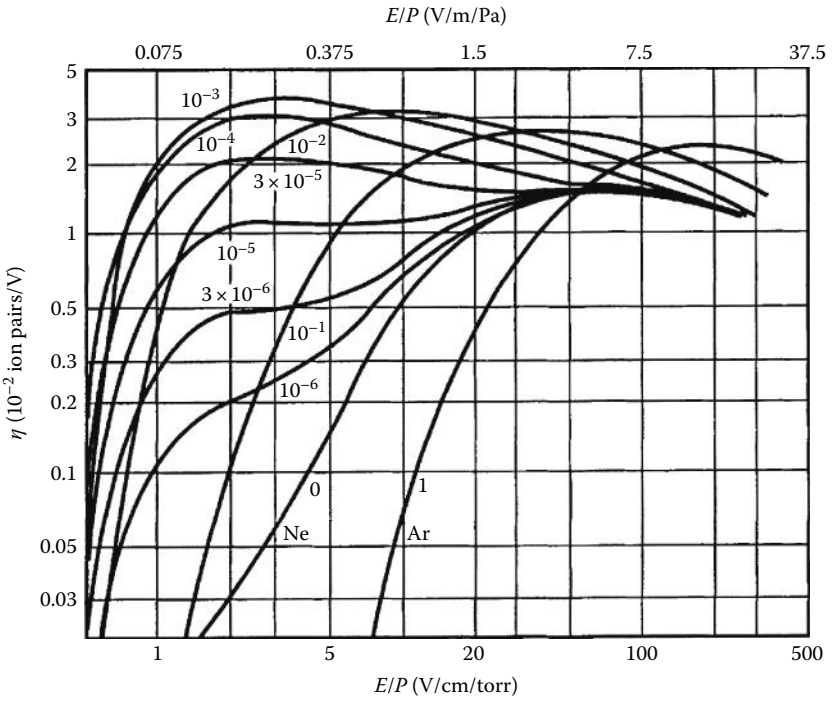


FIGURE 93.9 Ionizing collisions plotted vs. electric field strength divided by pressure. The numbers on each curve indicate the ratio of the Ar partial pressure to the total gas pressure. (From Brown, S., *Basic Data of Plasma Physics—The Fundamental Data on Electrical Discharges in Gas*, American Institute of Physics Press, New York, 1993. With permission.)

plasma pixels operate with *E/P* values near 18.8 V/m/Pa (25 V/cm/torr). Both charge pair production and luminous efficiency are then optimized with Ar gas fractions between 0.1% and 10%, depending upon the specifics of the pixel gas pressure, gap width, and driving voltage. For a given applied voltage, the product of the gas pressure (P) and the gas gap dimension (d) provides a measure of the balance between electron acceleration by the electric field and electron energy loss due to collisions with the background gas. Paschen curves, which plot the gas breakdown voltage vs. the Pd product, for several inert gas mixtures are shown in Figure 93.10 [26,33,34]. In each case, minimum voltage for breakdown occurs at a value of the Pd product which is dependent upon the ionization levels, collisionality, and energy channels within the gas. For example, in Ne atomic excitation and ionization processes dominate, while in air much of the energy absorbed by the gas goes into nonionizing molecular vibration, rotation, and dissociation. For fixed pressure, the Paschen curves show that increased gap dimension lowers the electric field strength per volt applied and a large voltage is required for breakdown. On the other hand, if d is reduced for a given pressure, the electric field strength can be large, but electrons transit the gap without initiating a sufficient number of collisions to drive the type of discharge avalanche shown in Figure 93.7. If the gas gap, d, is held fixed while pressure is varied, the shapes of the Paschen curves are again explained by electron acceleration and collisional processes. For high pressures, the mean free paths between electron collisions with the background gas atoms are short and electrons are unable to accelerate to energies sufficient to initiate ionization unless the electric field is especially strong. At low pressures, the electrons may be accelerated by the field to energies sufficient to initiate ionization, but few collisions with the background gas occur and, again, the avalanche is difficult to initiate. Penning processes are especially efficient at driving ionization. Introduction of 0.1% Ar into the neon gas lowers the minimum breakdown voltage from the value near 250 V, shown in Figure 93.10, to about 150 V. The minimum breakdown voltage occurs at a Pd product of 40 Pa m (30 torr cm) for this gas mixture.

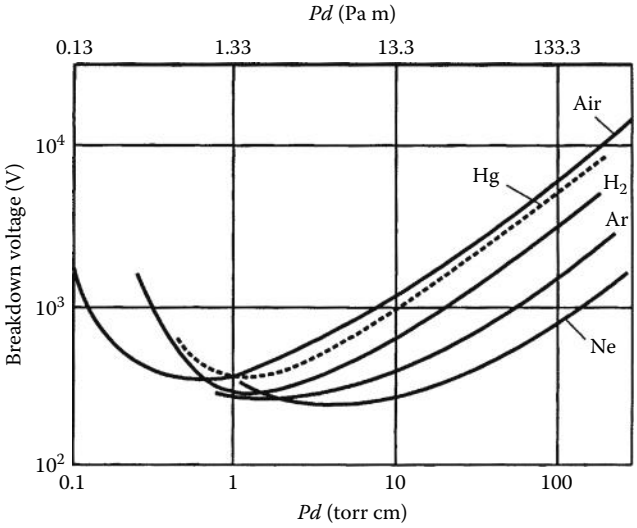


FIGURE 93.10 Breakdown voltage as a function of pressure—gas gap length product for various gases. (From Brown, S., *Basic Data of Plasma Physics—The Fundamental Data on Electrical Discharges in Gas*, American Institute of Physics Press, New York, 1993. With permission.)

93.2.2 Discharge Physics for Plasma Pixels

Within any discharge, electrons move very quickly, while the more massive ions move relatively slowly in comparison. In charge-neutral plasma that is subjected to an externally applied electric field, the mobile electrons quickly respond to the applied field and rush toward the anode. The inertia-laden ions, in a much slower fashion, begin their motion toward the cathode. Very quickly, a local charge imbalance is established as the electrons reach the anode faster than the rate of arrival of ions at the cathode. Poisson's equation

$$\nabla \cdot E(x) = 4\pi\rho(x) = 4\pi e(n_i(x) - n_e(x)) \tag{93.4}$$

shows that a local electric field is established in response to the net positive charge density, $\rho(x)$, in the plasma region. Here, $n_i(x)$ and $n_e(x)$ are the spatial profiles of the ion and electron densities, respectively, and e is the electron charge. The field established retards the rate at which electrons flow out of any volume within the plasma column and forces them to follow the net ion motion. The ion drift motion is correspondingly accelerated, but this acceleration is smaller by a factor proportional to the mass ratio of the electron to the ion. The net ion/electron motion is called ambipolar flow and is described in detail in many basic plasma physics texts [25–27].

In steady-state dc plasma pixel discharges, the amplitude of the current flowing in the circuit and in the gas gap is defined by the value of the applied voltage and the resistor in series with the pixel. Steady-state operation dictates that charge buildup within the gap region cannot occur. The rate at which charge particle pairs arrive at the electrodes must equal their rate of production due to ionization. At the same time, the rates at which ions and electrons leave the plasma volume, arriving at the cathode and anode, respectively, must be equal. Equilibrium is sustained by establishment of the spatial potential profile within the gas gap shown in Figure 93.11a. Due to the high electron mobility, the plasma is extremely efficient in shielding out externally applied electric fields. As a result, the potential profile is flat across the gas gap of a pixel sustaining a fully developed discharge. The entire potential drop is localized in a small zone called the sheath adjacent to each electrode. The spatial extent of the sheath is

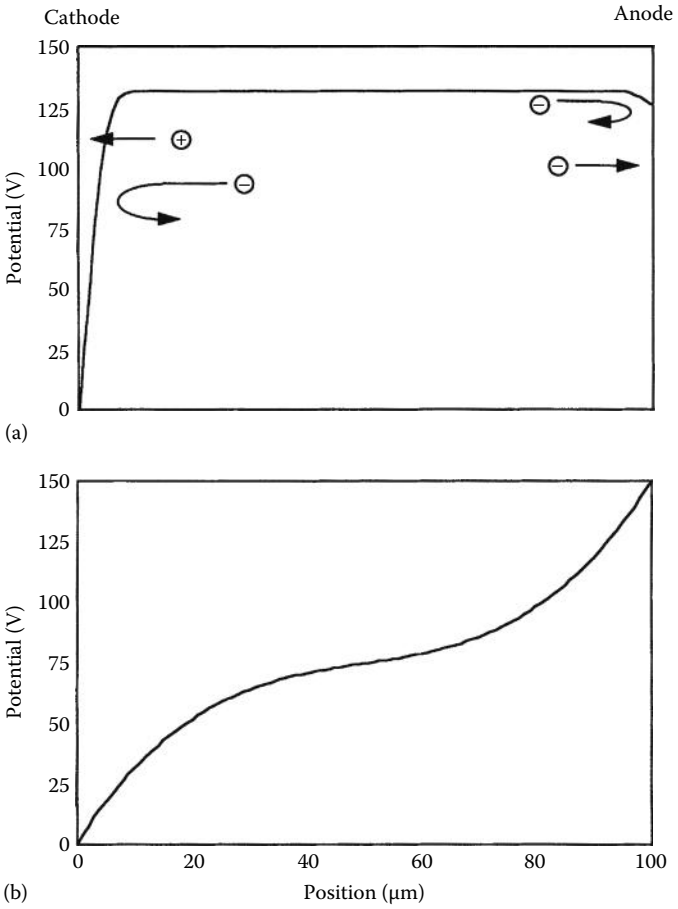


FIGURE 93.11 Potential profiles in the pixel gap region for (a) high-electron-density and (b) low-electron-density discharges.

determined by the effectiveness of the electron population in shielding out the electric fields produced by the electrode potentials. The Debye length

$$\lambda_D = \sqrt{\frac{kT_e}{4\pi e^2 n_e(x)}} \tag{93.5}$$

provides a measure of the shielding distance. The expression for λ_D implies that the sheath thickness increases with increasing electron temperature, T_e, and decreases as the electron density, n_e, increases. For fully developed plasma pixel discharges, the product of Boltzmann's constant and the electron temperature, kT_e, is at most a few electron volts, and n_e is of order $10^{16}/m^3$. Thus, the sheath thickness is roughly 5 μm. The potential within the plasma region adjusts, V_p, within the discharge volume rises to a value just above that of the applied voltage at the anode. Consequently, only the most energetic electrons can overcome the potential barrier at the anode which adjusts to a potential such that the rate of electron loss at the anode equals the rate of ion loss at the cathode. For ac plasma pixels, a similar potential profile is established, but changes dynamically as the pixel pulse evolves. Charge pairs incident upon the anode and cathode in ac pixels are trapped there by the dielectric film covering the conductor wires. Consequently, the potential at the discharge boundary is diminished as surface charge collects at each electrode, as shown in Figure 93.11b. Ultimately, the discharge terminates as the electric field produced

by the surface charge cancels that produced by the externally applied voltage. As the density of charge carriers is reduced near the termination of an ac pixel discharge pulse, the effectiveness of the electrons to shield out electric fields within the gap is diminished. In this situation, the sheath potential drop is small but the sheath region occupies a large fraction of the gap [35,36].

93.2.3 Plasma Surface Interactions

93.2.3.1 Ion-Induced Secondary Electron Emission

Ions arriving at the cathode sheath are accelerated by the sheath potential drop. Incident ions strike the cathode with kinetic energies equal to the plasma potential, V_p, which is just over 200 V in the example shown in Figure 93.12. Ions incident on the cathode quickly capture an electron, additionally depositing on the cathode surface energy equal to the recombination or ionization energy for that atom. Energy deposition on the cathode surface drives two important processes for plasma pixels—ion-induced secondary electron emission and sputtering. The first process significantly enhances the luminous efficiency of plasma pixels. The second shortens their operational lifetime as is discussed in subsequent sections.

Ion-induced secondary electron emission occurs when ion energy deposition on the surface results in electron ejection. Secondary electrons are exceptionally effective at driving discharge ionization since they gain large amounts of kinetic energy as they are accelerated across the cathode sheath and because they have ample opportunities for ionizing collisions as they traverse the entire width of the gas gap. The secondary electron emission coefficient, γ, is defined as the number of electrons ejected per incident ion [25,26]. As one would expect, γ varies with incident ion energy and with cathode material. Most ac plasma display panels take advantage of the strong secondary electron emission of MgO, which is also a good insulating material as required for surface charge storage in ac operation. Measurement of the MgO γ-value is difficult, especially for low-energy ion incidence (<500 eV), and is complicated by charge buildup on the samples during the measurements [37]. Most often, relative values of secondary electron yields for different materials are deduced from discharge intensity measurements [11,12,38–41].

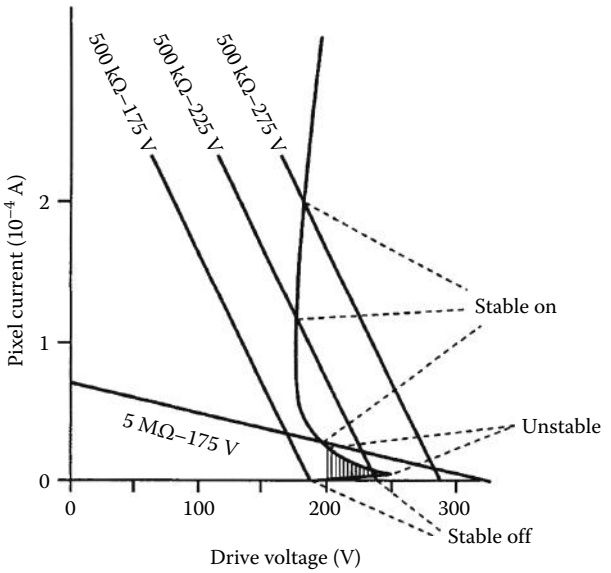


FIGURE 93.12 A representative current–voltage characteristic for gas breakdown. Load lines representative of plasma pixel operation are also shown. (From Weber, L.F., *Flat Panel Displays and CRTs*, Tannas, L.E., Jr., Ed., Van Nostrand Reinhold, New York, 1985. With permission.)

Chou directly measured the ion-induced secondary electron emission coefficient for MgO using a pulsed ion beam with sample surface neutralization between pulses. For ion incidence at 200 eV, he found $\gamma = 0.45$ and $\gamma = 0.05$ for Ne^+ and Ar^+, respectively [38]. Note, too, that photons and metastable atoms incident on the electrode surfaces are also capable of initiating secondary electron emission, as shown in Figure 93.7. Since neither photons nor metastables are influenced by the electric fields within the gas gap, they propagate isotropically throughout the gas volume and are often utilized as seed particles.

93.2.3.2 Sputtering

Ions accelerated across the sheath deposit energy on the cathode surface. This often initiates sputtering, whereby an atom within the cathode material is ejected from the surface. Sputtering processes erode the cathode surface and degrade pixel performance. Contamination of the discharge by sputtered surface impurities can lead to reduction in luminous efficiency due to visible emissions from the contaminant atoms or molecules which compromise the color purity of the pixel. Unwanted surface coatings from redeposited materials can also degrade the electrical characteristics of the pixel or, in color applications, shield the phosphors from VUV photons, and further degrading luminous efficiency. For argon ion, Ar^+, bombardment of MgO surfaces at 2 keV, the measured sputtering yield is slightly greater than one ejected atom per incident ion [42]. Data on sputtering yields at lower-energy ion incidence are difficult to obtain. Because yields are small, large incident ion currents are required to obtain measurable signals and sample charging is once again a problem. In spite of the lack of detailed data on low-energy MgO sputtering, manufactures of ac plasma panels have been able to demonstrate display lifetimes well in excess of 10,000 h [18,23]. Schone et al. [43] have demonstrated that Rutherford backscattering of high-energy (2.8 MeV) α-particle can be used to measure the thickness of MgO film on a PbO substrate. The film thickness accuracy obtained was ±1.5 nm. Because the technique requires a large (and expensive) particle accelerator, this technique is a very nice research tool but is ill suited for any fabrication line measurements.

93.3 Pixel Electrical Properties

93.3.1 Electrical Properties of DC Pixels

Figure 93.5 shows schematic diagrams and circuit models for dc and ac pixels. In the dc case, the pixel gas gap functions electrically as a variable impedance resistor. Prior to gas breakdown, the resistance is large and the pixel represents an open-circuit element. Once breakdown is initiated, the plasma is an excellent conductor and offers only modest resistance, R_P, to current flow. Since $R \gg R_P$, the circuit equation simplifies to

$$V_a = I(R + R_P) \approx IR \tag{93.6}$$

and the circuit current, I, is defined by the amplitude of the applied voltage and the size of the circuit series resistor, R. The externally applied voltage, V_a, is typically a 50 kHz square wave with a fast voltage rise time (~50 ns). The dc driving voltages range from 175 to 275 V and a typical value for the series resistor is $R = 500$ kΩ. Pixel currents then range from 0.35 to 0.55 mA. Note that without a large resistance in series with the pixel, the current is limited by some physical failures such as melting of the pixel electrodes.

Figure 93.12 shows the characteristic I–V behavior of a dc pixel which has a breakdown voltage of 250 V [4]. Only a very small current due to a few stray charge carriers flows across the gas gap as the voltage increases from 0 to 250 V and the pixel remains in the *off* state. At the breakdown voltage, the situation is dynamic with the current growing rapidly and the voltage across the gas gap dropping as a result. The steady-state operating point achieved is identified by the intersection of the load line, $V_a = IR$, with the discharge I–V characteristic as shown in Figure 93.12. For an applied voltage of 175 V the pixel is

always *off*, while for $V_a = 275$ V the pixel is always *on*. For a line resistance of 500 kΩ, the bimodal operation and memory of the dc pixel at $V = 225$ V is evident in the figure. If an applied voltage of 225 V is approached from the low-voltage direction, the pixel remains *off*. If, on the other hand, a large voltage is applied and subsequently lowered to $V_a = 225$ V, then the pixel will be in an *on* state. Note that the region where the 225 V/500 kΩ load-line intersects the negative resistance portion of the *I–V* characteristic is unstable. The pixel discharge will quickly transition to either the stable *on* or stable *off* operating point. As a practical matter, one should note that the negative resistance region of the *I–V* characteristic curve cannot be experimentally measured in a pixel circuit operating with a 500 kΩ series resistance. Instead, as shown in the figure, a much larger series resistor, $R = 5$ MΩ, provides a load line with slope small enough to produce stable operation in the negative resistance regime.

93.3.2 Electrical Properties of AC Pixels

The physical design of an opposed electrode ac pixel is shown in Figure 93.5b. Electrically, the pixel functions as a breakdown capacitor and is described by the circuit equation:

$$V_a(t) = I(t)R + \frac{1}{C}\int_0^t I(t')dt' = I(t)R + \frac{Q(t)}{C} \tag{93.7}$$

where

V_a is the externally applied voltage
I is the circuit current
C is the pixel capacitance
Q is the charge collected

For ac pixels the line resistance, R, is minimized in order to minimize power consumption and Equation 93.7 simplifies to

$$V_a(t) = \frac{1}{C}\int_0^t I(t')dt' = \frac{Q(t)}{C} \tag{93.8}$$

The capacitance for each pixel is the series summation of the capacitance for each dielectric film and for the gas gap:

$$\frac{1}{C} = \frac{1}{C_{PbO}} + \frac{1}{C_{MgO}} + \frac{1}{C_{gas}} \tag{93.9}$$

In each case,

$$C_i = \frac{e_i A}{d_i} \tag{93.10}$$

where

i is the material index
Surface area, A, is roughly equal to the square of the conductor wire width

As shown in Figure 93.5b, an ac pixel is typically constructed with a PbO film of thickness $d = 25$ μm, while the thin-film MgO has thickness $d = 50$–200 nm. The lead oxide has a dielectric constant of

roughly $\varepsilon_{PbO} = 15\varepsilon_0$, while that for MgO is $\varepsilon_{MgO} = 6\varepsilon_0$ with exact values dependent upon the film purity and microstructure [34]. Note that the MgO contribution to the total capacitance is negligible and that this material is incorporated into the design because of its excellent secondary electron emission properties. Prior to gas breakdown, the capacitance of the pixel is attributed largely to the gas gap. For 20 μm thick conductor wires, the capacitance of a pixel gas gap prior to breakdown is about 500 pF. The time derivative of Equation 93.8 gives the circuit current:

$$I(t) = C\frac{dV(t)}{dt} \tag{93.11}$$

This charge displacement current appears as the initial large amplitude current peak in Figure 93.13, which shows the temporal current response of a 45 × 45 ac pixel array to a single pulse within a 50 kHz square wave applied voltage pulse train. The electrical measurement shown was made using a simple induction loop probe to measure the current and a high impedance voltage probe (1 MΩ, 3 pF) to monitor the applied voltage. The signals were captured using a high-speed (300 MHz) oscilloscope.

If the applied voltage amplitude is below the gas breakdown threshold, only the capacitor charging displacement current, defined by Equation 93.11, is observed as shown in Figure 93.13a. If the voltage for gas breakdown is exceeded, a second current pulse due to the plasma discharge current within the gas gap is observed in the circuit (Figure 93.13b). The plasma pulse is accompanied, of course, by strong

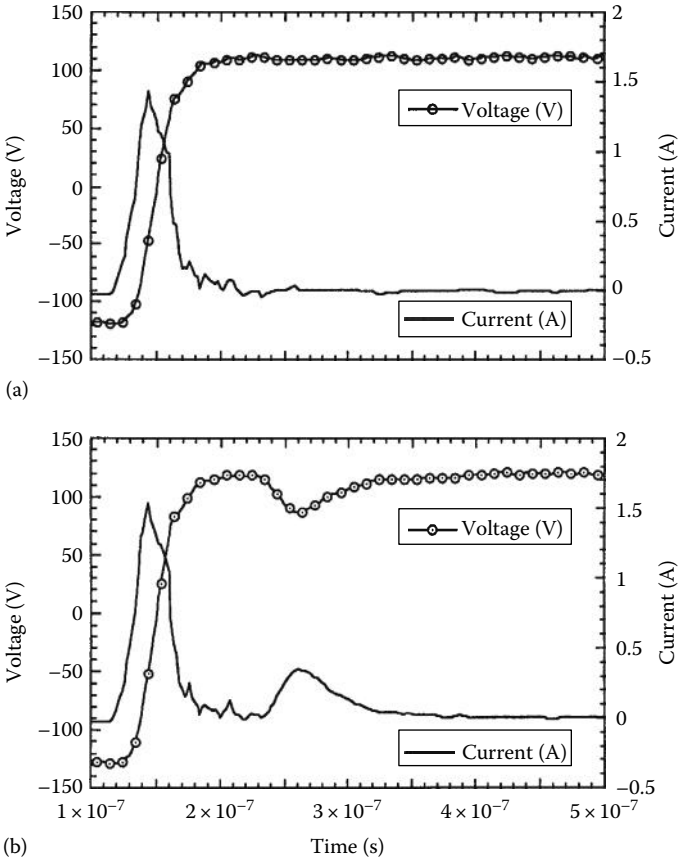


FIGURE 93.13 Voltage and current traces for a 45 × 45 array of ac plasma pixels in the (a) *on* and (b) *off* states. Drive voltage amplitudes were 117 and 127 V, respectively.

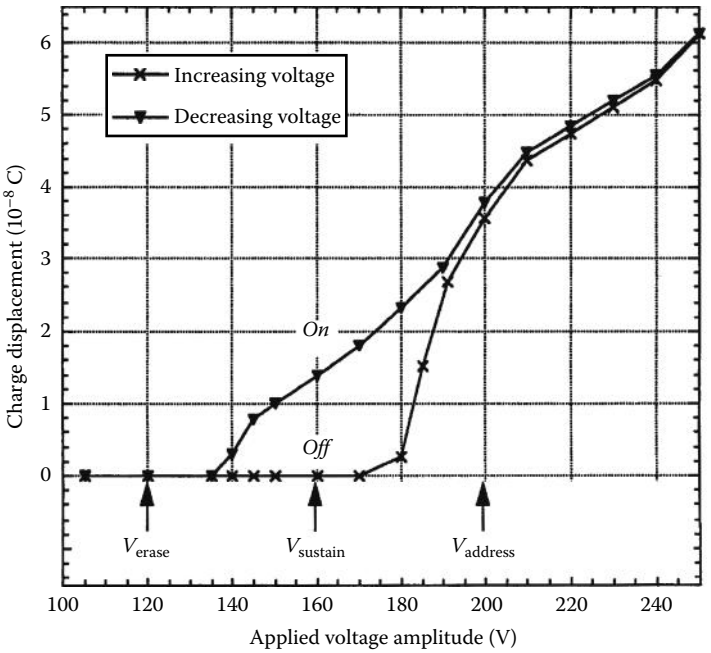


FIGURE 93.14 Discharge charge displacement for operation of a 45 × 45 array of ac-opposed electrode pixels with an He–Xe (2%) gas mixture at 53.3 kPa (400 torr).

photon emission from the gas gap region. The total charge displacement in the discharge pulse as a function of amplitude of the square wave–applied voltage is plotted in Figure 93.14 for a helium–xenon (2%) Penning gas mixture [44]. The hysteresis or inherent memory property of the ac pixel is apparent. As the applied voltage amplitude is increased from 0 to 180 V, no measurable current flows across the pixel gas gap. When no surface charge is present below 180 V, the electric field within the gap region is insufficient to drive the electron collisions into the avalanche regime. For any voltage amplitude in excess of 180 V, a gas discharge is initiated and the pixel turns *on*. If a pixel is subjected to a single voltage pulse with amplitude <135 V, the pixel turns *off* even if a surface charge is present.

In ac pixels, charge pairs produced during one discharge pulse collect on the surfaces of the dielectric films at the boundaries of the gas gap and are available to assist formation of the next discharge pulse in the sequence. In a fully developed ac pixel discharge, the surface charge accumulation on the dielectric film produces an electric field within the gas gap, which cancels the gap field produced by the externally applied voltage. This is shown in Figure 93.15, which is a composite representation of experimental current measurements and computational model predictions of the surface charge accumulation producing the surface or wall voltage [35,45]. When the polarity of the applied voltage is reversed, the potential drop due to the surface charge and that due to the applied voltage suddenly are additive as shown in the figure. The gas gap is momentarily subjected to an intense electric field which results from a potential drop roughly equal to twice the applied voltage. The presence or absence of surface charge results in the bimodal current–voltage behavior is shown in Figure 93.14.

Addressing of ac pixels is easily accomplished by taking advantage of the inherent memory of the pixel that results from this bimodal *I–V* behavior. For the pixel electrical properties shown in Figure 93.14, each pixel would be continuously supplied with an ac square wave–applied voltage pulse train with an amplitude of 160 V, called the sustain voltage, $V_{sustain}$. If the pixel is initially in an *off* state, it will remain so indefinitely since no surface charge is available to enhance the field produced by the sustain voltage. To turn the pixel *on*, a single high-amplitude voltage pulse, called an address (or write) pulse is delivered across the pixel electrodes. In this example, an address pulse of 200 V

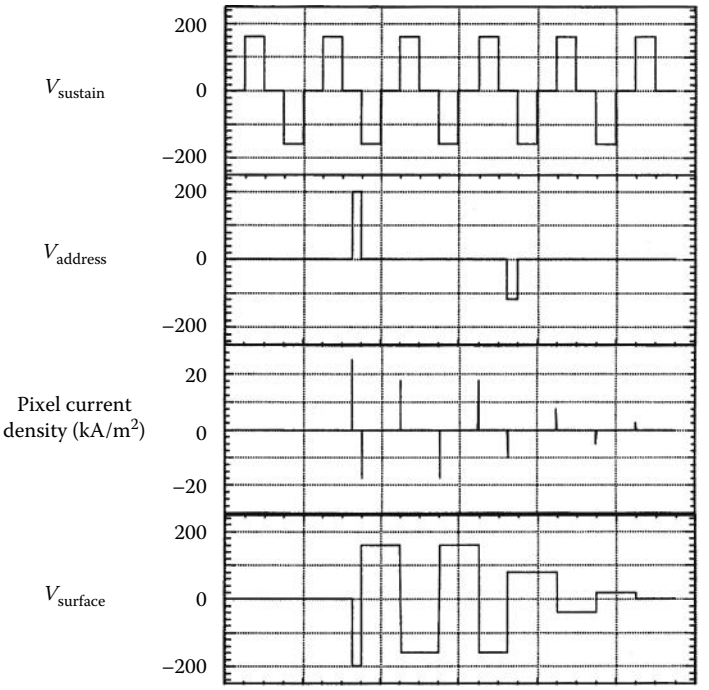


FIGURE 93.15 Sustain and address voltage waveforms for ac-driven plasma pixels. The amplitude of the pixel current density and wall voltage resulting from the surface charge buildup provide a measure of the discharge intensity.

initiates a discharge whose charge pairs collect on the internal dielectric surfaces of the pixel. The self-limiting nature of the ac pixel is such that the surface charge concentration produced for a fully developed pixel discharge completely shields the gap region from the externally applied field. When the next sustain polarity reversal occurs, the pixel gas gap experiences a voltage equal to the sum of the sustain voltage (160 V) plus the voltage due to the surface charge produced by the previous pulse, $V_{surface} = 200$ V in this case. The new gap voltage of 360 V is more than sufficient to initiate a second discharge and to establish a new surface charge whose polarity is opposite that of the preceding pulse. Once again, the surface charge adjusts to produce a voltage exactly canceling the field of the applied voltage. For this pulse, $V_{wall} = 160$ V, and the next sustain voltage polarity reversal subjects the gap to a potential difference of $V_{gap} = V_{sustain} + V_{address} = 160\ \text{V} + 160\ \text{V} = 320$ V, which is again sufficient to initiate a new discharge pulse. Consequently, the pixel remains in the *on* state until action is taken to eliminate or diminish the surface charge buildup accompanying each discharge. This is accomplished by application of a single low-voltage pulse called an erase pulse with amplitude $V_{erase} = 120$ V for the example shown in Figures 93.14 and 93.15. Application of the erase pulse produces a potential drop across the gas gap of $V_{gap} = V_{applied} + V_{surface} = V_{erase} + V_{surface} = 120\ \text{V} + 160\ \text{V} = 280$ V. The erase pulse produces a discharge of lower intensity which is insufficient to reestablish a reversed polarity surface charge. Consequently, the erase discharge diminishes the concentration of the surface charge so that no discharge is initiated during the next pulse in the sustained applied voltage train. Ideally, the erase pulse drives the surface charge concentration identically to zero, but this rarely occurs in practice and is not essential for ac pixel operation, as can be seen in Figure 93.15. Very low-intensity discharges with negligible photon production drive the pixel to its ideal *off* state within a few ac cycles. Fortunately, these minor deviations from the ideal *off* condition have little effect on subsequent write pulses for frequency-modulated ac operation, and therefore do not affect the timing of pixel addressing and refresh which is covered in the next section.

93.4 Display Priming, Addressing, Refresh, and Gray Scale

Pixel priming is necessary to provide the initial source of electrons, or the priming current, required to initiate a discharge avalanche. Metastable atoms or photons can also be used as priming particles since these produce electrons via ionization of the background gas. Pilot cell priming and self-priming are two options used in currently available commercial products. In pilot cell priming, a separate cell which generates electrons is located near the pixel to be addressed. Pilot cells are often located on the periphery of the display outside the viewing area, yet can produce seed electrons throughout the display. In self-priming, an auxiliary discharge created within each pixel provides the priming electron source for the main pixel discharge. These priming discharges are often hidden from view by positioning them behind opaque conductor wires. Introducing a trace amount of radioactive Kr^{85} into the gas mixture provides a passive priming option. The ionizing radiation from Kr^{85} generates priming electrons uniformly throughout the display interior. Because the required Kr concentration is low and because the β-radiation produced cannot penetrate the glass enclosure of the display, the radiation exposure risk to the display user is negligible. However, display manufacture using radioactive seeding involves potential health hazards associated with radioactive material handling. Consequently, this seeding approach, while very effective, is not at present employed in commercial products.

A simplistic scanning scheme for pixel illumination is shown in Figure 93.2, reproduced here from Ref. [4]. The scan switches on one axis open and close sequentially in a repetitive fashion, while the data switches on the other axis determine if the pixel is fired on a given scan. This simplistic refresh and data update method fails to take advantage of the discharge properties or inherent memory functions available with plasma pixels. High-resolution dynamic displays utilizing this address scheme would not be cost-competitive since display drivers constitute a significant portion of the total cost of plasma displays. Driver circuit costs also increase with required output voltage. Thus, it is desirable to design plasma displays with operating voltages as low as possible and which require the fewest number of driver chips. Designers strive then to maximize the number of pixels driven by a single chip.

For nonmemory dc pixels, one option for reducing the number of external drive switches required is to sweep the firing of priming discharges repetitively across each pixel row, such as in the self-scan circuitry developed by Burroughs [46,47]. More recently, NHK has developed a pulse memory drive scheme for its 102 cm (40 in.) diagonal dc HDTV plasma display, which is being widely used [48]. Sustain operation at high frequency is used to take advantage of residual charge pairs and metastable atoms present in the pixel gas volume as result of the preceding discharge [49]. In this fashion, each pixel is self-seeding, with seed particle populations dependent upon the time elapsed since the termination of the preceding discharge. The high-frequency operation is fast enough to take advantage of the short duration memory characteristic of the dc pixel. As the sustain voltage pulse train is applied to the electrode of a pixel, it will remain in the *on* or *off* state indefinitely until an address or erase pulse is supplied. In the NHK scheme, an auxiliary anode is used to assist in the *address* access operations. Figure 93.16a shows the block diagram of such as system, while Figure 93.16b shows the time sequences for the scheme [48]. Note that the pulses are dc and that *on*-state pulses have larger gap voltages than erase pulses. The timing sequence is critical to address a pixel selectively within the matrix. Implementation of this scheme requires (1) display anode drivers, (2) auxiliary anode drivers, (3) cathode drivers, and (4) and interfaces to decode the HDTV signals provided to the drivers.

For ac displays with memory, drivers need to provide (1) *address* (or *write*) pulses, (2) *sustain* pulses, and (3) *erase* pulses. A complex refresh scan signal is not required since a continuously supplied sustain signal, coupled with the ac pixels inherent memory, trivially maintains each pixel in either an *on* or *off* state. Pixels respond to address, sustain, and erase pulses as described in the preceding section. Similar to dc pixel dynamic control discussed earlier, ac pixel addressing requires well-timed application of voltage waveforms to the rows and columns of the display matrix so that the proper voltage appears across

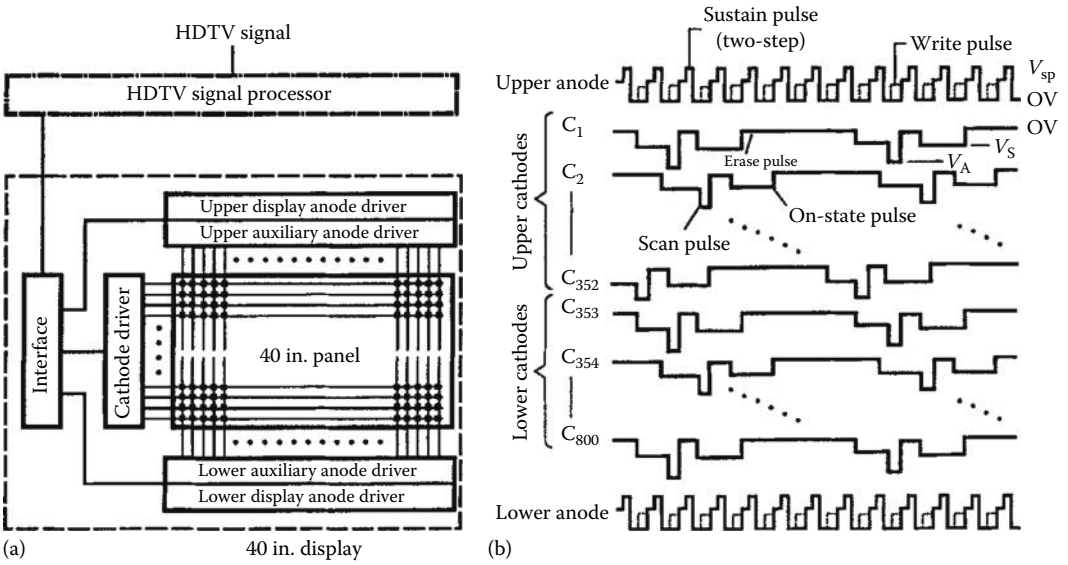


FIGURE 93.16 Pulsed memory operation of the NHK dc plasma display: (a) block diagram of the driver system and pixel array and (b) temporal sequences for pulsed memory operation. (From Yamamoto, T. et al., *SID '93 Symposium Digest*, 165–168, 1993. With permission.)

the electrodes for the pixel of interest without modifying the state of adjacent pixels. A more detailed discussion of ac pixel addressing can be found in Refs. [4,47].

Gray scale is achieved for dc or ac plasma displays either by modulation of the discharge current or by duty cycle modulation with fixed current. Modulating the applied voltage amplitude to vary the discharge current is not widely used because of practical limitations in effectively controlling the nonlinear response of the discharge current. However, duty cycle modulation is a viable technique both for pulse memory–driven dc displays and for ac memory displays. In either case, duty cycle modulation requires complex circuit design for the well-timed delivery of *on* and *off* pulses. Gray scale is achieved by varying the time a pixel is *on* compared with *off* during each refresh cycle. In 50 kHz operation, a sustain half cycle is 10 μs. VUV photon emission occurs usually in <1 μs. For color displays, the visible light emission persists much longer, with the fastest phosphors having 10% persistence times of about 5 μs. More typical phosphors have 10% persistence times in the 5–10 ms range [50]. If the image is updated every 40 ms, corresponding to a refresh rate of 25 images per second, then a 1/8 level brightness is achieved by having the pixel *on* for 5 ms and *off* for 35 ms during that refresh cycle. The time *on* is interspersed throughout the 40 ms refresh period by appropriate timing circuit design. For example, the NHK 102 cm (40 in.) display has a 2^8 or 256 levels of gray scale per color, providing a total of 16 million (256^3) color-scale levels [48].

93.5 Color Plasma Flat Panel Displays

93.5.1 Color Pixel Structures

In color plasma display panels, photoluminescent phosphors provide the primary RGB optical emissions required for full-color image display. In this case, visible emissions from the discharge itself must be suppressed in order to avoid color contamination. A common approach is to utilize xenon as the minority species constituent in the Penning gas mixture of the panel. The structure and phosphor layout of the 102 cm (40 in.) diagonal color dc plasma display available from NHK is shown in Figure 93.17, while that of the Fujitsu 53 cm (21 in.) diagonal ac color display is shown in Figure 93.6. Each uses

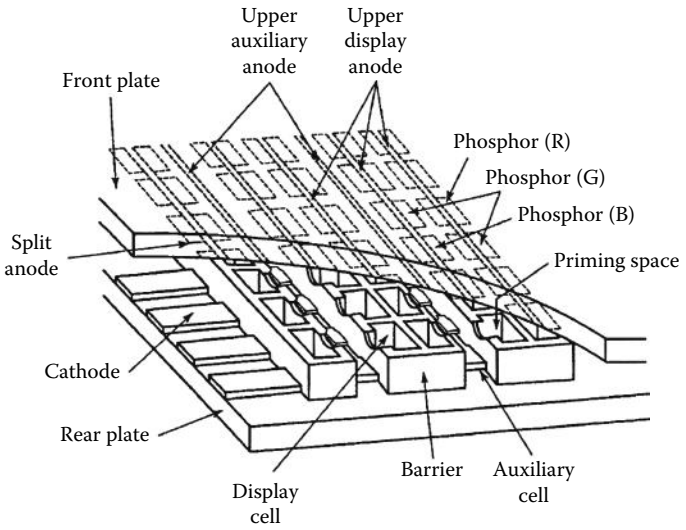


FIGURE 93.17 Structure of the 40 in. color display manufactured by NHK. (From Yamamoto, T. et al., *SID '93 Symposium Digest*, 165–168, 1993. With permission.)

screen printing and hard mask or abrasive-resistant lithographic processes for conductor wire deposition, barrier structure definition, and phosphor deposition [51]. In the NHK design, the fourth section within the honeycomb color pixel structure houses a redundant green phosphor subpixel to compensate for the lower photoluminance of green phosphors relative to that of either red or blue phosphors. In a similar honeycomb dc color pixel structure, Panasonic instead incorporates a series resistor in this fourth subpixel position [20]. Printing the series resistor for each pixel on the display glass substrate complicates panel manufacturing but simplifies design requirements for the drive electronics. In the Fujitsu, the opposed electrode ac color pixel structure shown in Figure 93.6, barrier or separation rib structures running between and parallel to each conductor wire are fabricated on the rear glass substrate. The barrier rib heights are typically 100–150 μm. The ac barrier rib structures and dc pixel honeycomb lattice structures are usually composed of the same PbO thick-film dielectric used to cover the conductor wires.

93.5.2 VUV Photon Production and Utilization for Color Plasma

93.5.2.1 Flat Panel Displays

Color plasma display gas mixtures utilizing xenon as the minority species are optimized for production of VUV emissions which are used to excite RGB photoluminescent phosphors. Both neon–xenon and helium–xenon combinations are popular. The ionization energy of xenon at 12.3 eV lies below the lowest excited atomic states of either neon or helium, as shown in Figure 93.8. Consequently, electrons accelerated by electric fields within the pixel volume preferentially impart their kinetic energy to the xenon atoms. In addition, the excited states of He or Ne produced readily transfer stored energy to the xenon atoms through ionizing Penning collision processes. Consequently, the red–orange visible emissions typical of Ne discharges are suppressed as Xe concentration is increased. Fujitsu utilizes an Ne–10% Xe working gas in its color display [18], while Photonics Imaging prefers to use He-based background gas in its panel [23] where suppression of unwanted optical emissions from the discharge can be accomplished at somewhat lower xenon concentrations.

The tendency of xenon to fill its outermost electronic shell results in the formation of the xenon dimer molecule, Xe_2^*, whose energy states are also shown in Figure 93.8 [52,53]. Radiative dissociation of the dimer produces photons with wavelengths near 173 and 150 nm. Figure 93.18 shows how

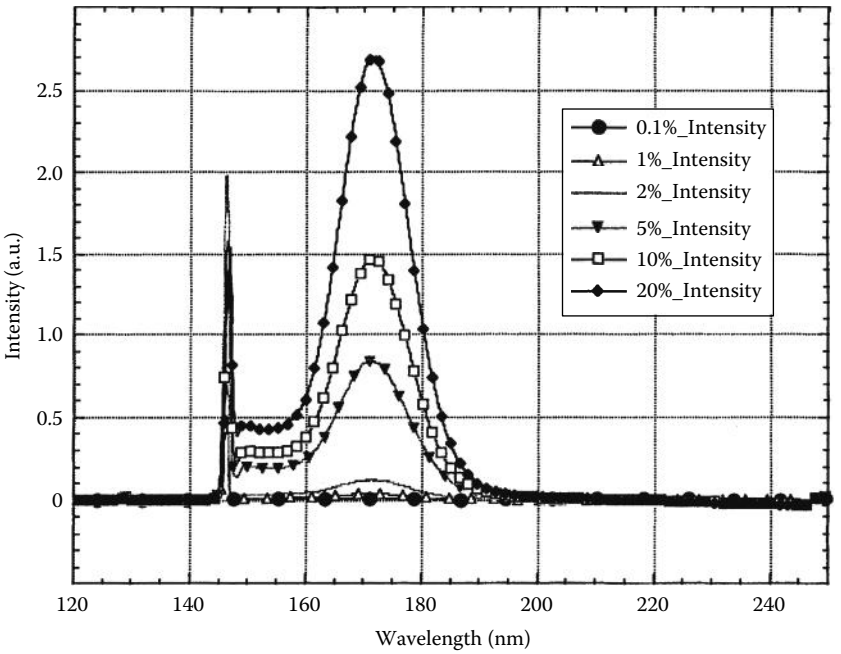


FIGURE 93.18 VUV emission spectra from opposed electrode ac plasma pixel discharges in He/Xe gas mixtures. Each spectrum was collected near the minimum sustain (or first *on*) voltage for that gas mixture, which ranged from 150 V for 0.1% Xe to 350 V for 20% Xe.

the dimer emissions dominate the VUV spectra from He–Xe gas pixel discharges as the fraction of Xe is increased. Since VUV photons are completely absorbed by glass, the spectra shown in the figure were measured by mounting opposed electrode pixels inside of a vacuum chamber filled with the gas mix of interest. The boundaries of the panel glass were not sealed which then allowed on-edge viewing of the pixel discharges with a McPherson 0.2 m monochromator operating in conjunction with a Princeton Instrument CCD optical multichannel analyzer [45]. The background gas mix was varied to obtain the various Xe concentrations in He shown while maintaining a total gas pressure of 53.3 kPa (400 torr). At low Xe concentrations, photons from the atomic Xe $1s^4$ and $1s^5$ states dominate the emission spectra producing lines at 147 nm and, with much less intensity, at 129 nm. The Tachibana laser-induced spectroscopic measurements show the spatial and temporal evolution of the $1s^4$ Xe atomic state in He/Xe plasma display discharges [54]. Both of these atomic lines experience significant resonant absorption and reemission. Thus, the measured line intensities are strong functions of photon path length traveled and of Xe partial pressure in the background gas [55]. For the emission spectra shown in Figure 93.18, the lithium fluoride (LiF) entrance window to the evacuated spectrometer chamber was positioned between 100 and 150 nm from the nearest pixel discharges, which is roughly the location of the phosphors relative to the discharge in an opposed electrode ac color display panel, for example, see Figure 93.5.

Recall that the optimal charge pair production per volt applied in Penning gas discharges occurs at minority species concentrations as low as 0.1% (see Figure 93.9). However, color plasma pixels must optimize usable photon production per watt while maintaining stringent color purity requirements. Consequently, color plasma pixels typically operate with xenon concentrations ranging between 2% and 10%. Figure 93.18 shows that increased xenon concentration results in significant dimer formation and radiative emission from dimer dissociation. Since the dimer dissociation is a three-body process involving a photon and two xenon atoms, the momentum and energy conservation equations do not demand unique solutions. Consequently, emission lines produced cover a broad spectral range spanning several

tens of nanometers. Increased dimer emission is accompanied by the suppression of xenon atomic emission as energy within the atomic manifolds continues to flow toward the lowest available atomic levels (see Figure 93.8). Note, too, that the dimer emission lines are not subject to resonant absorption. Therefore, the measured intensities shown reflect dimer emission from all pixels rows within the line of site of the spectrometer (four for the data of Figure 93.18). In contrast, due to the strong resonance absorption of the atomic lines, >90% of the measured intensity of the 147 nm line is produced in the pixel row adjacent to the spectrometer window [45,55]. Care must be taken to account for these large variations in photon mean free paths when analyzing emission data.

93.5.2.2 Phosphor Excitation and Emission for Color Plasma Flat Panels

A variety of photoluminous phosphors are commercially available. Efficiencies for conversion of VUV photons to visible emissions have a complex dependence on excitation photon wavelength as can be seen in Figure 93.19, which shows quantum conversion efficiencies relative to a sodium salicylate standard for some of the available green phosphors [50]. Conversion efficiencies for red, blue, and other green phosphors can be found in Refs. [56,57]. Table 93.2 provides the compositions of some selected commercially available phosphors and lists their relative quantum efficiencies for the principal emission lines of xenon discharges. Note that quantum efficiencies listed are relative values and that phosphors that convert 8.4 eV photons to visible photons near 2.3 eV have absolute efficiencies of only 27%. In principle, it is possible to produce two or more visible photons from a single high-energy photon, but to date no such phosphors have been developed [58]. Table 93.2 also lists the chromaticity diagram coordinates which provide a measure of the color purity of the visible RGB emission spectra produced. The chromaticity diagram can be found in many references including Ref. [59]. Another consideration in the plasma display phosphor selection is persistence. Most of the phosphors listed in Table 93.2 require

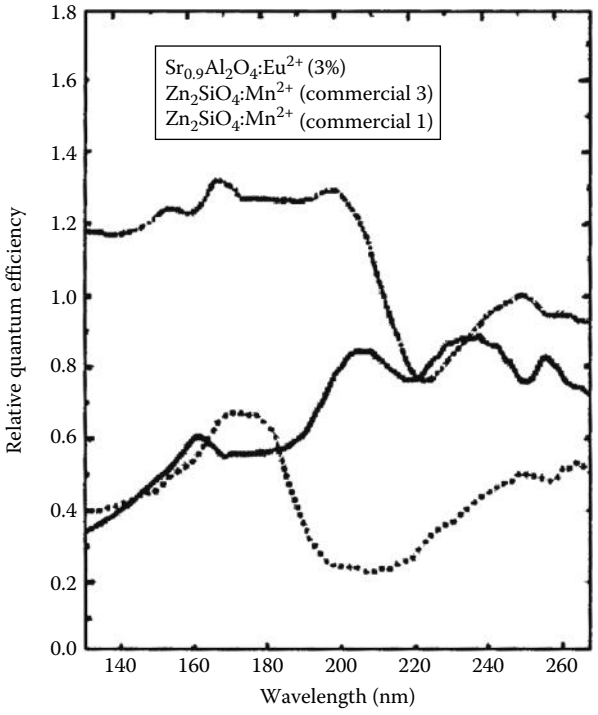


FIGURE 93.19 Relative quantum efficiencies of a Tb-activated lanthanum phosphate compared to that of yttrium and gadolinium phosphate prepared by Sarnoff Research Center. (From Yocum, N. et al., *J. SID*, 4/3, 169, 1996. With permission.)

TABLE 93.2 Relative Quantum Efficiencies (QE) and Chromaticity Coordinates for Selected Phosphors

Phosphor	Rel. QE[a] (174 nm)	Rel. QE[a] (170 nm)	Lifetime (10%) (ms)	*x*	*y*
NTSC green				0.21	0.71
$(La_{0.87}Tb_{0.13})PO_4$	1.1	1.4	13	0.34	0.57
$(La_{0.6}Ce_{0.27}Tb_{0.13})PO_4$	1.1	1.5	12	0.33	0.59
$(Y_{0.85}Ce_{0.1}Tb_{0.05})PO_4$	1.1	1.1	—	—	—
$(Y_{0.6}Ce_{0.27}Tb_{0.13})PO_4$	1.35	1.35	—	—	—
$(Gd_{0.87}Ce_{0.1}Tb_{0.03})PO_4$	1.0	1.1	10	0.34	0.58
$(Gd_{0.6}Ce_{0.27}Tb_{0.13})PO_4$	1.35	1.45	—	—	—
(Ce, Tb)$MgAl_{11}O_{19}$	0.9	1.4	—	—	—
$Sr_{0.9}Al_2O_4:Eu^{2+}$ (3%)	0.4	0.7	0.01	0.26	0.59
Zn_2SiO_4:Mn	1.2	1.3	12.5	0.21	0.72
Zn_2SiO_4:Mn	1.1	1.5	9.8	—	—
Zn_2SiO_4:Mn	0.45	0.55	5.4	—	—
Zn_2SiO_4:Mn	1.0	1.1	5	—	—
Zn_3SiO_4:Mn	1.0			0.21	0.72
$BaAl_{l2}O_{19}$:Mn	1.1			0.16	0.74
$BaMgAl_{14}O_{23}$:Mn	0.92			0.15	0.73
$SrAl_{12}O_{19}$:Mn	0.62			0.16	0.75
$ZnAl_{12}O_{19}$:Mn	0.54			0.17	0.74
$CaAl_{12}O_{19}$:Mn	0.34			0.15	0.75
YBO_3:Tb	1.1			0.33	0.61
$LuBO_8$:Tb	1.1			0.33	0.61
$GdBO_3$:Tb	0.53			0.33	0.61
$ScBO_8$:Tb	0.36			0.35	0.60
$Sr_4\ Si_8O_sCl_4$:Eu	1.3			0.14	0.33
NTSC red				0.67	0.33
Y_2O_3:Eu	0.67			0.65	0.34
Y_2SiO_s:Eu	0.62			0.66	0.34
$Y_3Al_5O_{12}$:Eu	0.47			0.63	0.37
$Zn_8(PO_4)_2$:Mn	0.34			0.67	0.33
YBO_3:Eu	1.0			0.65	0.35
(Y,Gd)BO_8:Eu	1.2			0.65	0.35
$GbBO_3$:Eu	0.94			0.64	0.36
$ScBO_3$:Eu	0.94			0.61	0.39
$LuBO_8$:Eu	0.74			0.63	0.37
NTSC blue				0.14	0.08
$CaWO_4$:Pb	0.74			0.17	0.17
Y_2SiO_5:Ce	1.1			0.16	0.09
$BaMgAl_{14}O_{23}$:Mn	1.6			0.14	0.09

Sources: Yocum, N. et al., *J. SID*, 4/3, 169, 1996; Koike, J. et al., *J. Electrochem. Soc. Solid-State Sci. Tech.*, 1008, June 1979.

[a] QEs above double rule are relative to sodium salicylate, those below relative to Zn_2SiO_4:Mn.

5–13 ms for the emission intensity to decay to 10% of maximum value. For ac pixel operation at 50 kHz, each sustain voltage half cycle lasts only 10 μs while the discharge produces VUV emissions for only a small fraction of that time. Efforts are continuing for development of phosphors with faster response times. For example, Eu^{2+} green phosphors with 10% decay times of only 5–10 μs and with good quantum efficiencies near 173 nm have been developed [50].

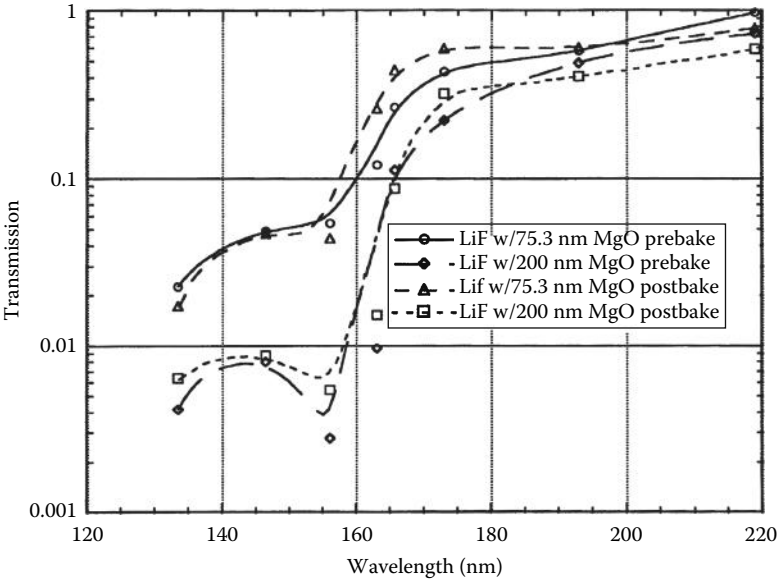


FIGURE 93.20 Photon transmission through MgO films before and after bake-out to remove water vapor.

93.5.2.3 Color Plasma Display Lifetime Considerations

Phosphors for plasma flat panel displays must be tolerant of the harsh environment produced by the pixel discharge. Photoluminous phosphor degradation mechanisms are at present not well understood. Contamination of the discharge by phosphor materials is a serious concern. Discharge modeling indicates that damage results principally from the accumulated fluence of photon and metastable bombardment, although fringe electric fields and prolonged surface charge accumulations could also result in ion bombardment [36]. Most ac plasma displays take advantage of MgO for enhancement of the discharge intensity by coating dielectric surfaces above the electrodes with an additional thin film of MgO. For ease of fabrication, the MgO is most often deposited using electron beam evaporation as one of the final manufacturing steps before glass seal and gas fill. If no mask is used, the MgO can also cover the phosphors. While providing the phosphors with a protective coating, the MgO film also attenuates the intensity of the VUV photon flux available to excite the phosphors. Figure 93.20 shows VUV photon transmission as a function of wavelength through MgO films [45,60]. The measurements show that the primary atomic Xe emission lines at 129 and 147 nm are nearly 95% attenuated by an MgO film only 75 nm thick. In contrast, the dimer emission lines centered near 173 nm are much less attenuated with transmission factors of 30%–50%, respectively, for MgO films of 200 and 75 nm. Consequently, ac plasma display designers must often achieve a balance among discharge enhancement due to MgO secondary electron emission, discharge degradation due to MgO photon absorption, and fabrication complexity associated with MgO deposition, which impacts the final cost of the display. Additionally, the designer must consider the aging or brightness degradation with time of the display, which is influenced in part by the rate of MgO sputter erosion discussed briefly later.

93.6 Inspection and Metrology

Plasma display panel manufacturing uses some process steps typically found in semiconductor device manufacturing. Although the feature sizes are much larger than those required in semiconductor manufacturing, the dimensions over which feature integrity must be assured are also much larger. Confirmation that a display manufacturing process is under control or rapid quantitative characterization of the deviation from acceptable process limits is essential to producing high-quality displays with

TABLE 93.3 Equipment for Noncontact Profiling and Chemical Characterization of Thin Films

	CyberOptics Point Range Sensor	Zygo NewView 200	Leitz FTM400
Physical principle	Laser triangulation	Scanning white light interferometry	Laser interferometry
Maximum thickness/step height	150 μm (high-resolution head)	100 μm (standard) (to 5 mm with z-drive)	70 μm
Vertical resolution	0.38 μm	100 pm	150 pm or 0.5%
Spatial resolution	>0.5 μm	0.22 μm	0.5 μm
Spot diameter/FOV	5.1 μm	140 × 110 μm	20 μm
Scan rate	10 points/s	2 or 4 μm/s	75 mm/s max.
Maximum sample size	Based on table	10 cm × 15 cm	47 cm × 37 cm
Maximum sample weight	Based on table	4.5 kg	4.5 kg (including holder)
Elemental detection	NA	NA	NA

high yield. Information on a variety of length scales is required. Blanket deposition of materials onto substrates should be uniform; large deviations in thickness of the deposited material from one point on the substrate to a widely separated point should be avoided. Similarly, roughness of the deposited layers must be kept within defined limits. Finally, chemical compositions of the deposited layers and their spatial homogeneity need to be measured and controlled.

The three commercial profile measuring devices, the CyberOptics, Zygo, and Leitz units, are based on laser or white light triangulation or interferometry (see Table 93.3) [61–63]. These machines are well suited for determining the profile of metallization after patterning. They do not appear to be useful for measuring the thickness of opaque layers and cannot, for example, measure the thickness of the blanket metallization before patterning. The CyberOptics unit is in wide use in the electronic packaging industry because of the low noise in its signal. Improvements in vertical resolution would increase its value for display manufacturing process control. The Zygo instrument has better vertical resolution, but has no real scanning capability. The Leitz unit is designed specifically for metrology of unfinished liquid crystal display components, and is presumably optimized for that application. Both the CyberOptics and Zygo units have a variety of heads with differing resolutions and fields of view. However, the vertical distance over which the interference can take place in the Zygo unit is very limited, so it may not be suitable for measuring features with large vertical dimensions.

Depending on system requirements, a useful laser metrology system can be purchased for as little as $20,000. Full-featured systems can easily cost $100,000. The bulk of the cost is, however, in the precision transport tables used for moving samples in a precisely controlled manner beneath stationary measurement heads. Since precision transport tables meeting the needs of the metrology equipment may already be part of the display production process or would require only minor upgrades to be brought in line with requirements, the cost to bring the metrology capability online may be much smaller than the figures mentioned earlier. Since laser head costs are low, it may also be desirable to use multiple fixed heads to speed up inspection and metrology.

Listed in Table 93.4 are the characteristics of two advanced measurement techniques for noncontact profiling and chemical analysis. To the authors' knowledge, systems of this type are not yet commercially available. They are being utilized, however, by researchers at Sandia National Laboratories and show promise for commercial usage. In β-backscattering measurements, the energy spectra of the backscattered β-particles provide an accurate measure of the elemental composition of a surface. This technique has been used to measure the thin-film thickness of trace-deposited metals on large-area surfaces (several square meters) in tokamaks [64]. With a depth resolution of about 100 pm, β-backscatter could serve as the physical basis for a high-sensitivity surface-profiling device. X-ray fluorescence measurements, routinely performed in air, can give information on film thickness and composition, and performance parameters for this technique are also listed in Table 93.4.

TABLE 93.4 Advanced Techniques for Noncontact Profiling/Chemical Characterization of Thin Films

Physical Principle	β-Backscatter	X-Ray Microfluorescence
Maximum thickness/step height	A few mm	50 μm
Vertical resolution	About 0.1 nm	0.1 μm
Spatial resolution	Depends on pinhole size	50 μm–2 mm
Spot diameter/FOV	Depends on pinhole size	100 μm
Scan rate	<1 min/measurement	1 spot/min
Maximum sample size	No restriction	24 cm × 21.5 cm
Maximum sample weight	No restriction	4.5 kg
Elemental detection	NA	<10–300 ppm

Vacuum must be used when studying light elements like carbon, a potential undesired product of phosphor binder burnout, because the x-rays produced are "soft." Although x-ray fluorescence equipment may have good lateral resolution, a larger spot size may prove useful when more averaged information such as film thickness is needed. Finally, eddy current measurements may provide a fast and reliable method for determining the thickness of the opaque, blanket metallization before patterning. This technique should be capable of providing 10% thickness measurement accuracy for conductor layers as thin as 1 μm.

Many of the devices for quantitative characterization of the thickness and chemistry of the deposited layers suffer from the fact that they have a rather limited field of view. Characterization techniques are needed that will allow rapid identification of visual defects in the blanket-deposited layers and in the patterned layers produced from them. Visual inspection equipment is available for defect identification in liquid crystal displays. Manufacturers include Minato Electronics, Advantest, ADS, Photon Dynamics, and Teradyne. All use CCD devices and special algorithms for the identification of line and area defects. The defects found during high-volume plasma display manufacturing may be sufficiently similar in appearance to those found in liquid crystal display manufacturing that this equipment will prove useful with appropriate adjustments. These devices, however, are expensive.

Although researchers and equipment manufacturers believe that the equipment and techniques described earlier will be suitable for online process control during plasma display manufacturing, all agree that more development is required. Real parts will need to be characterized extensively on specific commercial equipment in production environments before conclusions can be drawn about the suitability of the equipment and techniques for the intended application. This is particularly true for more difficult measurements like that of the thickness of dielectric above conductor lines. It is clear that there is no single instrument that will meet all film and feature dimensional measurement requirements and those for chemical characterization. A number of instruments will be needed to measure confidently the parameters needed for inspection, and characterize and control the manufacturing processes for plasma displays.

References

1. D.L. Bitzer and H.G. Slottow, The plasma display panel—A digitally addressable display with inherent memory, 1966 Fall Joint Computer Conference, *AFIPS Conference Proceedings*, Washington, DC, 29, 541, 1966.
2. H.G. Slottow, Plasma displays, *IEEE Trans. Electron. Devices*, 23, 760–772, 1976.
3. T.N. Criscimagna and P. Pleshko, AC plasma display. Chapter 3, in *Topics in Applied Physics*, Vol. 40, *Display Devices*, Springer-Verlag, Berlin, 91–150, 1980.
4. L.F. Weber, Plasma displays, Chapter 10, in *Flat Panel Displays and CRTs*, L.E. Tannas, Jr., Ed., Van Nostrand Reinhold, New York, 1985.

5. J. Nolan, Gas discharge display panel, *1969 International Electron Devices Meeting*, Washington, DC, 1969.
6. H.J. Hoehn and R.A. Martel, A 60 line per inch plasma display panel, *IEEE Trans. Electron Devices*, 18, 659–663, 1971.
7. O. Sahni and W.P. Jones, Spatial distribution of wall charge density in AC plasma display panels, *IEEE Trans. Electron Devices*, 25, 223–226, 1979.
8. O. Sahni and C. Lanza, Origin of the bistable voltage margin in the AC plasma display panel, *IEEE Trans. Electron. Devices*, 24, 853–859, 1977.
9. C. Lanza and O. Sahni, Numerical calculation of the characteristics of an isolated ac gas discharge display panel cell, *IBM J. Res. Dev*., 22, 641–646, 1978.
10. O. Sahni and M.O. Aboelfotoh, The pressure dependence of the bistable voltage margin of an AC plasma panel cell, *Proc. SID*, 22, 212–218, 1981.
11. O. Sahni and C. Lanza, Importance of the secondary electron emission coefficient on E/Po for Paschen breakdown curves in AC plasma panels, *J. Appl. Phys*., 47, 1337, 1976.
12. O. Sahni and C. Lanza, Influence of the secondary electron emission coefficient of argon on Paschen breakdown curves in AC plasma panels for neon + 0.1% argon mixture, *J. Appl. Phys*., 47, 5107, 1976.
13. C. Lanza, Analysis of an ac gas display panel, *IBM J. Res. Dev*., 18, 232–243, 1974.
14. O. Sahni, C. Lanza, and W.E. Howard, One-dimensional numerical simulation of ac discharges in a high-pressure mixture of Ne + 0.1% Ar confined to a narrow gap between insulated metal electrodes, *J. Appl. Phys*., 49, 2365, 1978.
15. O. Sahni and C. Lanza, Failure of Paschen's scaling law for Ne–0.1% Ar mixtures at high pressures, *J. Appl. Phys*., 52, 196, 1981.
16. H. Wakabayshi, Display market projections report from Namora Research Institute, paper presented at *Imaging 2001: The U.S. Display Consortium Business Conference*, San Jose, CA, January 28, 1997.
17. K. Werner, Plasma hits the ground running, *Inf. Display*, 12(12), 30–34, 1996.
18. Anonymous, *Indium Tin Oxide (ITO Product Data Sheet)*, Indium corporation, http://www.indium.com/inorganic-compounds/indium-compounds/indium-tin-oxide/#whitepapers, accessed on August 23, 2013.
19. H. Murakami and R. Toyonaga, A pulse discharge panel display for producing a color TV picture with high luminance and luminous efficiency, *IEEE Trans. Electron Devices*, 29, 988, 1982.
20. M. Ushirozawa, Y. Motoyama, T. Sakai, K. Wani, and K. Takahashi, Color dc-PDP with an improved resistor design in each cell, paper 33.2, *SID '94 Symposium Digest*, 719–722, 1994.
21. S. Mikoshiba, Color plasma displays: Where are we now? *Inf. Display*, 10(10), 21, 1994.
22. Y. Takano et al., A 40-in, DC-PDP with new pulse memory drive scheme, Digital Technical Papers, *SID International Symposium*, San Jose, CA, Vol. 25, pp. 731–734, 1994.
23. P.S. Friedman, Are plasma display panels a low-cost technology? *Inf. Display*, 11(10), 22–25, 1995.
24. L.F. Weber, History of the plasma display panel, *IEEE Transactions on Plasma Science*, Vol. 34, No. 2, pp. 268–278, 2006.
25. A. von Engle, *Ionized Gases*, AIP Press, New York, 1994.
26. S. Brown, *Basic Data of Plasma Physics—The Fundamental Data on Electrical Discharges in Gas*, American Institute of Physics Press, New York, 1993.
27. F.F. Chen, *Introduction to Plasma Physics*, Plenum Press, New York, 1977.
28. D. Miller, J. Ogle, R. Cola, B. Caras, and T. Maloney, An improved performance self-scan I panel design, *Proc. SID*, 22, 159–163, 1981.
29. M.H. Miller and R.A. Roig, Transition probabilities of XeI and XeII, *Phys. Rev. A*, 8, 480, July 1973.
30. G.C. Rodrigues, P. Indelicato, J.P. Santos, P. Patte, and F. Parente, Systematic calculation of atomic energies of ground state configurations, *Atomic Data and Nuclear Data Tables*, Vol. 86, No. 2, Academic Press, New York, pp. 117–233, 2004.
31. R.S. Van Dyck, C.E. Johnson, and H.A. Shugart, Lifetime lower limits for the $^3P_0$ and $^3P_2$ metastable states of neon, argon and krypton, *Phys. Rev. A*, 5, 991–993, 1972.

32. A.A. Kruithof and F.M. Penning, The Townsend ionization coefficient and some elementary processes in neon with small admixtures of argon, *Physica*, 4(6), 450, 1937.
33. F.L. Jones and W.R. Galloway, The sparking potential of mercury vapor, *Proc. Phys. Soc. Lond.*, 50, 207–212, 1938.
34. *CRC Handbook of Chemistry and Physics*, 94th ed., W.M. Haynes, Eds., CRC Press, Boca Raton, FL, 2013.
35. R. Veerasingam, R.B. Campbell, and R.T. McGrath, One-dimensional single and multipulse simulations of the ON/OFF voltages and bistable margin for He, Xe, and He/Xe filled plasma display pixels, *IEEE Trans. Plasma Sci.*, 24(6), 1399–1410, 1996.
36. R. Veerasingam, R.B. Campbell, and R.T. McGrath, Two-dimensional simulations of plasma flow and charge spreading in ac plasma displays, *IEEE Trans. Plasma Sci.*, 24, 1411–1421, 1996.
37. M. Scott, Los Alamos National Laboratory, Private communication, 1997.
38. H. Uchike, K. Miura, N. Nakayama, T. Shinoda, and Y. Fukushima, Secondary electron emission characteristics of dielectric materials in AC-operated plasma display panels, *IEEE Trans. Electron. Devices*, 23, 1211–1217, 1976.
39. N.J. Chou and O. Sahni, Comments on "Secondary Emission Characteristics of Dielectric Materials in AC Operated Plasma Display Panels," *IEEE Trans. Electron. Devices*, 25, 60–62, 1978.
40. T. Shinoda, H. Uchike, and S. Andoh, Low-voltage operated AC plasma-display panels, *IEEE Trans. Electron. Devices*, 26, 1163–1167, 1979.
41. M. Aboelfotoh and J.A. Lorenzen, Influence of secondary-electron emission from MgO surfaces on voltage-breakdown curves in Penning mixtures for insulated-electrode discharges, *J. Appl. Phys.*, 48, 4754–4759, 1977.
42. T. Nenadovic, B. Perraillon, Z. Bogdanov, Z. Djordjevic, and M. Milic, Sputtering and surface topography of oxides, *Nucl. Instrum. Methods Phys. Res.*, B48, 538–543, 1990.
43. H. Schone, D. Walsh, R.T. McGrath, and J.H. Burkhart, Microbeam Rutherford backscattering measurements of flat panel display thin film erosion, *Nuclear Instrum. Methods Phys. Rev. B*, 130(1–4), 543–550, 1998.
44. P. Hines, Spectroscopic and electrical measurements on opposed electrode ac plasma display pixels, Honors engineering senior thesis, The Pennsylvania State University, University Park, Pennsylvania, 1997.
45. R.T. McGrath, R. Veerasingam, J.A. Hunter, P.D. Rockett, and R.B. Campbell, Measurements and simulations of VUV emissions from plasma flat panel display pixel microdischarges, *IEEE Trans. on Plasma Sci.*, 26(5), 1998.
46. S. Matsumoto, *Electronic Displays Devices*, John Wiley & Sons, New York, 1990.
47. G.E. Holz, The primed gas discharge cell—A cost and capability improvement for gas discharge matrix displays, *Proc. SID*, 13(2), 1972.
48. T. Yamamoto, T. Kuriyama, M. Seki, T. Katoh, H. Murakami, K. Shimada, and H. Ishiga, A 40 in. diagonal HDTV plasma display, *SID '93 Symposium Digest*, 165–168, 1993.
49. J. Deschamps and H. Doyeus, Plasma displays, *Phys. World*, June 1977.
50. N. Yocum, R.S. Meltzer, K.W. Jang, and M. Grimm, New green phosphors for plasma displays, *J. SID*, 4/3, 169–172, 1996.
51. H. Fujii, H. Tanabe, H. Ishiga, M. Harayama, and M. Oka, A sandblasting process for fabrication of color PDP phosphor screens, paper 37.5, 728–731, *SID '92 Symposium Digest*, 1992.
52. M.C. Castex, Experimental determination of the lowest excited Xe_2 molecular states from VUV absorption spectra, *J. Chem. Phys.*, 74, 759–771, 1981.
53. M.R. Flannery, K.J. McCann, and N.W. Winter, Cross sections for electron impact ionization of metastable rare gas excimers (He_2^*, Kr_2^*, Xe_2^*), *J. Phys. B Mol. Phys.*, 14, 3789–3796, 1981.
54. K. Tachibana, Spatio-temperal measurement of excited Xe($1s^4$) atoms in a discharge cell of a plasma display panel by laser induced spectroscopic microscopy, *Appl. Phys. Lett.*, 65, 935–937, 1994.
55. T. Holstein, Imprisonment of resonance radiation in gases, *Phys. Rev.*, 72, 1212, 1947.

56. D.E. Husk and S.E. Schnatterly, Quantum efficiency and linearity of 16 phosphors in the soft x-ray regime, *J. Opt. Soc. Am. B*, 9, 660–663, 1992.
57. J. Koike, T. Kojima, and R. Toyonaga, New tricolor phosphors for gas discharge display, *J. Electrochem. Soc. Solid-State Sci. Tech*., 1008, June 1979.
58. P.N. Yocum, Future requirements of display phosphors from an historical perspective, *J. SID*, 4/3, 149, 1996.
59. L.E. Tannas, Jr., Ed., *Flat Panel Displays and CRTs*, Van Hostrand Reinhold, New York, 1985.
60. S.G. McLean and W.W. Duley, VUV absorption in thin MgO films, *J. Phys. Chem. Solids*, 45, 223, 1984.
61. R. Kuntz, CyberOptics, Private communication, 1-800-746-6315.
62. Zygo, Private communication, 1-860-347-8506, www.zygo.com.
63. A. Machura, Leica, Private communication, 49 (0) 6441 29-2316, Andreas.Machura@lmw.leica.com.
64. B.E. Mills, D.A. Buchenauer, A.E. Pontau, and M. Ulrickson, Characterization of deposition and erosion on the TFTR bumper limiter and wall, *J. Nucl. Mater*., 162–164, 343–349, 1989.

94

Electroluminescent Displays

William A. Barrow
Planar Systems, Inc.

94.1 Introduction

Electroluminescence (EL) is the nonthermal generation of light resulting from the application of an electric field to a substance. The light-emitting substance is generally a luminescent crystal. Most commercially available monochrome EL displays utilize ZnS:Mn as the luminescent material. EL displays have become very important in certain display markets. These include medical instrumentation, industrial control equipment, portable instrumentation, and military vehicles. The attributes of EL displays that make them attractive in these types of applications are mechanical ruggedness, relative insensitivity to ambient temperature, rapid display response, essentially unlimited viewing angle, compactness, and lightweight.

There are four types of EL devices: ac thin film, ac powder, dc thin film, and dc powder. The ac thin-film display is by far the dominant device type. Some liquid crystal displays use ac powder EL for backlights. There is currently little or no commercial application of dc EL devices, either thin

film or powder [1]. The focus here is on ac thin-film EL (ACTFEL) devices. While there are no widely established standard measurement techniques for the other EL device types, measurements similar to those described herein for ACTFEL devices could be applied with appropriate modifications to the other methods of device excitation.

94.2 Device Structure and Operation

A schematic, cross-sectional representation of the basic ACTFEL device structure [2] is shown in Figure 94.1. The supporting substrate is usually made of very low sodium glass. If a suitable ion barrier layer is deposited between the substrate and lower electrode of the EL device, soda-lime glass can be used. The lower electrode, usually a transparent conductor such as indium tin oxide (ITO), is deposited next. The ITO is usually between 350 and 120 nm in thickness, providing sheet resistance in the range of 5–15 Ω/square. On top of the ITO electrode a lower insulator is deposited. This layer is typically SiON or aluminum titanium oxide (ATO). The thickness is normally around 200 nm. The phosphor layer, typically ZnS:Mn, is deposited between the lower insulator and a similar upper insulator. The phosphor thickness is typically in the range 200–1000 nm thickness, depending on the application. The upper electrode, typically aluminum, is deposited on top of the upper insulator. The aluminum is generally 100–200 nm thick.

A matrix-addressed monochrome display is created by dividing the upper and lower electrodes into orthogonal arrays of electrode stripes. The EL device is then excited locally by applying a voltage between a pair of crossing electrodes, causing an electric field to exist between them, which excite the phosphor. A color display can be created by dividing the phosphor into stripes of red-, green- and blue-emitting phosphors which are aligned with one of the sets of electrode stripes [3,4]. This is shown schematically in Figure 94.2. Color displays can also be created by using an unpatterned phosphor which emits a broad spectrum including red, green, and blue and filtering the emission using either a patterned color filter which is aligned with the electrode stripes or a frame sequential color filter.

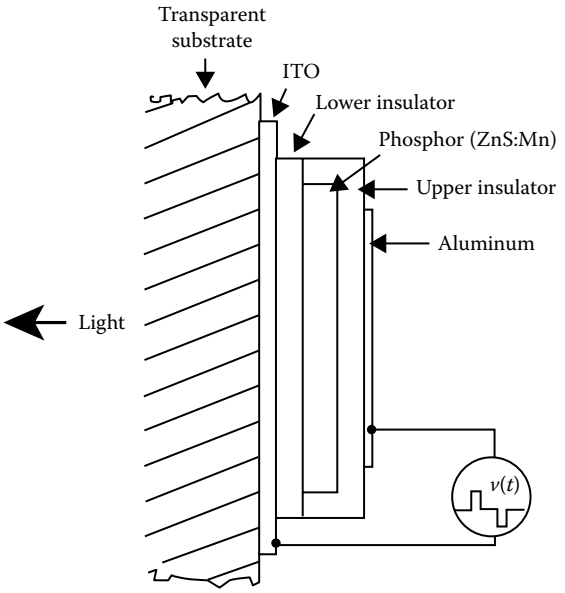


FIGURE 94.1 Schematic, cross-sectional diagram of the basic ACTFEL device structure.

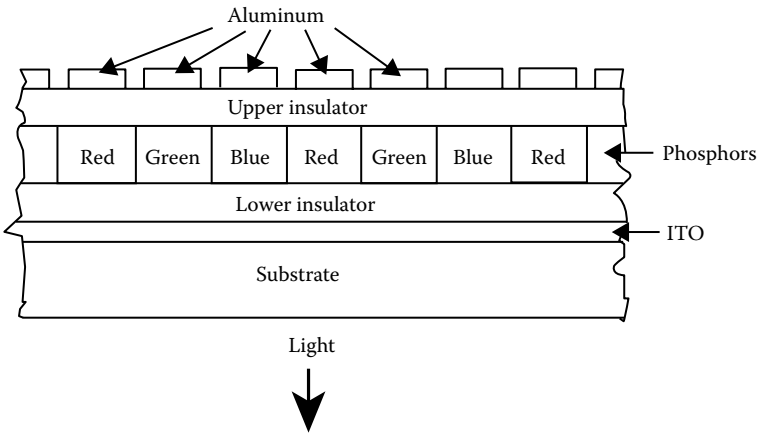


FIGURE 94.2 Schematic, cross-sectional diagram of a color ACTFEL device with the red, green, and blue primary colors produced by patterned color phosphor stripes. The color phosphor stripes are end on in this view.

94.3 Device Fabrication

94.3.1 Thin-Film Deposition Methods

A wide variety of deposition techniques are used by various manufacturers and researchers in the fabrication of ACTFEL devices. The electrode materials ITO and aluminum are usually deposited by physical vapor deposition (PVD) techniques. ITO is almost universally deposited by dc magnetron sputtering from either a conductive ITO target or a metal alloy target [5]. Optimum ITO conductivity is obtained by postdeposition annealing in a very low oxygen environment. Al is deposited either by dc magnetron sputtering from an Al metal target or by electron beam evaporation of Al metal. The insulator and phosphor layers are deposited by chemical vapor deposition (CVD) as well as PVD. Phosphor layers have been deposited by thermal as well as e-beam evaporation, sputtering, metal–organic CVD (MOCVD), and atomic layer epitaxy (ALE). Insulator layers have been deposited by e-beam evaporation, radio frequency (RF) sputtering, plasma-enhanced CVD (PECVD), and ALE.

94.3.2 Thin-Film Patterning Methods

Patterning of the electrodes is generally accomplished either through etching or liftoff, although some early workers patterned the upper aluminum electrodes by evaporating through a shadow mask. Aluminum is easily etched wet or dry. There are commercially available wet etches for Al. Dry etching of Al can be carried out using standard chlorine chemistries, e.g., Cl_2/BCl_3. Al can also be patterned by evaporating onto a reversed photoresist pattern and lifting off excess metal. ITO can be etched wet or dry. ITO wet etches generally consist of mixtures of HCl and HNO_3. ITO can be dry-etched in HI or HBr. Patterning of the phosphor layers is more problematic. Phosphor etches exist for most ACTFEL phosphor materials, but they are generally proprietary. Thin-film phosphors are difficult to etch because they are usually water-sensitive, and some color phosphors contain heavy metals which are difficult to volatilize in a dry-etch process.

94.4 Device Operation

94.4.1 Luminescence Mechanisms

As the device structure shown in Figure 94.1 indicates, ACTFEL devices are capacitively coupled. Since only displacement current can flow across the insulator layers, the drive signal must be an alternating polarity waveform. A typical alternating polarity, trapezoidal waveform is shown in Figure 94.3.

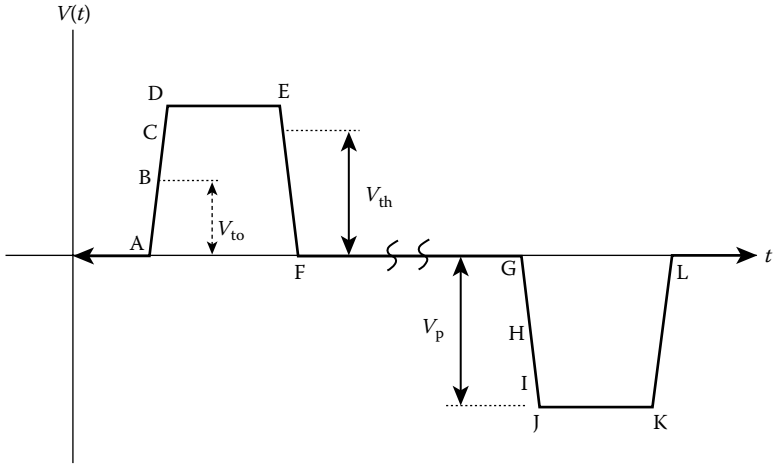


FIGURE 94.3 Typical alternating polarity, pulse drive waveform. Letters A–L mark points on the drive waveform which are referenced later in Figures 94.5 and 94.6. The pulse width would generally be about 30 μs and the frequency would be between 60 and 500 Hz for a passive matrix-addressed display.

If the peak voltage, V_P, in Figure 94.3 is larger than the *threshold voltage* of the device, V_{th}, then, when the positive pulse of the waveform is applied between the Al and ITO electrodes of the device structure shown in Figure 94.1, the energy band diagram of the ACTFEL device will be as shown in Figure 94.4. Electrons, which are the majority carriers in ACTFEL devices, tunnel out of the interface states on the left and into the conduction band. Once in the conduction band the electrons are accelerated by the large electric field, which is ~1 MV/cm = 100 kV/mm. The conduction electrons drift across the ZnS:Mn layer until they impact excite an Mn^{2+} center, transferring some energy to one of its electrons and causing it to undergo a transition to an excited state. The conduction electron may undergo additional collisions,

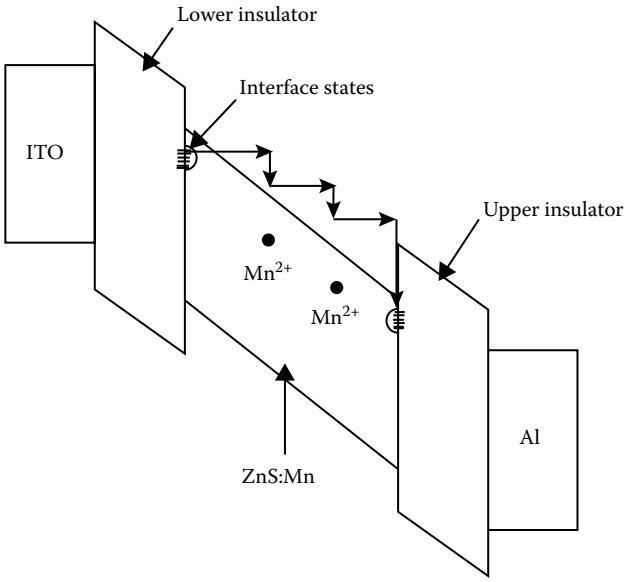


FIGURE 94.4 Energy band diagram of an ACTFEL device during the peak of the applied voltage pulse. Electrons tunnel out of insulator/phosphor interface states into the conduction band, are swept across the phosphor layer, and impact exciting emission centers as they go until they are finally thermalized and trapped at the opposite interface.

eventually reaching the right interface and getting trapped in interface states there until the next voltage pulse, which is of the opposite polarity. This pulse causes the electrons to tunnel out and drift back across the ZnS:Mn layer, transferring energy to Mn^{2+} centers along the way, until eventually they are trapped again at the left interface. This transfer of charge back and forth between the interface states continues as long as the alternating polarity drive signal with peak amplitude above the threshold voltage of the device continues to be applied. Light emission occurs when the Mn^{2+} centers, which have been impact excited by the hot electrons, relax [6–8]. The light emission thus results from transitions of the electrons within the Mn^{2+} centers, rather than electron–hole pair recombination near a *p-n* junction as occurs in a light-emitting diode (LED).

94.4.2 Description of Charge Flow

If the external charge, *Q*, flowing into the ACTFEL device is plotted vs. the externally applied voltage, *V*, the resulting curve is called a *QV* loop [9,10]. If the amplitude of the applied voltage pulses is less than the threshold voltage of the device, the *QV* loop is simply a straight line with slope equal to the total capacitance of the insulator/phosphor/insulator stack. If the amplitude of the applied voltage pulses is greater than the threshold voltage of the device, the *QV* loop opens up. *QV* loops below and above threshold are shown in Figure 94.5. Above threshold, power is dissipated in the ACTFEL device. The area encompassed by the *QV* loop is equal to the energy delivered to the device per period of the drive waveform.

The *QV* loop is measured directly. A theoretical extension of the *QV* loop that is sometimes used by researchers studying the physics of ACTFEL devices is a plot of the actual charge flow across the phosphor layer, Q_p, vs. the electric field across the phosphor layer, F_p. The quantities required for a plot of Q_pF_p can be calculated from the *QV* data if the thicknesses and dielectric constants of the insulator layers and phosphor layer are known. The actual charge flow across the phosphor is calculated by subtracting the reactive charge from the total charge. The field in the phosphor layer is calculated by adding the externally applied field to the internal *polarization field* due to the actual flow of charges across the phosphor layer. A Q_pF_p loop

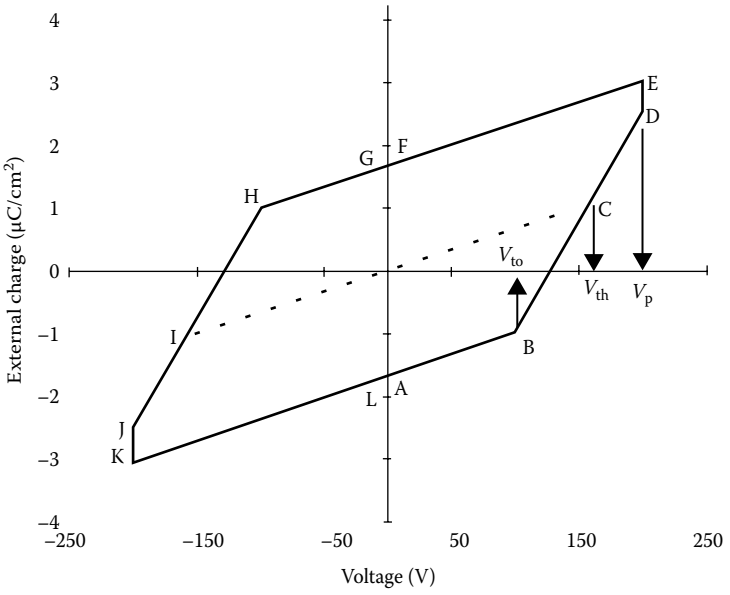


FIGURE 94.5 Idealized *QV* loop with no charge leakage from interface states between drive pulses. Letters A–L mark points on the *QV* loop which are coincident in time with the points labeled A–L on the drive waveform in Figure 94.3. The dashed line is the *QV* loop for the case just below threshold. The solid line is the open loop above threshold. The area of the *QV* loop represents the energy dissipated by the device per cycle of the drive waveform.

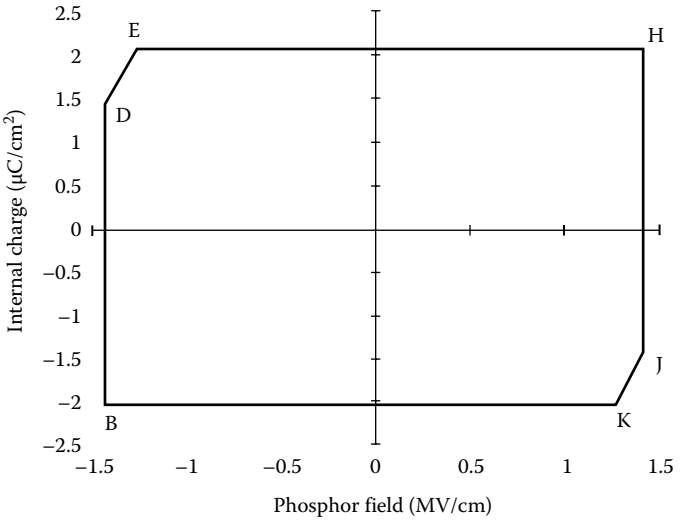


FIGURE 94.6 Q_pF_p loop corresponding to the above threshold *QV* loop in Figure 94.5. This loop represents the charge flow across the phosphor layer as a function of the electric field across the phosphor layer.

corresponding to the above threshold *QV* loop of Figure 94.5 is shown in Figure 94.6. The Q_pF_p loop is useful because it expresses the internal electrical characteristics of the phosphor layer during device operation.

94.4.3 Device Excitation

Whether the device under test is a test dot or a matrix display, the drive waveform must be ac-coupled. Passively addressed matrix displays are scanned one row at a time. Data are loaded into the column drivers for a single row of pixels and the selected pixels in the row are all turned on simultaneously. The row pulse brings the voltage across each pixel in the row to a level just below threshold. The columns of selected pixels are then pulsed to bring the voltage across selected pixels to a level above threshold and the pixels are turned on. During the following frame, the voltage polarities across the pixels are reversed. Each individual pixel is subjected to a drive signal similar to that shown in Figure 94.3. To activate a pixel fully requires the voltage to be held above V_{th} for 10–20 μs. Since each row must be scanned in sequence, a typical display with ~500 rows requires at least 5 ms to scan one frame. The maximum frame rate is thus ~200 Hz. Displays that are addressed by an active matrix of transistors are not scanned a line at a time as passively addressed displays are, but instead are bulk driven. They have one unpatterned common electrode, usually an upper layer of ITO, to which an ac drive signal is applied. Each pixel is connected to ground through a transistor which drops a portion of the applied drive voltage when the pixel is not selected. In this type of display, the drive waveform can be any ac signal of the appropriate voltage. Sine, trapezoid, and triangle waveforms have been used. An active matrix display can be driven with a frame rate higher than a passive matrix-addressed display by a factor equal to the number of rows in the display, since row-at-a-time scanning is not required. Since a light pulse is emitted for each voltage pulse, the average luminance is proportional to the frame rate, resulting in much higher luminance capability for active matrix-addressed displays.

94.5 Standard Measurements

Measurable quantities of interest include luminance, luminous efficiency, emission spectrum, latent image, and defects. The luminance of an ACTFEL display is a function of the peak voltage and frequency of the drive waveform, the intrinsic efficiency of the insulator/phosphor/insulator stack, and

the operating history (aging). The efficiency is a function of the drive waveform shape and frequency as well as various device parameters. The emission spectrum is primarily determined by the phosphor host material and activators, although it is also affected by deposition and anneal conditions and in some cases by the drive voltage and frequency. *Latent image* is the burning in of a permanent image of a fixed pattern which is displayed for long periods. It can appear as a faint dark image superimposed on a bright background or as a faint bright image superimposed on a dark background. Formation of latent image is affected primarily by drive waveform symmetry and device processing parameters. Display defects include pixel, line, and mura defects.

94.6 Time-Resolved Measurements

Measurements of luminance, efficiency, and emission spectra as introduced so far involve time-averaged light emission. There are two fairly common time-resolved measurements of ACTFEL light emission: light emission decay time and time-resolved light emission spectroscopy. The *light emission decay time*, τ, is the time it takes for the light emission from one excitation pulse to fall to $1/e$ times its initial value. The measurement of τ is started just after the trailing edge of the excitation pulse. This is necessary in order that the measured value of τ not be affected by the continuing excitation of additional emission centers, so that it represents the intrinsic relaxation time of the emission center. In devices with evaporated ZnS:Mn phosphor, τ is a strong function of the Mn concentration. It can thus be used as an analytical technique to determine the Mn concentration. Time-resolved spectroscopy is the measurement of the emission spectrum occurring during specific portions of the excitation and emission process. An example of the application of this technique is the study of the separate light pulses emitted during the leading and trailing edges of the excitation pulse with the blue-emitting phosphor SrS:Ce. Both of these techniques are frequently used to help elucidate the excitation and emission mechanisms in ACTFEL phosphors.

94.7 Test Dot Characterization

94.7.1 Luminance and Efficiency

A schematic representation of a measurement system for collecting luminance and efficiency vs. voltage data on a test dot is shown in Figure 94.7. An arbitrary waveform generator provides a drive signal with the waveshape, frequency, and peak voltage determined by a control computer. The waveform generator output signal is amplified from ±5 to ±300 V range. A photometer, also under computer control, measures the luminance, L, of the test dot. An oscilloscope measures the voltage across the test dot and the voltage across the sense resistor. The current is calculated by dividing the voltage across the sense resistor by its resistance. The control computer can thus adjust the peak voltage up and down and collect the luminance data, as well as the waveforms representing the voltage across the test dot and the current through it. The energy dissipated, E_p, during each drive pulse is calculated as

$$E_p = \int_D I(t) \times V(t)\, dt \tag{94.1}$$

where

$I(t)$ and $V(t)$ are the current and voltage waveforms, respectively
D is the duration of either the positive or negative drive pulse

This integration can be carried out by most oscilloscopes or the $I(t)$ and $V(t)$ data can be transferred to the control computer for the calculation, although this approach is generally slower. The average energy dissipated per period of the drive waveform, E, is the average of E_p for a positive pulse and

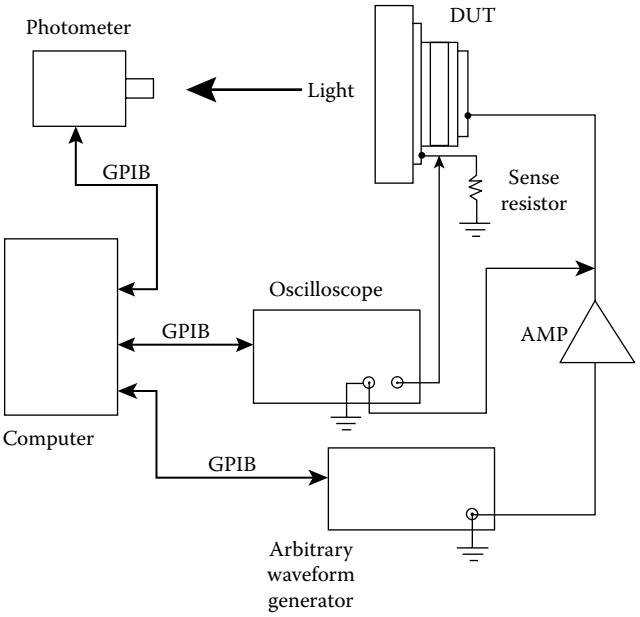


FIGURE 94.7 Schematic diagram of a system for measuring luminance and efficiency as functions of the peak drive voltage. The voltage waveform, $V(t)$, is measured by the oscilloscope at the Al electrode of the device under test. The current waveform, $I(t)$, is measured by the oscilloscope as the voltage across the sense resistor divided by its resistance. The luminance is measured by the photometer.

a negative pulse. The average power dissipated, P, can be calculated by multiplying E by the frequency of the drive waveform. The efficiency, η, is calculated as follows:

$$\eta = \frac{\pi LA}{P} \tag{94.2}$$

where

L is the luminance in cd/m^2
A is the area of the test dot in m^2
P is the average power in watts

Values of L and η are collected for peak voltages ranging from 10 V below threshold to 40 or 50 V above threshold. Plots of L and η vs. peak voltage for a typical device are shown in Figure 94.8.

94.7.2 Charge Flow and Electric Field

The QV loop is measured using a circuit identical to that in Figure 94.7, except that the sense resistor is replaced by a sense capacitor and the photometer is not used. The sense capacitor value is chosen to be much larger than the capacitance of the ACTFEL device so that the voltage dropped by the sense capacitor is small. Since the sense capacitor and the ACTFEL device are in series, the charges stored on them are equal. The charge, Q, on the ACTFEL device is thus

$$Q = C_S V_S \tag{94.3}$$

where

C_S is the capacitance of the sense capacitor
V_S is the voltage appearing across it

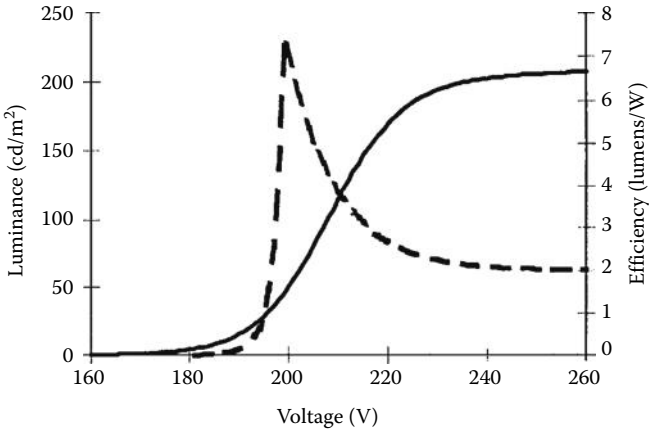


FIGURE 94.8 Plots of luminance and efficiency as functions of peak drive voltage. The solid line is the luminance and the dashed line is the efficiency.

When $Q(t)$ is plotted vs. $v(t)$, the QV loop results. In the idealized QV loop shown in Figure 94.5, V_{th} (C) is the threshold voltage, V_{to} (B) is the turn-on voltage, and V_p (D, E) is the peak voltage of the drive waveform. Threshold voltage is the voltage at which the first knee in the LV curve occurs. If C_T is the total capacitance of the device below threshold, this is also the voltage at which the line $Q = C_T V$ intersects the open, above threshold QV loop (C). In practice, it is sometimes defined as the voltage at which a certain luminance value occurs at a given frequency, e.g., the voltage at which the luminance is 1 cd/m^2 at 60 Hz. V_{to} is the voltage at which the slope of the QV loop changes from C_T to C_I, where C_I is the capacitance of the insulator layers.

Since a differential element of energy delivered to the ACTFEL device is $dE = V(t)dQ$, then

$$E = \int V(t)\,dQ \tag{94.4}$$

The energy delivered per period of the drive waveform is thus equal to the area encompassed by the QV loop. The power dissipated is just the energy per period multiplied by the frequency. In practice, the area of the QV loop is measured by numerical integration. The calculation can be carried out on the oscilloscope if it has analysis capabilities or the data can be transferred to the control computer for integration.

Generation of the Q_pF_p loop does not require any electrical measurements other than those required for the QV loop. Q_p is the charge separation across the phosphor layer and F_p is the electric field across the phosphor layer. If the thicknesses and dielectric constants of the insulator and phosphor layers are known, Q_p and F_p can be calculated from the values of $Q(t)$ and $V(t)$ on the QV loop. This is accomplished by applying the following equations [11,12]:

$$Q_p(t) = \frac{C_i + C_p}{C_i} Q(t) - C_p V(t) \tag{94.5}$$

and

$$F_p(t) = \frac{1}{d_p}\left(\frac{Q(t)}{C_i} - V(t)\right) \tag{94.6}$$

where

C_i is the capacitance of the insulators
C_p is the capacitance of the phosphor layer
d_p is the thickness of the phosphor layer

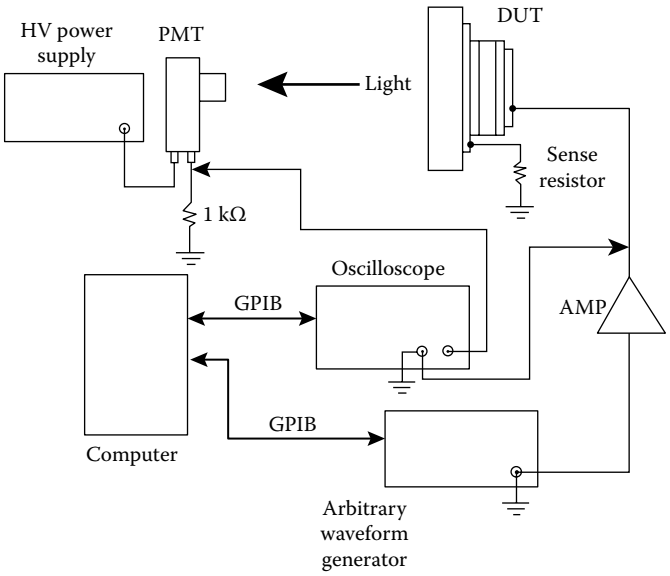


FIGURE 94.9 Apparatus for measuring the luminescent decay time of the ACTFEL phosphor. The light signal is measured by the oscilloscope across the sense resistor on the output of the PMT. The oscilloscope is triggered on the drive pulse.

94.7.3 Time-Resolved Measurements

The apparatus for measuring τ is shown in Figure 94.9. A photomultiplier tube (PMT) is used to detect the light emission as a function of time. The drive system is set to provide relatively narrow drive pulses, typically 10 μs pulse width, at relatively low frequency, typically 60 Hz. This approach works well for phosphors with relatively long decay times, on the order of 100 μs to a few milliseconds. This is the case for many common ACTFEL phosphors such as ZnS:Mn and ZnS:Tb. Phosphors such as SrS:Ce, however, have very fast decay times and cannot be measured in this manner. In such cases, the photoluminescent decay time must be measured using a pulsed laser to excite the phosphor and appropriately low RC response time of the light detection system.

The general setup for measuring time-resolved emission spectra from ACTFEL devices is shown schematically in Figure 94.10. The oscilloscope is triggered on the drive waveform and the signal integration period is set to the region of interest. A boxcar integrator can also be used to integrate the light signal during the desired time window. The monochrometer wavelength is scanned and the emission spectrum is collected for the selected portion of the emission process.

94.7.4 Aging

ACTFEL devices tend to stabilize after a few tens of hours of burn-in, but can exhibit complex aging behavior during the burn-in process. The luminance vs. voltage curves for ZnS:Mn devices in which the phosphor layer is deposited by evaporation, for example, tend to shift to slightly higher voltage during burn-in [2]. Luminance vs. voltage curves for devices in which the ZnS:Mn is deposited by ALE tend to shift to slightly lower voltage [13]. Aging data are collected by measuring luminance vs. voltage at selected time intervals during aging. The measurement is carried out as described earlier for luminance vs. voltage measurements. The aging is done by continuously operating the device at a fixed voltage or at a fixed voltage above threshold. The aging process can be accelerated by operating the device at higher frequency.

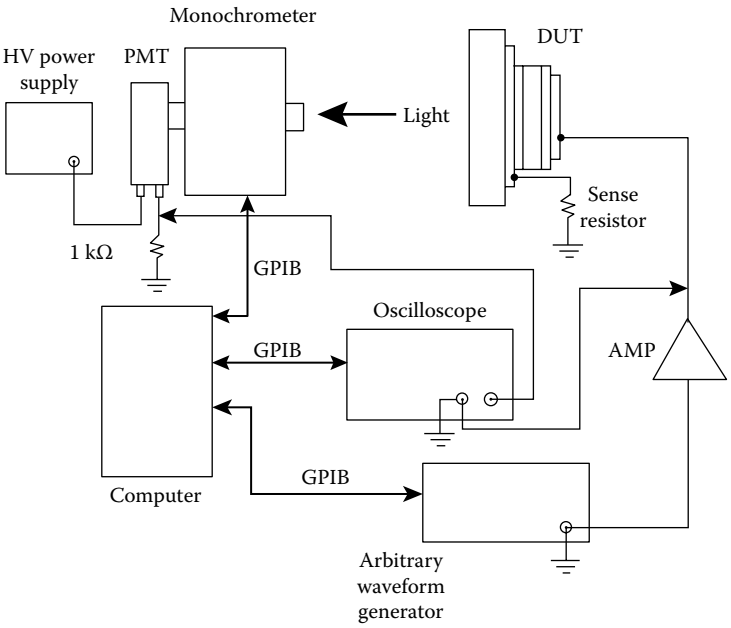


FIGURE 94.10 System for measuring the time-resolved emission spectrum of an ACTFEL device. The oscilloscope is set to integrate the light signal from the PMT during a selected time window. The monochrometer wavelength is scanned over the entire spectral range and the emission intensity data are collected from the oscilloscope.

94.8 Characterization of Matrix-Addressed Displays

The characterization of matrix-addressed displays differs from characterization of test dots because less control of the drive waveform is readily available. The row and column drivers do not provide great flexibility, although some control can be exercised by varying the composite drive waveform and the control signals to the driver chips. These types of modifications, however, require detailed knowledge of the addressing and control electronics involved and are best left to the original display manufacturer. Measurements, which are more accessible and of more general interest, involve characterization of the luminance, chromaticity, uniformity, display life, latent image, and defects. For these measurements, the display under test is controlled by a computer through a standard video output or a custom video display interface provided with the display. The display operating voltage and frequency are fixed. The display is placed in a dark room or enclosure. Luminance is measured with a photometer. Chromaticity is measured with a spectrophotometer or a photometer equipped with tristimulus filters. Uniformity can be measured by mounting either the display or the photometer/spectrophotometer on a translation system and collecting data at points in a regular array of locations throughout the display surface. Display life is characterized by operating the display in a full-on pattern, cycling through checkerboard patterns, etc., and making luminance measurements at exponentially increasing time intervals. Latent image is formed by displaying small blocks in various fixed locations and various fixed gray levels (if available) continuously for long periods. It is characterized by setting the entire screen to each gray level available and measuring the luminance in the locations which had blocks displayed during aging as well as unaged areas nearby. Latent image is the percent luminance difference between aged blocks and nearby unaged areas. Latent image is typically measured after aging the block pattern for 1000 h. Defects are classified into pixel, line, and mura defects. They are characterized by visual inspection with the aid of an eye loupe or by an automated flat panel inspection system.

94.9 Excitation and Measurement Equipment

94.9.1 Excitation of Test Dots

Test dots can be excited by any signal source that provides bipolar pulses up to a peak voltage of 300 V and sufficient current sourcing and sinking capability to charge and discharge the device capacitance. Bipolar pulse drivers with sufficient voltage and current output can be built with commercially available components. A hybrid circuit op amp, produced by Apex Microtechnology, provides sufficient voltage output and frequency response. A current boosting stage can be added to the output if the DUT capacitance is too large to drive directly. This amplifier approach is very flexible since any waveform that can be generated by the arbitrary waveform generator can be used. There are also some commercial amplifiers available that are effective for driving test dots for some measurements. They tend to have limited bandwidth and must be used with caution in situations in which the measurement is sensitive to the exact shape of the drive pulse. This would be the case, for example, in *QV* measurements and often in efficiency measurements.

94.9.2 Excitation of Matrix-Addressed Displays

Excitation of matrix-addressed displays is straightforward since the drive electronics are integrated with the display. Generally, only a standard ac power outlet and a computer with an appropriate video card are required. Test patterns for measurement can be created using simple computer programs.

94.10 Measurement Instruments

94.10.1 Measurement of Drive Voltage and Current

Digital oscilloscopes are used for all of the electrical measurements described in this chapter. Suitable instruments are available from several manufacturers, including Tektronix and Hewlett-Packard. A bandwidth of 100 MHz is more than sufficient. Waveform analysis capabilities are very helpful but not absolutely necessary. RS 232 or IEEE 488 interfaces are required for computer control and data transfer.

94.10.2 Measurement of Emitted Light

Several types of instruments are used for measuring light emission from ACTFEL displays. Photometers are used for luminance measurements. Spectrophotometers are used for measuring the emission spectrum and with suitable software can also provide luminance measurements. Time-resolved measurements are accomplished by using PMTs or photodiode detectors. Table 94.1 lists

TABLE 94.1 Light Measurement Instruments

Instrument Type	Model	Manufacturer
Photometer	PR880	Photo Research
Photometer/spectroradiometer	PR650	Photo Research
	Pritchard 1980B	Photo Research
	GS-1280 RadOMAcam	Gamma Scientific (EG&G)
Photomultiplier tubes	—	Oriel Corp.
	—	Hamamatsu
Photodiode	PIN 10AP	UDT Sensors, Inc.
Flat panel inspection system	FIS 250	Photon Dynamics, Inc.

TABLE 94.2 Manufacturers of Light Measurement Instruments

Photo Research 9330 DeSoto Avenue, P.O. Box 2192 Chatsworth, CA 91313-2192 (818) 341-5151	**Hamamatsu Photonics Systems Corp.** 360-T Foothill Road, P.O. Box 6910 Bridgewater, NJ 08807-0910 Tel: (908) 231-0960
Gamma Scientific (EG&G) 8581 Aero Dr. San Diego, CA 92123-1876 (619) 279-8034	**UDT Sensors, Inc.** 12525 Chadron Ave. Hawthorne, CA 90250 (310) 978-0516
Oriel Corp. 252 Long Beach Blvd., P.O. Box 872 Stratford, CT 06497-0872 (203) 380-4200	**Photon Dynamics, Inc.** 6325 San Ignacio Ave. San Jose, CA 95119 (408) 226-9900

some examples of photometers, spectroradiometers, photomultipliers, and photodiodes along with the names of the companies that manufacture them. A relatively new development for characterizing the light emission characteristics of flat panel displays, including ACTFEL displays, is the flat panel inspection system. This is a large measurement system comprising a CCD camera detector, light-tight enclosure, control computer, image processor, and specialized software. This type of system images an entire flat panel display on the CCD camera and measures luminance, chromaticity, and various defects by analyzing the image. These systems are intended for high-throughput manufacturing environments and cost several hundred thousand dollars. An example of this type of system is also included in Table 94.1. Contact information for the companies listed in Table 94.1 is provided in Table 94.2.

Defining Terms

Electroluminescence: The nonthermal generation of light resulting from the application of an electric field to a substance, usually a luminescent crystal.

Latent image: The ghost image of a previously displayed pattern which can sometimes be seen in a full field on an electronic display screen.

Light emission decay time: The time it takes for the light emission from one excitation pulse to fall to $1/e$ times its initial value.

Polarization charge: The charge trapped at the phosphor/insulator interface following the application of a drive pulse.

Polarization field: The field across the phosphor layer resulting from the polarization charge.

***Q_pF_p* loop:** The closed curve which results from plotting the internal charge flow across the phosphor layer (Q_p) vs. the electric field across the phosphor layer (F_p).

***QV* loop:** The closed curve which results from plotting the external charge (Q) flowing into a TFEL device vs. the externally applied voltage (V).

Threshold voltage: The voltage amplitude of the drive waveform above which current flows across the phosphor layer and light is emitted from a TFEL device.

Turn-on voltage: The voltage corresponding to the first knee in the *QV* loop of a TFEL device. This is the voltage at which charge begins to flow across the phosphor layer. This voltage is generally less than the threshold voltage because the internal field across the phosphor layer is enhanced by the polarization field once the polarization charge has built up in the steady state.

References

1. Y.A. Ono, *Electroluminescent Displays*, Singapore: World Scientific, 1995.
2. T. Inoguchi, M. Takeda, Y. Kakihara, Y. Nakata, and M. Yoshida, Stable high-brightness thin-film electroluminescent panels, *Digest of the 1974 SID International Symposium*, Los Angeles, CA, Vol. 84, 1974.
3. C.N. King, R.E. Coovert, and W.A. Barrow, Full-color 320 × 240 TFEL display panel, *Eurodisplay '87*, London, U.K., Vol. 14, 1987.
4. W.A. Barrow, R.E. Coovert, C.N. King, and M.J. Ziuchkovski, Matrix-addressed full-color TFEL display, *Digest of the 1988 SID International Symposium*, Anaheim, CA, Vol. 284, 1988.
5. R. Tueta and M. Braguier, Fabrication and characterization of indium tin oxide thin films for electroluminescent applications, *Thin Solid Films*, 80: 143, 1981.
6. J.D. Davidson, J.F. Wager, and I. Khormaei, Electrical characterization and SPICE modeling of ZnS:Mn ACTFEL devices, *Digest of the 1991 SID International Symposium*, Anaheim, CA, Vol. 77, 1991.
7. A.A. Douglas and J.F. Wager, ACTFEL device response to systematically varied pulse waveforms, in *Electroluminescence—Proceedings of the Sixth International Workshop on Electroluminescence*, V.P. Singh and J.C. McClure, Eds., El Paso, TX: Cinco Puntos Press, 1992.
8. D.H. Smith, Modeling AC thin-film electroluminescent devices, *J. Lumin*., 23: 209, 1981.
9. P.M. Alt, Thin-film electroluminescent displays: Device characteristics and performance, *Proc. SID*, 25: 123, 1984.
10. Y.A. Ono, H. Kawakami, M. Fuyama, and K. Onisawa, Transferred charge in the active layer and EL device characteristics of TFEL cells, *Jpn. J. Appl. Phys*., 26: 1482, 1987.
11. E. Bringuier, Charge transfer in ZnS-type electroluminescence, *J. Appl. Phys*., 66: 1314, 1989.
12. A. Abu-Dayah, S. Kobayashi, and J.F. Wager, Internal charge-phosphor field characteristics of alternating-current thin-film electroluminescent devices, *Appl. Phys. Lett*., 62: 744, 1993.
13. A. Mikami, K. Terada, K. Okibayashi, K. Tanaka, M. Yoshida, and S. Nakajima, Aging characteristics of ZnS:Mn electroluminescent films grown by a chemical vapor deposition technique, *J. Appl. Phys*., 72: 773, 1992.

95

Light-Emitting Diode Displays

Mohammad A. Karim
The University of Tennessee, Knoxville

95.1 Introduction

A light-emitting diode (LED) display manifests itself as a flat panel display that uses LEDs to display video. An LED is a particular solid-state *p-n* junction diode that gives out light upon the application of a bias voltage. The luminescence process in this case is electroluminescence, which is associated with emission wavelengths in the visible and infrared regions of the spectrum. When a forward bias is applied to the *p-n* junction diode, carriers are injected into the depletion region in large numbers. Because of their physical proximity, the electron–hole pairs undergo a recombination that is associated with the emission of energy. Depending on the semiconductor band-gap characteristics, this emitted energy can be in the form of heat (as phonons) or light (as photons).

95.2 LED Bandgaps

The solution of the Schrödinger equation for a typical crystal reveals the existence of Brillouin zones. A plot between the energy E of an electron in a solid and its wave vector $\mathbf{k}$ represents the allowed energy bands. It may be noted that the lattice structure affects the motion of an electron when k is close to $n\pi/l$ (where n is any integer and l is the crystal periodicity) and the effect of this constraint is to introduce an energy band gap between the allowed energy bands. Figure 95.1a shows portions of two E vs. k curves for neighboring energy bands within the regions $k = \pi/l$ and $k = -\pi/l$ (also known as the reduced zone).

While the upper band of Figure 95.1 represents the energy of conduction band electrons, the curvature of the lower band can be associated with electrons having negative effective mass. The concept of negative effective mass can readily be identified with the concept of holes in the valence band. While the majority of the electrons are identified with the minima of the upper E–k curve, the majority of the holes are identified with the maxima of the lower E–k curve. The minimum value of the conduction band and the maximum value of the valence band in Figure 95.1a both have identical k values. A semiconductor having such a characteristic is said to have a direct band gap, and the associated recombination in such a semiconductor is referred to as direct.

The *direct recombination* of an electron–hole pair always results in the emission of a photon. In a direct band-gap semiconductor, the emitted photon is not associated with any change in momentum

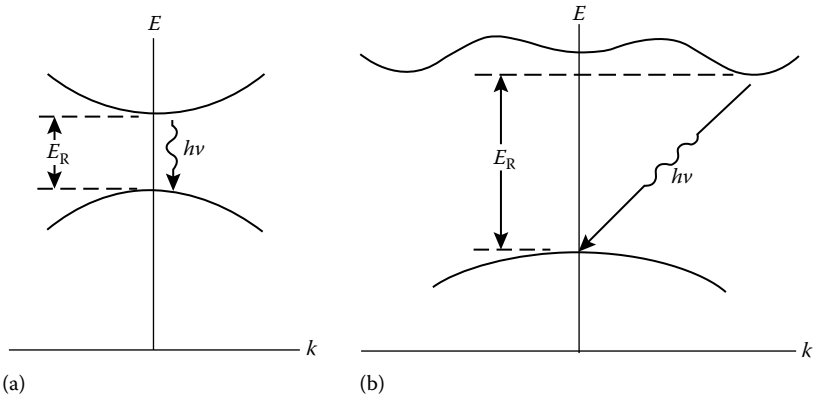


FIGURE 95.1 *E* vs. *k* for semiconductors having (a) a direct band gap and (b) an indirect band gap.

(given by $hk/2\pi$) since $\Delta k = 0$. However, for some semiconducting materials, the *E* vs. *k* curve may be somewhat different, as shown in Figure 95.1b. While the minimum conduction band energy can have a nonzero *k*, the maximum valence band energy can have $k = 0$. The electron–hole recombination in such a semiconductor is referred to as indirect.

An *indirect recombination* process involves a momentum adjustment. Most of the emission energy is thus expended in the form of heat (as phonons). Very little energy is left for the purpose of photon emission, which in most cases is a very slow process. Furthermore, since both photons and phonons are involved in this energy exchange, such transitions are less likely to occur. The interband recombination rate is basically given by

$$\frac{dn}{dt} = B_r np \tag{95.1}$$

where B_r is a recombination-dependent constant which for a direct band-gap semiconductor is $\sim 10^6$ times larger than that for an indirect band-gap semiconductor. For direct recombination, B_r value ranges from 0.46×10^{-10} to 7.2×10^{-10} cm^3/s.

All semiconductor crystal lattices are alike, being dissimilar only in terms of their band characteristics. Si and Ge both have indirect band transitions, whereas GaAs, for example, is a semiconductor that has a direct band transition. Thus, while Si and Ge are preferred for fabrication of transistors and integrated circuits, GaAs is preferred for the fabrication of LEDs.

95.3 LED Wavelengths

The direct recombination (when *k* = constant) results in a photon emission whose wavelength (in micrometers) is given by

$$\lambda = \frac{hc}{E_g} = \frac{1.24}{E_g}\,(\text{eV}) \tag{95.2}$$

where E_g is the band-gap energy. The LEDs under proper forward-biased conditions can operate in the ultraviolet, visible, and infrared regions. For the visible region, however, the spectral luminous efficiency curves of Figure 95.2, which account for the fact that the visual response to any emission is a function of wavelength, should be of concern. It is unfortunate that there is not a single-element semiconductor suitable for fabrication of LEDs, but there are many binary and ternary compounds that can be used for fabrication of LEDs. Table 95.1 lists some of these binary semiconductor materials. The ternary semiconductors include GaAlAs, $CdGeP_2$, and $ZnGeP_2$ for infrared region operation, $CuGaS_2$ and $AgInS_2$ for visible region

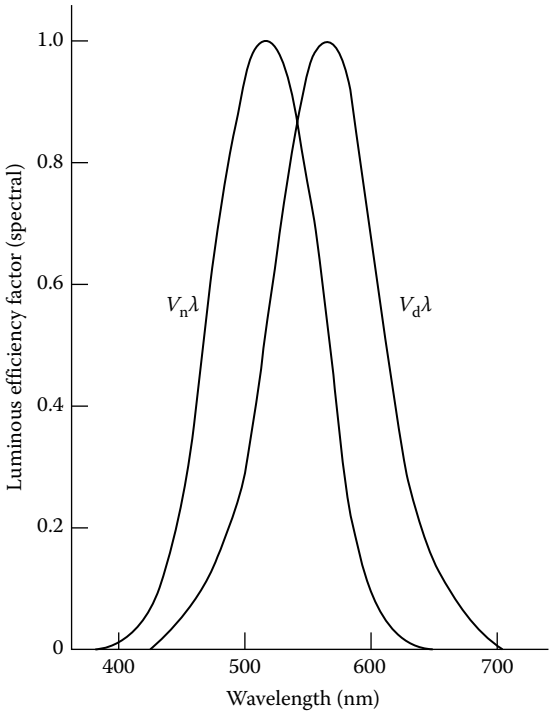


FIGURE 95.2 Spectral luminous efficiency curves. The photopic curve $V_d\lambda$ corresponds to the daylight-adapted case while the scotopic curve $V_n\lambda$ corresponds to the night-adapted case. (From Boyd, R.W.: *Radiometry and the Detection of Optical Radiation*. 1983. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reprinted with permission.)

TABLE 95.1 Binary Semiconductors Suitable for LED Fabrication

	Material	E_g (eV)	Emission Type
III–V	GaN	3.5	UV
II–VI	ZnS	3.8	UV
II–VI	SnO_2	3.5	UV
II–VI	ZnO	3.2	UV
III–VII	CuCl	3.1	UV
II–VI	BeTe	2.8	UV
III–VII	CuBr	2.9	UV–visible
II–VI	ZnSe	2.7	Visible
III–VI	In_2O_3	2.7	Visible
II–VI	CdS	2.52	Visible
II–VI	ZnTe	2.3	Visible
III–V	GaAs	1.45	IR
II–VI	CdSe	1.75	IR–visible
II–VI	CdTe	1.5	IR
III–VI	GaSe	2.1	Visible

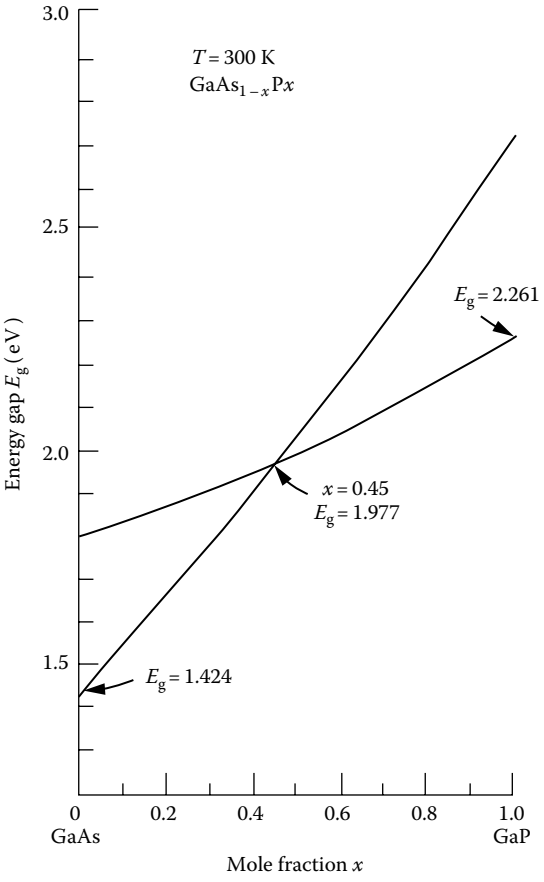


FIGURE 95.3 Band-gap energy vs. x in $GaAs_{1-x}P_x$. (From Casey, H.J., Jr. and Parish, M.B., eds., *Heterostructure Lasers*, Academic Press, New York, 1978. With permission.)

operation, and $CuAlS_2$ for ultraviolet region operation. Ternary semiconductors are used because their energy gaps can be tuned to a desired emission wavelength by picking appropriate composition.

Of the ternary compounds, gallium arsenide–phosphide (written as $GaAs_{1-x}P_x$) is an example that is basically a combination of two binary semiconductors, namely, GaAs and GaP. The corresponding band-gap energy of the semiconductor can be varied by changing the value of x. For example, when $x = 0$, $E_g = 1.43$ eV. E_g increases with increasing x until $x = 0.44$ and $E_g = 1.977$ eV, as shown in Figure 95.3. However for $x \geq 0.45$, the band gap is indirect. The most common composition of $GaAs_{1-x}P_x$ used in LEDs has $x = 0.4$ and $E_g \cong 1.3$ eV. This band-gap energy corresponds to an emission of red light. Calculators and watches often use this particular composition of $GaAs_{1-x}P_x$.

Interestingly, the indirect band gap of $GaAs_{1-x}P_x$ (with $1 \geq x \geq 0.45$) can be used to output light ranging from yellow through green provided the semiconductor is doped with impurities such as nitrogen. The dopants introduced in the semiconductor replace phosphorus atoms which, in turn, introduce electron trap levels very near the conduction band. For example, $x = 0.5$, the doping of nitrogen increases the LED efficiency from 0.01% to 1%, as shown in Figure 95.4. It must be noted, however, that nitrogen doping shifts the peak emission wavelength toward the red. The shift is comparatively larger at and around $x = 0.05$ than $x = 1.0$. The energy emission in nitrogen-doped $GaAS_{1-x}P_x$ devices is a function of both x and the nitrogen concentration.

Nitrogen is a different type of impurity from those commonly encountered in extrinsic semiconductors. Nitrogen, like arsenic and phosphorus, has five valence electrons, but it introduces no net charge

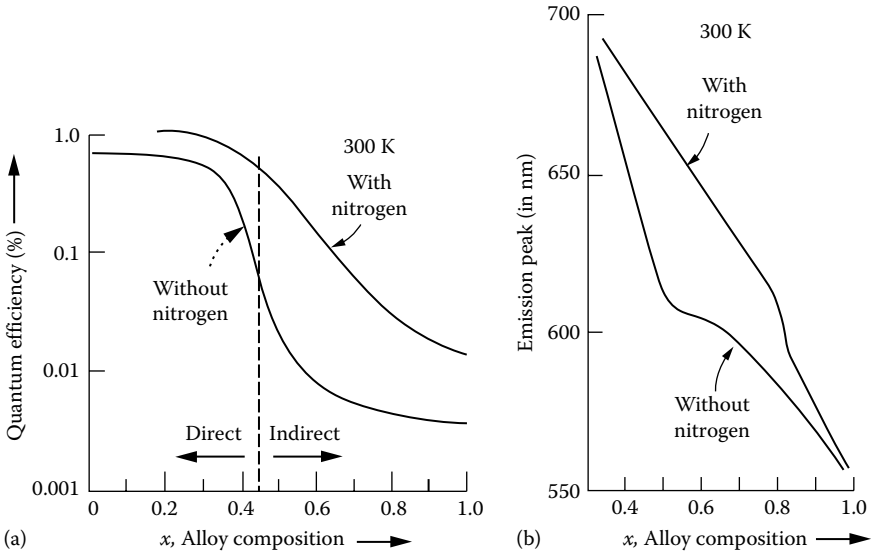


FIGURE 95.4 The effects of nitrogen doping in $GaAS_{1-x}P_x$: (a) quantum efficiency vs. x and (b) peak emission wavelength vs. x.

carriers in the lattice. It provides active radiative recombination centers in the indirect band-gap materials. For an electron, a recombination center is an empty state in the band gap into which an electron falls and, then, thereafter, falls into the valence band by recombining with a hole. For example, while a GaP LED emits green light (2.23 eV), a nitrogen-doped GaP LED emits yellowish green light (2.19 eV), and a heavily nitrogen-doped GaP LED emits yellow light (2.1 eV).

The injected excess carriers in a semiconductor may recombine either radiatively or nonradiatively. Whereas nonradiative recombination generates phonons, radiative recombination produces photons. Consequently, the internal quantum efficiency η, defined as the ratio of the radiative recombination rate R_r to the total recombination rate, is given by

$$\eta = \frac{R_r}{(R_r + R_{nr})} \tag{95.3}$$

where R_{nr} is the nonradiative recombination rate. However, the injected excess carrier densities return to their value exponentially as

$$\Delta p = \Delta n = \Delta n_o e^{-t/\tau} \tag{95.4}$$

where

τ is the carrier lifetime

Δn_o is the excess electron density at equilibrium

Since $\Delta n/R_r$ and $\Delta n/R_{nr}$ are, respectively, equivalent to the radiative recombination lifetime τ_r and the nonradiative recombination lifetime τ_{nr}, we can obtain the effective minority carrier bulk recombination time τ as

$$\left(\frac{1}{\tau}\right) = \left(\frac{1}{\tau_r}\right) + \left(\frac{1}{\tau_{nr}}\right) \tag{95.5}$$

such that $\eta = \tau/\tau_r$. The reason that a fast recombination time is crucial is that the longer the carrier remains in an excited state, the larger the probability that it will give out energy nonradiatively. In order for the internal quantum efficiency to be high, the radiative lifetime τ_r needs to be small. For indirect band-gap semiconductors, $\tau_r \gg \tau_{nr}$ so that very little light is generated, and for direct band-gap semiconductors, τ_r increases with temperature so that the internal quantum efficiency deteriorates with the temperature.

95.4 LED Response Time

As long as the LEDs are used as display devices, it is not too important to have fast response characteristics. However, LEDs are also used for the purpose of optical communications, and for those applications it is appropriate to study their time response characteristics. For example, an LED can be used in conjunction with a photodetector for transmitting optical information between two points. The LED light output can be modulated to convey optical information by varying the diode current. Most often, the transmission of optical signals is facilitated by introducing an optical fiber between the LED and the photodetector.

There can be two different types of capacitances in diodes that can influence the behavior of the minority carriers. One of these is the *junction capacitance*, which is caused by the variation of majority charge in the depletion layer. While it is inversely proportional to the square root of bias voltage in the case of an abrupt junction, it is inversely proportional to the cube root of bias voltage in the case of a linearly graded junction. The second type of capacitance, known as the *diffusion capacitance*, is caused by the minority carriers.

Consider an LED that is forward biased with a dc voltage. Consider further that the bias is perturbed by a small sinusoidal signal. When the bias is withdrawn or reduced, charge begins to diffuse from the junction as a result of recombination until an equilibrium condition is achieved. Consequently, as a response to the signal voltage, the minority carrier distribution contributes to a signal current.

Consider a one-dimensional *p*-type semiconducting material of cross-sectional area *A* whose excess minority carrier density is given by

$$\frac{\delta \Delta n_p}{\delta t} = \frac{D_n \delta^2 \Delta n_p}{\delta x^2} - \frac{\Delta n_p}{\tau} \tag{95.6}$$

As a direct consequence of the applied sinusoidal signal, the excess electron distribution fluctuates about its dc value. In fact, we may assume excess minority carrier density to have a time-varying component as described by

$$\Delta n_p(x,t) = \langle \Delta n_p(x) \rangle + n'_p(x) e^{j\omega t} \tag{95.7}$$

where $\langle \Delta n_p(x) \rangle$ is a time-invariant quantity. By introducing Equation 95.7 into Equation 95.6, we get two separate differential equations:

$$\frac{\delta^2}{\delta x^2}(\langle \Delta n_p(x) \rangle) = \frac{\langle \Delta n_p(x) \rangle}{(L_n)^2} \tag{95.8a}$$

and

$$\frac{\delta^2}{\delta x^2}[\Delta n'_p(x)] = \frac{\Delta n'_p(x)}{[L_n^*]^2} \tag{95.8b}$$

where

$$L_n^* = \frac{L_n}{(1+j\omega\tau)^{1/2}} \tag{95.9a}$$

and

$$L_n = (D_n\tau)^{1/2} \tag{95.9b}$$

The dc solution of Equation 95.8a is well known. Again, the form of Equation 95.8b is similar to that of Equation 95.8a and, therefore, its solution is given by

$$\Delta n_p'(x) = \Delta n_p'(0)e^{-x/L} \tag{95.10}$$

Since the frequency-dependent current $I(\omega)$ is simply a product of eAD_n and the concentration gradient, we find that

$$\begin{aligned} I(\omega) &= \left| \frac{eAD_n dn_p'(x)}{dx} \right|_{x=0} \\ &= \frac{I(0)}{(1+\omega^2\tau^2)^{1/2}} \end{aligned} \tag{95.11}$$

where $I(0)$ is the intensity emitted at zero modulation frequency. We can determine the admittance next by dividing the current by the perturbing voltage. The real part of the admittance, in this case, will be equivalent to the diode conductance, whereas its imaginary part will correspond to the diffusion capacitive susceptance.

The modulation response as given by Equation 95.11 is, however, limited by the carrier recombination time. Often an LED is characterized by its modulation bandwidth, which is defined as the frequency band over which signal power (proportional to $I^2(\omega)$) is half of that at $\omega = 0$. Using Equation 95.11, the 3 dB modulation bandwidth is given by

$$\Delta\omega \approx \frac{1}{\tau_r} \tag{95.12}$$

where the bulk lifetime has been approximated by the radiative lifetime. Sometimes the 3 dB bandwidth of the LED is given by $I(\omega) = \frac{1}{2}I(0)$, but this simplification contributes to an erroneous increase in the bandwidth by a factor of 1.732.

Under conditions of thermal equilibrium, the recombination rate is proportional to the product of initial carrier concentrations, n_o and p_o. Then, under nonequilibrium conditions, additional carriers $\Delta n = \Delta p$ are injected into the material. Consequently, the recombination rate of injected excess carrier densities is given by initial carrier concentrations and injected carrier densities as

$$\begin{aligned} R_{\Delta r} &= [B_r(n_o + \Delta n)(p_o + \Delta p) - B_r n_o p_o] \\ &= B_r(n_o + p_o + \Delta n)\Delta n \end{aligned} \tag{95.13}$$

where B_r is the same constant introduced in Equation 95.1. For p-type GaAs, for example, $B_r = 1.7 \times 10^{-10}$ cm^3/s when $p_o = 2.4 \times 10^{18}$ holes/cm^3. Equation 95.13 is used to define the radiative carrier recombination lifetime by

$$\tau_r = \frac{\Delta n}{R_{\Delta r}} = [B_r(n_o + p_o + \Delta n)]^{-1} \tag{95.14}$$

In the steady-state condition, the excess carrier density can be calculated in terms of the active region width d by

$$\Delta n = \frac{J\tau_r}{ed} \tag{95.15}$$

where J is the injection current density.
The radiative recombination lifetime is found by solving Equation 95.14 after having eliminated Δn from it using Equation 95.15:

$$\tau_r = \frac{\left[\left\{(n_o + p_o)^2 + (4J/B_r ed)\right\}^{1/2} - (n_o + p_o)\right]}{(2J/ed)} \tag{95.16}$$

Thus, while for the low carrier injection (i.e., $n_o + p_o \gg \Delta n$), Equation 95.16 reduces to

$$\tau_r \approx [B_r(n_o + p_o)]^{1/2} \tag{95.17a}$$

for the high carrier injection (i.e., $n_o + p_o \ll \Delta n$), it reduces to

$$\tau_r \approx \left(\frac{ed}{JB_r}\right)^{1/2} \tag{95.17b}$$

Equation 95.17a indicates that in highly doped semiconductors, τ_r is small. But the doping process has its own problem, since in many of the binary LED compounds higher doping may introduce nonradiative traps just below the conduction band, thus nullifying Equation 95.12. In comparison to Equation 95.17a, Equation 95.17b provides a better alternative whereby τ_r can be reduced by decreasing the active region width or by increasing the current density. For the case of p-type GaAs, the radiative lifetimes vary between 2.6 and 0.35 ns, respectively, when p_o varies between 1.0×10^{18} and 1.5×10^{19} holes/cm^3.

Usually, LEDs are operated at low current (≈10 mA) and low voltages (≈1.5 V), and they can be switched on and off in the order of 10 ns. In addition, because of their small sizes, they can be reasonably treated as point sources. It is, therefore, not surprising that they are highly preferred over other light sources for applications in fiber-optic data links.

95.5 LED Designs

Two particular LED designs are popular: *surface emitters* and *edge emitters*. They are shown in Figure 95.5. In the former, the direction of major emission is normal to the plane of the active region, whereas in the latter the direction of major emission is in the plane of the active region. The emission pattern of the surface emitters is very much isotropic, whereas that of the edge emitters is highly directional.

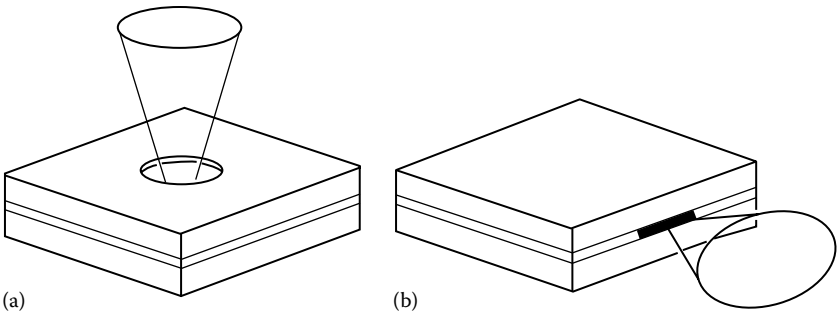


FIGURE 95.5 LED type: (a) surface emitter and (b) edge emitter.

As the LED light originating from a medium of refractive index n_1 goes to another medium of refractive index $n_2 (n_2 < n_1)$, only a portion of incident light is transmitted. In particular, the portion of the emitted light corresponds to only that which originates from within a cone of semiapex angle θ_c, such that

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{95.18}$$

In the case of an LED, n_1 corresponds to the refractive index of the LED medium and n_2 corresponds to that of air (or vacuum). Light originating from *beyond* angle θ_c undergoes a total internal reflection. However, the light directed from *within* the cone of the semiapex angle θ_c will be subjected to Fresnels loss. Thus, the overall transmittance T is given by

$$T = 1 - \left\{\frac{(n_1 - n_2)}{(n_1 + n_2)}\right\}^2 \tag{95.19}$$

Accordingly, the total electrical-to-optical conversion efficiency in LEDs is given by

$$\begin{aligned}
\eta_{\text{LED}} &= T\left[\frac{(\text{solid angle within the cone})}{(4\pi)}\right] \\
&= \left(\frac{T}{2}\right)(1 - \cos\theta_c) \\
&= \left(\frac{T}{4}\right)\sin^2\theta_c \\
&= \left(\frac{T}{4}\right)\left(\frac{n_2}{n_1}\right)^2\left[1 - \left\{\frac{(n_1 - n_2)}{(n_1 + n_2)}\right\}^2\right]
\end{aligned} \tag{95.20}$$

Only two schemes increase the electrical-to-optical conversion efficiency in an LED. The first technique involves guaranteeing that most of the incident rays strike the glass-to-air interface at angles less than θ_c. It is accomplished by making the semiconductor–air interface hemispherical. The second method involves schemes whereby the LED is encapsulated in an almost transparent medium of high refractive index. The latter means is comparatively less expensive. If a glass of refractive index 1.5 is used for encapsulation, the LED efficiency can be increased by a factor of 3. Two of the possible encapsulation arrangements and the corresponding radiation patterns are illustrated in Figure 95.6.

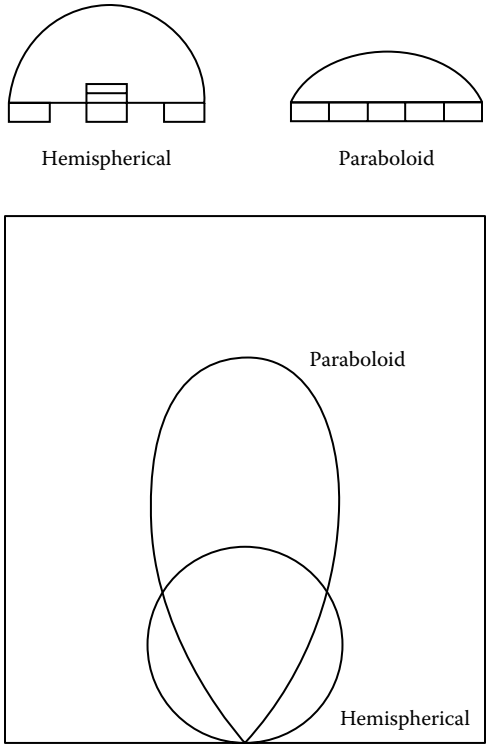


FIGURE 95.6 LED encapsulation geometries and their radiation patterns.

LEDs are often used in conjunction with a phototransistor to function as an optocoupler. The optocouplers are used in circumstances when it is desirable to have a transmission of signals between electrically isolated circuits. They are used to achieve noise separation by eliminating the necessity of having a common ground between the two systems. Depending on the type of coupling material, these miniature devices can provide both noise isolation as well as high voltage isolation. Figure 95.7 shows a typical case

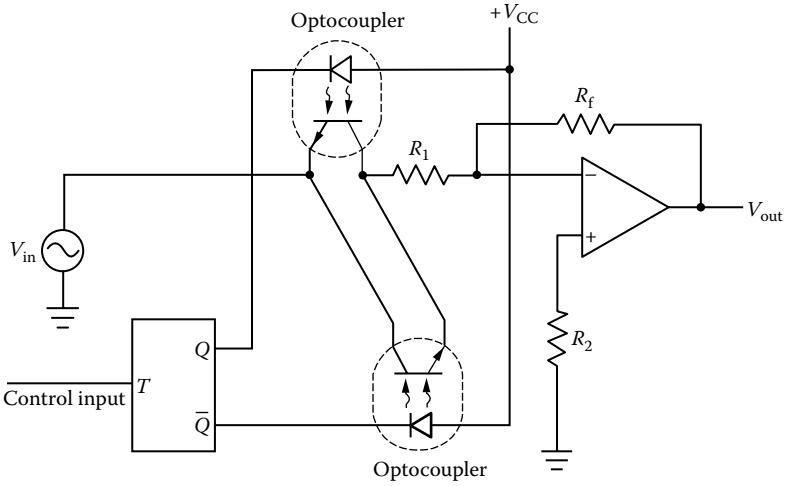


FIGURE 95.7 A chopping circuit with an amplifier.

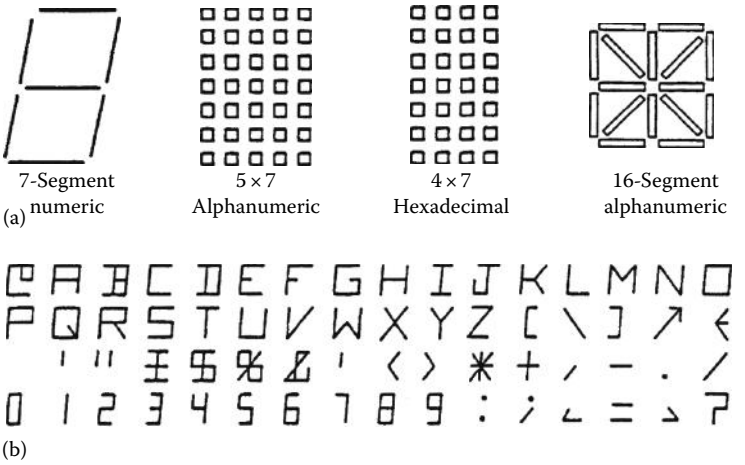


FIGURE 95.8 (a) LED display formats and (b) displayed alphanumeric characters using 16-segment displays.

where two optocouplers are used to attain a chopper circuit. The two optocouplers chop either the positive or the negative portion of the input signals with a frequency of one half that of the control signal that is introduced at the *T* flip-flop. The operational amplifier provides an amplified version of the chopped output waveform. In comparison, a chopper circuit that uses simple bipolar transistors produces noise spikes in the output because of its inherent capacitive coupling.

The visible LEDs are best known for their uses in displays and indicator lamps. In applications where more than a single source of light is required, an LED array can be utilized. An LED array is a device consisting of a row of discrete LEDs connected together within or without a common reflector cavity. Figure 95.8a shows different LED arrangements for displaying hexadecimal numeric and alphanumeric characters, whereas Figure 95.8b shows, for example, the possible alphanumeric characters using 16-segment displays. In digital systems, the binary codes equivalent to these characters are usually decoded and, consequently, a specific combination of LED segments are turned on to display the desired alphanumeric character.

The dot matrix display provides the most desirable display font. It gives more flexibility in shaping characters and has a lower probability of being misinterpreted in case of a display failure. However, these displays involve a large number of wires and increased circuit complexity. LED displays, in general, have an excellent viewing angle, high resonance speed (≈10 ns), long life, and superior interface capability with electronics with almost no duty cycle limitation. LEDs with blue emission are not available commercially. When compared with passive displays, LED displays consume more power and involve complicated wiring with at least one wire per display element.

An alternative version of the LED displays can be realized using organic LED (OLED) that relies instead on carbon-based semiconductors. OLEDs create white by combining red, green, and blue semiconductor films—either stacked on top of each other or arranged in the form of alternating stripes. In OLED, the electroluminescent emissive layer of the diode is an organic compound in either crystalline form or polymers. OLEDs can provide for thin displays with a low driving voltage, wide viewing angle, and high contrast and color gamut.

References

1. R.W. Boyd, *Radiometry and the Detection of Optical Radiation*, New York: John Wiley & Sons, 1983.
2. H.J. Casey, Jr. and M.B. Parish (eds.), *Heterostructure Lasers*, New York: Academic Press, 1978.

Bibliography

1. C. Hilsum, Flat-panel electronic displays: A triumph of physics, chemistry and engineering, *Phil. Trans. Roy. Soc*., 368, 1027–1082, March 2010.
2. M.A. Karim, *Electro-Optical Devices and Systems*, Boston, MA: PWS-Kent Publishing, 1990.
3. M.A. Karim, *Electro-Optical Displays*, New York: Marcel Dekker, 1992.
4. J.H. Lee, D.N. Liu, and S.T. Wu, *Introduction to Flat Panel Displays*, New York: John Wiley & Sons, February 2009.
5. G.P. Crawford, Flexible flat pane display technology, in *Flexible Flat Panel Displays*, G.P. Crawford (ed.), Wiley, Chichester, West Sussex, U.K., pp. 1–11, 2005.
6. J. Wilson and J.F.B. Hawkes, *Optoelectronics: An Introduction*, 3rd edn., Englewood Cliffs, NJ: Prentice Hall PTR, 1998.

96

Magnetic and Optical Recording

Yufeng Li
Samsung Information Systems America

96.1 Introduction

The heart of recording technology is for the process of information storage and retrieval. In addition to its obvious importance in different branches of science and engineering, it has become indispensable to our daily life. When we make a bank transaction, reserve an airplane ticket, use a credit card, watch a movie from a video tape, or listen to music from a CD, we are using the technology of recording. The general requirements for recording are information integrity, fast access, and low cost. Among the different techniques, the most popularly used ones are magnetic and optical recording.

Typical recording equipment consists of a read/write head, a medium, a coding/decoding system, a data access system, and some auxiliary mechanical and electronic components. The head and medium are for data storage and retrieval purposes, and the coding/decoding system is for data error correction. The data access system changes the relative position between the head and the medium, usually with a servo mechanism for data track following and a spinning mechanism for on-track moving. While the data access system and the auxiliary components are important to recording equipment, they are not considered essential in this chapter to the understanding of recording technology, and will not be covered. Interested readers are referred to Ref. [1].

96.2 Magnetic Recording

At present, magnetic recording technology dominates the recording industry. It is used in the forms of hard disk, floppy disk, removable disk, and tape with either digital or analog mode. In its simplest form, it consists of a magnetic head and a magnetic medium, as shown in Figure 96.1. The head is made of a piece of magnetic material in a ring shape (core), with a small gap facing the medium and a coil away from the medium. The head records (writes) and reproduces (reads) information, while the medium stores the information. The recording process is based on the phenomenon that an electric current i generating a magnetic flux ϕ as described by Ampere's law. The flux ϕ leaks out of the head core at the gap,

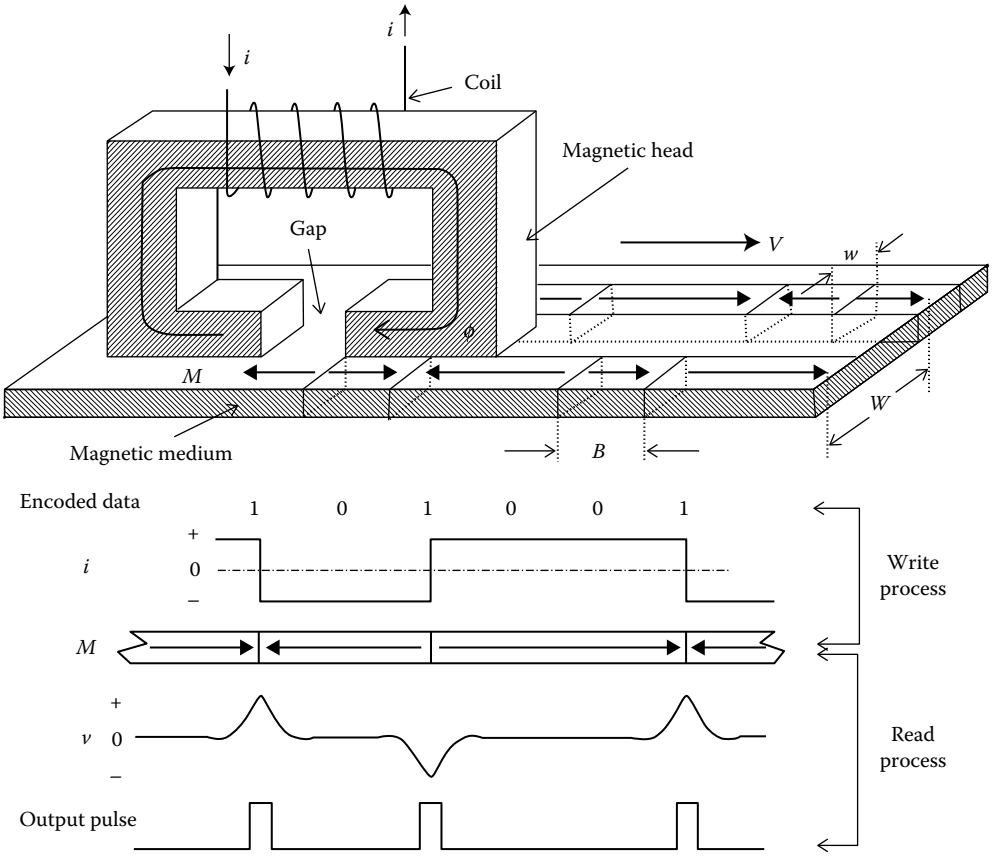


FIGURE 96.1 Conceptual diagrams illustrating the magnetic recording principle (a) and recording/reproducing process (b).

and magnetizes the magnetic medium which moves from left to right with a velocity V under the head gap. Depending on the direction of the electric current i, the medium is magnetized with magnetization M pointing either left or right. This pattern of magnetization is retained in the memory of the medium even after the head moves away.

Two types of head may be used for reproducing. One, termed the *inductive head*, senses magnetic flux change rate, and the other, named the *magnetoresistive* (MR) *head*, senses the magnetic flux. When an inductive head is used, the reproducing process is just the reverse of the recording process. The flux coming out of the magnetized medium surface is picked up by the head core. Because the medium magnetization under the head gap changes its magnitude and direction as the medium moves, an electric voltage is generated in the coil. This process is governed by Faraday's law. Figure 96.1b schematically shows the digital recording/reproducing process. First, all user data are encoded into a binary format—a serial of 1s and 0s. Then a write current i is sent to the coil. This current changes its direction whenever a 1 is being written. Correspondingly, a change of magnetization, termed a *transition*, is recorded in the medium for each 1 in the encoded data. During the reproducing process, the electric voltage induced in the head coil reaches a peak whenever there is a transition in the medium. A pulse detector generates a pulse for each transition. These pulses are decoded to yield the user data.

The minimum distance between two transitions in the medium is the flux change length B, and the distance between two adjacent signal tracks is the track pitch W, which is wider than the signal track width w. The flux change length can be directly converted into bit length with the proper

code information. The reciprocal of the bit length is called *linear density*, and the reciprocal of the track pitch is termed *track density*. The information storage areal density in the medium is the product of the linear density and the track density. This areal density roughly determines how much information a user can store in a unit surface area of storage medium, and is a figure of merit for a recording technique. Much effort has been expended to increase the areal density. For example, it has been increased 50 times during the last decade in hard disk drives, and is expected to continue increasing 60% per year in the foreseeable future. At present, state-of-the-art hard disk products feature areal densities of >7 Mbits/mm^2 ($B < 0.1$ μm and $W < 1.5$ μm). This gives a total storage capacity of up to 6 Gbytes for a disk of 95 mm diameter.

96.2.1 Magnetism and Hysteresis Loop

Magnetism is the result of uncompensated electron spin motions in an atom. Only transition elements exhibit this property, and nearly all practical interest in magnetism centers on the first transition group of elements (Mn, Cr, Fe, Ni, and Co) and their alloys. The strength of magnetism is represented by magnetization M, and is related to magnetic field H and magnetic flux density B by

$$B = \mu_0(H + M) \tag{96.1}$$

where

μ_0 is the permeability of vacuum
since M is a property of a magnetic material, it does not exist outside the magnetic material
H represents the strength acting on a magnetic material from a magnetic field which is generated either by a magnetic material or by an electric current
B is the flux density which determines the induced electric voltage in a coil. The ratio of B with and without a magnetic material is the relative permeability μ of that magnetic material

When a magnetic field H is applied to a piece of demagnetized magnetic material, the magnetization M starts increasing with H from zero. The rate of increase gradually slows down and M asymptotically approaches a value M_s at high H. If H is reduced to zero, then M is reduced to a lower value M_r. Continuous reduction of H to a very high negative value will magnetize the material to $-M_s$. In order to bring the material to demagnetized state, a positive field H_c is required. Further increase in the H field will bring the trace of M to a closed loop. This loop is the major hysteresis loop, as shown in Figure 96.2.

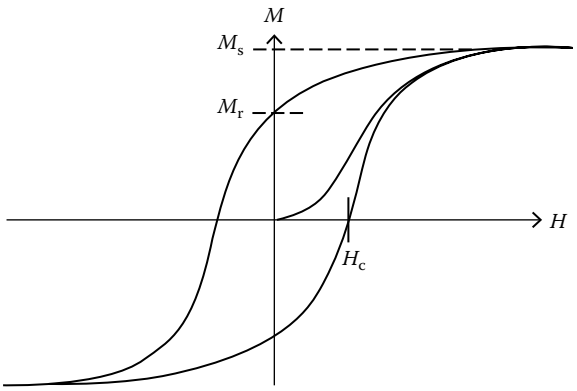


FIGURE 96.2 Hysteresis loop of a magnetic material shows the nonlinear relationship between M and H that results in magnetic memory.

The hysteresis loop shows that a magnetic material has memory. It is this memory that is used in the medium for storing information. H_c is the coercivity, indicating the strength of magnetic field required to erase the memory of a magnetic material. Magnetic materials with high H_c are "hard" magnets, and are suitable for medium applications if they have high M_r. On the other hand, magnetic materials with low H_c are "soft" magnets, and are candidates for head core materials if they have high M_s and high μ. M_r and M_s are the remanent and saturation magnetization, respectively, and their ratio is the remanent squareness. The flux density corresponding to M_s is B_s.

96.2.2 Magnetic Media

Magnetic media are used to store information in a magnetic recording system. In order to increase the areal density, we need to reduce flux change length B and track width w. Since B is limited by the term $M_r\delta/H_c$, where δ is the magnetic layer thickness, we can reduce B by either decreasing $M_r\delta$ or increasing H_c. However, the amplitude of the magnetic signal available for reproducing head is proportional to the term $M_r\delta w$. If we reduce track width w to increase areal density, we must increase $M_r\delta$ to avoid signal deterioration. In addition, if the magnetic layer is so thin that it causes thickness nonuniformity, more noise will appear in the reproducing process. Therefore, the major requirements for magnetic layer are high H_c, high M_r, and ease of making a uniform thin layer. Additional requirements include good magnetic and mechanical stability.

There are two groups of magnetic media. The first group is called particulate media because the magnetic materials are in the form of particles. This group includes iron oxide (γ-Fe_2O_3), cobalt-modified iron oxide (γ-Fe_2O_3+Co), chromium dioxide (CrO_2), metal particles, and barium ferrite ($BaFe_{12}O_{19}$). Some of these have been used in the magnetic recording for several decades. More recently, another group of media has been developed largely due to the ever-increasing demand for higher storage capacity in the computer industry. This group of media is the thin-film media, where the magnetic layer can be made as a continuous thin film. Most materials in this group are cobalt-based metal alloys. Compared with particulate media, the thin-film media usually have a higher coercivity H_c, a higher remanence M_r, and can be deposited in a very thin continuous film. Table 96.1 lists H_c and M_r for some of the most popularly used particulate and thin-film media. Note that magnetic properties are affected by the fabrication process and film structure. Therefore, their values can be out of the ranges of Table 96.1 if different processes are used.

Magnetic media can be classified into three general forms of applications. Tape is the oldest form and remains an important medium today. It is central to most audio, video, and instrumentation recording, although it is also used in the computer industry for archival storage. Tape is economical and can hold a large capacity, but suffers slow access time. Hard disk is primarily used as the storage inside a computer, providing fast data access for the user, but having poor transportability. Flexible disk is designed for easy data transportation, but is limited in capacity. Besides these three general forms of applications, a hybrid of

TABLE 96.1 Remanence (M_r) and Coercivity (H_c) Values of Some Commonly Used Magnetic Media

Group	Material	M_r (kA/m)	H_c (kA/m)	Application
Particulate	γ-Fe_2O_3	56–140	23–32	Floppy disk, audio, video, and instrumentation tapes
	γ-Fe_2O_3 + Co	60–140	44–74	Floppy disk, audio, video, and instrumentation tapes
	CrO_2	110–140	38–58	Floppy disk, audio, video, and instrumentation tapes
	$BaFe_{12}O_{19}$	56	58	Floppy disk
Thin film	Co–Ni	600–1100	30–85	Hard disk
	Co–Fe	1100–1500	60–150	Hard disk
	Co–P	600–1000	36–120	Hard disk
	Co–Ni–Pt	600–1100	60–175	Hard disk
	Co–Cr–Ta	350–900	55–190	Hard disk
	Co–Cr–Pt	300–750	56–200	Hard disk

Source: Some Values Are from Maller, V.A.J. and Middleton, B.K., *IERE Conf. Proc.*, 26, 137, 1973.

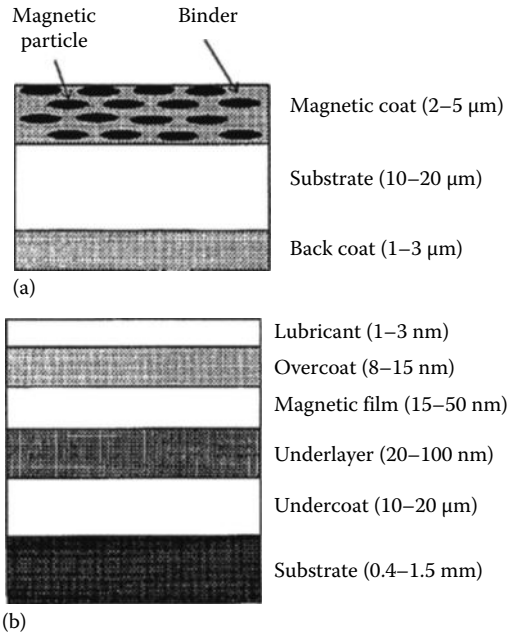


FIGURE 96.3 Cross-sectional views of a particulate magnetic tape (a) and a thin-film hard disk (b).

flexible and hard disk is being gradually accepted. It is a removable rigid disk capable of holding up to several gigabytes of digital data. In addition, magnetic stripes are getting wide use in different forms of cards.

The magnetic layer alone cannot be used as a medium. It needs additional components to improve its chemical and mechanical durability. Typical cross sections of a particulate magnetic tape and a thin-film hard disk are shown in Figure 96.3. In the case of tape application, iron particles with typical size of 0.5 μm long and 0.1 μm wide are dispersed in a polymeric binder, together with solvents, lubricants, and other fillers to improve magnetic and mechanical stability. This dispersed material is then coated on an abiaxially oriented polyethylene terephthalate substrate. An optional back coat may also be applied to the other side of the substrate. The cross section of a hard disk is more complex. A high-purity aluminum–magnesium (5 wt%) substrate is diamond turned to a fine surface finish, and then electrolessly plated with a nonmagnetic nickel–phosphorus (10 at%) undercoat. This layer is used to increase the hardness, reduce the defects, and improve the finish of the Al–Mg alloy, and is polished to a super surface finish. Next, an underlayer of chromium is sputtered to control the properties of the magnetic film, followed by sputtering the magnetic film. Finally, a layer of hydrogenated or nitrogenated carbon is overcoated on the magnetic film, and an ultrathin layer of perfluorinated hydrocarbon liquid lubricant is applied on top. The carbon and lubricant layers are used to improve the corrosion and mechanical resistance of the disk. For a 95 mm disk, the finished product should have a surface flatness better than 10 μm and a tightly control surface roughness. In some applications, an arithmetic average roughness (R_a) of <0.5 nm is required.

96.2.3 Magnetic Heads

Magnetic heads have three functions: recording, reproducing, and erasing. Usually for stationary head applications such as tape drives, multiple heads are used to perform these functions. For moving head applications such as disk drives, a single head is employed because of the requirements of simple connections and small head mass for fast data access. Most of these heads are the inductive type, where the fundamental design is an inductive coil and a magnetic core. The general requirements for the core materials are high relative permeability μ, high saturation flux density B_s, low coercivity H_c, high electric

resistivity ρ, and low magnetostriction coefficient λ. Some of the properties for the commonly used core materials are listed in Table 96.2.

The evolution of the magnetic head follows the selection of core materials, as shown in Figure 96.4. Early heads used laminated molybdenum Permalloy (Ni–Fe–Mo, 79–17–4 wt%). These heads are inexpensive to make, and have low H_c and high μ and B_s. The primary drawbacks are frequency limitation, gap dimension inaccuracy, and mechanical softness. Frequency limitation is caused by the difficulty of making the

TABLE 96.2 Relative Permeability (μ), Saturation Flux Density (B_s), Coercivity (H_c), and Resistivity (ρ) Values of Some Commonly Used Magnetic Head Materials at Low Frequency

Material	μ	B_s (T)	H_c (A/m)	ρ (μΩ cm)	Application
Ni–Fe–Mo	11,000	0.8	2.0	100	Audio tape
Ni–Zn	300–1,500	0.4–0.46	11.8–27.6	10^{11}	Floppy and hard disk drives, video and instrumentation tapes
Mn–Zn	3,000–10,000	0.4–0.6	11.8–5.8	10^{6}	Floppy and hard disk drives, video and instrumentation tapes
Fe–Si–Al	8,000	1.0	2.0	85	Floppy and hard disk drives, video and instrumentation tapes
Ni–Fe	2,000–4,000	1.0	<10	20	Hard disk drives

Source: Some Values Are from Maller, V.A.J. and Middleton, B.K., *IERE Conf. Proc.*, 26, 137, 1973.

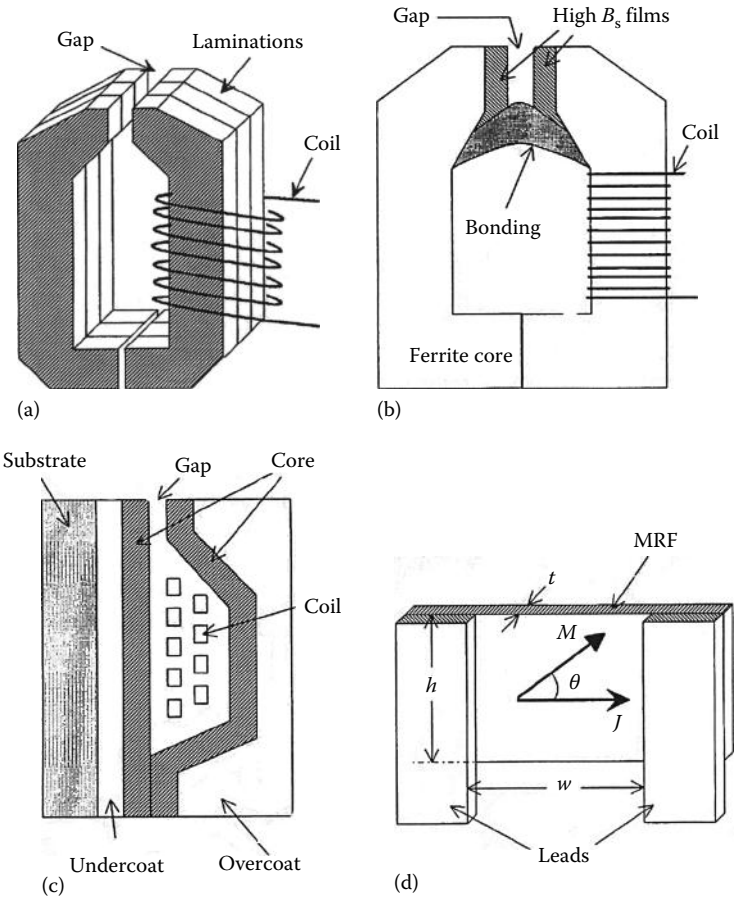


FIGURE 96.4 Schematic illustrations of (a) a laminated head, (b) cross section of a MIG head, (c) cross section of a thin-film head, and (d) an MR sensor with leads.

lamination layer thinner than 25 μm. Eddy current loss, which is proportional to layer thickness and square root of frequency, reduces the effective permeability. As a result, laminated heads are seldom used for applications exceeding 10 MHz. Gap dimension inaccuracy is associated with the head fabrication process, and makes it unsuitable for high areal density applications. Lack of mechanical hardness reduces its usable life.

One way to reduce eddy current loss is to increase core material electric resistivity. Two types of ferrite material have high resistivity (four to nine orders higher than Permalloy) and reasonable magnetic properties: Ni–Zn and Mn–Zn. These materials are also very hard, elongating head life during head/medium contacts. The major deficiency of ferrite materials is their low B_s values. In order to record in high H_c media, high flux density B is needed in the head core. When the flux density in the core material reaches its saturation B_s, it will not increase despite the increase of recording current or coil turns. This saturation starts from the corners of the gap due to its geometry. To remedy this deficiency, a layer of metallic alloy material with much higher B_s is deposited on the gap faces. This type of head is called the metal-in-gap (MIG) head. Sendust (Fe–Si–Al, 85–9.6–5.4 wt%) is one of the materials used for the deposition. MIG heads are capable of recording up to 100 MHz frequency and 180 kA/m medium coercivity.

Thin-film heads capitalize on semiconductor-like processing technology to reduce the customized fabrication steps for individual heads. The core, coil, gap, and insulator layers are all fabricated by electroplating, sputtering, or evaporation. Due to the nature of the semiconductor process, the fabrication is accurate for small dimensions. Small gap dimensions are suitable for high linear and track density, and small core dimensions allow the use of high B_s Permalloy material (Ni–Fe, 80–20 wt%) as core with low inductance for high data rate applications. Thin-film heads are used for high medium H_c, high areal density applications. The high cost of the semiconductor-like process is offset by high throughput: a 150 × 150 mm wafer can produce 16,000 nanoslider heads. One disadvantage is the limited-band recording capability because the small pole length limits low-frequency response and introduces undershoots. A second disadvantage, the Barkhausen noise, is caused by the relatively small number of magnetic domains in the core. At present, thin-film heads are used up to frequencies of 80 MHz and medium coercivity of 200 kA/m. MIG thin-film heads are also being used for high-coercivity applications.

An inductive head is often used for both recording and reproducing. The optimal performance cannot be achieved because recording and reproducing have contradictory requirements for head design. To solve this problem, the MR head has been developed. The MR head is for reproducing only, and an inductive head is used for recording. As schematically shown in Figure 96.4, an MR head has a magnetoresistive element (MRE) and two electric leads. The MRE is a Permalloy stripe (Ni–Fe, 80–20 wt%), with thickness t, width w, and height h. An electric current, with density J, passes through the MRE through the leads. The electric resistivity of the MRE is a function of the angle θ between J and MRE magnetization M:

$$\rho_\theta = \rho\left(1 + \frac{\Delta\rho}{\rho}\cos^2\theta\right) \tag{96.2}$$

where

ρ_θ is the resistivity at θ
ρ is the resistivity at $\theta = 90°$
$\Delta\rho$ is the resistivity difference between $\theta = 0°$ and $\theta = 90°$
$\Delta\rho/\rho$ is the anisotropic MR ratio (AMR) of the MRE

Usually, a transverse magnetic field is applied to the MRE so that $\theta = \theta_0$ when the head is not reading a magnetic signal. Assume that a magnetic signal from the medium rotates M from θ_0 to θ, then an electric voltage change v will be detected across the MRE as the output signal:

$$v = Jw\rho\frac{\Delta\rho}{\rho}(\sin^2\theta_0 - \sin^2\theta) \tag{96.3}$$

where θ_0 is the bias angle and is set to 45° for good linearity. In practice, a longitudinal bias is also used along the MRE width direction to stabilize the magnetic domain and reduce large Barkhausen noise. To compare the output between an MR head and an inductive head, we write the inductive head output using Faraday's equation:

$$v = -nV\frac{d\phi}{dx} \tag{96.4}$$

where

n is the number of the coil turns
V is the medium velocity
ϕ is the magnetic flux
x is the coordinate axis fixed on the medium surface

Equations 96.3 and 96.4 tell us that while inductive head output is proportional to medium velocity and not suitable for low-velocity applications, the MR head can be used for either high- or low-velocity applications.

96.2.4 Recording Process

We can imagine the recording process in two steps. First, the magnetic flux flowing in the head core generates a fringe magnetic field around the gap. Then the magnetic field magnetizes the magnetic medium and leaves a magnetization transition in it. Partly due to the nonlinear nature of the hysteresis loop, the recording process is so complex that there has been no rigorous explanation. However, we can still obtain significant insights into the recording process by using simple models if we keep in mind the limitations.

If we set the origin of a coordinate system at the center of the gap with x axis on the head surface and y axis pointing away from the head, then the longitudinal magnetic field H_x and perpendicular magnetic field H_y of this head can be expressed by the Karlqvist approximation [2]:

$$H_x = \frac{ni}{\pi(g + lA_g/\mu A_c)}\tan^{-1}\left[\frac{yg}{x^2 + y^2 - (g^2/4)}\right] \tag{96.5}$$

$$H_y = \frac{ni}{2\pi(g + lA_g/\mu A_c)}\ln\left[\frac{(x - g/2)^2 + y^2}{(x + g/2)^2 + y^2}\right] \tag{96.6}$$

where

n is the number of coil turns
i is the current in the coil
g is the gap length
l is the core length
A_g is the core cross-sectional area
μ is the relative permeability of core material
A_c is the gap cross-sectional area

Both Equations 96.5 and 96.6 give accurate results for points $0.25g$ away from gap corners. Since longitudinal recording mode dominates the magnetic recording industry, we will focus on the field H_x. Equation 96.5 shows that the contours of constant H_x field are circles nesting on the two gap corners, as shown in Figure 96.5. The greater the diameter of the circle is the weaker the magnetic field. Assume a magnetic medium, moving from left to right with a distance d above the head, has a thickness δ and a magnetization M pointing to right. At some instant, the recording current turns on and generates

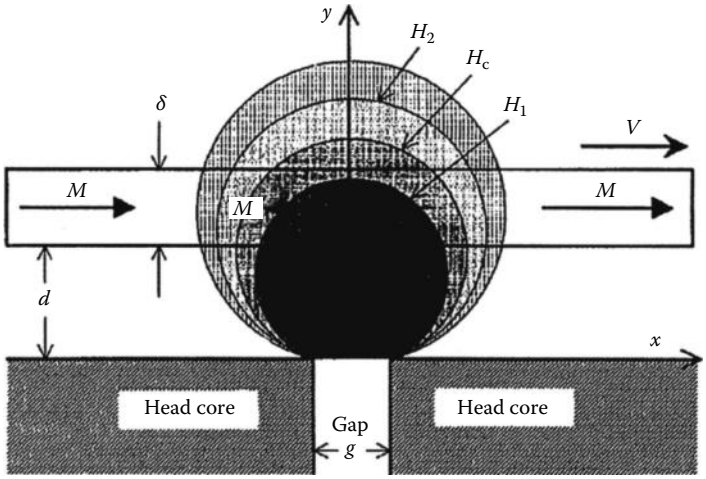


FIGURE 96.5 The constant horizontal fields of the Karlqvist approximation are circles resting on the gap corners of a head, and the change of magnetization in the medium is gradual.

the magnetic field H_x above the gap as depicted in Figure 96.6. On the circumference of $H_x = H_c$, half of medium material has its magnetization reversed and half remains the same, resulting in a zero total magnetization. Since H_x has a gradient, the medium closer to the gap (inside a smaller circle) gets its magnetization reversed more completely than the medium farther away from the gap (outside a bigger circle). Therefore, magnetic transition is gradual in the medium even if the change of recording current follows a step function. Assume the original magnetization is M_r and the completely reversed magnetization is $-M_r$, this gradual change of magnetization for an isolated transition can be modeled by [3]:

$$M = \frac{2}{\pi} M_r \tan^{-1} \frac{x}{a} \tag{96.7}$$

where

x is the distance from the center of transition

a is a parameter characterizing the sharpness of the transition as shown in Figure 96.6

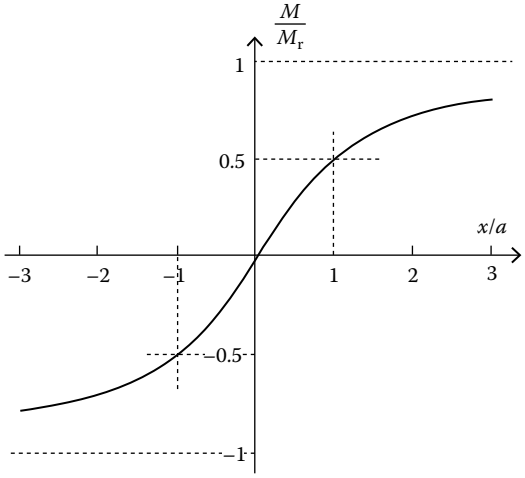


FIGURE 96.6 An isolated arctangent magnetization transition from negative M_r to positive M_r.

Assuming a thin-film medium and using the Karlqvist approximation for the head field, a is found to be [4–6]

$$a = \frac{(1-S^*)(d+\delta/2)}{\pi Q} + \sqrt{\left[\frac{(1-S^*)(d+\delta/2)}{\pi Q}\right]^2 + \frac{M_r\delta(d+\delta/2)}{\pi Q H_c}} \tag{96.8}$$

where

S^* is the medium loop squareness
Q is the head-field gradient factor

For a reasonably well-designed head, $Q \approx 0.8$. It is obvious that we want to make parameter a as small as possible so that we can record more transitions for a unit medium length. If the head gap length g and medium thickness δ are small compared with head/medium separation d, and the medium has a squareness of 1, then the minimum possible value of a is [7]

$$a_m = \begin{cases} \dfrac{M_r\delta}{2\pi H_c} & \dfrac{M_r\delta}{4\pi H_c d} \geq 1 \\ \sqrt{\dfrac{M_r\delta d}{\pi H_c}} & \dfrac{M_r\delta}{4\pi H_c d} < 1 \end{cases} \tag{96.9}$$

In order to decrease the value of a and therefore increase areal density, we need to reduce medium remanence M_r, thickness δ, head/medium separation d, and to increase coercivity H_c.

96.2.5 Reproducing Process

In contrast to the recording process, the reproducing process is well understood. The flux density induced in the head core is on the order of a few millitesla, yielding a linear process for easier mathematical treatment. The head fringe field is the Karlqvist approximation (Equation 96.5) and the foundation is the reciprocity theorem. For an isolated transition (Figure 96.6) with a thin magnetic layer $\delta \ll d$, the induced electric voltage v for an inductive head is [7]

$$v(x) = \frac{-2\mu_0 V w M_r \delta n}{\pi(g + lA_g/\mu A_c)}\left[\tan^{-1}\left(\frac{g/2+x}{a+d}\right) + \tan^{-1}\left(\frac{g/2-x}{a+d}\right)\right] \tag{96.10}$$

where

μ_0 is the permeability of vacuum
μ is the relative permeability of the core
V is the medium velocity
w is the track width
n is the number of coil turns
g is the head gap length
d is the head/medium separation
x is the distance between the center of the medium transition and the center of the head gap

The term $lA_g/\mu A_c$ is closely related to g for head efficiency. When a transition passes under the head, its voltage starts with a very low value, reaches a peak, then falls off again, as shown in Figure 96.7, where the following typical values for a hard disk drive are used: V = 20 m/s, w = 3.5 μm, M_r = 450 kA/m, δ = 50 nm, n = 50, $lA_g/\mu A_c = 0.1g$. The effects of g and $a + d$ are shown in Figure 96.7. Since a greater peak voltage and a narrower spatial response are desired for the reproducing process, smaller g and $a + d$ values are helpful.

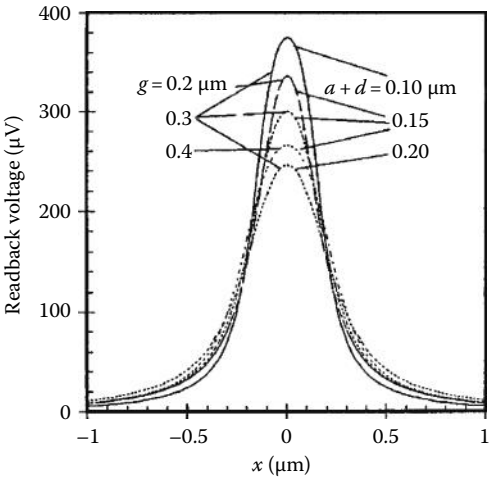


FIGURE 96.7 The reproducing voltage of an inductive head over an isolated arctangent transition shows the effects of gap length *g*, parameter *a*, and head/medium separation *d*.

When an MR head is used for reproducing, the MRE is usually sandwiched between two magnetic shields to increase its spatial resolution to medium signals, as shown in Figure 96.8. Since the MR head is flux-sensitive, the incident flux ϕ_i, on the bottom surface of the MRE, should be derived as a function of the distance (x) between the center of MRE and the center of the transition [7,8]:

$$\phi_i(x) = \frac{2\mu_0 w M_r \delta(a+d)}{\pi g}\left\{ f\left[\frac{x+(g+t)/2}{a+d}\right] - f\left[\frac{(x+t)/2}{a+d}\right] + f\left[\frac{x-(g-t)/2}{a+d}\right] - f\left[\frac{(x-t)/2}{a+d}\right]\right\} \tag{96.11}$$

where

g is the distance between the MRE and the shield

t is the MRE thickness

$$f(\beta) = \beta \tan^{-1}\beta - \ln\sqrt{1+\beta^2} \tag{96.12}$$

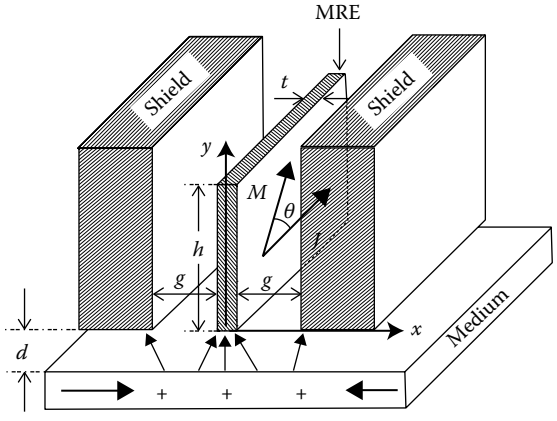


FIGURE 96.8 Schematic diagram of a shielded MR head with a shield to MRE distance *g*.

The angle between magnetization and current varies along the MRE height h. To find out the variation, we need to calculate the signal flux decay as a function of y by

$$\phi_s(y) = \phi_i \frac{\sin h[(h-y)/l_c]}{\sin h(h/l_c)} \tag{96.13}$$

where

$$l_c = \sqrt{\frac{\mu g t}{2}} \tag{96.14}$$

Then the bias angle θ_0 and signal angle θ can be calculated by

$$\sin\theta_0(y) = \frac{\phi_b(y)}{M_s t} \tag{96.15}$$

and

$$\sin\theta(y) = \frac{\phi_s(y) + \phi_b(y)}{M_s t} \tag{96.16}$$

where

ϕ_b is the biasing flux in the MRE

M_s is the saturation magnetization of the MRE

Application of Equations 96.15 and 96.16 to Equation 96.3 and integration over height h lead to

$$v = Jw\rho \frac{\Delta\rho}{\rho} \frac{1}{h} \left[\int_0^h \sin^2\theta_0(y)\,dy - \int_0^h \sin^2\theta(y)\,dy \right] \tag{96.17}$$

For an MR head with a 45° bias at the center and small height $h \ll l_c$, the peak voltage is [6]

$$v_P \approx \frac{9\Delta\rho JwM_r\delta(g+t)}{8\sqrt{2}\ tgM_s} \tan^{-1}\left[\frac{g}{2(a+d)}\right] \tag{96.18}$$

The general shape of the reproducing voltage from an MR head is similar to that in Figure 96.7.

The study of an isolated transition reveals many intrinsic features of the reproducing process. However, transitions are usually recorded closely in a magnetic medium to achieve high linear density. In this case, the magnetization variation in the medium approaches a sinusoidal wave. That is,

$$M(x) = M_r \sin\frac{2\pi}{\lambda}x \tag{96.19}$$

where λ is the wavelength. The reproducing voltage in an inductive head becomes [9,10]

$$v(x) = \frac{-\mu_0 VwM_r ng}{g + lA_g/\mu A_c} (e^{-2\pi d/\lambda})(1 - e^{-2\pi\delta/\lambda}) \left(\frac{\sin(\pi g/\lambda)}{\pi g/\lambda}\right) \cos\frac{2\pi}{\lambda}x \tag{96.20}$$

This equation presents all the important features of the high linear density–reproducing process. The term $\exp(-2\pi d/\lambda)$ is the spacing loss. It shows that the reproducing voltage falls exponentially

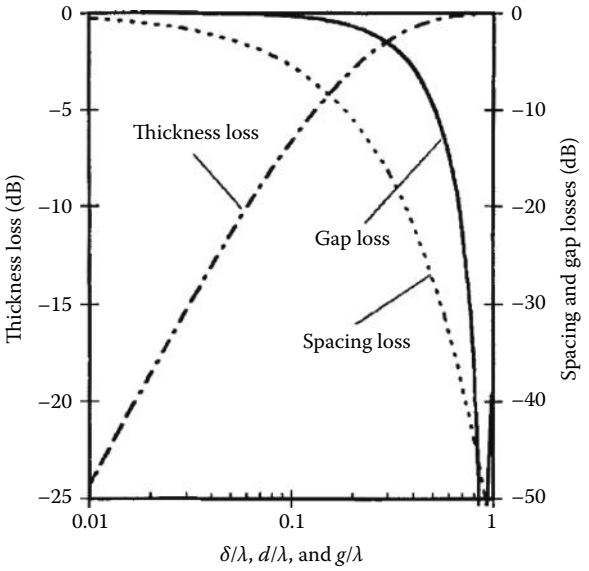


FIGURE 96.9 Spacing, thickness, and gap losses of an inductive head versus frequency for the reproducing of a sinusoidal medium magnetization.

with the ratio of head/medium spacing to wavelength. The second term $1 - \exp(-2\pi\delta/\lambda)$ is the thickness loss. The name of this term is misleading because its value increases with a greater medium thickness. However, the rate of increase diminishes for thicker medium. In fact, 80% of the maximum possible value is achieved by a medium thickness of 0.25λ. The last term $\sin(\pi g/\lambda)/(\pi g/\lambda)$ is the gap loss. This term is based on the Karlqvist approximation. If a more accurate head fringe field is used, this term is modified to $\sin(\pi g/\lambda)/(\pi g/\lambda)\cdot(1.25g^2 - \lambda^2)/(g^2 - \lambda^2)$ [11]. It shows a gap null at $\lambda = 1.12g$, and limits the shortest wavelength producible. These three terms are plotted in Figure 96.9. The most significant loss comes from the spacing loss term, which is 54.6 dB for $d = \lambda$. Therefore, one of the biggest efforts spent on magnetic recording is to reduce the head/medium spacing as much as possible without causing mechanical reliability issues. For an MR head, the reproducing voltage is [11]

$$v \propto \frac{4M_r i\Delta\rho w\lambda}{ht}(e^{-2\pi d/\lambda})(1-e^{-2\pi\delta/\lambda})\left(\frac{\sin(\pi g/\lambda)}{\pi g/\lambda}\right)\sin\frac{\pi(g+t)}{\lambda}\cos\frac{2\pi}{\lambda}x \tag{96.21}$$

96.2.6 Digital versus Analog Recording

Due to the nonlinearity of the hysteresis loop, magnetic recording is intrinsically suitable for digital recording, where only two states (1 and 0) are to be recognized. Many physical quantities, however, are received in analog form before they can be recorded, such as in consumer audio and instrumentation recording. In order to perform such recording, we need to either digitize the information or use the analog recording technique. In the case of digitization, we use an analog-to-digital converter to change a continuous signal into binary numbers. The process can be explained by using the example shown in Figure 96.10. An electric signal V, normalized to the range between 0 and 1, is to be digitized into 3 bits. The signal is sampled at time $t = 1, 2, \ldots, 6$. At each sampling point, the first bit is assigned a 1 if the value of the continuous signal is in the top half (>0.5), otherwise assigned a 0. The second bit is assigned a 1 if the value of the continuous signal is in the top half of each half ($0.25 \le V < 0.5$, or >0.75), otherwise assigned a 0. The third bit is assigned similarly. The first bit is the most significant bit (MSB), and the last bit is the least significant

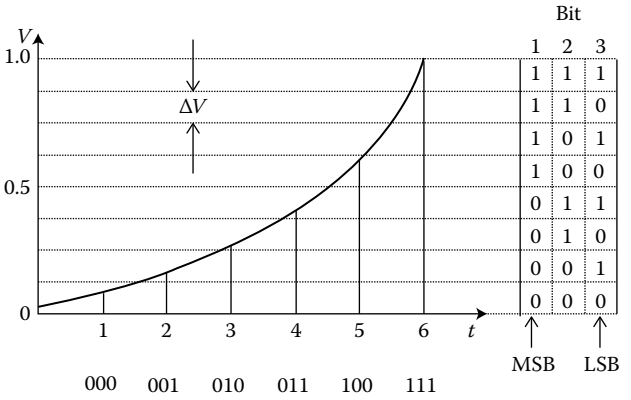


FIGURE 96.10 Schematic illustration of the quantization of a continuous signal to 3 bits.

bit (LSB). The converted binary numbers are listed below each sampling point in Figure 96.10. This process of digitization is termed *quantization*. In general, the final quantization interval is

$$\Delta V = \frac{V}{2^N} \tag{96.22}$$

where

V is the total voltage range
N is the number of bits

Because we use a finite number of bits, statistically there is a difference between the continuous signal and the quantized signal at the sampling points. This is the quantization error. It leads to a signal-to-quantization-noise ratio [12]:

$$\text{SNR} = \frac{12P_s}{\Delta V^2} \tag{96.23}$$

where P_s is the mean square average signal power. For a signal with uniform distribution over its full range, this yields

$$\text{SNR} = 2^{2N} \tag{96.24}$$

For a sinusoidal signal, it changes to

$$\text{SNR} = 1.5 \times 2^{2N} \tag{96.25}$$

The SNR can be improved by using more bits. This improvement, however, is limited by the SNR of the incoming continuous signal. The quantized signal is then pulse code modulated (PCM) for recording.

For analog recording, a linear relationship between the medium magnetization and the recording signal is required. This is achieved through the anhysteretic magnetization process. If we apply an alternating magnetic field and a unidirectional magnetic field to a previously demagnetized medium, and then reduce the amplitude of the alternating field to zero before we remove the unidirectional field, the remanent magnetization shows a pseudo-linear relationship with the unidirectional field strength H_u up to some level. Figure 96.11 shows such an anhysteretic curve. The linearity deteriorates as H_u gets greater. In applications, the recording signal current is biased with an ac current of greater amplitude

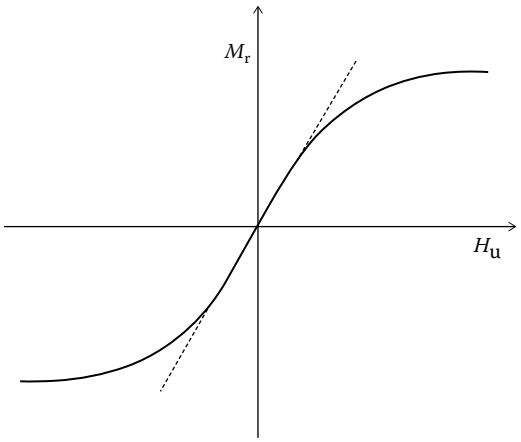


FIGURE 96.11 An anhysteretic remanent magnetization shows a pseudo-linear relationship with the applied unidirectional magnetic field to some H_u level.

and higher frequency. Therefore, it is also termed ac-biased recording. Analog recording is easy to implement, at the price of a lowered SNR because remanent magnetization is limited to about 30% of the maximum possible M_r to achieve good linearity.

96.2.7 Recording Codes

PCM is a scheme of modifying input binary data to make them more suitable for a recording and reproducing channel. These schemes are intended to achieve some of the following goals: (1) reducing the dc component, (2) increasing linear density, (3) providing self-clocking, (4) limiting error propagation, and (5) achieving error-free detection. There is numerous code schemes, only three of the ones developed early are shown in Figure 96.12. The earliest and most straightforward one is the return-to-zero (RZ) code. In this scheme, a positive and negative pulse is used to represent each 1 and 0, respectively, of the data. The main drawback is that direct recording over old data is not possible due to the existence of zero recording current between two data. It also generates two transitions for each bit, therefore reducing the linear density. In addition, it only uses half of the available recording current range for a transition. The non-return-to-zero-invert (NRZI) method was developed to alleviate some of these problems. It changes the recording current from one direction to the other for each 1 of the data, while making no changes for all 0s. However, it has a strong dc component and may lose reproducing synchronization if

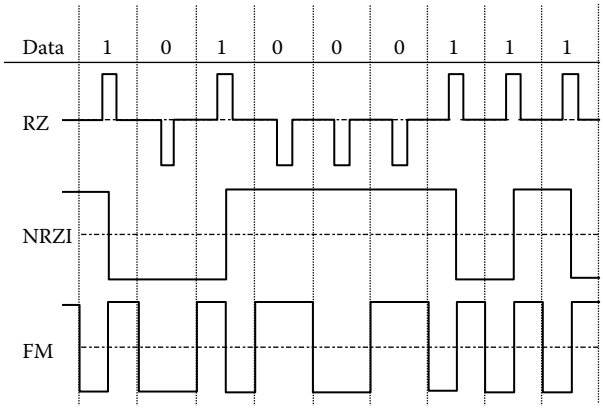


FIGURE 96.12 Comparison of some early developed codes.

TABLE 96.3 ½(2, 7) Code

Before Coding	After Coding
10	0100
11	1000
000	000100
010	100100
011	001000
0010	00100100
0011	00001000

there is a long string of 0s in the input data. In addition, reproducing circuits are usually not designed for dc signal processing. In frequency modulation (FM) code, there is always a transition at the bit-cell boundary which acts as a clock. There is also an additional transition at the bit-cell center for each 1 and no transition for 0s. It reduces the dc component significantly. The primary deficiency is the reduction of linear density since there are two transitions for each 1 in the data.

The most popularly used codes for magnetic recording are the run length–limited (RLL) codes. They have the general form of $m/n(d, k)$. In these schemes, data are encoded in groups. Each input group has m bits. After encoding, each group contains n bits. In some schemes, multiple groups are coded together. d and k are the minimum and maximum 0s, respectively, between two consecutive 1s in the encoded sequence. While d is used to limit the highest transition density and intersymbol interference, k is employed to ensure adequate transition frequency for reproducing clock synchronization. The encoding is carried out by using a lookup table, such as Table 96.3 for a ½(2, 7) code [13].

96.2.8 Head/Medium Interface Tribology

As expressed in Equation 96.20, the most effective way to increase signal amplitude, therefore areal density, is to reduce head/medium spacing d. However, wear occurs when a moving surface is in close proximity to another surface. The amount of wear is described by Archard's equation:

$$V = k\frac{WS}{H} \tag{96.26}$$

where

V is the volume worn away
W is the normal load
S is the sliding distance
H is the hardness of the surface being worn away
k is a wear coefficient

In order to increase medium hardness H, hard Al_2O_3 particles are dispersed in particulate media and a thin layer of hard carbon (≈10 nm) with either hydrogenation or natrogenation is coated on thin-film media of hard disks. A layer of liquid lubricant, typically perfluoropolyethers with various end groups and additives, is applied on top of the carbon film to reduce the wear coefficient k. Load is minimized to reduce wear while keeping adequate head/medium dynamic stability. For applications where the sliding distance s is modest over the lifetime of the products such as floppy disk drives and consumer tapes drives, the head physically contacts the medium during operations. In the case of hard disk application, heads are separated nominally from the media by a layer of air cushion. The head is carried on a slider, and the slider uses air-bearing surfaces (ABS) to create the air film based on hydrodynamic lubrication theory. Figure 96.13 shows two commonly used ABS. Tapers are used to help the slider takeoff

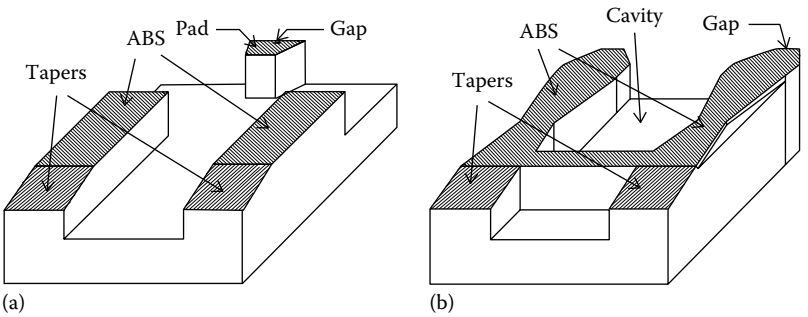


FIGURE 96.13 The ABS of (a) a tripad slider for pseudo-contact recording and (b) a SAP slider for conventional flying recording.

and maintain flying stability. ABS generates higher-than-ambient pressure to lift the slider above the medium surface during operations. The tripad slider is for pseudo-contact applications while the sub-ambient pressure (SAP) slider is for flying (such as MR head) applications. Because the relative linear velocity between the slider and the medium changes when the head moves to different radii, a cavity region is used in the SAP slider to generate suction force to reduce flying height variation. The ABS is designed based on the modified Reynolds equation:

$$\frac{\partial}{\partial X}\left(PH^3Q\frac{\partial P}{\partial X}\right)+\frac{\partial}{\partial Y}\left(PH^3Q\frac{\partial P}{\partial Y}\right)=\Lambda_x\frac{\partial(PH)}{\partial X}+\Lambda_y\frac{\partial(PH)}{\partial Y}+\sigma\frac{\partial(PH)}{\partial T} \tag{96.27}$$

where

X and Y are coordinates in the slider longitudinal and transverse directions normalized by the slider length and width, respectively
P is the hydrodynamic pressure normalized by the ambient pressure
H is the distance between the ABS and medium surface normalized by the minimum flying height
Q is the molecular slip factor
T is time normalized by the characteristic squeeze period
Λ_x and Λ_y are the bearing numbers in the x and y directions, respectively
σ is the squeeze number

A full derivation and explanation of the Reynolds equation can be found in Ref. [14]. At present, high end hard disk drives feature a flying height on the order of 20–50 nm.

When power is turned off, the slider in the popularly used Winchester-type drives rests on the medium surface. Although the ABS and medium surface look flat and smooth, they really consist of microscopic peaks and valleys. If we model an ABS/medium contact by a flat surface and a sphere tip, the liquid lubricant on the medium surface causes a meniscus force F_m as depicted in Figure 96.14 [15]:

$$F_m=\frac{4\pi R\gamma\cos\theta}{1+y/(h-y)} \tag{96.28}$$

where

R is the radius of the sphere
γ is the surface tension of the lubricant
θ is the contact angle between the lubricant and the surfaces
y is the sphere to flat surface distance
h is the lubricant thickness

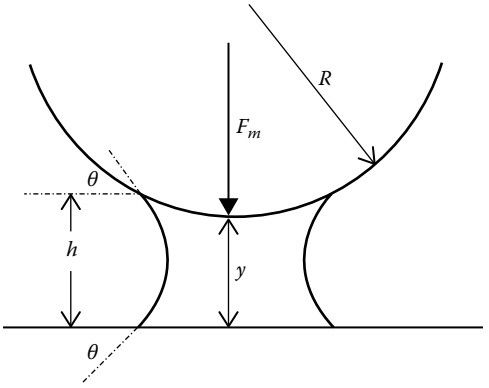


FIGURE 96.14 Formation of meniscus between a sphere tip and a flat surface.

Detailed analysis [16] shows that the static friction F at a head/medium interface is a function of several parameters:

$$F = f(h, R, A, \eta, \gamma, \theta, E', \phi, \sigma, s) \tag{96.29}$$

where

A is the ABS area
η is the peak density
E' is the effective modulus of elasticity
ϕ is the peak height distribution
σ is the rms peak roughness
s is the solid-to-solid shear strength

If friction F is too large, either the drive cannot be started or the head/medium interface is damaged. While friction can be reduced practically by reducing A, γ, and increasing θ, the most effective ways are to control h, σ, η, and ϕ. Too thin a lubricant layer will cause wear, and too thick will induce high friction. This limits h to the range of 1–3 nm. σ is controlled by surface texture. Historically, texture is created by mechanical techniques using either free or fixed abrasives. This leaves a surface with a random feature and is unsuitable for controlling η and ϕ. Recently, people started to use lasers [17]. This technique generates a surface texture with well-defined η and ϕ to improve wear and friction performance. Figure 96.15 shows AFM images of a mechanical and a laser texture.

96.3 Optical Recording

The major obstacle to achieving higher areal density in magnetic recording is the spacing loss term. It is a great engineering challenge to keep heads and media in close proximity while maintaining the head/medium interface reliable and durable. Care must also be taken in handling magnetic media since even minute contamination or scratches can destroy the recorded information. In addition, the servo technique of using magnetic patterns limits the track density to about one order lower than the linear density. Optical recording, on the other hand, promises to address all these concerns.

Optical recording can be categorized into three groups. In the first group, information is stored in the media during manufacturing. Users can reproduce the information, but cannot change or record new information. CD-ROM (compact disk–read only memory) belongs to this group. The second group is WORM (write once read many times). Instead of recording information during manufacturing, it leaves this step to the user. This is usually achieved by creating physical holes or blisters in the media during

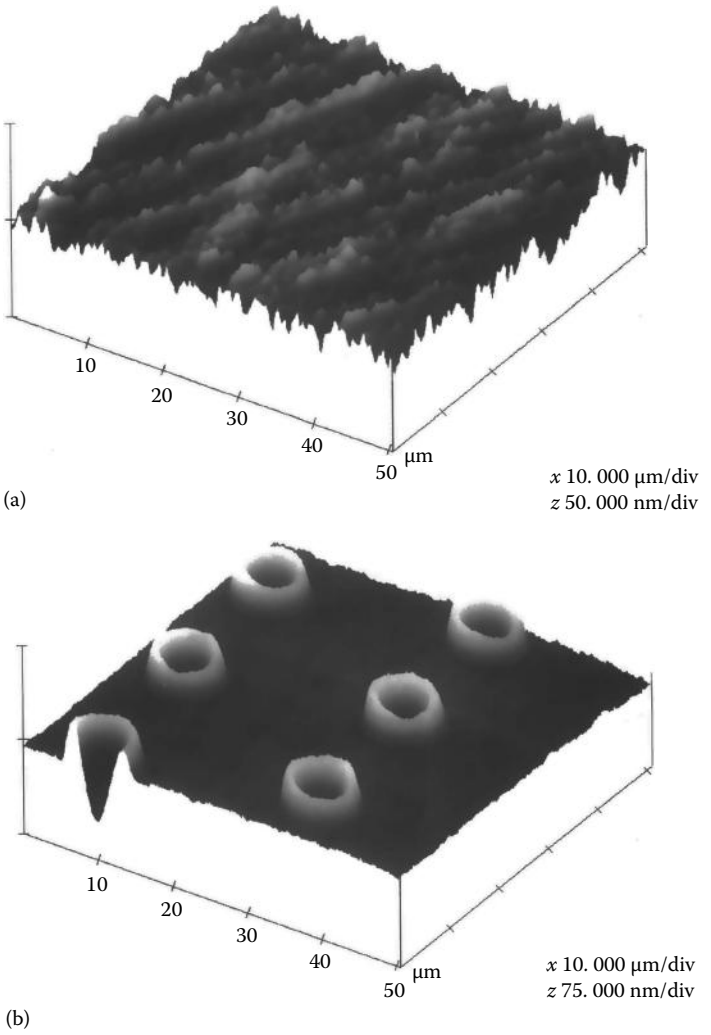


FIGURE 96.15 AFM images of (a) a mechanical texture and (b) a laser texture. (Courtesy of J. Xuan.)

the recording process. Once it is recorded, however, the medium behaves like the first group: no further recording is possible. The third group is similar to magnetic recording. Recording can be performed infinitely on the media by changing phase or magnetization of the media. The most noticeable example in this group is the magneto-optic (MO) technique [18]. Only CD-ROM and the MO recording are described in the following sections.

96.3.1 CD-ROM

Figure 96.16 shows the CD-ROM reproducing principle. Data are pressed as physical pits and lands on one surface of a plastic substrate, usually polycarbonate. A layer of aluminum is deposited on this surface to yield it reflective. Lacquer is then coated to protect the aluminum layer. During the reproducing process, an optical lens is used to focus a laser beam on the reflective pit and land surface. The diameter of the lens is D, the distance between the lens and the substrate is h_3, and the substrate thickness is h_2. The diameter of the laser beam is d_2 when entering the substrate, and becomes d_1 when focused on the reflective surface. The width of the pits is designed smaller than d_1. The reflected light consists of two

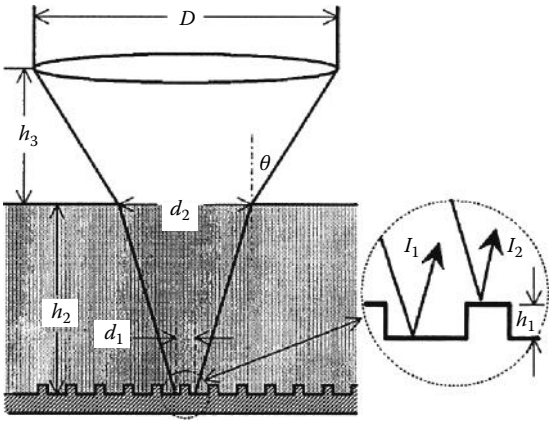


FIGURE 96.16 Schematic representation of the CD-ROM reproducing principle.

portions: I_1 from the land and I_2 from the pit. According to the theory of interference, the intensity of the reflected light is

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\frac{4\pi h_1}{\lambda} \tag{96.30}$$

where

λ is the wavelength of the laser

$h_1 \approx \lambda/4$ is the pit height

This leads to

$$I = \begin{cases} I_1 + I_2 - 2\sqrt{I_1 I_2} & \text{if there is a pit } (h_1 = \lambda/4) \\ I_1 + I_2 + 2\sqrt{I_1 I_2} & \text{if there is no pit } (h_1 = 0) \end{cases} \tag{96.31}$$

This change of light intensity is detected and decoded to yield the recorded data. The reflected light is also used for focusing and track following.

The fundamental limit on optical recording density is the focused beam diameter d_1. For a Gaussian (TEM_{00}) laser, this is the diffraction-limited diameter at which the light intensity is reduced to $1/e^2$ of the peak intensity:

$$d_1 \approx \frac{2\lambda}{\pi\theta} \tag{96.32}$$

where θ is the aperture angle. The following values are typical for a CD-ROM system: λ (gallium arsenide laser) = 780 nm, $\theta = 27°$, $h_2 = 1.2$ mm, $D = 5$ mm, $h_3 = 4.2$ mm. This yields $d_1 \approx 1.0$ μm and $d_2 \approx 0.7$ mm, and sets the areal density limit of optical (including MO) recording to about 1 Mbit/mm^2. For most CD-ROM applications, the areal density is smaller than this limit, and a disk with 120 mm diameter holds about 600 Mbyte information. In order to increase areal density, we can either reduce light wavelength or increase numerical aperture. Much of the effort has been to adopt a new light source with short wavelength such as a blue laser. Increasing numerical aperture is more difficult because increasing lens diameter is cost prohibitive and reducing h_2 or h_3 is reliability limited. Note that although the beam

size on the focus plane is on the order of 1 μm (d_1), it is two to three orders greater at the air/substrate interface (d_2). This means that optical recording can tolerate disk surface contamination and scratches much better than magnetic recording. However, the performance of optical recording does not match magnetic recording in general. The data transfer rate of CD-ROM drives is expressed as multiple (×) of 150 kB/s. Even for a 12× CD-ROM drive, the data access time and data transfer rate are still on the order of 100 ms and 1.8 MB/s, respectively, while for a high-performance rigid disk drive these values are less than 10 ms and greater than 30 MB/s, respectively.

96.3.2 Magneto-optic Recording

The primary drawback of a CD-ROM to an end user is its inability to record. This deficiency is remedied by MO recording technology, as depicted in Figure 96.17. A linearly polarized laser beam is focused on a layer of magnetic material, and a coil provides a dc magnetic field on the other side of the medium. This dc magnetic field is too weak to affect the medium magnetization at normal temperature. The recording process utilizes the thermomagnetic property of the medium, and the reproducing process is achieved by using the Kerr effect. During recording, the medium is initially magnetized vertically in one direction, and the dc magnetic field is in the opposite direction. The laser heats up the medium to its curie temperature, at which the coercivity becomes zero. During the cooling process, the dc magnetic field aligns the medium magnetization of the heated region to the magnetic field direction. In the process of reproducing, the same laser is used with a smaller intensity. The medium is heated up to its compensation temperature, at which the coercivity becomes extremely high. Depending on the direction of the magnetization, the polarization of the reflected light is rotated either clockwise or counterclockwise (Kerr rotation). This rotation of polarization is detected and decoded to get the data. The main disadvantage of MO recording is that a separate erasing process is needed to magnetize the medium in one direction before recording. Recently, some technologies have been developed to eliminate this separate erasing process at the cost of complexity.

The medium used in MO recording must have a reasonable low curie temperature (<300°C). The materials having this property are rare earth transition metal alloys, such as $Tb_{23}Fe_{77}$ and $Tb_{21}Co_{79}$. Unfortunately, the properties of these materials deteriorate in an oxygen and moisture environment. To protect them from air and humidity, they are sandwiched between an overlayer and an underlayer, such as SiO, AlN, SiN, and TiO_2. Another issue with the rare earth transition metal alloys is their small Kerr rotation, about 0.3°. To increase this Kerr rotation, multiple layers are used. In the so-called quadrilayer structure (Figure 96.17b), the overlayer is about a half-wavelength thick and the underlayer is about a quarter-wavelength thick [18]. The MO layer is very thin (≈3 nm). Light reflected from the reflector is out of phase with the light reflected from the surface of the MO layer, and is in-phase with the light reflected from the inside of the MO layer. As a result, the effective Kerr rotation is increased several times.

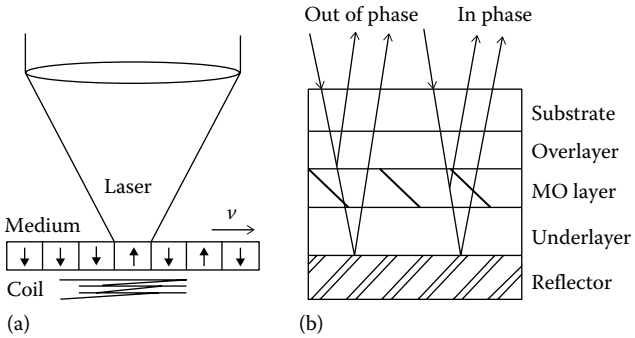


FIGURE 96.17 Schematic illustrations of (a) MO recording/reproducing and (b) quadrilayer medium cross section.

TABLE 96.4 Digital Magnetic and Optical Storage Devices

Description	Manufacturers	Approximate Price, $
Thin-film head for hard disk drive	AMC, Read-Rite, SAE	6.00–9.00
MR head for hard disk drive	AMC, Read-Rite, SAE, Seagate	8.00–12.00
Thin-film hard disk	Akashic, HMT, Komag, MCC, Stormedia	7.00–10.00
Hard disk drive	IBM, Maxtor, Quantum, Samsung, Seagate, WD	0.02–0.20/Mbytes
Floppy drive	Panasonic, Sony	20.00–40.00
Floppy disk	3M, Fuji, Memorex, Sony	0.15–0.50
Removable rigid disk drive	Iomega, Syquest	100.00–400.00
Removable rigid disk	Iomega, Maxell, Sony	5.00–20.00/100 Mbytes
Tape drive	Exabyte, HP, Seagate	100.00–400.00
Backup tape	3M, Sony, Verbatim	4.00–25.00/Gbytes
8 × CD-ROM drive	Goldstar, Panasonic	100.00–200.00
Recordable optical drive	JVC, Philips	300.00–500.00
Recordable optical disk	3M, Maxell, Memorex	3.00–15.00/650 Mbytes

Compared with magnetic recording, optical recording has the intrinsic advantages of superior reliability and portability. However, its performance is inferior due to slower data access time and transfer rate. Another advantage of optical recording, higher areal density, has been disappearing or even reversing to magnetic recording. Both magnetic and optical recording will be continuously improved in the near future, probably toward different applications. Currently, there are some emerging techniques that try to combine the magnetic and optical recording techniques [19], such as amorphous rare earth materials. Table 96.4 is a short list of representative magnetic and optical devices for digital recording.

References

1. C.D. Mee and E.D. Daniel, Eds., *Magnetic Storage Handbook*, 2nd ed., New York: McGraw-Hill, 1996.
2. O. Karlqvist, Calculation of the magnetic field in the ferromagnetic layer of a magnetic drum, *Trans. R. lnst. Technol. (Stockholm)*, 86, 3–28, 1954.
3. J.J. Miyata and R.R. Tartel, The recording and reproduction of signals on magnetic medium using saturation-type recording, *IRE Trans. Elec. Comp*., EC-8, 159–169, 1959.
4. M.L. Williams and R.L. Comstock, An analytical model of the write process in digital magnetic recording, *A.I.P. Conf. Proc*., 5, 738–742, 1972.
5. V.A.J. Maller and B.K. Middleton, A simplified model of the writing process in saturation magnetic recording, *IERE Conf. Proc*., 26, 137–147, 1973.
6. H.N. Bertram, *Theory of Magnetic Recording*, Cambridge, U.K.: Cambridge University Press, 1994.
7. C.D. Mee and E.D. Daniel, Eds., *Magnetic Recording Technology*, 2nd ed., New York: McGraw-Hill, 1996.
8. R. Potter, Digital magnetic recording theory, *IEEE Trans. Magn*., MAG-10, 502–508, 1974.
9. W.K. Westmijze, Studies on magnetic recording, *Philips Res. Rep*., 8, 161–183, 1953.
10. G.J. Fan, A study of the playback process of a magnetic ring head, *IBM J. Res. Dev*., 5, 321–325, 1961.
11. J.C. Mallinson, *The Foundations of Magnetic Recording*, 2nd ed., San Diego, CA: Academic Press, 1993.
12. M.S. Roden, *Analog and Digital Communication Systems*, 4th ed., Upper Saddle River, NJ: Prentice Hall, 1996.
13. P.A. Franaszek, Run-length-limited variable length coding with error propagation limitation, U.S. Patent No. 3, 689, 899, 1972.
14. W.A. Gross, L. Matsch, V. Castelli, A. Eshel, T. Vohr, and M. Wilamann, *Fluid Film Lubrication*, New York: Wiley, 1980.

15. J.N. Israelachvili, *Intermolecular and Surface Forces*, London, U.K.: Academic Press, 1985, 224.
16. Y. Li and F.E. Talke, A model for the effect of humidity on stiction of the head/disk interface, *Tribol. Mech. Magn. Storage Syst*., SP-27, 79–84, 1990.
17. R. Ranjan, D.N. Lambeth, M. Tromel, P. Goglia, and Y. Li, Laser texturing for low-flying-height media, *J. Appl. Phys*., 68, 5745–5747, 1991.
18. J. Watkinson, *The Art of Data Recording*, Oxford, U.K.: Focal Press, 1994.
19. A. Lyberatos, J. Earl, and R.W. Chantrell, Model of thermally activated magnetization reversal in thin films of amorphous rare-earth-transition-metal alloys, *Phys. Rev*., 53(9), 1580–1587, 1996.

97

Liquid Crystal Displays

Kalluri R. Sarma
Honeywell, Inc.

97.1 Introduction

Liquid crystals (LC) are an important class of materials with applications ranging from display devices, television screens, optoelectronic devices, sensors, and biological and structural materials. The focus of this chapter will be on LCs for display applications. In general, most substances have a single melting point where a solid possessing a positional and orientational order changes upon melting to an isotropic liquid that has neither positional nor orientational order. However, some materials when melted from the solid state change into a cloudy liquid with orientational order at one temperature, and upon further heating change into an isotropic liquid that has no order, as shown in Figure 97.1. Thus, an LC is a mesophase existing between the melting temperatures, T_m, of a crystalline phase, and clearing point T_c, of the liquid phase, i.e., below T_m, the material has a crystalline phase, above T_c, it has a liquid (isotropic) phase, and between T_m and T_c, it has an LC phase. This type of LC in which the mesophase is defined by the temperature (between T_m and T_c) is called a *thermotropic* LC. When the mesophase is defined by a solvent concentration, it is called a *lyotropic* LC. Thermotropic LCs are used for display applications. The orientational order in LC materials results in important physical properties, such as birefringence, that make these materials useful for display devices. Because LCs have the attributes of low drive voltage, low power consumption, thin form factor (flat panel displays), lightweight, full color, gray scale with a wide dynamic range, full motion video, superior image quality, and high reliability, LC displays (LCDs) are the preferred approach for battery-powered (portable) applications ranging from wristwatch displays and handheld TVs to laptop computer displays. They are also replacing cathode ray tubes (CRTs) in select applications such as avionic displays because of their high brightness and readability in sunlight.

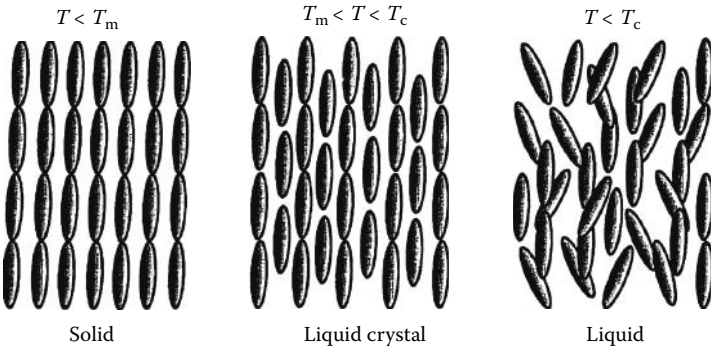


FIGURE 97.1 Illustration of a solid, an LC, and a liquid. A solid has an orientational as well as a positional order for the molecules. An LC has an orientational order only. A liquid phase is isotropic with neither positional nor orientational order.

LCDs are also being used in projection display devices for head-mounted display (HMD) and for large-screen display applications. They are widely accepted and used in computer monitors, televisions, instrument panels, vehicle dashboards, aircraft cockpits, clocks, watches, calculators, gaming devices, telephones, consumer goods, Global Positioning Systems (GPS), and many more applications. The following will discuss the various types of LC materials, their properties, LCD materials and fabrication processes, various LCD modes, and display addressing methods. There are many good general references, for example, Refs. [1–7], on LCs and LCDs. At the time of this writing, LCD technology is advancing very rapidly with respect to technology development for LCD products with improved viewing angles, improved image quality, lower power consumption, and larger display sizes. The purpose of this chapter, however, is to present the basic LCD principles and technologies, as opposed to reviewing the current state of the art.

97.2 Types of LC Materials

Most of the LC materials are organic compounds which consist of rod-shaped or disk-shaped molecules. For display applications, LC materials with rod-shaped molecules are the most commonly used. The LC materials are broadly classified into three types (phases)—smectic, nematic, and cholesteric—according to their molecular order, as shown in Figure 97.2. In a smectic LC, the rod-shaped molecules are arranged in layers with molecules parallel to each other. There are many different smectic phases, but smectic A and smectic C are the most common. In the smectic A LC the molecular axis (director) is perpendicular to the layers as shown in Figure 97.2a, and in smectic C it is tilted at an angle from the layer normal as shown in Figure 97.2b. Also, in the nematic LC, the rod-shaped molecules are parallel to each other, but the individual molecules move relatively easily in the direction along their axis

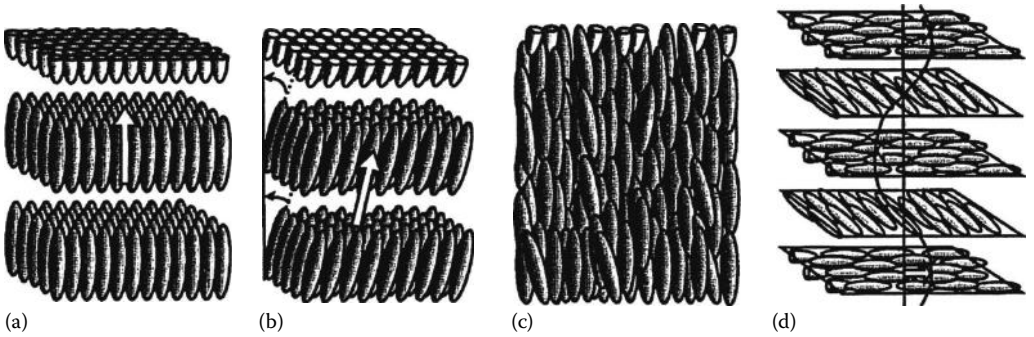
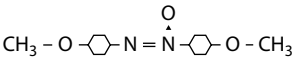


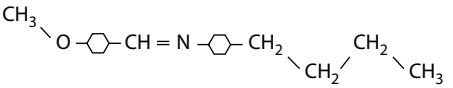
FIGURE 97.2 Molecular orientation in (a) smectic A, (b) smectic C, (c) nematic, and (d) cholesteric LC phases.

without a layer structure, as shown in Figure 97.2c. In the cholesteric LC, the molecules are arranged in a layered fashion as in smectic LC, but the molecular axis is in the plane of each layer as shown in Figure 97.2d. In addition, the cholesteric LC shows a helical structure in which the director ***n*** changes from layer to layer. The same LC material may have different LC phases at different temperatures. For example, the LC material may have a smectic C phase at a lower temperature, and as the temperature increases it may change to a smectic A phase and then to a nematic phase, before changing to an isotropic liquid phase at T_c.

The nematic LC is the basis for most widely used active matrix (AM)–addressed twisted nematic (TN) LCDs, and passive matrix (PM)–addressed supertwisted nematic (STN) LCDs. An example of a smectic C LCD is a PM-addressed ferroelectric LCD [8]. An example of a cholesteric display is a PM-addressed stabilized cholesteric texture (SCT) display with bistability [9]. A classic example of a nematic LC material is *p*-azoxyanisole (PAA) with a nematic phase in the range of 117 °C–136 °C.



In PAA, the two benzene rings are nearly coplanar and the rigid rod is about 20 Å long and 5 Å wide. Another historical example of a nematic LC is *N*-(*p*-methoxybenzylidene-*p*-butylaniline) (MBBA) with a nematic phase in the range of 22 °C–47 °C.



MBBA has a central group that connects the two ringlike cores firmly and serves to maintain the linearity of the entire LC molecule. The terminal groups and conjugated bonds in the core are largely responsible for the dielectric, optical, and other anisotropic properties of the material. Azoxy and Schiff's base compounds were among the materials used earliest in LCDs. Because of environmental stability problems, they were replaced by biphenyl materials. More recently, phenylcyclohexane, bicyclohexane, and estercyclohexane compounds are developed to satisfy the requirements of broad temperature operation and enhanced electro-optical characteristics. The LC materials used in current LCDs are highly developed mixtures of various compounds tailored to meet the requirements of environmental stability, wide operating temperature range, proper response to the applied electric field, high electrical resistivity for matrix addressing, and fast response time.

The degree of order in the LC is an important parameter and is defined by the *order parameter*, *S*, given by

$$S = \frac{1}{2}\langle 3\cos^2\theta - 1\rangle \tag{97.1}$$

where

θ is the angle between the molecular axis and the predominant molecular orientation ***n***

$\langle\rangle$ represents averaging over the whole space

The predominant molecular orientation, which is known as the LC *director* ***n***, is defined as the average alignment direction of the long molecular axis. For a perfect orientational order, i.e., when all the molecules align parallel to the director, as in a perfect crystal, $\theta = 0$ and thus $S = 1$. For no orientational order, i.e., for a completely random molecular orientation, as in an isotropic liquid, $S = 0$. In a typical nematic LC S is in the range of 0.3–0.7, and in a typical smectic LC it is in the range of 0.7–0.8, with higher values at lower temperatures. Figure 97.3 shows the temperature dependence of S for the LC material PAA [10]

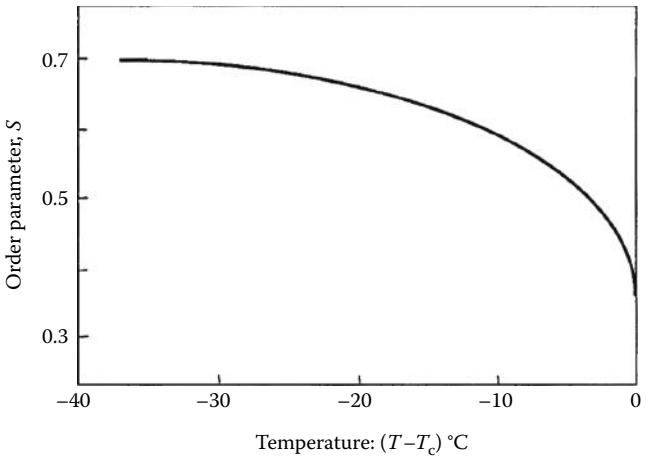


FIGURE 97.3 Temperature dependence of the order parameter S of the nematic LC P = azoxy anisole (PPA) [10]. S decreases with increasing temperature T, and rapidly approaches 0 near the clearing temperature, T_c.

as an example. The order parameter decreases rapidly to a value of around 0.3, close to the clearing point T_c, and becomes zero in the isotropic state. The exact dependence of S on temperature T depends on the type of molecules considered. However, the following analytical expression derived from theory has been shown to be useful [11]:

$$S = \left(1 - y \cdot \frac{T}{T_c}\right)^{\beta} \tag{97.2}$$

where

T_c corresponds to nematic-isotropic transition
y is of the order of 0.98
β is an exponent in the range of 0.13–0.18 depending on the material in question

97.3 Physical Properties of LCs

Because of ordered structure with the molecules aligned with their long axis parallel to each other, LC molecules exhibit anisotropic properties. That is, various physical properties, such as dielectric constant ε, refractive index n, magnetic susceptibility χ, conductivity σ, and viscosity η, have different values in the direction parallel (||) and perpendicular ($\perp$) to the molecular axis. The anisotropic physical properties, in conjunction with the ease of controlling the initial orientation (boundary condition) by surface alignment and the ease of reorienting the molecular axis by applying a voltage, is the basis for application of LCs for displays. The following will discuss the anisotropy of the dielectric constant and refractive index, elastic constants, and electro-optic characteristics which are important in the use of LCs for displays.

97.3.1 Dielectric Anisotropy

LC molecules exhibit dielectric anisotropy because of their permanent and induced dipoles. The dielectric anisotropy, $\Delta\varepsilon$, is expressed as

$$\Delta\varepsilon = \varepsilon_{\parallel} - \varepsilon_{\perp} \tag{97.3}$$

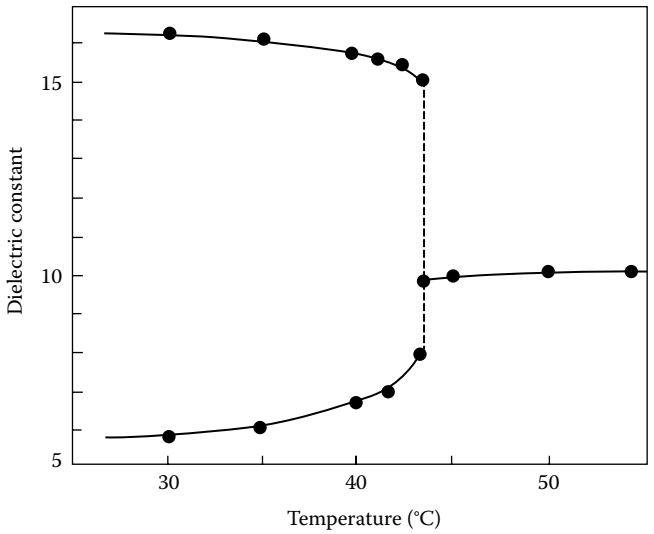


FIGURE 97.4 Temperature dependence of dielectric constant of the nematic LC cyanobiphenyl at a frequency of 100 kHz [13] exhibiting positive dielectric anisotropy ($\varepsilon_{||} > \varepsilon_{\perp}$). $\varepsilon_{||}$ decreases with temperature, whereas $\varepsilon_{\perp}$ increases with temperature until both are equal at the clearing temperature corresponding to the isotropic liquid.

where $\varepsilon_{||}$ and $\varepsilon_{\perp}$ are the dielectric constants measured parallel and perpendicular to the LC director. Materials that exhibit positive dielectric anisotropy ($\Delta\varepsilon > 0$) are referred to as *p*-type materials and the materials that exhibit negative dielectric anisotropy ($\Delta\varepsilon < 0$) are referred to as *n*-type materials. The *p*-type LC materials tend to align themselves with their molecular axis parallel to the applied electric field, whereas *n*-type materials align themselves with their molecular axis perpendicular to the applied field. Generally, the dielectric constant $\varepsilon_{||}$ decreases with increasing frequency [12] due to the relaxation phenomenon. However, $\varepsilon_{\perp}$ is independent of frequency over a large range of frequencies. At the crossover frequency, f_c, where $\varepsilon_{||} = \varepsilon_{\perp}$, the LC material becomes isotropic. Depending on the material, this frequency, f_c, falls in the range of 100 kHz to >1 MHz. The dielectric constants $\varepsilon_{||}$ and $\varepsilon_{\perp}$ also change as a function of temperature [13], as shown in Figure 97.4 for the nematic LC cyanobiphenyl. The two dielectric constants rapidly converge as the temperature approaches T_c, where $\varepsilon_{||} = \varepsilon_{\perp} = \varepsilon_{\text{isotropic}}$.

97.3.2 Refractive Index Anisotropy

An LC is birefringent with anisotropic refractive indices. It has two principal refractive indices, n_o and n_e, as shown in Figure 97.5. For the ordinary refractive index n_o, the electric field vector of the light beam oscillates perpendicular to the optic axis, and for the extraordinary refractive index n_e, the electric field vector oscillates parallel to the optic axis. In the nematic and smectic LCs, the direction of the LC director $\boldsymbol{n}$ is the optic axis of the uniaxial crystal and therefore the refractive indices for light rays with oscillations in the directions parallel and perpendicular to the director are $n_{||}$ and $n_{\perp}$, respectively, i.e., $n_o = n_{\perp}$, $n_e = n_{||}$, and the optical anisotropy or birefringence, Δn, is given by

$$\Delta n = n_{||} - n_{\perp} = n_e - n_o \tag{97.4}$$

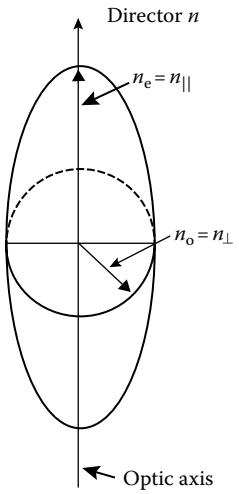


FIGURE 97.5 Refractive index anisotropy of uniaxial LC.

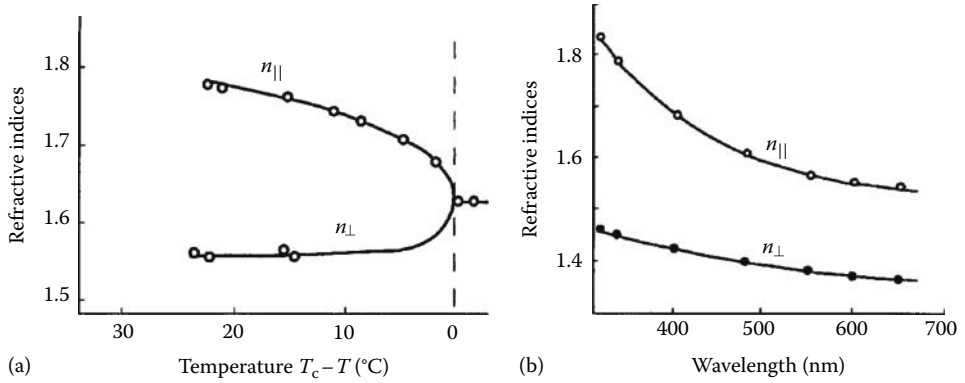


FIGURE 97.6 (a) Temperature dependence of the refractive indices $n_{||}$ and $n_\perp$ for a nematic LC MBBA at $\lambda = 546$ nm and (b) wavelength dependence of refractive indices of the nematic LC 4-butoxyphenyl ester of 4′-hexyloxybenzoic acid at 80°C.

Figure 97.6a and b shows the temperature dependence [2] and wavelength dependence [2] of the refractive indices of typical LC materials. The dependence of refractive index on wavelength λ is generally expressed by the so-called Cauchy equation:

$$n_{o,e} = n_\infty + \frac{\alpha_{o,e}}{\lambda^2} \tag{97.5}$$

where

n_∞ is the refractive index extrapolated to infinite wavelength
α is a material-specific coefficient

97.3.3 Elastic Constants

In uniaxial LCs, the preferred or equilibrium orientation of the LC molecule is given by the director ***n***, which may be imposed by the surface treatments at the boundary conditions or by an external field.

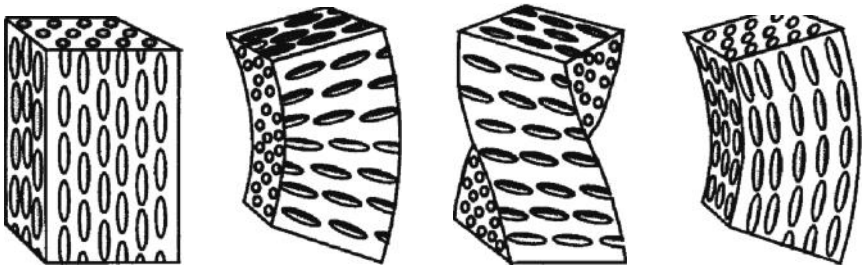


FIGURE 97.7 Deformation of nematic LC molecules from equilibrium configuration is shown in (a). Three types of deformations—(b) splay, (c) twist, and (d) bend can describe all possible types of deformations.

When the LC is perturbed from an equilibrium condition by application or removal of an external field, the elastic and electrical forces determine the static deformation pattern of the LC. The transition of the director from one direction to the other induces curvature strain in the medium. Frank [14] showed that an arbitrary deformation state can be envisaged as the combination of three basic operations: *splay*, *twist*, and *bend*, denoted by the elastic constants K_{11}, K_{22}, and K_{33} following the notation of the Oseen–Frank theory. Figure 97.7 illustrates the *splay*, *twist*, and *bend* deformations. The elastic part of the internal free energy, F, of a perturbed LC is given by the equation

$$F = \frac{1}{2}[K_{11}(\nabla \cdot n)^2 + K_{22}(n \cdot \nabla \times n)^2 + K_{33}(n \times \nabla \times n)^2] \tag{97.6}$$

The free energy density is thus a quadratic function of the curvature strains with the elastic constants appearing as constants of proportionality. The elastic constants K_{11}, K_{22}, and K_{33} are temperature-dependent, and decrease with increase in temperature. The magnitudes of the elastic constants, K_{ii}, can be approximated [15] by

$$K_{ii}\alpha S^2 \tag{97.7}$$

97.3.4 Electro-optic Characteristics

When an electric field is applied to an LC with an initial molecular (director) orientation, it will change to a new molecular orientation due to the dielectric anisotropy ($\Delta\varepsilon = \varepsilon_{\|} - \varepsilon_{\perp}$) of the LC. This change in molecular orientation is accompanied by a change in the optical transmission/reflection characteristics of the LC which forms the basis for LCDs. This phenomenon of an electrically driven optical modulation is known as the electro-optic effect of the LC. When an electric field E is applied to an LC, it produces an electric energy, f_e, given by

$$f_e = -\frac{1}{2}\varepsilon_{\perp}E^2 - \frac{1}{2}\Delta\varepsilon(n \cdot E)^2 \tag{97.8}$$

The initial molecular orientation (boundary condition achieved by surface alignment) of the LC molecules (director $\boldsymbol{n}$) is either parallel (for $\Delta\varepsilon > 0$) or perpendicular (for $\Delta\varepsilon < 0$) to the plane of the two parallel electrodes of the display. When a field is applied across the parallel electrodes with the LC material in between, the director $\boldsymbol{n}$ orients parallel to the electric field E in *+ve* $\Delta\varepsilon$ materials, and it orients perpendicular to the field in *−ve* $\Delta\varepsilon$ materials. The total free energy, F, of the LC when the initial undeformed

molecular orientation undergoes deformation due to the applied field is given by the sum of the electric energy f_e and the elastic energy. This transition from an undeformed state, known as *Freedericksz transition*, occurs as the field is increased to a critical field E_c. The Freedericksz transition is simply a transition from a uniform director configuration to a deformed director configuration, i.e., at any point in the LC, the order of the molecules relative to one another remains the same. The threshold electric field, E_c, is calculated by a free energy minimization technique [3], and it is given by

$$E_c = \left(\frac{\pi}{d}\right)\left(\frac{K_{ii}}{|\Delta\varepsilon|}\right)^{1/2} \tag{97.9}$$

Thus the threshold voltage V_{th} of the LC electro-optic effect is given by

$$V_{th} = E_c \cdot d = \pi\left(\frac{K_{ii}}{|\Delta\varepsilon|}\right)^{1/2} \tag{97.10}$$

In Equations 97.9 and 97.10, d the thickness of the LC and K_{ii} is the appropriate elastic constant. When the field is perpendicular to the initially homogeneous orientation of the director, $K_{ii} = K_{11}$ or K_{22}. When the field is parallel to the initially homogeneous orientation, $K_{ii} = K_{33}$. In the case of a twisted orientation, $K_{ii} = K_{11} + (K_{33} - 2K_{22})/4$.

97.4 LCD Materials and Fabrication Processes

There are several types of LCDs utilizing different LC materials and LCD modes which are discussed in the next section. However, the general display assembly processes and materials are very similar for all these LCD modes. Figure 97.8a shows the plan view and Figure 97.8b shows the cross-sectional view of a PM-addressed LCD. The display fabrication can be broadly divided into three parts: (1) lower and upper glass fabrication processes, (2) cell assembly processes, and (3) polarizer and driver attachment and module assembly processes as illustrated in Figure 97.9. The following will describe the various display materials and the assembly processes.

97.4.1 Glass Substrate

The quality of the glass substrate is important with regard to its chemical compatibility with the LC materials, surface flatness, defects, and dimensional stability under processing temperatures associated with various display fabrication steps. With a typical LCD cell gap in the range of 5–10 μm, the importance of the glass flatness and surface quality is clear. Glass substrate defects such as voids, scratches, streaks, and attached particles can cause electrode defects and hinder uniform LC cell spacing. Therefore, depending on the type of display, glass substrates are sometimes polished to achieve the required surface quality. Typical display glass materials include borosilicate (e.g., Corning 7059) and aluminosilicate glasses (e.g., Corning 1737), with a thickness of 0.7 or 1.1 mm.

97.4.2 Color Filters

In full-color LCDs, color most often is generated by use of red, green, and blue (R, G, B) color filters fabricated at each pixel, as shown in Figure 97.8b, using a white backlight system. The color filter requirements include proper spectral transmission characteristics and chemical, thermal, and dimensional stability. The display color gamut is a function of the spectral characteristics of the backlight used and the color filter transmission characteristics. By a suitable choice of these parameters, an LCD can achieve a color gamut comparable to that of a high-quality CRT. However, trade-offs are sometimes

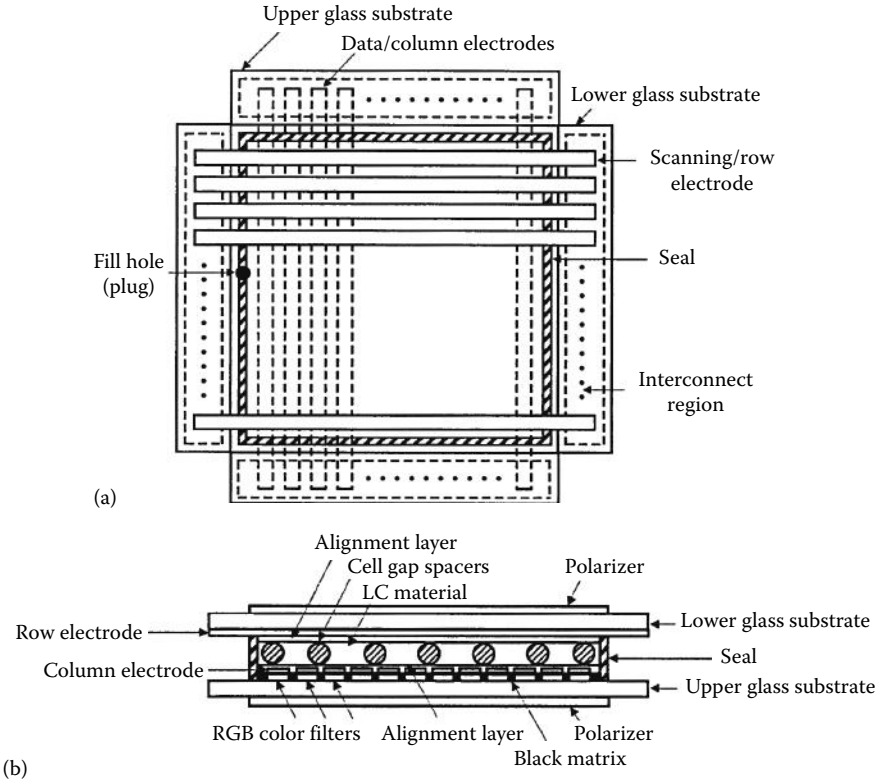


FIGURE 97.8 Plan (a) and cross-sectional view (b) of a passive matrix–addressed LCD.

made between color purity and transmission (brightness) characteristics of the LCD. Typical color filter thickness is about 2 μm. Color filter materials include dye- and pigment-dispersed polyimides and photoresists. The color filter materials are applied on the display glass by various processes, such as spin coating, printing, electrodeposition, and photolithography. First, color filter material of the first color is applied and photolithographically patterned. Then, the color filter material for the second color is processed, and then the third. A black matrix material is also applied and patterned between the color filters to block the light transmission from the interpixel regions. In some cases, a passivation layer such as low-temperature SiO_2 dielectric is deposited on the color filters to act as a barrier for impurities and to achieve a smooth surface for the subsequent transparent electrode deposition.

97.4.3 Transparent Electrodes

Most often, indium tin oxide (ITO) with a typical concentration of 90% In_2O_3 and 10% SnO_2 is used as the transparent conducting electrode material. The typical transmission of ITO is about 90%. It is generally deposited by e-beam evaporation or sputtering in an oxygen-containing atmosphere. The film thickness is typically 50–300 nm depending on the required sheet resistance and transmission. Typical resistivity is in the range of 2×10^{-4} Ω cm. Light transmission through ITO is not linearly proportional to the film thickness because of light interference.

97.4.4 Alignment Materials and Techniques

After the transparent electrodes are patterned, an LC alignment coating is applied. Depending on the display mode of interest, either a homogeneous (parallel), tilted, or a homeotropic (vertical) alignment is

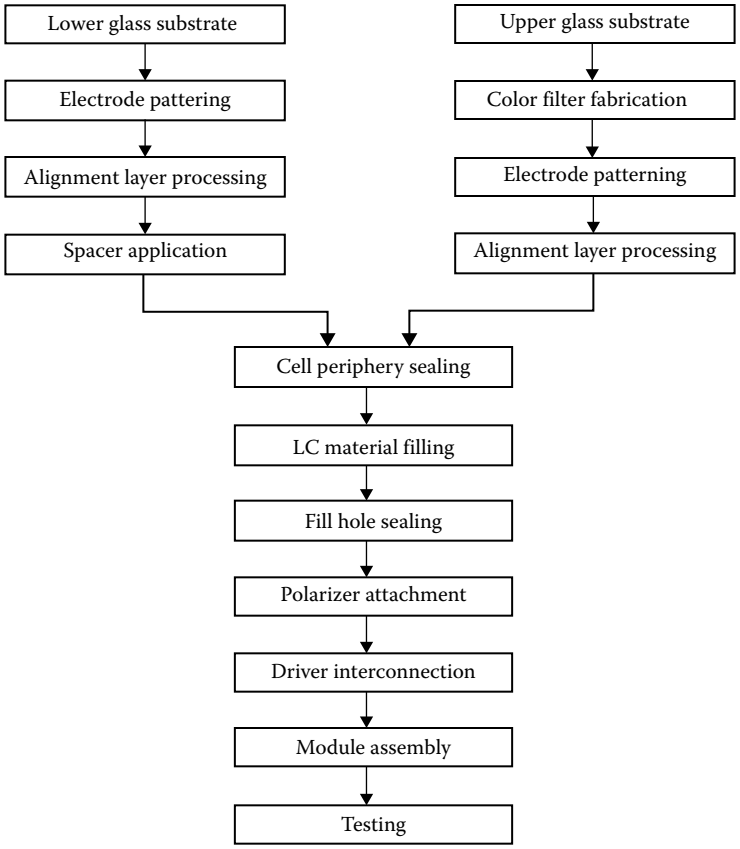


FIGURE 97.9 LCD assembly flowchart.

achieved using an appropriate alignment material and process. The most commonly used TN LCD requires a tilted alignment to eliminate reverse tilt disclinations. Inorganic films such as an obliquely evaporated SiO, as well as organic materials such as polyimide, can be used as alignment materials. Polyimide is most commonly used as an alignment layer. It is spin-coated or printed to achieve a layer thickness of about 50 nm, and cured around 150 °C–200 °C. It is then buffed (rubbed) using a roller covered with cotton or a synthetic fabric. For TN LCD, the alignment process is selected to achieve a pretilt angle of around 2°–5°.

97.4.5 Cell Spacing and Sealing

The cell assembly process starts after the alignment layer treatment of the two substrates by rubbing. To control the LC cell spacing (thickness) accurately, spacing materials are applied in the active area of the display as well as in the peripheral seal area, where the two glass substrates are sealed. Typical spacing materials include glass or plastic fibers (cylinders) or spheres (balls), with a tightly controlled size (diameter) distribution. A sealing adhesive material such as an epoxy is then applied at the seal area with an opening (fill hole) left for LC material injection after the sealing process. The seal is typically 1–3 mm wide and is a few millimeters from the active display area. The requirement for the seal material is that it must not chemically react with (contaminate) the LC material and must be a barrier against moisture and contamination from outside. After the spacers and seal materials are applied on the first (active) substrate, it is precisely aligned to the second (color filter) substrate, and the seal is cured either by heating it to a higher temperature, typically 100 °C–150 °C, or by ultraviolet exposure depending on the type of epoxy seal material used.

97.4.6 LC Material Filling

After the empty cell is fabricated, it is filled with the LC material. Because the cell thickness (spacing) is small (~5–10 μm), it is filled using special techniques. The most popular filling method is by evacuating the cell in a vacuum chamber, dipping the fill hole into a vessel containing the LC material, and increasing the pressure in the chamber. As the chamber pressure is raised, the cell gets filled by capillary action. After the cell is completely filled, the fill hole is capped by using an epoxy adhesive that is chemically compatible with the LC material.

97.4.7 External Components

The external components of an LCD include polarizers, reflectors, display drivers, and a backlight assembly. A reflective LCD uses a reflector at the back of the display, works by modulating the ambient light, and does not require backlighting. In the most commonly used transmissive mode TN LCDs, a polarizer is attached on the front as well as back surfaces of the LCD after the cell assembly is complete. Also, in the most commonly used NW mode TN LCD, the polarizers are attached with their polarization axis crossed and along the rubbing directions of the alignment layers. The polarizer is a three-layer composite film with a stretched iodine-doped polyvinyl alcohol (PVA) polarizing film in the center, and two outer films of triacetyl cellulose (TAC) for protecting the PVA film from the ambient (moisture, temperature, and harsh environment) conditions. A typical transmission range of a polarizer is 41%–45%, with polarization efficiency in the range of 99.9%–99.99%.

The display row and column IC drivers are attached to the row and column bond pads of the display either by tape-automated bonding (TAB) using an ACA (anisotropic conductive adhesive), or chip on glass (COG) approaches. For backlighting transmissive LCDs, a fluorescent lamp is generally used. The R, G, B emission spectrum of the backlight and transmission spectrum of the R, G, B color filters are tuned together to achieve the desired color coordinates for the primary colors. Also, a diffuser is used to achieve uniform backlighting of the display. The backlight system may also use brightness enhancement films to tailor the light intensity distribution in the viewing cone. In addition to the earlier components, LCDs for specialized applications requiring enhanced performance may use a cover glass with EMI and antireflection coatings at the front, and a heater glass at the back side (between the backlight and the LCD) that facilitates low-temperature operation.

97.5 LCD Modes

LCDs based on many different modes of operation have been developed. Historically, the phase change (PC) effect was discovered first in 1968 by Wysocki et al. [16]. The same year, dynamic scattering (DS) mode [17] and guest–host (GH) mode [18] were announced by Heilmeier et al. Then in 1971, the TN mode [19] was reported by Schadt and Helfrich and electrically controlled birefringence (ECB) was reported by Schiekel and Fahrenschon [20] and Hareng et al. [21]. The physical effects and the various display modes based on these effects include

Current effects	• DS effect
Electric field effects	• TN effect
	• STN effect
	• ECB
	• GH effect
	• PC effect
Thermal effects	• Smectic effect

The DS effect is based on the anisotropy of the conductivity. Because of higher voltage operation and higher power consumption, the DS mode is not currently used. The TN and STN effects are most widely used among all the LCDs. The following sections will discuss various display modes.

97.5.1 TN Effect

Figure 97.10 shows a schematic of a display based on the TN effect. It consists of nematic LC material with a positive dielectric anisotropy ($\Delta\varepsilon > 0$) with a layer thickness of about 5 μm, sandwiched between two transparent substrates with transparent electrodes. The surfaces of the transparent electrodes are coated with a polyimide alignment layer and rubbed to orient LC molecules at the substrate surfaces along the rubbing direction with a small (~3°) pretilt angle. The molecules on the two substrates are oriented 90° from each other as shown in Figure 97.10a, i.e., the LC molecular axis rotates (twists) continuously through 90° from the first substrate to the second substrate. The TN display can be fabricated to operate in a normally black (NB) or normally white (NW) mode based on how the polarizers are attached to the outer surface of the two glass substrates. Figure 97.10a and b shows the on- and off-state of a NW mode TN LCD with crossed (orthogonal) polarizers attached with their polarization direction parallel to the LC director orientation on that substrate. Since the pitch of the twist is sufficiently large compared with the wavelength of the visible light, the direction of polarization of linearly polarized

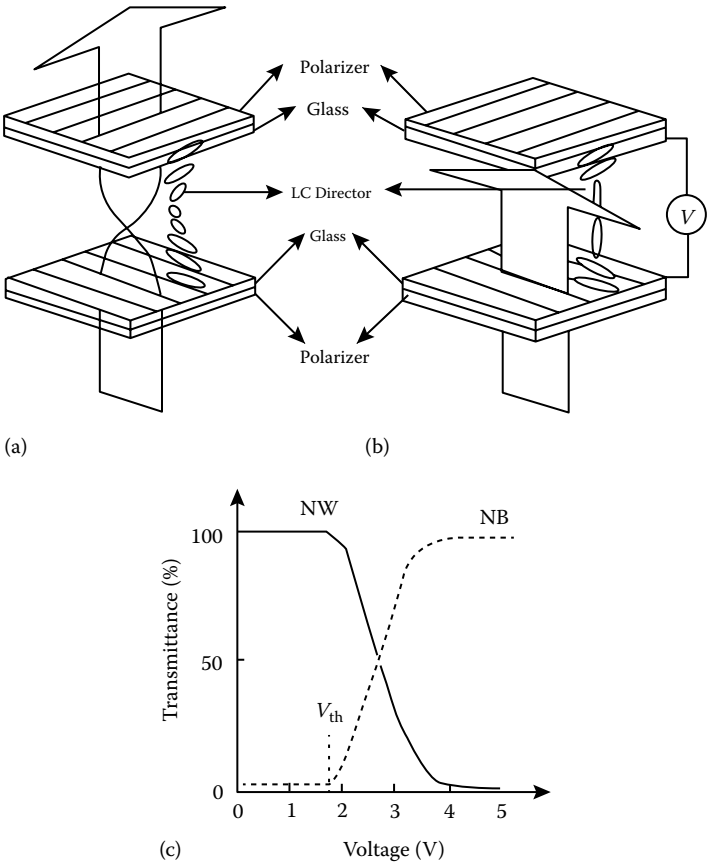


FIGURE 97.10 Illustration of TN effect: (a) in the off-state, the incident polarized light is transmitted through the entrance polarizer, the 90° TN LC, and the crossed exit polarizer; (b) in the on-state; (c) the solid line shows the voltage-transmission behavior for the NW configuration, shown in (a) and (b), with crossed polarizers. The dashed line is for an NB configuration with parallel polarizers.

light incident normally on one surface of the display rotates through 90° by the twist of the LC molecules as it propagates through the cell and exits through the second polarizer. When a voltage is applied to the TN cell, the molecules align parallel to the direction of the field as shown in Figure 97.10b, and the 90° optical rotatory power is eliminated. Thus, the incident polarized light from the first polarizer is not rotated as it goes through the LC cell and gets blocked by the crossed exit polarizer.

Figure 97.10c shows the voltage-transmission characteristics of a TN cell for NB and NW modes of operation. For an NW mode, when a sufficiently high voltage is applied, LC molecules are aligned homeotropically (parallel to the field) and there is no rotation of the electrical field vector of the polarized light. This results in complete suppression of transmission regardless of the wavelength of light. The transmission in the on-state is, however, wavelength-dependent, but this does not have a significant effect on the contrast ratio, although it can influence the color balance. In the NB mode of operation, the transmission is suppressed to zero [22] for the off-state only for a monochromatic light of wavelength, $\lambda = 2d\Delta n/\sqrt{3}$. Therefore, in a practical display using a broadband backlight, there is a small amount of light leakage which lowers the display contrast ratio. In a color display, the cell gaps for the R, G, B pixels can be optimized to eliminate the light leakage and improve contrast [22].

The threshold voltage of a TN mode LCD is given by

$$V_{\text{th}} = \pi \sqrt{\left\{ \frac{[K_{11} + (K_{33} - 2K_{22})/4]}{\varepsilon_0 \cdot \Delta\varepsilon} \right\}} \tag{97.11}$$

V_{th} depends on the dielectric anisotropy and elastic constants, and is generally in the range of 2–3 V, with the maximum operating voltage being in the range of 5–8 V. This low-voltage driving, coupled with low current due to high resistivity of LC materials, contributes to the very low power consumption (~1 μW/cm^2) of LCDs. The response times measured by the rise and decay times of the display τ_r, τ_f are given by [23]

$$\tau_{\text{d}} = \gamma \cdot \frac{d^2}{\Delta\varepsilon \left(V^2 - V_{\text{th}}^2 \right)} \tag{97.12}$$

$$\tau_{\text{d}} = \gamma \cdot \frac{d^2}{\left(\Delta\varepsilon \cdot V_{\text{th}}^2 \right)} \tag{97.13}$$

where γ is the rotational viscosity coefficient. The earlier equations show that rise time can be improved by using a thinner cell gap d, a higher $\Delta\varepsilon$, and a higher drive voltage V. Similarly, the decay time can be improved by reducing the cell gap. The turn-on time τ_r is usually shorter than the turn-off time τ_d. At room temperature, these times are of the order of 10 ms, which is adequate for many common applications such as computer and TV displays. In a TN LCD, gray scale is generated by varying the voltage using the electro-optic curve shown in Figure 97.10c. The shallow slope of the electro-optic curve works well for the gray-scale generation in AM-addressed displays (see next section). However, in the case of PM addressing, the shallow slope greatly limits the multiplexibility (number of addressable rows) of the display, which led to the development of the STN effect.

97.5.2 STN Effect

In a PM-addressed display, the addressability or the maximum number of addressable rows N is given by the Alt and Pleshko [24] limit:

$$\frac{V_{\text{on}}}{V_{\text{off}}} = \sqrt{\frac{(N^{1/2} + 1)}{(N^{1/2} - 1)}} \tag{97.14}$$

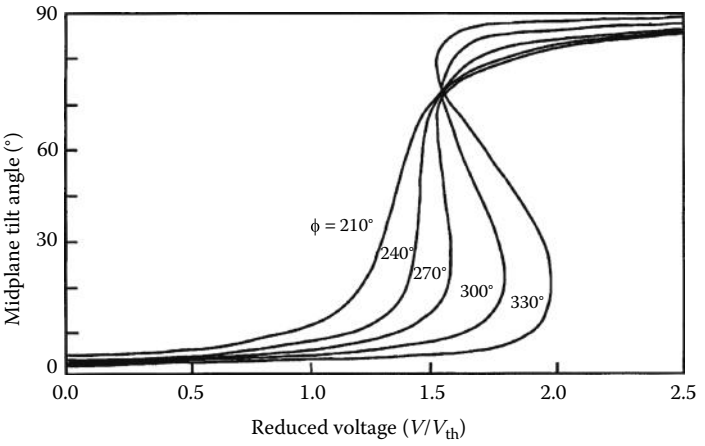


FIGURE 97.11 Calculated curves of tilt angle of local directors in the midplane of an STN cell as a function of reduced voltage V/V_{th}, where V_{th} is the Freedericksz threshold voltage of a nontwisted layer with a zero pretilt angle. The steepness of the curves increases as the twist angle, Φ, is increased, and bistability is achieved when $\Phi > 240°$.

where V_{on} and V_{off} are the rms voltages at the select and nonselect pixels. Equation 97.14 shows that as N increases V_{off} approaches V_{on} and the contrast ratio becomes 1, which makes the display not viewable. For $N = 100$, $V_{on} = 1.11\ V_{off}$, i.e., select voltage is only 11% higher than the nonselect voltage. This will result in a very low contrast ratio when using a TN mode with a shallow turn-on curve (Figure 97.10c). STN displays have been developed [25,26] to achieve a steep electro-optic curve, so that large numbers of rows can be multiplexed. The STN effect uses a twist angle of 180° to about 270° with a relatively high pretilt angle alignment. Figure 97.11 [25] illustrates the STN effect. The figure shows the voltage dependence of the midplane director tilt of a chiral nematic layer with a pretilt of 28° at both substrates. Bistability is achieved when a twist angle, φ, greater than 245° is used. In highly multiplexed displays, twist angles in the range of 240°–275° and tilt angles in the range of 5°–30° are generally used. High pretilts ensure that competing distortional structure which has 180° less twist is eliminated. For a 270° left-handed twist, optimum results are achieved when the front polarizer is oriented with its polarization axis at 30° with the LC director and the rear polarizer is oriented at an angle of 60° with the projection of the director at the rear substrate. Due to interference of the optical normal modes propagating in the LC layer, the display has a yellow birefringence color in the nonselect state (yellow mode). Rotation of one of the polarizers by 90° results in a complementary image with a bright colorless state, and a blue nonselect state (blue mode). White-mode STN displays are made using retardation films. The response time of a typical STN display is on the order of 150 ms. These displays typically have a lower contrast ratio and a narrow viewing angle.

97.5.3 Electrically Controlled Birefringence

This display technique is based on controlling the birefringence of the LC cell by application of an electric field. There are a number of types of this display depending on the molecular orientation of the LC cell used; examples include DAP type, homogeneous type, hybrid-aligned nematic (HAN) type, and in-plane switching (IPS) type. A DAP type (homeotropic orientation) is made using an LC with a negative $\Delta\varepsilon$, sandwiched between transparent electrode substrates and placed between crossed polarizers, as shown in Figure 97.12. In the off-state (with no electric field), the incident polarized light does not see birefringence when passing through the cell, and thus gets blocked by the crossed exit polarizer. In the on-state (when a voltage is applied), the molecular axis of the LC is inclined at an angle θ (as shown in Figure 97.12), so the linearly polarized light becomes elliptically polarized as it passes through the cell

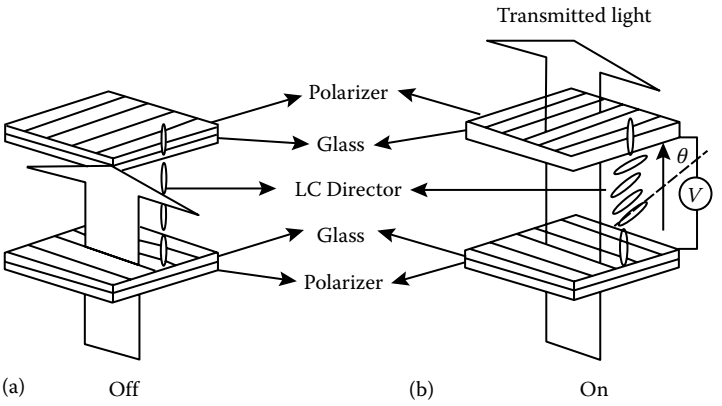


FIGURE 97.12 Illustration of an ECB display: (a) with homeotropic alignment and crossed polarizers, the off-state is black and (b) in the on-state, the output light through the LC is elliptically polarized due to the LC birefringence and the light is transmitted through the crossed polarizer.

due to birefringence. Hence, a portion of the light passes through the crossed exit polarizer; the intensity I of the transmitted light through the cell is given by [20]

$$I = I_0 \sin^2 2\theta \sin^2\left(\frac{\pi d\Delta n(V)}{\lambda}\right) \tag{97.15}$$

where

I_0 is the intensity of the incident light
θ is the angle between the direction of polarization of the incident light and the direction of oscillation of the ordinary light within the cell
d is the cell spacing
$\Delta n(V)$ is the birefringence of the cell
$d\Delta n$ is the optical phase difference
λ is the wavelength of the incident light

The equation shows that I depends on the applied voltage and λ.

In case of the homogeneous technique, the LC cell uses positive material with a homogeneous orientation. With this method, the threshold voltage is obtained by replacing the bend elastic coefficient K_{33}, with splay elastic coefficient K_{11}. The HAN cell is characterized by a hybrid orientation cell in which the molecules are aligned perpendicular to one substrate, but parallel to the second substrate. In this mode, both positive and negative $\Delta\varepsilon$ materials can be used, and since there is no clear threshold voltage, it has the advantage of a very low drive voltage. Recently, an ECB mode based on an IPS-type display has been developed to produce LCDs with extremely wide viewing angles. In the IPS mode displays, LC is homogeneously aligned and switched between on- and off-states using interdigitated electrodes fabricated on one of the display substrates [27]. The IPS mode displays use an NB mode with either positive or negative $\Delta\varepsilon$ LC materials.

97.5.4 GH Type

Some organic dyes show anisotropy of light absorption, i.e., they absorb more light in a specific wavelength band when the **E** vector of the light is parallel to the optic axis of the dye molecules, than they do when it is perpendicular. LCDs based on this principle are called GH displays. In these displays, a small

amount of a dichroic dye (guest) is mixed in the LC material (host). These dye molecules get aligned to the LC molecules; hence, their orientation can be changed (by changing the orientation of the LC molecules) by application of an electric field. When an LC material with a positive $\Delta\varepsilon$ is used, in the off-state, the **E** vector of the polarized light coincides with the light absorption axis of the dichroic dye; hence, light is absorbed and transmitted light is colored. When a voltage is applied for the on-state, the **E** vector of the polarized light is orthogonal to the absorption axis of the dye; hence, no absorption takes place, and transmitted light is not colored (white). GH LCD requires only one polarizer. Further, because the optical effect is based on absorption, the display provides a better viewing angle than a TN mode LCD.

97.5.5 Phase Change Type

The PC type of display is based on a change in molecular orientation from a helical cholesteric phase to a homeotropic nematic phase, and vice versa. For this technique, a cholesteric LC with a long helical pitch with a positive or negative $\Delta\varepsilon$ is used. No polarizers are used in this display. In the off-state of this display, the incident light passing through the cholesteric cell with a focal conic orientation is optically dispersed (scattered), and the cell looks cloudy. However, when a voltage is applied, helical structure of the cholesteric phase changes to a nematic phase with a homeotropic orientation, and the cell becomes transparent.

97.5.6 Thermal Effects

Thermal effect is based on a change in the electro-optic behavior due to a change in the molecular orientation of the LC when it is heated or cooled. This effect is utilized with smectic LCs with a homeotropic alignment. When this cell is heated until the isotropic phase is reached and cooled, and then if the cooling is sudden, the cell becomes cloudy, whereas, if the cooling is gradual, the cell becomes transparent. These cloudy and transparent states correspond, respectively, to the focal conic and homeotropic orientations of the smectic A LC. This effect is used for large-size memory-type displays, in which a laser beam is used to write the image thermally. The heating can also be accomplished by one of the transparent electrodes, while the other transparent electrode is used as a signal electrode.

97.6 Display Addressing

Display addressing (driving) techniques has a major influence on the LCD image quality. The addressing techniques can be classified in three essential types, namely, direct (static) addressing, PM addressing, and AM addressing. In the case of low-information-content displays such as numeric displays, bar graph displays, and other fixed pattern displays, using segmented electrodes, direct addressing is used. A common example of direct-addressed displays is a numeric display using seven segmented electrodes for each digit. Each of these segmented electrodes on the front substrate and the common electrode on the back substrate are directly connected to drive signals. A voltage is selectively applied to each of the segments so that any of the digits between 0 and 9 can be displayed. For high-information-content displays, this approach becomes impractical because of the huge number of interconnects, and hence, either PM or AM addressing is used.

97.6.1 PM Addressing

A PM display comprising an LC between a matrix of transparent conducting row and column electrodes (Figure 97.8) is the simplest and least expensive matrix-addressed LCD to manufacture. An example of a PM LCD is a color (R, G, B) video graphics array (VGA) display using the STN effect, with a pixel format of 640 ($\times$ 3 = 1920) $H \times$ 480 V, with 1920 columns and 480 rows, for a total of 2400 interconnects used for addressing a display containing 921,600 pixels. In PM addressing, the row voltages are scanned in succession with a voltage, V_r, while all the columns in a given row are driven in parallel, during the

row time, with a voltage of $\pm V_c$ depending on whether the pixel is selected to be ON or OFF. As discussed earlier under Section 97.5.2, the contrast ratio of PMLCDs is influenced by the Alt and Pleshko [24] addressability limitation. To enhance the operating margin for improved contrast ratio, dual-scan STN (DSTN) configuration is used in higher-information-content displays. In a DSTN, the display is separated into two halves, and the rows in each half are scanned simultaneously and synchronously, to essentially double the duty ratio of the ON pixels to increase the contrast ratio. One of the major shortcomings of the PM-addressed STN display is the slow response time of the LC, which is of the order of 150 ms. This slow response time is not adequate for video applications and is barely fast enough for the graphical interface of a computer. The response time of the STN LCDs can be improved by active addressing or multiline addressing techniques [28,29]. These techniques involve simultaneous addressing of several rows of a display to suppress the frame response problems of conventional STN LCDs.

97.6.2 AM Addressing

AM addressing removes the multiplexing limitations [24] of the PM LCDs by incorporating a nonlinear control element in series with each pixel, and provides 100% duty ratio for the pixel using the charge stored at the pixel during the row addressing time. Figure 97.13a illustrates an AM array with row and column drivers and the associated display module electronics. Figure 97.13b shows a magnified view of the AM array in the AM LCD panel. In the figure, C_{LC} and C_S represent the pixel capacitance and the pixel storage capacitance. Typically a storage capacitor, C_S, is incorporated at each pixel to reduce the pixel voltage offset (see Equation 97.16) and for a broad temperature operation. Figure 97.14 shows the cross section through an AM LCD illustrating various elements of the display. Figure 97.15 shows a typical AM LCD pixel, showing the gate and data buses, thin-film transistor (TFT), ITO pixel electrode, and the storage capacitor. Fabrication of the AM substrate is one of the major aspects of AM LCD manufacturing. Both two-terminal devices such as back-to-back diodes, and metal–insulator–metal (MIM) diodes as well as three-terminal TFTs are developed for AM addressing. While two-terminal devices are simple to fabricate and cost less, their limitations include lack of uniform device performance (breakdown voltage/threshold voltage) over a large display area, and lack of total isolation of the pixel when neighboring pixels are addressed. For a superior image quality AM LCDs use TFT for the AM device, which provides a complete isolation of the pixel from the neighboring pixels. Large-area AM LCDs use amorphous silicon (a-Si) TFTs [6], while polysilicon TFTs with integrated row and column drivers are used in small high-resolution LCDs [30].

Figure 97.16 shows the electric equivalent of a TFT-LCD pixel, display drive waveforms, and the resulting pixel voltage. As in most matrix-addressed displays with line-at-a-time addressing, the rows (gates) are scanned with a select gate pulse $V_{g,sel}$, during the frame time t_f, while all the pixels in a row are addressed simultaneously with the data voltage $\pm V_d$ during the row time t_r (= t_f/N). During the row time the select gate voltage, $V_{g,sel}$, "turns on" the TFT and charges the pixel and the storage capacitor to the data voltage V_d. After the row time, the TFT is "switched off" by application of the nonselect gate voltage, $V_{g,nonsel}$; hence, the voltage (charge) at this pixel is isolated from the rest of the matrix structure until it is time to charge the pixel during the next frame time. Note that the LC pixel must be driven in an ac fashion with $+V_d$ and $-V_d$, during alternate frame periods, with no net dc across the pixel. A net dc voltage across the pixel results in flicker and image sticking effects [31], resulting from LC conductivity. Large and sustained dc voltages also degrade the LC material due to electrolysis. The shift in pixel voltage, ΔV_p, shown in Figure 97.16, at the end of the row time is due to the parasitic gate-to-drain capacitance, C_{gd}, of the TFT. When the gate voltage is switched, the distribution of the charge from the TFT gate dielectric causes the pixel voltage shift ΔV_p, given by

$$\Delta V_p = \frac{(\Delta V_g) C_{gd}}{(C_{gd} + C_{lc} + C_s)} \tag{97.16}$$

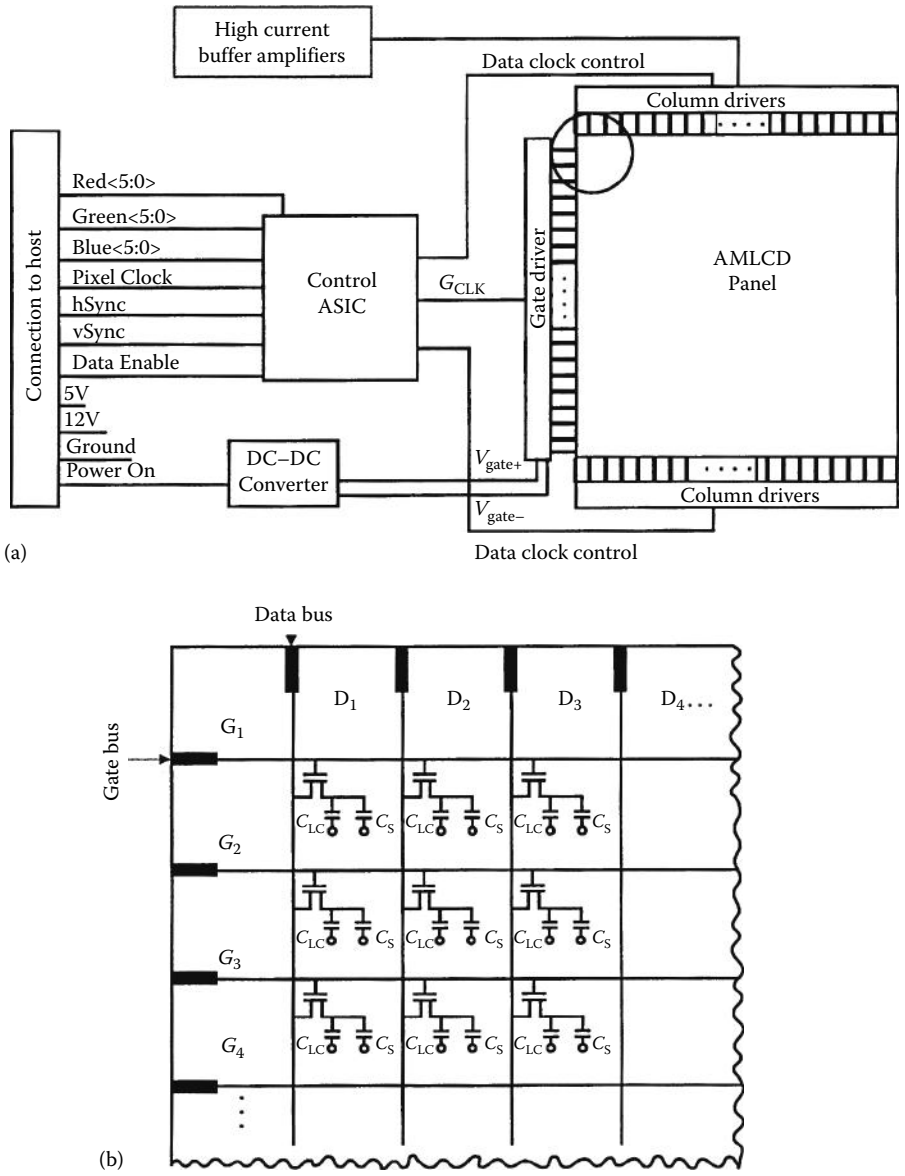


FIGURE 97.13 (a) AM LCD module electronics block diagram and (b) magnified view of the region shown in (a) illustrating active matrix TFT array.

For the *n*-channel enhancement mode a-Si TFT, this voltage shift ΔV_p is negative for both the positive and negative frames; thus, it helps pixel charging in the negative frame and hinders it in the positive frame. Further, due to increased gate bias during the negative frame, the pixel attains the data voltage much more rapidly during the addressing period. Hence, the TFT is designed for the worst-case positive frame conditions. ΔV_p is reduced by minimizing C_{gd} by decreasing the source drain overlap area of the TFT and by using a storage capacitor. Further, ΔV_p is compensated by adjusting the common electrode voltage V_{com} as shown in Figure 97.16. Note that C_{LC} is a function of the V_p (V_{LC}) due to the dielectric anisotropy of the LC; hence, adjustment to V_{com} alone does not eliminate dc for all gray levels, and modification of the gray-scale voltages is required to compensate for the dielectric anisotropy of the LC.

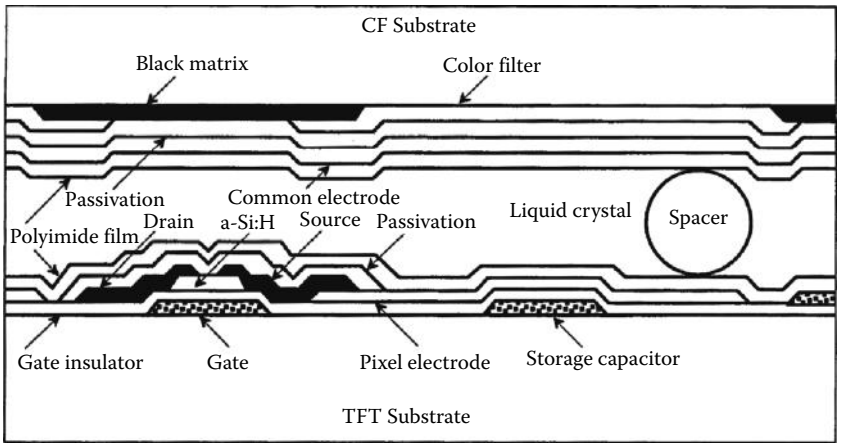


FIGURE 97.14 Cross-sectional view through an AM LCD showing TFT, pixel electrode, storage capacitor, polyimide alignment layers, color filter, and black matrix.

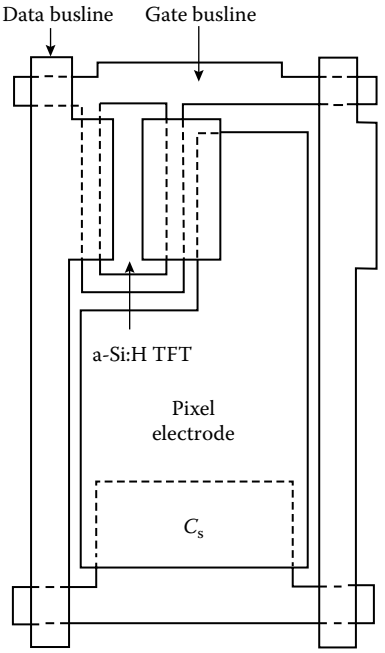


FIGURE 97.15 Layout of a pixel in a TFT-LCD showing the TFT, pixel electrode, and storage capacitor.

97.6.3 Display Module Electronics

Figure 97.13 shows a block diagram for an AM LCD module electronics. The control block and power supply generation means are separately mounted on a PC board and connected to the row and column drivers of the LCD on one side and to the host controller on the other. The control block may include level shifters, timing generators, and analog functions in some cases; the control block takes in digital data from the host system, which is typically a graphics controller chip, and converts it into timing and signal levels required by the row and column drivers. The architecture and design of the module

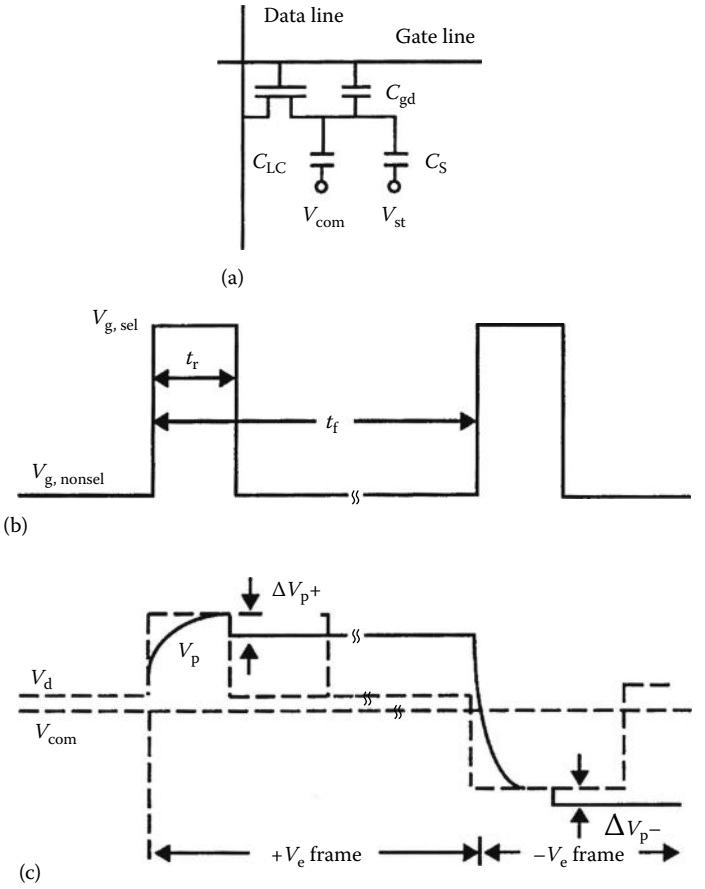


FIGURE 97.16 (a) Electric equivalent of a TFT-LCD pixel, (b) gate drive voltage waveform, and (c) data voltage waveform and pixel charging behavior in the positive and negative frames.

electronics encompassing row and column drivers have a significant impact not only on the display system cost and power consumption, but also on the image quality. The LC material typically requires about 5 V to achieve optical saturation (see Figure 97.10c). Considering the need for an ac drive, the required voltage swing across the LC material is about 10 V. To achieve this 10 V swing across the LC material, the column drivers typically use 12 V power supplies. Column driver voltage can be reduced by using a V_{com} modulation drive method. In this method, the V_{com} node (which is connected to all pixels in the display) is driven above and below a 5 V range of the column drivers. Each and every row time, the V_{com} node is alternated between a voltage above and a voltage below the 5 V output ranges of the column drivers. This achieves 10 V across the LC material using 5 V column drivers. This method requires additional components and consumes additional power due to the oscillation of the V_{com} node. In addition, to avoid capacitive injection problems, the row drivers usually have their negative supply modulated with the same frequency as the V_{com} node. Note, however, that compared to 10 V column drivers, 5 V column drivers consume less power, and are simpler to design and fabricate using small-geometry CMOS. The V_{com} modulation drive method can be used with a row (polarity) inversion scheme only (for elimination of pixel flicker) which results in some horizontal cross talk. However, column inversion and pixel inversion schemes provide better image quality with much-reduced cross talk, but they cannot be used with the V_{com} modulation drive.

References

1. P.J. Collings, *Liquid Crystals: Nature's Delicate Phase Matter*, Princeton University Press, Princeton, NJ, 1990.
2. L.M. Blinov and V.G. Chigrinov, *Electro-Optical Effects in Liquid Crystal Materials*, Springer, New York, 1996.
3. P.G. De Gennes and J. Prost, *The Physics of Liquid Crystals*, 2nd ed., Clarendon Press, Oxford, U.K., 1993.
4. B. Bahadur, Ed., *Liquid Crystals: Applications and Uses*, World Scientific, Singapore, 1990.
5. E. Kaneko, *Liquid Crystal TV Displays: Principles and Applications of Liquid Crystal Displays*, D. Reidel Publishing, Boston, MA, 1987.
6. T. Tsukada, *TFT/LCD: Liquid Crystal Display Addressed by Thin Film Transistors*, Gordon & Breach, Canada, 1996.
7. E. Lueder, Fundamentals of passive and active addressed liquid crystal displays, *Short Course S-1*, *SID Conference*, San Diego, CA, May 12, 1996.
8. P.W. Ross, L.K.M. Chan, and P.W. Surguy, Ferroelectric LCD: Simplicity and versatility, *SID '94 Digest*, 147–150, 1996.
9. J.W. Doane, D.K. Yang, and Z. Yaniv, Front-lit flat panel display for polymer stabilized cholesteric textures, *Proceedings of 12th International Display Conference*, *Japan Display '92*, Hiroshima, Japan, October 12–14, 73–76, 1992.
10. J.C. Rowell, W.D. Phillips, L.R. Melby, and M. Panar, NMR studies of some liquid crystal systems, *J. Chem. Phys*., 43, 3442–3454, 1965.
11. L. Pohl and E. Merck, Physical properties of liquid crystals, in *Liquid Crystals: Applications and Uses*, B. Bahadur, Ed., World Scientific, Singapore, 1990, 139–170.
12. P.R. Kishore, N.V.S. Rao, P.B.K. Sarma, T.F.S. Raj, M.N. Avadhanlu, and C.R.K. Murty, Field and frequency effects in nematic mixtures of negative and positive dielectric anisotropy, *Mol. Cryst. Liq. Cryst*., 45(3/4), 231–241, 1978.
13. D. Lippens, J.P. Parneix, and A. Chapoton, Study of 4-heptyl 4′-cyanobiphenyl using the analysis of its dielectric properties, *J. Phys*., 38, 1465, 1977.
14. F.C. Frank, On the theory of liquid crystals, *Discuss. Faraday Soc*., 25, 19, 1958.
15. W.H. de Jeu, *Physical Properties of Liquid Crystalline Materials*, Gordon & Breach, New York, 1980.
16. J.J. Wysocki, A. Adams, and W. Haas, Electric-field-induced phase change in cholesteric liquid crystals, *Phys. Rev. Lett*., 20, 1024, 1968.
17. G.H. Heilmeier, L.A. Zanoni, and L.A. Barton, Dynamic scattering: A new electro-optic effect in certain classes of nematic liquid crystals, *Proc. IEEE*, 56, 1162, 1968.
18. G.H. Heilmeier and L.A. Zanoni, Guest–host interaction in nematic liquid crystal—A new electro-optic effect, *Appl. Lett*., 13, 91, 1968.
19. M. Schadt and W. Helfrich, Voltage-dependent optical activity of a twisted nematic liquid crystal, *Appl. Phys. Lett*., 18, 127, 1971.
20. M.F. Schiekel and K. Fahrenschon, Deformation of nematic liquid crystal with vertical orientation in electric fields, *Appl. Phys. Lett*., 19, 391, 1971.
21. M. Hareng, G. Assouline, and E. Leiba, Liquid crystal matrix display by electrically controlled birefringence, *Proc. IEEE*, 60, 913, 1972.
22. C.H. Gooch and H.A. Tarry, The optical properties of twisted nematic liquid crystal structures with twist angles ≤90°, *J. Phys. D Appl. Phys*., 8, 1575–1584, 1975.
23. E. Jakeman and E.P. Raynes, Electro-optical response times in liquid crystals, *Phys. Lett*., 39A, 69–70, 972.
24. P.M. Alt and P. Pleshko, Scanning limitations of liquid–crystal displays, *IEEE Trans. Electron Dev*., ED-21, 146, 1974.

25. T.J. Scheffer and J. Nerring, A new highly multiplexable liquid crystal display, *Appl. Phys. Lett*., 45, 1021, 1984.
26. T.J. Scheffer, *Super Twisted Nematic (STN) LCDs*, *SID '95 Seminar Notes*, Vol. I, M-2, 1995.
27. M. Oh-e, M. Ohta, S. Aratani, and K. Kondo, Principles and characteristics of electro-optical behavior of in-plane switching mode, in *Proceedings of 15th International Display Research Conference*, *Asia Display '95*, Hamamatsu, Japan, 577–580, 1995.
28. T. Scheffer and B. Clifton, Active addressing method for high-contrast video-rate STN displays, *SID '92 Digest*, 228, 1992.
29. H. Muraji, H. Koh, T. Kuwata, and Y. Nakagawa, A 9.4-in. color VGA F-STN display with fast response time and high contrast ratio by using MLS method, *SID '94 Digest*, 61–64, 1994.
30. S. Higashi, Y. Matsueda, S. Takenaka, M. Miyasaka, I. Yudasaka, and H. Ohshima, A 1.8-in poly-Si TFT-LCD for HDTV projectors with 5V fully integrated driver, *SID '95 Digest*, 81, 1995.
31. Y. Nano, Y. Mino, E. Tekada, and S. Nagata, Characterization of sticking effects in TFT-LCD, *SID '90 Digest*, Las Vegas, NV, pp. 404–407, 1990.

98

Cathode Ray Tube Displays

Christopher J. Sherman
Merrimack

98.1 Introduction

The cathode ray tube (CRT) is unequaled in its ability to produce dynamic, quality, high-information-content imagery at high resolution. Even more impressive is that it achieves this for a lower cost per pixel than any other comparable electronic display technology. For the instrument designer requiring a high-information-content display, it offers numerous advantages. As a raw image tube, it is commonly available as an off-the-shelf item with a broad infrastructure of vendors, integrators, support, and part suppliers. Interface standards are well established and as a complete system, ready to take a standard signal input, it is available worldwide in a variety of performance ranges. To meet different application requirements, it is available in diagonal sizes from 12 mm to over 1 m with resolution from thousands of pixels to over 5 million pixels per frame. Tube characteristics improve on a yearly basis, and prices continue to decrease. Although CRTs have been superseded by other display technologies, they are still used in many applications in laboratories, photography, transportation industry, and in military. They are used in medical applications for color comparisons [1,2].

98.2 History

The CRT has a rich and distinguished history. The roots of today's CRT technology extend back more than 100 years to the latter half of the nineteenth century. Eugene Goldstein first introduced the term *cathode rays*; John W. Hittorf, Heinrich Geissler, Julius Plücker, and others made important contributions specific to CRT technology [3]. Throughout the nineteenth century, researchers were interested in the nature of a luminous gas discharge and shadowy rays that occurred when a high voltage potential was applied between two electrodes in a vacuum. Sir William Crookes was an active experimentalist in this field, and early cathode ray devices came to be known as Crookes tubes. Crookes noted that a glow was generated by the surface that the rays struck; the rays themselves were not the source of light. By producing tubes with a vacuum of 1.3×10^{-4} Pa (10^{-6} torr), he eliminated the luminous gas discharge and worked directly with the cathode rays [4]. Crookes demonstrated the following: luminance depended directly upon the material properties of the surface the rays struck; a magnetic field would deflect the path of the rays; the deflection was proportional to the strength of the magnetic field; and a magnetic field could be used to focus the rays into a beam. He also suggested the rays emitted by the cathode were a stream of charged tiny particles, which were soon to be identified by Joseph John Thomson as electrons. Continuing CRT experimentation lead to the discovery of x-rays by Wilhelm Konrad Röntgen in 1895.

Ferdinand Braun was the first person to envision the CRT as a tool for the display of information and is generally credited with the invention of the first device to be a direct forerunner to the modern CRT [5]. Braun in 1896 designed a CRT "indicator tube" for monitoring high frequencies in power-generating equipment. His design contained all of the same elements as today's CRTs. It utilized a cathode as an electron source, two control coils for vertical and horizontal deflection, an anode for electron acceleration and beam control, and a focusing slit. He also incorporated a phosphor screen normal to the beam for tracing its path. Although crude by later standards, this was the direct prototype for CRT oscilloscopes. In 1903–1905, Arthur Wehnelt added several very significant advances to Braun's design. Wehnelt developed and implemented a hot oxide-coated cathode and a beam control grid [5]. This lowered the voltage necessary to generate the electron stream and provided for much finer control of the beam current. These developments were the forerunner of the modern electron gun.

The CRT remained mostly a laboratory device until three important applications ushered it onto the center stage of information display: oscilloscopes, television, and military radar. By the early 1900s, oscilloscopes were being widely used for the study of time-varying electric circuits in communication. Television was developed in the 1920s and in July of 1930 the National Broadcasting Company began the experimental broadcast of television in New York City. During World War II, the CRT underwent a second wave of maturation with the advent of radar technology. By the end of the 1940s, the TV-CRT was available as a consumer product. In the 1950s, RCA perfected shadow mask technology and before long color became a new standard, gaining wide consumer acceptance in the 1960s. While the principal components of a CRT system have not fundamentally changed in many decades, the CRT has continued to improve.

98.3 Image Formation with a CRT

Although there are many different CRT designs, some of which are quite intricate, the process of forming an image using a CRT is straightforward. The procedure can be divided into four basic stages: beam formation, beam focusing, beam deflection, and energy conversion. These stages occur in four different regions of the CRT and follow in order from the rear of the tube (the neck) to the faceplate. As shown in Figure 98.1, the elements at the rear of the tube are collectively called the electron gun.

98.3.1 Beam Generation with the Electron Gun

The cathode generates the stream of electrons used to form the image-writing beam. The traditional cathode is a metal conductor such as nickel coated with a thin layer of oxide, typically a barium strontium compound.

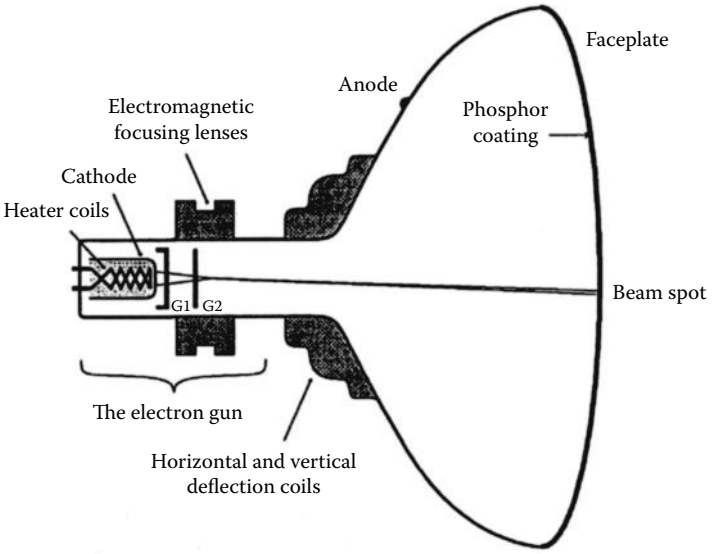


FIGURE 98.1 Example of a typical CRT with electromagnetic focus, and electromagnetic deflection. This type of tube design is normally used for demanding applications such as bright, high-resolution image projection. Tubes like this are available from a number of manufacturers such as Sony, Matsushita, and Thomson.

To reduce the voltage required to generate electron emission, the cathode is heated to 700 °C–1200 °C. Applications that require high brightness often use more advanced and expensive dispenser cathode designs to increase the beam current while maintaining reasonable cathode life. These designs incorporate complex cathode structures and materials such as barium ceramics, molybdenum, rhenium, and tungsten. Readers interested in advanced cathode designs and additional information on CRT materials technology should refer to References [6–10]. The flow of electrons from the cathode is controlled by varying the potential between the cathode and a series of control grids commonly known as G1 (the control grid) and G2 (the acceleration grid). A voltage potential of 100–1000 V between G2 and the cathode creates the potential necessary to pull a stream of electrons off the cathode, forming the beam. The beam amplitude can be controlled and even completely shut off by varying the potential on G1. Thus, the voltage at G1 controls brightness because the brightness is proportional to beam current. The design of the cathode with respect to impedance and loading influences the maximum rate at which the beam can be modulated. The cathode and its associated control grids can be designed to produce a crossover flow of electrons or a laminar flow. In crossover gun designs, the emitted electrons converge to a point in front of the cathode. By using electron optics, this beam spot is imaged onto the phosphor screen. Due to inherent advantages, the crossover design is widely used. A crossover beam is narrow, making it easier to deflect than a thicker beam, and the spot can be very small, improving resolution at the screen. In theory, a laminar flow design provides for the possibility of higher beam current from a similar-sized cathode. In practice, the improvement is not usually advantageous enough to offset the added difficulty of controlling a wider beam.

98.3.2 Electron Beam Focusing

Beam focusing and beam current are critical in determining the final spot size and thus the resolution of the CRT. Focusing a beam of electrons is directly analogous to focusing a beam of light; the discipline is called electron optics. Concerns familiar to optical imaging such as magnification, spherical aberration, and astigmatism also confront electron optics. As CRTs become larger, and operate at higher deflection angles, spot control becomes critical. Beam focusing is achieved using either electrostatic focusing grids or electromagnetic focusing coils. Electrostatic focus is the most extensively used technique. It can be

found in use in applications from television to desktop computer monitors. Electrostatic focus is achieved by applying a succession of potentials across a complex sequence of focusing grids built into the electron gun. As designers seek to improve performance further, grid designs have become intricate [11,12]. Magnetic focus is the system of choice for all high-performance systems where resolution and brightness are design objectives. Magnetic lenses are better at producing a small spot with few aberrations. External coils in a yoke around the neck of the tube control the beam. Since it provides superior performance, electromagnetic focus is common on high-resolution commercial systems. Magnetic focus can also be achieved using permanent magnets and specialized hybrid electrostatic/magnetic focus components. Due to the tremendous impact focus has on resolution, tube suppliers continue to improve focus control [13,14]. For an excellent and comprehensive treatment of electron physics in CRTs, beam control, detailed design discussions, and other aspects of CRT devices, refer to Sol Sherr's textbook [15].

98.3.3 Electron Beam Deflection

The beam deflection system is responsible for controlling the position of the spot on the front face of the CRT. As with focusing, beam deflection techniques can be electromagnetic or electrostatic. A magnetic deflection system consists of two electromagnetic yolks on opposite sides of the CRT neck: a horizontal deflection coil and a vertical deflection coil. The position of the beam is easily controlled; the amount of beam deflection is directly proportional to the current in the coil. Deflection coil design also influences spot size and shape. Coil designs can incorporate correction for coma, convergence errors, and pincushion distortion [16]. Because of its low cost and efficiency, CRTs for both moderate- and high-resolution applications normally use magnetic deflection. Electrostatic deflection provides faster beam displacement but less spot size control. It is typically used in oscilloscope systems, where resolution requirements are moderate and deflection speed is paramount.

98.4 CRT Addressing: Raster Scanning vs. Stroke

Raster scanning is the most common CRT addressing technique. In raster scanning, the electron beam writes each frame one line at a time. This means the horizontal deflection system requires a higher bandwidth than the vertical deflection system. High-performance raster CRTs have a horizontal bandwidth of 15–150 kHz, and a vertical deflection bandwidth of 30–180 Hz. The information is written onto the front face of the CRT from left to right and top to bottom, one line at a time. Raster scanning can be interlaced or progressive (noninterlaced). In interlace scanning, each frame of information is decomposed into two fields. The first field consists of all of the odd-numbered lines in the original frame, and the second field contains all of the even-numbered lines. Each field is scanned onto the CRT at twice the frame rate. Commercial television is interlaced. In the United States, the frame rate is 30 Hz and the field rate is 60 Hz; this scheme is known as the NTSC standard. PAL and SECAM are two other well-known commercial formats in use today. There are numerous other standards, both analog and digital, in use worldwide, each with slightly different scan rates, resolution/addressing formats, luminance and color encoding, and timing protocols [17,18]. Interlace scanning is an engineering compromise that conserves transmission bandwidth, electronics bandwidth, and CRT bandwidth, while maintaining acceptable performance with respect to moving imagery, resolution, and flicker. Computer monitors employ a progressive raster scan. Each frame has one field. The display is generated line by line, in order from the first pixel to the last pixel. Frame rates for the typical desktop monitor vary from 60 to 85 Hz to over 180 Hz for specialized systems.

Stroke is an alternative addressing technique that was once quite common. The name for the technique comes from the phrase "the stroke of a pen." Stroke is a point-to-point addressing system. The beam is directed to the starting point of a line, turned on, and then moved directly to the end of the line. Because there are no raster lines, stroke CRT resolution is independent of direction and is limited primarily by spot size. Stroke addressing is excellent for low-information-content screens with detailed graphic characters or simple vector graphics. It provides extremely high-quality drawing, clean lines,

precise detail, and high speed while conserving power and bandwidth. Stroke systems address the CRT only where there is information, not the entire screen as in raster scanning. This technique is uncommon today because it is too slow and computationally intensive for systems that have high information content, gray scale, or imagery.

98.5 Phosphor Screen

The phosphor functions as the CRT transducer. It converts the energy of the electron beam into light. This conversion process is called cathodoluminescence. CRT phosphors are inorganic crystalline materials doped with one or more impurities called activators and coactivators. Phosphors emit light in two ways: fluorescence and phosphorescence. Fluorescence is the emission of light by the phosphor material while it is under bombardment by the electron beam. The continued emission of light after the bombardment has ceased is called phosphorescence. The length of time phosphorescence lasts is known as persistence. Persistence can vary from tens of nanoseconds to many minutes, or even hours. CRTs take advantage of both forms of cathodoluminescence. Briefly, cathodoluminescence occurs when the electron beam excites the electrons of the phosphor into higher, unstable energy states available due to the presence of the activators. Light is emitted, when the electrons transition back to their stable states. The choice of phosphor depends on the requirements of the application with respect to wavelength characteristics (narrow emission spectra or broadband emission), color, brightness, resolution, and persistence. Commercial television CRTs typically make use of the following phosphor powders: ZnS:Ag:Cl (blue), Zn(Cd)S:Cu:Al or ZnS:Cu:Au:Al (green), and Y_2O_2S:Eu [19]. Television CRTs use moderately short-persistence phosphors. This ensures that a new frame does not exhibit blurring due to the previous frame. Traditionally, phosphors for radar displays, where the screen is refreshed infrequently, had a mix of short and long persistence. However, with the advent of today's digital systems, the use of long-persistence phosphors has declined. The Electronics Industries Association maintains an information database of commercial phosphors for CRTs. The interested reader should refer to its publication TEP116-C [20]. Figure 98.2 illustrates the spectral characteristics of several typical phosphors.

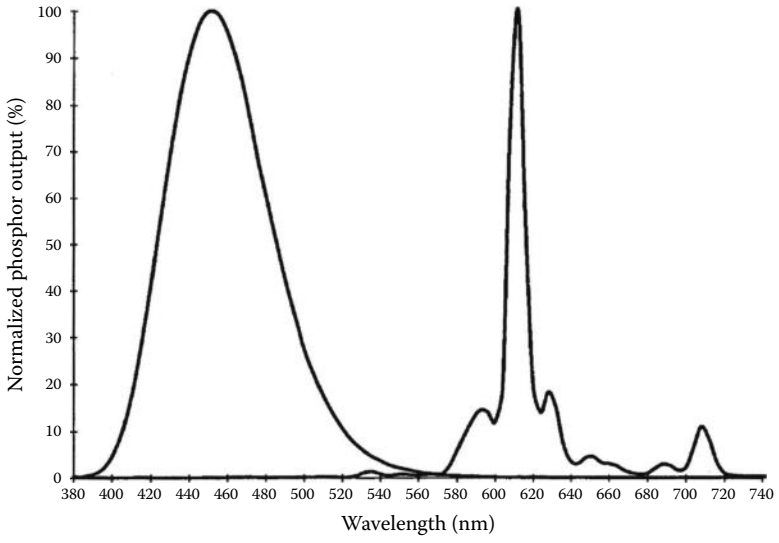


FIGURE 98.2 Phosphors can have a wide range of spectral characteristics. This graph illustrates the spectral profiles of two representative cases. The blue source is an example of a phosphor with a wide and smooth profile. The maximum output for this phosphor is in the blue at a wavelength of 452 nm. The red phosphor shows a sharper profile with several narrow peaks. The largest peak for this phosphor is at 612 nm.

98.6 Color CRTs Using Shadow Masks

Color has been a part of the CRT world since the early 1950s. A sound understanding of human color perception is important for a detailed understanding of color in CRTs. Readers interested in an excellent comprehensive treatment of color science should refer to Wyszecki and Stiles [21] or Robertson and Fisher [22] for a brief overview. The most common method of introducing color to the CRT is the three-gun shadow mask technique. The shadow mask is a metal grid of many tiny apertures held in place immediately before the phosphor screen. The phosphor screen is an array of three different phosphor subpixels. Three electron guns are placed closely together in the neck of the tube, one gun for each one of the primary colors (red, green, and blue). The guns make a small angle with respect to each other. The shadow mask and beams are aligned so that each beam falls on the appropriate phosphor dot. The dots are placed closely together so the eye spatially integrates the subpixels into a continuous blend of color. There are two common configurations: delta and in-line. Delta, the traditional method, arranges the electron guns and the phosphor dots in a triad. This requires a shadow mask consisting of a grid of circular apertures. More recently, Sony introduced the Trinitron™ design. In the Trinitron tube the guns are placed in line, the shadow mask consists of vertical slots, and the phosphors are in vertical stripes. The design offers improved vertical resolution and is easier to converge, but the mask is slightly more difficult to support and is subject to thermal stress problems.

Trinitron continues to be an important and popular design [23]. The shadow mask technique of producing a color image has three main drawbacks. First, the shadow mask typically absorbs >75% of the electron beam energy, limiting brightness and causing undesirable thermal effects. Second, both require precise beam control to converge the three different images from the electron guns. Third, the resolution of the display is limited by the requirements of spatial color; three subpixels are needed to make each full-color pixel.

98.7 Alternative Techniques for Realizing Color Using CRTs

There are alternative approaches to producing color with CRTs. In its own way, each seeks to improve upon one or more of the enormously successful shadow mask designs. The three most noteworthy challengers are beam index tubes, penetration phosphor tubes, and field sequential color (FSC) systems. All three have been around for decades, but only FSC systems are commercially available.

98.7.1 Beam Index and Penetration Phosphor Tube Designs

A beam index tube uses a slotted mask and vertical phosphor stripes. One electron beam is focused to less than the width of a phosphor stripe. A reflected ultraviolet signal from the mask provides a feedback signal. This signal is used to index the position of the beam very accurately. Color selection is achieved by precise beam positioning and rapid beam modulation. Although beam indexing offers perfect convergence, few guns, energy savings, and high resolution, practical problems in control and manufacturing have left this approach unrealized. Penetration tubes are potentially even more advantageous since color is produced with one electron gun and no mask is required. The phosphor screen consists of a layering of several different phosphors. There are many approaches to the layering structure, but the basic principle is that each layer produces a different color. To select among the different colors, the beam energy is varied, altering its layer penetration. In theory, multiple layers of phosphors could be used to produce a full-color display. Unfortunately, this design requires high switching voltages, and the color purity is poor because of the dilution caused by leakage from the unselected phosphor layers. In the past, a few two-color systems have been produced on a limited basis with stroke writer systems.

98.7.2 Field Sequential Color

FSC is a different approach to realizing color using a monochrome CRT and color filters [24]. The shadow mask relies on the ability of the eye to perform a *spatial integration on a group of color elements*. FSC exploits

the ability of the eye to perform a *temporal integration of color fields*. FSC is not a new concept; it was an early design suggestion for commercial color television. To implement FSC, each full-color frame is decomposed into primary color fields (red, green, and blue). Each of these color fields is displayed sequentially on a monochrome CRT while simultaneously a color filter is placed over the front of the CRT. The filters must be rapidly changed to stay in sequence with the individual fields which are displayed at three times the desired frame rate. This sequence is so quick that the eye fuses the individual color fields into a steady blend of continuous color. FSC systems have all of the advantages of a monochrome CRT, namely, superior resolution and simplicity. The major historical drawbacks to FSC have been the difficulty in controlling a bulky, spinning, mechanical filter wheel, and two-thirds of the light energy of the phosphor is thrown away in each field. In the past, some viewers have reported the ability to perceive the individual color fields intermittently; this artifact is known as "color breakup." FSC systems have seen a commercial rebirth due to several factors. The luminous efficiency of several phosphors has improved to the point where the light loss penalty is no longer a severe price to pay. Awkward, mechanical filter wheels have been replaced by several different types of compact, low-power, liquid crystal shutters. The use of faster frame rates and digital sampling techniques has reduced color artifacts. Finally, monochrome CRTs using FSC meet the growing market need for small high-resolution displays.

98.8 Image Quality and Performance

For any display technology, the ultimate test of performance and quality is how the image looks to the viewer in a real-world environment. This is highly subjective, but critical none the less. Table 98.1 highlights some important parameters associated with performance and image quality. Resolution is the most

TABLE 98.1 Parameters That Influence CRT Performance and Image Quality

Resolution	Spot size and shape
	Focus accuracy
	Line width
	Addressing/scan format
	Shadow mask design
Luminance	Dynamic range
	Maximum brightness
	Gray scale in high and low ambient light
	Gray scale vs. display brightness
	Phosphor luminous efficiency
Contrast	Contrast under high ambient illumination
	Contrast under low ambient illumination
	Large area contrast
	Pixel-to-pixel contrast (small area)
Image fidelity	Frame rate
	Flicker/refresh
	Uniformity and linearity
	Aberrations (linearity, pincushion, barrel distortion, keystone, and focus)
Color	Phosphor selection
	Convergence accuracy
	Color gamut and saturation
	Color uniformity
	Color accuracy
Bandwidth	Video bandwidth (cathode design)
	Horizontal deflection design
	Vertical deflection design
	Bandwidth of the CRT supporting electronics
	Bandwidth of the signal delivery electronics

critical characteristic of any CRT. A note of caution, resolution should not be confused with addressing. A system can have a precise addressing format, but if the spot is too large, the CRT resolution capability will be poor despite the addressing format. Alternatively, if the spot is precise and the addressing format is coarse, the *image* will be low resolution, but the CRT resolution will be high. Resolution is affected by almost every system parameter: spot size, focus, deflection, raster format, the mask, and beam current. For a perfectly designed system, line width and spot size will ultimately determine the resolution. Spot diameter is limited by the quality of the electron optics, the particle size and thickness of the phosphor layer, and beam current. The cross-section energy distribution of the electron beam is approximately Gaussian. The higher the current is the larger the beam cross section. Scattering by individual phosphor particles also increases the spot size. Thus, high resolution is much more difficult to achieve for high-brightness applications. The champions of small spot sizes are the small 12–25 mm monochrome CRTs. Several commercial systems achieve ≤25 μm spot sizes, and sizes ≤15 μm have been reported [25]. Monochrome CRTs in the 53 cm class (21 in.) have spot diameters ≤150 μm. Color CRTs with shadow mask pitches ≤250 μm are available in sizes up to 21 in. at commodity prices, and this will continue to improve. Spot size does not tell the whole resolution story. In most systems, something other than spot size limits performance. The scan format may be well below the performance of the CRT, the bandwidth of the support electronics is frequently less than the tube bandwidth, and, for color, the shadow mask will limit resolution more than spot size. Engineers can still design a CRT that maintains its performance even with state-of-the-art electronics. Modulation transfer function (MTF) is the standard metric for resolution. MTF compares the modulation of the displayed output with the modulation of a known input signal. As the spatial frequency of the input increases, the performance of the output will show a steady roll-off in modulation depth and quality. Readers interested in an introduction to this complex topic should refer to Infante's review [26]. Contrast and its related companion, gray scale, are also important to CRT image quality. Since contrast can be measured and stated in so many different terms, gray scale is probably a more practical indicator of performance from a user's standpoint. As it attempts to replace traditional film, the medical imaging community demands the most of the monochrome CRT. Medical CRTs can produce up to 256 shades of gray with 8 bit controllers. Systems are available that will do 10 bit monochrome and color. Although frequently quoted, *shades of gray* are not a true indication of CRT performance and should not be confused with true *gray-scale* capability. Shades of gray are a characteristic of the image source and the capability of the electronics; gray scale reflects the capability of the CRT. Gray scale is typically defined as a series of steps in brightness in increments of $2^{1/2}$ from black to maximum CRT output. Thus, gray scale is a realistic indicator that combines CRT dynamic range, contrast ratio, halation, and ambient reflections. More than 12 steps of gray scale are considered good performance.

The subject of display image quality and evaluation is an important topic. In addition to resolution and gray scale, users and designers need to consider many other metrics, such as luminance, dimmability, contrast, readability, color, persistence, and convergence. Users may also be aware of visual distortions such as pincushion, barrel, and keystone. A discussion of these topics and the other issues raised by Table 98.1 is beyond the scope of a single chapter; indeed, it has been the subject of books. There are good sources of information in this area [15,27]; in particular, Keller's [28] text is superb. It is essential to remember that no matter which metrics are used, the most single most important metric is the subjective determination of the end users. Although many factors will affect final display performance, the bottom line remains *how good does the display look* in the real-world application environment it was designed to meet.

98.9 CRT and Phosphor Lifetime

The lifetime of a CRT depends on how it is driven. The higher the beam current, the more rapidly its performance will degrade. In general, degradation is due to a failure of three parts of the CRT—the cathode, the phosphor, and the glass. Brighter displays require higher beam current, and this, in turn, requires a higher level of cathode loading. Oxide cathodes can comfortably operate with a loading range

of 0.1–10 A/cm^2. This will provide a typical lifetime in the range of 10,000 h. Dispenser cathode designs can improve cathode life two to three times, but again individual cathode life depends on how much brightness the user regularly demands from the tube. Phosphor degradation is the second mechanism affecting CRT lifetime and performance. All phosphors degrade under constant electron bombardment, an effect called phosphor burning. The rate of degradation is related to beam current, anode voltage, and material parameters specific to each phosphor. Today, there is a broad range of good phosphor materials with greatly improved aging characteristics. The third mechanism is glass browning. This effect is a complicated interaction between the high energy of the electron beam and the molecular structure of the glass [10]. In recent years, glass suppliers have made great strides in providing CRT manufacturers with improved glass materials. However, phosphor aging and glass browning are still important concerns, especially in extremely bright (high beam current) applications, such as CRT projection.

98.10 CRT Strengths and Weaknesses

With a century of development behind it, the CRT has abundant advantages that make it an attractive display option for the system designer dealing with high information content or dynamic imagery. Table 98.2 lists some attributes, positive and negative, which the system designer should consider in evaluating whether or not to use CRT technology. The most notable traits of the CRT are flexibility and cost. CRTs can operate over a wide range of resolution formats from commercial television to the strict requirements of medical imaging and digital-to-film recording. One unit can easily be designed to accept a myriad of scan formats and resolutions from static text to dynamic imagery with megapixels at frame rates of hundreds of frames per second. This flexibility means the system can be upgraded later without having to redesign the display subsystem. Interfacing with the CRT is also straightforward; there are well-established standards and hardware for many scan formats from the television NTSC to the RGB standards of the personal computer industry. Finally, cost is important to emphasize; CRTs are inexpensive. The desktop CRT delivers its performance for $\leq$\$0.00005 per color resolution element! Currently, no other high-information-content display technology can match this cost.

TABLE 98.2 Strength and Weakness of CRT Technology

Advantage	Weakness
Highest resolution technology	High power consumption
Versatile addressing formats	Weight of glass
Excellent image fidelity	Size of footprint
High speed	Emits EM radiation
Bandwidth ($\geq$350 MHz)	Must be refreshed
Excellent contrast/gray scale	
Good selection of phosphors	
High luminous efficiency	
Bright display (up to 20,000 fL)	
Excellent color gamut	
Simplicity of interface	
Universal interface standards	
Good design flexibility	
Broad application base	
Long life and reliability	
Mature and broad knowledge base	
Worldwide sources of tubes and parts	
Inexpensive—low cost per resolution element	

TABLE 98.3 Companies That Manufacture Cathode Ray Tubes

Clinton Electronics Corporation
6701 Clinton Road
Rockford, Illinois 61111
Tel. 815-633-1444, Fax. 815-633-8712

Hitachi, Ltd.
New Marunouchi Bldg.
Marunouchi 1-chrome
Chiyoda-ku, Tokyo 100, Japan
Tel. 81-3-3212-1111, Fax. 81-3-3212-3857
U.S. address:
Hitachi America, Ltd., Electron Tube Division
3850 Holcomb Bridge Road, Suite 300
Norcross, Georgia 30092-2202
Tel. 770-409-3000, Fax. 770-409-3028

Hughes Lexington, Inc.
A subsidiary of Hughes Electronics Company
1501 Newtown Pike
Lexington, Kentucky 40511
Tel. 606-243-5500, Fax. 606-243-5555

Image Systems Corporation
11595 K-tel Drive
Hopkins, Minnesota 55343
Tel. 612-935-1171, Fax 612-935-1386

Imaging & Sensing Technology Corporation
300 Westinghouse Circle
Horseheads, New York 14845-2299
Tel. 607-796-4400, Fax. 607-796-4482

ITPO Institute of Surface Engineering and Optoelectronics
Teslova 30
1000 Ljubljana, Slovenia
Tel. 386-61 1264 592/111, Fax. 386-61 1264 593

L.G. Electronics
20 Yoido-dong Youngdungpo-gu
Seoul, Korea
E-mail display2@www.goldstar.co.kr or monitor@www.goldstar.co.kr
U.S. address:
L G Electronics, Monitors Division
1000 Silvan Avenue
Englewood Cliffs, New Jersey 07632
Tel. 201-816-2000, Fax. 201-816-2188

Matsushita Electric Industrial Ltd.
Twin 21 National Tower
1-61, Shiromi, 2-Chome, Chuo-ku
Osaka 540, Japan
Tel. (06) 908-1121
U.S. address:
Panasonic Industrial Company
Computers and Communications Division
2 Panasonic Way
Secaucus, New Jersey 07094
Tel. 201-392-4502, Fax. 201-392-4441

Phillips Netherlands, BV
Phillips Components and Semiconductors
Bldg. VB Postbus 90050
5600 PB Eindhoven, The Netherlands
Tel. 040-783749, Fax. 040-788399
U.S. address:
Discrete Products Division
2001 West Blue Heron Blvd.
Riviera Beach, Florida 33404
Tel. 407-881-3200 or 800-447-3762, Fax. 407-881-3300

Rank Brimar, Ltd.
Greenside Way
Middleton, Manchester M24 1SN, England
Tel. 0161-681 7072, Fax. 0161-682-3818
U.S. address:
25358 Avenue Stanford
Valencia, California 91355-1214
Tel. 805-295-5770, Fax. 805-295-5087

Sony Corporation
Display Systems
16550 Via Esprillo
San Diego, California 92127
Tel. 619-487-8500

Thomas Electronics
100 Riverview Drive
Wayne, New Jersey 07470
Tel. 201-696-5200, Fax. 201-696-8298

Thomson Tubes Electronics
13, avenue Morane Saulnier
Bâtiment Chavez—Vélizy Espace
BP 121/F-78148 VELIZY CEDEX France
Tel. 33-1 30 70 35 00, Fax. 33-1 30 70 35 35
U.S. address:
Thompson Components and Tubes Corporation
40 G Commerce Way
Totowa, New Jersey 07511
Tel. 201-812-9000, Fax. 201-812-9050

Toshiba Corporation
Electronic Components, Cathode Ray Tube Division
1-1 Shibaura 1-Chrome, Minato-KU
Tokyo 105, Japan
Tel. 03-457-3480, Fax. 03-456-1286
U.S. address:
Toshiba America, Inc.
1220 Midas Way
Sunnyvale, California 94086-4020
Tel. 408-737-9844

98.11 CRT Selection

The design requirements for (1) resolution, (2) brightness, and (3) screen size will quickly steer the system designer to a particular class of CRTs. If the design requirements are close to either television or desktop computer monitors, the choice is easy. There are literally hundreds of vendors offering integrated CRTs and electronic drivers as off-the-shelf items. Color shadow mask systems with >40 elements per cm are commonly available. If higher resolution is required, then medical imaging and film recorder CRTs may meet the requirements. Physically large display requirements can be supplied by projection CRTs or by tiled CRTs using video wall processors. Small to miniature CRTs provide resolutions up to 500 to 600 elements/cm and portability. If the application falls out of the mainstream, the designer will need to speak directly with CRT vendors. Table 98.3 is a partial list of CRT suppliers. Since CRTs have been designed to satisfy a broad range of specialty applications, the chances are excellent that an acceptable design will already exist. If not, the CRT is so mature that vendors working from a proven experience base can design a tube directly to customer specifications with little in the way of design problems or surprises.

98.12 Future Trends in CRT Technology

CRT technology is mature, and the industry does not foresee major surprises on the horizon. The CRT will continue to evolve, resolution will improve, screens will be flatter, deflections angles will be larger, and footprints will be reduced. The computer world continues to demand more resolution. In 1987, the 36 cm (14 in.) VGA class monitor (640 × 480) with its 300,000 pixels and 16 colors was the new standard. In 1998, the 43 cm (17 in.) desktop monitor is now capable of 1600 × 1200 performance (1.9 megapixels) with 24 bit color. The 53 cm (21 in.) monitor class, once an expensive custom device, is also widely available. The trend of increasing resolution and performance for decreasing cost shows every sign of continuing. The medical community is quickly moving to monochrome systems with 12 bit gray scale at a resolution of 2560 × 2048 (5 megapixels). All of this means the instrument designer has better resolution and fidelity for less money.

There is a quiet revolution taking place in the information display industry, which may have a significant effect on CRTs and all information display technology. There is a design trend to decouple the resolution and fidelity of the display device from its physical size. There are a number of applications motivating this change: communications, military, portable computing, simulation, and virtual reality. This is why miniature CRT technology has been more active in recent years. This is also the engine driving some of the major competitors to the miniature CRT, such as on-chip field emission displays (FED), deformable mirror devices (DMD), ferroelectric liquid crystal integrated circuit displays, and on-chip active matrix liquid crystal display (LCD) technology. However, the CRT backed by its formidable 100 years of design evolution and maturity is not standing still; its assets and market remains. Although it is one of the few vacuum tubes still in use today, the traditional CRT is not doomed for obsolescence any time in the immediate future.

References

1. M. Gazziro, M. Fernandes, L. Almeida, P. Matias, and J. Slaets, Expanding the horizontal capabilities of CRT monitors using artificial inter-pixel steps for neuroscience experiments, *Digital Signal Processing*, 22(2): 367–375, 2012.
2. G. Schueller et al., Comparison of 5-megapixel cathode ray tube monitors and 5-megapixel liquid crystal monitors for soft-copy reading in full-field digital mammography, *European Journal of Radiology*, 76(1): 68–72, October 2010.
3. G. Shires, Ferdinand Braun and the cathode ray tube, *Sci. Am*., 230(3): 92–101, March 1974.
4. N. H. Lehrer, The challenge of the cathode-ray tube, in L. E. Tannas, Jr., Ed., *Flat Panel Displays and CRTs*, New York: Van Nostrand Reinhold, 1985.

5. P. Keller, *The Cathode-Ray Tube, Technology, History, and Applications*, New York: Palisades Press, 1991.
6. W. Harris, How CRT and LCD monitors work, Bit-Tech, http://www.bit-tech.net/hardware/2006/03/20/how_crt_and_lcd_monitors_work/1, accessed on August 23, 2013.
7. L. R. Falce, CRT dispenser cathodes using molybdenum rhenium emitter surfaces, *Society for Information Display International Symposium Digest of Technical Papers*, 23: 331–333, 1992.
8. J. H. Lee, J. I. Jang, B. D. Ko, G. Y. Jung, W. H. Kim, K. Takechi, and H. Nakanishi, Dispenser cathodes for HDTV, *Society for Information Display International Symposium Digest of Technical Papers*, 27: 445–448, 1996.
9. T. Nakadaira, T. Kodama, Y. Hara, and M. Santoku, Temperature and cutoff stabilization of impregnated cathodes, *Society for Information Display International Symposium Digest of Technical Papers*, 27: 811–814, 1996.
10. W. Kohl, *Materials Technology for Electron Tubes*, New York: Reinhold Publishing, 1951.
11. S. Sugawara, J. Kimiya, E. Kamohara, and K. Fukuda, A new dynamic-focus electron gun for color CRTs with tri-quadrupole electron lens, *Society for Information Display International Symposium Digest of Technical Papers*, 26: 103–106, 1995.
12. J. Kimiya, S. Sugawara, T. Hasegawa, and H. Mori, A 22.5 mm neck color CRT electron gun with simplified dynamically activated quadrupole lens, *Society for Information Display International Symposium Digest of Technical Papers*, 27: 795–798, 1996.
13. D. Imabayashi, M. Santoku, and J. Karasawa, New pre-focus system structure for the trinitron gun, *Society for Information Display International Symposium Digest of Technical Papers*, 27: 807–810, 1996.
14. K. Kato, T. Sase, K. Sasaki, and M. Chiba, A high-resolution CRT monitor using built-in ultrasonic motors for focus adjustment, *Society for Information Display International Symposium Digest of Technical Papers*, 27: 63–66, 1996.
15. S. Sherr, *Electronic Displays*, 2nd edn., New York: John Wiley & Sons, 1993.
16. N. Azzi and O. Masson, Design of an NIS pin/coma-free 108° self-converging yoke for CRTs with super-flat faceplates, *Society for Information Display International Symposium Digest of Technical Papers*, 26: 183–186, 1995.
17. J. F. Fisher and R. G. Clapp, Waveforms and spectra of composite video signals, in K. Benson and J. Whitaker, *Television Engineering Handbook, Featuring HDTV Systems*, New York: McGraw-Hill Reinhold, 1992.
18. D. Pritchard, Standards and recommended practices, in K. Benson and J. Whitaker, *Television Engineering Handbook, Featuring HDTV Systems*, New York: McGraw-Hill Reinhold, 1992.
19. A. Rapaport, J. Millez, M. Bass, A. Cassanho, and H. Jemssen, Review of the properties of up-conversion phosphors for new emissive display, *Journal of Display Technology*, 7(2): 77–83, 2011.
20. Optical characteristics of cathode ray tube screens, EIA publication TEP 116-C, February 1993.
21. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd edn., New York: John Wiley & Sons, 1982.
22. A. Robertson and J. Fisher, Color vision, representation, and reproduction, in K. Benson and J. Whitaker, Eds., *Television Engineering Handbook, Featuring HDTV Systems*, New York: McGraw-Hill Reinhold, 1992.
23. M. Maeda, Trinitron technology: Current status and future trends, *Society for Information Display International Symposium Digest of Technical Papers*, 27: 867–870, 1996.
24. C. Sherman, Field sequential color takes another step, *Inf. Display*, 11(3): 12–15, March, 1995.
25. L. Ozawa, Helmet mounted 0.5 in. crt for SVGA images, *Society for Information Display International Symposium Digest of Technical Papers*, 26: 95–98, 1995.
26. D. P. Chakrabotry, J. Fan, and H. Roehrig, Measuring CRT display image quality: Effects of phosphor type, pixel contrast and luminance, *Proceedings of the SPIE*, 5034: 153–163, 2003.

27. J. Whitaker, *Electronic Displays, Technology, Design, and Applications*, New York: McGraw-Hill, 1994.
28. P. Keller, *Electronic Display Measurement, Concepts, Techniques, and Instrumentation*, New York: John Wiley & Sons, 1997.

Further Information

Electronic Industries Association (EIA), 2500 Wilson Blvd., Arlington, VA 22201 (Internet: www.eia.org). The Electronic Industries Association maintains a collection of over 1000 current engineering publications and standards. The EIA is an excellent source for information on CRT engineering, standards, phosphors, safety, market information, and electronics in general.

Ozawa, L., *Cathodoluminescence: Theory and Applications*, New York: Kodansha, 1990.

The Society for Information Display (SID), 1526 Brookhollow Dr., Suite 82, Santa Ana, CA 92705-5421 (Internet: www.display.org). The Society for Information Display is a good source of engineering research and development information on CRTs and information display technology in general.

Wandell, B., The foundations of color measurement and color perception, *Society for Information Display International Symposium, Conference Seminar M-I*, 1993. A nice brief introduction to color science (31 pages).

Zworykin, V. K. and G. A. Morton, *Television: The Electronics of Image Transmission in Color and Monochrome*, New York: John Wiley & Sons, 1954.

Internet Resources

The following is a brief list of places to begin looking on the World Wide Web for information on CRTs and displays, standards, metrics, and current research. Also many of the manufacturers listed in Table 98.3 maintain websites with useful information.

The Society for Information Display
www.display.org

The Society of Motion Picture and Television Engineers
www.smpte.org

The Institute of Electrical and Electronics Engineers
www.ieee.org

The Electronic Industries Association
www.eia.org

National Information Display Laboratory
www.nta.org

The International Society for Optical Engineering
www.spie.org

The Optical Society of America
www.osa.org

Electronics & Electrical Engineering Laboratory
www.eeel.nist.gov

National Institute of Standards and Technology (NIST)
www.nist.gov

The Federal Communications Commission
www.fcc.gov

Appendix: Units and Conversions

B.W. Petley
National Physical Laboratory

This appendix contains several tables that list the SI base units (Table A.1), define the SI base units (Table A.2), list their derived units (Table A.3), list their prefixes (Table A.4), and list their conversion units (Table A.5).

TABLE A.1 SI Base Units

Base Quantity	Name of Base Unit	Symbol
Length	Meter	m
Mass	Kilogram	kg
Time	Second	s
Electric current	Ampere	A
Thermodynamic temperature	Kelvin	K
Amount of substance	Mole	mol
Luminous intensity	Candela	cd

TABLE A.2 International Definitions of the SI Base Units[a]

Unit of length (m)

Meter is the length of the path traveled by light in vacuum during a time interval of 1/299 792 458 of a second (17th CGPM,[b] 1983, Resolution 1).

Note: The original international prototype, made of platinum–iridium, is kept at the BIPM[c] under conditions specified by the 1st CGPM in 1889.

Unit of mass (kg)

Kilogram is the unit of mass: it is equal to the mass of the international prototype of the kilogram (3rd CGPM, 1901).

Unit of time (s)

Second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom (13th CGPM, 1967, Resolution 1).

Unit of electric current (A)

Ampere is that constant current that, if maintained in two straight parallel conductors of infinite length, of negligible cross section, and placed 1 m apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} N/m of length (CIPM, 1946, Resolution 2 approved by the 9th CGPM, 1948).

Note: The expression "MKS unit of force" that occurs in the original text has been replaced here by "newton," a name adopted for this unit by the 9th CGPM (1948), Resolution 7.

Unit of thermodynamic temperature (K)

Kelvin, unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water (13th CGPM, 1967, Resolution 4).

The 13th CGPM (1967, Resolution 3) also decided that the unit kelvin and its symbol K should be used to express an interval or a difference in temperature.

Note: In addition to the thermodynamic temperature (symbol T), expressed in kelvin, use is also made of Celsius temperature (symbol t) defined by the equation.

$t = T - T_0$

where $T_0 = 273.15$ K by definition. To express Celsius temperature, the unit "degree Celsius" that is equal to the unit "kelvin" is used; in this case, "degree Celsius" is a special name used in place of "kelvin." An interval or difference of Celsius temperature can, however, be expressed in kelvins as well as degrees Celsius.

Unit of amount of substance (mol)

1. Mole is the amount of substance of a system that contains as many elementary entities as there are atoms in 0.012 kg of carbon-12.
2. When the mole is used, the elementary entities must be specified and may be atoms, molecules, ions, electrons, other particles, or specified groups of such particles.

In the definition of the mole, it is understood that unbound atoms of carbon-12, at rest and in their ground state, are referred to.

Note: This definition specifies at the same time the nature of the quantity whose unit is the mole.

Unit of luminous intensity (cd)

Candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} Hz and that has a radiant intensity in that direction of (1/683) watt per steradian (16th CGPM, 1979, Resolution 3).

[a] The United States denotes the unit of length by "meter" in place of the international usage of "meter."

[b] CGPM, Conférence Général de Poids et Mesures; CIPM, Comité International des Poids et Mesures.

[c] BIPM: Bureau International des Poids et Mesures.

TABLE A.3 SI-Derived Units with Special Names[a]

Derived Quantity	Name	Symbol	Expressed in Terms of Other Units	Expressed in Terms of SI Base Units
Plane angle	Radian	rad		$m\ m^{-1}$
Solid angle	Steradian	sr		$m^2\ m^{-2}$
Frequency	hertz	Hz		s^{-1}
Force	newton	N		$m\ kg\ s^{-2}$
Pressure, stress	pascal	Pa	$N\ m^{-2}$	$m^{-1}\ kg\ s^{-2}$
Energy, work, quantity of heat	joule	J		$m^2\ kg\ s^{-2}$
Power, radiant flux	watt	W		$m^2\ kg\ s^{-3}$
Electric charge, quantity of electricity	coulomb	C		s A
Electric potential, potential difference, electromotive force	volt	V	W/A	$m^2\ kg\ s^{-3}\ A^{-1}$
Capacitance	farad	F	C/V	$m^{-2}\ kg^{-1}\ s^4\ A^2$
Electric resistance	ohm	Ω	V/A	$m^2\ kg\ s^{-3}\ A^{-2}$
Electric conductance	siemens	S	A/V	$m^{-2}\ kg^{-1}\ s^3\ A^2$
Magnetic flux	weber	Wb	V s	$m^2\ kg\ s^{-2}\ A^{-1}$
Magnetic flux density	tesla	T	Wb/m^2	$kg\ s^{-2}\ A^{-1}$
Inductance	henry	H	Wb/A	$m^2\ kg\ s^{-2}\ A^{-2}$
Celsius temperature	Degree Celsius	°C		K
Luminous flux	Lumen	lm	cd sr	$cd\ m^2\ m^{-2} = cd$
Illuminance	Lux	lx	m^{-2} cd sr	m^{-2} cd
Activity (referred to a radio nuclide)	becquerel	Bq		s^{-1}
Absorbed dose, specific energy imparted, kerma	gray	Gy	J/kg	$m^2\ s^{-2}$
Dose equivalent, ambient dose equivalent, organ equivalent dose	sievert	Sr	J/kg	$m^2\ s^{-2}$

[a] Note that when a unit is named after a person, the *symbol* takes a capital letter and the name takes a lowercase letter.

TABLE A.4 SI Prefixes[a]

Factor	Prefix	Symbol	Factor	Prefix	Symbol
10^{24}	Yotta	Y	10^{-1}	Deci	d
10^{21}	Zetta	Z	10^{-2}	Centi	c
10^{18}	Exa	E	10^{-3}	Milli	m
10^{15}	Peta	P	10^{-6}	Micro	μ
10^{12}	Tera	T	10^{-9}	Nano	n
10^{9}	Giga	G	10^{-12}	Pico	p
10^{6}	Mega	M	10^{-15}	Femto	f
10^{3}	Kilo	k	10^{-18}	Atto	a
10^{2}	Hecto	h	10^{-21}	Zepto	z
10	Deca	da	10^{-24}	Yocto	y

[a] The 11th CGPM (1960, Resolution 12) adopted a first series of prefixes and symbols of prefixes to form the names and symbols of the decimal multiples and submultiples of SI units. Prefixes for 10^{-15} and 10^{-18} were added by the 12th CGPM (1964, Resolution 8), those for 10^{15} and 10^{18} were proposed by the 15th CGPM (1975, Resolution 10), and those for 10^{21}, 10^{24}, 10^{-21}, and 10^{-24} were proposed by the CIPM (1990) for approval by the 19th CGPM (1991).

TABLE A.5 Conversion Factors from English Measures to SI Units[a]

Unit	Equivalent
1. *Acceleration*	
Acceleration of free fall, standard gravity	9.806 65 m s^{-2}
1 ft/s^2	0.304 8 m s^{-2}
1 gal	0.01 m s^{-2}
2. *Angle*	
1 second (″)	$4.484\ 81 \times 10^{-6}$ rad
1 minute (′)	$2.908\ 9 \times 10^{4}$ rad
1 degree (°)	0.0174 532 rad
1 rad	206 264.8″
3. *Area*	
1 barn (b)	10^{-28} m^2
1 $in.^2$	$6.451\ 6 \times 10^{-4}$ m
1 ft^2	0.092 903 04 m^2
1 yd^2	0.836 127 36 m^2
1 are	100 m^2
1 acre [43560 (statute ft)^2]	4046.86 m^2
1 hectare	10,000 m^2
1 mi^2	$2.590\ 0 \times 10^{6}$ m^2
1 square mile (based on U.S. survey foot)	2.589 998 km^2
4. *Concentration, density, mass density*	
1 grain/gal (United States)	0.017 118 kg/m^3
1 lb/ft^3	16.018 46 kg/m^3
1 lb/gal (United States)	119.826 4 kg/m^3
1 short ton/yd^3	1186.6 kg/m^3
1 long ton/yd^3	1328.9 kg/m^3
1 oz (avdp)/$in.^3$	1730.0 kg/m^3
1 oz (avd)/gal (United States)	7.489 152 kg/m^3
1 $lb/in.^3$	27,680 kg/m^3
5. *Energy*	
1 ft lbf	1.355 818 J
1 cal_{th} (thermochemical calorie)	4.184 J
1 cal_{15} (15 °C calorie)	4.185 5 J
1 cal_{IT}[b]	4.186 8 J
1 kilocalorie (nutrition)[c]	4.186.8 J
1 watt second (W s)	1 J
1 watt hour (W h)	3600 J
1 therm (EC)	$1.055\ 06 \times 10^{8}$ J
1 therm (United States)	$1.054\ 804 \times 10^{8}$ J
1 ton TNT (equivalent)	4.184×10^{9} J
1 BT_{th}	1 054.350 J
1 Btu_{15}	1 054.728 J
1 Btu_{ST}	1 055.055 852 62 J
1 quad (=10^{15} Btu)	≈10^{18} J = 1 EJ

TABLE A.5 (continued) Conversion Factors from English Measures to SI Units[a]

Unit	Equivalent
6. *Force*	
1 dyne	10^{-5} N
1 ounce-force	0.278 013 9 N
1 pound-force	4.448 222 N
1 kilogram-force	9.806 65 N
1 kip (1000 lbf)	4448.222 N
1 ton-force (2000 lbf)	8.896 443 N
7. *Fuel consumption*	
1 gallon (United States) per horsepower hour	$1.410\,089 \times 10^{-9}$ m^3/J
1 gallon (United States)/mile	2.352 15 1/km
1 gallon (United Kingdom)/mile	2.824 81 1/km
1 mile/gallon (United States), mpg	0.425 144 km/L
1 mile/gallon (United Kingdom)	0.354 006 km/L
1 pound per horsepower	$1.689\,659 \times 10^{-7}$ kg/J
1 1/(100 km)	235.215/(mpg) (United States)
8. *Length*	
1 fermi	10^{-15} m = 1 fm
1 angstrom (Å)	10^{-10} m
1 microinch	2.54×10^{-8} m
1 mil	2.54×10^{-5} m
1 point (pt) [0.013837 in.][d]	—
1 pica (12 pt)	4.217 5 mm
1 inch (in.)	0.025 4 m
1 hand (4 in.)	0.101 6 m
1 foot (12 in.) (0.999998 statute ft.)	0.304 8 m
1 foot (U.S. survey)	0.304 800 6 m
1 statute foot [(1200/3937) m]	0.304 800 6 m
1 yard (yd)	0.914 4 m
1 fathom (6 ft, U.S. survey)	1.828 8 m
1 rod (16.5 statute ft)	5.029 2 m
1 chain (4 rod)	20.116 8 m
1 furlong (10 chain)	201.168 m
1 mile (8 furlong, 5280 ft)	1609.344 m
1 statute mile (8 furlong, 5280 statute ft)	1609.347 2 m
1 nautical mile (international)[e]	1852 m
1 light-year[f]	$9.640\,73 \times 10^{15}$ m
9. *Light*	
1 foot-candle	10.763 91 lx
1 phot	10,000 lx
1 $cd/in.^2$	1550.003 cd/m^2
1 footlambert	3.426 259 cd/m^2
1 lambert	3183.099 cd/m^2
1 stilb	10,000 cd/m^2

(*continued*)

TABLE A.5 (continued) Conversion Factors from English Measures to SI Units[a]

Unit	Equivalent
10. *Mass*	
1 pound (avdp.) (lb) (7000 g)	0.453 592 37 kg
1 pound (troy) (5760 g)	0.373 241 721 6 kg
1 grain (g)	64.798 91 mg
1 scruple (20 g)	1.296 0 g
1 pennyweight (24 g)	1.555 174 g
1 dram (60 g)	3.887 9 g
1 ounce (avdp) (437.5 g)	28.349 52 g
1 ounce (troy) (480 g)	31.103 48 g
1 carat (metric)	0.2 g
1 stone (14 lb)	6.350 29 kg
1 slug	14.593 9 kg
1 hundredweight (long)	50.802 35 kg
1 ton (short) (2000 lb)	907.184 7 kg
1 ton (long) (2240 lb)	1016.047 kg
	1.016 047 t
Mass per unit length	
1 tex	10^{-6} kg/m
1 denier	$1.111\,111 \times 10^{-7}$ kg/m
1 pound per foot	1.488 164 kg/m
1 pound per inch	17.857 97 kg/m
1 ton/mile	0.631 342 mg/km
1 ton/1000 yd	1.111 6 kg/m
1 lb/ft	1.488 16 kg/m
Mass per unit area	
1 ton/mile^2	3.922 98 kg/ha
1 ton/acre	2510.71 kg/ha
1 oz/yd^2	33.905 7 g/m^2
Mass carried × distance (traffic factor)	
1 ton mile	1635.17 kg km
Mass carried × distance/volume (traffic factor)	
1 ton mile/gal (United States)	431.967 6 mg km/m^3
11. *Power*	
1 erg/s	10^{-7} W
1 ft lbf/h	$3.766\,161 \times 10^{-4}$ W
(1 Btu_{ST})	1.000 669 Btu_{th}
1 metric horsepower (force de cheval)	735.498 8 W
1 horsepower (550 ft lbf/s)	745.70 W
1 electric horsepower	746 W
12. *Pressure, stress*	
1 standard atmosphere	101 325 Pa
1 dyne/cm^2	0.1 Pa
1 torr [(101th 325/760) Pa]	133.322 4 Pa
1 N/cm^2	10,000 Pa
1 bar	100,000 Pa
1 lbf/ft^2	47.880 26 Pa
1 lbf/in.^{2} (psi)	6894.8 Pa

TABLE A.5 (continued) Conversion Factors from English Measures to SI Units[a]

Unit	Equivalent
1 kgf/cm^2	98 066.5 Pa
1 cm water (4 °C)	98.063 7 Pa
1 mm of mercury (0 °C)	133.322 4 Pa
1 in of water (39.2 °F)	249.082 Pa
1 in of mercury (60 °F)	3376.85 Pa
1 ft water (39.2 °F)	2988.98 Pa
13. *Thermal quantities*	
Fixed points	
Triple point of natural water T_{tp}	273.16 K
Zero Celsius (= $T_0 = t_{F,0}$)	273.15 K = 32 °F
Temperature conversions	
Kelvin to Rankine (T_R)	$T = (5/9)\ T_R$
Kelvin to Celsius	$t = T - T_0$
Kelvin to Fahrenheit	$t_F = (9/5)(T - T_0) + t_{F,0}$
Celsius to Fahrenheit	$t_F = (9/5)\ t + t_{F,0}$
[Numerically, $5(\{t_F\} + 40) = 9(\{t\} + 40)$, where $\{t\}$ and $\{t_F\}$ are the numerical values of the Celsius and Fahrenheit temperatures, respectively.]	
Temperature interval conversions	
1 degree centigrade	1 degree Celsius, denoted 1 °C
1 °C	1 K
1 °F	(1/1.8) K
1 °R	(1/1.8) K
Other thermal quantities	
1 Btu$_{th}$/h	0.292 875 W
1 Btu$_{IT}$/h	0.293 071th 1 W
1 cal$_{IT}$/s	4.186 8 W
1 cal$_{th}$/s	4.184 W
1 cal$_{IT}$/(g°C)	4186.8 J/(kg K)
1 Btu ft/(ft^2 h°F)	1.730 735 W m^{-1} K^{-1}
1 Btu in./(ft^2 s°F)	519.220 4 W m^{-1} K^{-1}
1 clo	0.155 m^2 K/kW
1 °F h ft^2/Btu	0.176 110 2 K m^2/W
1 °F h ft^2/Btu in	6.933 472 K m/W
1 Btu/lb°F ≡ 1 cal$_{ST}$/g°C	4186.8 J/kg.K
14. *Torque, moment of force*	
1 dyne·cm	10^{-7} N m
1 kgf·m	9.806 65 N m
1 ozf·in.	0.007 061 552 N m
1 lbf·in.	0.112 984 8 N m
1 lbf·ft	1.355 818 N m
15. *Velocity (includes speed)*	
1 foot per hour	8.466 667 × 10^{-5} m/s
1 foot per minute	0.005 08 m/s
1 knot (nautical mile per hour)	0.514 444 m/s
1 mile per hour (mi/h)	0.447 04 m/s

(*continued*)

TABLE A.5 (continued) Conversion Factors from English Measures to SI Units[a]

Unit	Equivalent
16. *Viscosity*	
1 poise	0.1 Pa s
1 ft^2/s	0.092 903 04 m^2/s
1 lb/(ft s)	1.488 164 Pa s
1 lb/(ft h)	$4.133\,789 \times 10^{-4}$ Pa s
1 lbf s/ft^2	47.880 26 Pa s
1 lbf·s/$in.^2$	6894.757 Pa s
1 rhe	10 $Pa^{-1}\ s^{-1}$
1 slug/ft s	47.880 26 Pa s
1 stokes, St	1.0×10^{-4} m^2/s
17. *Volume (includes capacity)*	
1 stere, st	1 m^3
1 liter[g]	0.001 m^3
1 ft^3	0.028 316 8 m^3
1 $in.^3$	$1.638\,7 \times 10^{-5}$ m^3
1 board foot	$2.359\,7 \times 10^{-3}$ m^3
1 acre-foot	1233.48 m^3
1 dram (U.S. fluid)	$3.696\,7 \times 10^{-6}$ m^3
1 gill (United States)	$1.182\,941 \times 10^{-4}$ m^3
1 ounce (U.S. fluid)	$2.957\,353 \times 10^{-5}$ m^3
1 teaspoon (tsp)[h]	$4.9288\,922 \times 10^{-6}$ m^3
1 tablespoon (tbsp)	$1.4787\,676 \times 10^{-5}$ m^3
1 pint (U.S. fluid)	$4.731\,765 \times 10^{-4}$ m^3
1 quart (U.S. fluid)	$9.463\,529 \times 10^{-4}$ m^3
1 gallon (U.S. liquid) [231 $in.^3$]	$3.785\,412 \times 10^{-3}$ m^3
1 wine barrel (bbl) [31.5 gal (United States)]	0.119 240 m^3
1 barrel (petroleum, 42 gal, United States), bbl	0.158 987
1 ounce (U.K. fluid)	$2.841\,3 \times 10^{-5}$ m^3
1 gill (Canada and United Kingdom)	$1.420\,6 \times 10^{-4}$ m^q
1 gallon (Canada and United Kingdom)	$4.546\,09 \times 10^{-3}$ m^3
	1.200 950 gal (United States)
1 pint (U.S. dry)	$5.506\,105 \times 10^{-4}$ m^3
1 quart (U.S. dry)	$1.101\,221 \times 10^{-3}$ m^3
1 gallon (U.S. dry)	$4.404\,884 \times 10^{-3}$ m^3
1 peck	$8.809\,768 \times 10^{-3}$ m^3
1 bushel (United States) [2150.42 $in.^3$]	$3.523\,907 \times 10^{-2}$ m^3

[a] The conversion factor for a compound unit is usually not given here if it may easily be derived from simpler conversions; e.g., the conversion factors for "ft/s" to "m/s" or "ft/s^2" to "m/s^2" are not given, since they may be obtained from the conversion factor for "ft." Values are given to five or six significant digits except for exact values, which are usually indicated in bold type. A few former cgs measures are also included.

[b] The International Steam Table calorie of 1956.

[c] In practice, the prefix kilo is usually omitted. The kilogram calorie or large calorie is an obsolete term for the kilocalorie that is used to express the energy content of foods.

[d] Typographer's definition, 1886.

[e] Originally, in 1929, the international nautical mile.

[f] Based on 1 day = 86,400 s and 1 Julian century = 36,525 days.

[g] Post-1964 value, SI symbol l or L. Between 1901 and 1964, the liter was defined as 1.000th 028 dm^3.

[h] Although often given, it is doubtful whether normal usage justifies this accuracy. In Europe and elsewhere, teaspoon and tablespoon are usually exactly 5 and 15 mL, respectively.

Bibliography

Anderton P. and Bigg P. H., *Changing to the Metric System*, HMSO, London, U.K., 1980.

CIPM, Proces-Verbaux CIPM, 49th Session, 1960, pp. 71–72; *Comptes Rendues, 11th CGPM*, 1960, p. 85.

Cohen E. R., *The Physics Quick Reference Guide*, American Institute of Physics Press, New York, 1995.

Standard for Use of the International System of Units (*SI*): *The Modern Metric System, IEEE/ASTM SI 10–1997*, IEEE Standards Coordinating Committee 14 (Revision and redesignation of ANSI/IEEE Std 268–1992 and ASTM E380), IEEE, New York, 1997.

Taylor B. N., *Interpretation of the SI for the United States and Metric Conversion Policy for Federal Agencies*, Natl. Inst. Stand. Technol., Spec. Publ. 814, U.S. Government Printing Office, Washington, DC, 1991.

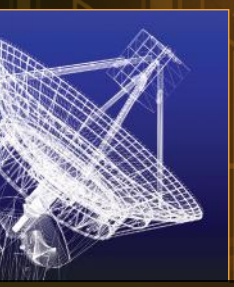
Taylor B. N., *Guide for the Use of the International System of Units*, 1995 ed., Natl. Inst. Stand. Technol., Spec. Publ. 811, U.S. Government Printing Office, Washington, DC, 1995.

The International System of Units, 7th edn., BIPM, Sèvres, France, 1998.

The International System of Units (*SI*), Natl. Inst. Stand. Technol., Spec. Publ. 330, 1991 ed., U.S. Government Printing Office, Washington, DC, 1991.

The NIST Guide for the Use of the International System of Units. http://www.nist.gov/pml/pubs/sp811/index.cfm (last accessed August 23, 2013).

Electrical Engineering



SECOND EDITION

Measurement, Instrumentation, and Sensors Handbook

Electromagnetic, Optical, Radiation, Chemical, and Biomedical Measurement

The **Second Edition** of the bestselling **Measurement, Instrumentation, and Sensors Handbook** brings together all aspects of the design and implementation of measurement, instrumentation, and sensors. Reflecting the current state of the art, it describes the use of instruments and techniques for performing practical measurements in engineering, physics, chemistry, and the life sciences and discusses processing systems, automatic data acquisition, reduction and analysis, operation characteristics, accuracy, errors, calibrations, and the incorporation of standards for control purposes.

Organized according to measurement problem, the **Electromagnetic, Optical, Radiation, Chemical, and Biomedical Measurement** volume of the **Second Edition**:

- Contains contributions from field experts, new chapters, and updates to all 98 existing chapters
- Covers sensors and sensor technology, time and frequency, signal processing, displays and recorders, and optical, medical, biomedical, health, environmental, electrical, electromagnetic, and chemical variables

A concise and useful reference for engineers, scientists, academic faculty, students, designers, managers, and industry professionals involved in instrumentation and measurement research and development, **Measurement, Instrumentation, and Sensors Handbook, Second Edition: Electromagnetic, Optical, Radiation, Chemical, and Biomedical Measurement** provides readers with a greater understanding of advanced applications.

K12210

CRC Press
Taylor & Francis Group
an informa business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
711 Third Avenue
New York, NY 10017
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

ISBN: 978-1-4398-4891-3
90000
9 781439 848913