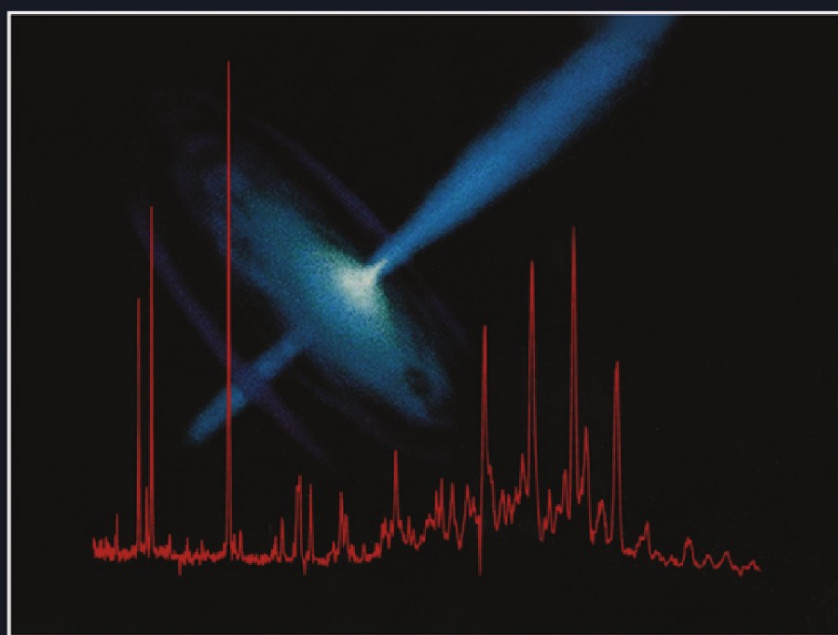


Saas-Fee Advanced Course 30
2000

Swiss Society
for Astrophysics and Astronomy

S. M. Kahn P. von Ballmoos
R. A. Sunyaev

High-Energy Spectroscopic Astrophysics



S.M. Kahn P.von Ballmoos R.A.Sunyaev

High-Energy Spectroscopic Astrophysics

Saas-Fee Advanced Course 30
2000

Swiss Society for Astrophysics and Astronomy
Edited by M. Güdel and R. Walter

With 90 Figures

 Springer

Steven M. Kahn

Department of Physics
Stanford University
Stanford, CA 94305-4060, USA

Rashid A. Sunyaev

Max-Planck-Institut für Astrophysik
Karl-Schwarzschild-Str. 1
85748 Garching, Germany

Peter von Ballmoos

Centre d'Etude Spatiale des Rayonnements
9 avenue du Colonel-Roche
31028 Toulouse, France

Volume Editors:

Manuel Güdel

Paul Scherrer Institut
Würenlingen and Villigen
5232 Villigen PSI, Switzerland

Roland Walter

Integral Science Data Centre
1290 Versoix, Switzerland

This series is edited on behalf of the Swiss Society for Astrophysics and Astronomy:
Société Suisse d'Astrophysique et d'Astronomie

Observatoire de Genève, ch. des Maillettes 51, 1290 Sauverny, Switzerland

Cover picture: The background picture illustrates an artist's view of an accretion disk and jets in an active galactic nucleus (courtesy of G. Dana Berry, Space Telescope Science Institute, Baltimore). The red curve shows a calculated soft X-ray spectrum of a hot cosmic plasma as recorded by a grating spectrometer.

Library of Congress Control Number: 2005921507

ISBN 3-540-40501-1 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springeronline.com

© Springer-Verlag Berlin Heidelberg 2005
Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready copy from authors/editors
Data conversion by TechBooks
Cover design: *design & production* GmbH, Heidelberg

Printed on acid-free paper 55/3141/mh-543210

Preface

The 30th Saas-Fee Advanced Course, held at Les Diablerets, Switzerland, between 3 and 8 April, 2000, symbolizes the beginning of a new era in high-energy astrophysics. Only 8 months earlier, NASA's Chandra X-ray Observatory began operations, and only 4 months before the course, ESA's cornerstone mission XMM-Newton was launched successfully. The first results were presented during the lectures, comprising splendid pictures and the first high-resolution spectra from cosmic X-ray sources. Soon to come were a suite of complementary high-energy missions, covering the adjacent hard X-ray and gamma-ray regimes. ESA's INTEGRAL mission was under construction, and so was NASA's High-Energy Solar Spectroscopic Imager, HESSI. Both satellites have been launched meanwhile and both provide excellent data from hard X-ray and gamma-ray sources. These new observatories bring to maturity many fields of research related to high-energy processes across the universe. After a few years, all missions have met our boldest expectations, with many new discoveries being made on a regular basis.

The timing of this Saas Fee course was ideal. After three decades of intense research in X-ray and gamma-ray astronomy, the time was ripe to summarize basic knowledge on X-ray and gamma-ray spectroscopy for interested students and researchers ready to become involved in the new missions. The main purpose of this course was to communicate the scientific basics and methods of high-energy spectroscopic astrophysics. These methods are surprisingly similar in and common to all of its disciplines, illuminating our common interest to understand energetic processes in the universe in general. The emphasis was therefore on physical principles and observing methods rather than on discussions of particular classes of high-energy objects. In this spirit, the three speakers presented excellent lectures discussing topics from physical processes to instrumentation.

Steven M. Kahn's lectures on soft X-ray spectroscopy reviews the large field of atomic physics in low-density cosmic plasmas from a strict quantum mechanics point of view. He discusses details of ionization and recombination processes, atomic transitions, and equilibria relevant for the interpretation of soft X-ray spectra from cosmic sources.

Peter von Ballmoos in his series of lectures presents the basic science of detector and telescope systems for high-energy astrophysics. Probably in no

other area of astronomy is the precise understanding of detector characteristics as important as in the field of gamma-ray astronomy where incoming photons transform the telescope and detector structures themselves into radiating sources.

Rashid Sunyaev presents a comprehensive review of fundamental processes in high-energy plasmas, concentrating on radiation processes in extreme environments such as magnetospheres of neutron stars, accretion disks around black holes, or plasma in active galactic nuclei. Much emphasis is put on comptonization mechanisms. We deeply regret that Prof. Sunyaev was not willing to send us his complete version of the manuscript. The version printed in this book unfortunately lacks all the figures foreseen for the article.

The end of the course also marked the 100th anniversary of a very significant event without which these lectures would not have happened: often forgotten nowadays, on April 9, 1900, the French physicist Paul Villard found that radium emitted some very penetrating radiation; he had discovered (and named) the gamma rays!

The immediate interest in the present course and the current esteem for the presented topics was reflected in the very large number of participants that reached the capacity limits of the Saas Fee course. A total of 142 persons (speakers and organizers included) registered for this course, and despite a few cancellations more than 130 people came to Les Diablerets. Several participants also took the opportunity to combine this course with the “INTEGRAL spring school” that was organized at the same place during the preceding week.

The organizers wish to thank the three speakers for their great enthusiasm and their brilliant lectures. A magnificent concert with Italian music of the 17th century was given at Vers l’Église on Tuesday night. We are much indebted to the excellent performers. Organizing such a course is impossible without the help of a number of people. We extend our warmest thanks to our secretary, Martine Logossou, for her management of correspondence, registrations, and the budget. Pascal Favre provided invaluable help with editing part of the source text. We thank Marc Audard for the photographs taken during the course and reproduced in this volume with permission. The Eurotel-Victoria hotel has as usual provided a splendid environment to make course participation a pleasure. The rich banquet dinner remains unforgettable. And last but not least we thank the Swiss Academy of Natural Sciences for its substantial financial contributions, and the Swiss Society for Astrophysics and Astronomy for its continuing support of the Saas-Fee course series.

Zürich and Versoix
February 2005

*Manuel Güdel
Roland Walter*



Contents

Soft X-Ray Spectroscopy of Astrophysical Plasmas	
<i>S.M. Kahn</i>	3
1 Introduction	3
1.1 The Role of X-Ray Spectroscopy in Astrophysics	4
1.2 Characteristics of Cosmic X-ray Sources	5
2 Classical and Quantum Radiation Theory	9
2.1 Introduction	9
2.2 Overview of the Classical Equations	10
2.3 Electromagnetic Waves	11
2.4 The Classical Multipole Expansion	12
2.5 The Classical Oscillator	15
2.6 Quantum Radiation Theory – Overview	17
2.7 The Radiation Hamiltonian	19
2.8 Bound-Free Absorption (Photoionization)	20
2.9 Bound-Bound Transitions	21
2.10 The Quantum Multipole Expansion	22
2.11 Spontaneous Emission	23
3 The Structure of Multi-Electron Atoms	24
3.1 Introduction	24
3.2 Hydrogen-like Ions	25
3.3 Scaling with Nuclear Charge	28
3.4 Relativistic Corrections	30
3.5 The Central Field Approximation and Quantum Indistinguishability	31
3.6 Electron Exchange – Helium-like Atoms	33
3.7 Approximation Techniques for Multi-Electron Atoms	35
3.8 LS, jj and Intermediate Coupling	37
3.9 Spectroscopic Notation and Ground-State Configurations	38
3.10 Configuration Interaction	40
3.11 Selection Rules for Radiative Transitions	40
4 Electron-Ion Collisional Processes	41
4.1 Overview	41
4.2 Collisional Excitation – Scattering Theory	44
4.3 Collisional Excitation – Classical Estimate	47

4.4	Collisional Ionization	49
4.5	Radiative Recombination	51
4.6	Dielectronic Recombination and Autoionization	53
5	Types of Equilibria	57
5.1	Properties of LTE	58
5.2	Coronal Equilibrium	60
5.3	X-Ray Photoionization Equilibrium	62
5.4	Thermal Instability in Photoionized Plasmas	67
6	Discrete Line Diagnostics	70
6.1	Lyman Series Transitions in H-like Ions	71
6.2	He-like Transitions	74
6.3	Iron L-Shell Transitions	75
6.4	The Iron K-Shell Complex	78
7	Concluding Remarks	80
	References	81

Instruments for Nuclear Astrophysics

<i>P. von Ballmoos</i>	83
1 Introduction	83
1.1 The Instrumental Development of Gamma-Ray Astrophysics .	84
1.2 From Gamma-Ray Astronomy to Nuclear Astrophysics	91
1.3 Requirements on Instruments for Gamma-Ray Spectroscopy ..	96
2 Interaction of High Energy Photons with Matter	98
2.1 Photoelectric Effect	102
2.2 Scattering from Free Electrons	103
2.3 Scattering from Bound Electrons	105
2.4 Optical Properties of Materials: Reflection and Refraction....	112
2.5 Pair Production	114
2.6 The Spectral Signatures of Energy Loss Processes	117
2.7 Characterizing the Detector Response	119
3 Detectors.....	121
3.1 Gas-filled Detectors	122
3.2 Scintillators	133
3.3 Semiconductor Detectors	143
4 The Instruments for Nuclear Astronomy.....	149
4.1 Geometric Optics: Modulating Aperture Systems	149
4.2 Quantum Optics: Compton Telescopes.....	168
4.3 Wave Optics: Focusing Telescopes.....	180
References	193

Hard X-Ray and Gamma Ray Spectroscopy

<i>R. Sunyaev and S. Sazonov</i>	199
1 Fundamentals of Compton Scattering	200
1.1 Photon Frequency Shift upon Scattering from a Free Electron	200

1.2	Scattering Cross Section	201
1.3	Radiation Force	204
1.4	Energy Exchange Between Plasma and Radiation	209
2	Comptonization in Infinite Homogeneous Media	212
2.1	Analytic Approximations for the Compton Scattering Kernel .	213
2.2	Kompaneets Equation	217
2.3	Plasma Heating and Cooling	221
2.4	Analytic Results for the Homogeneous Problem	221
2.5	Induced Compton Scattering	223
2.6	Photon Production Mechanisms	228
3	Comptonization in Bounded Plasma Clouds	234
3.1	Spatial Problem	235
3.2	Distribution of Photons over the Escape Time	235
3.3	Solution of the Stationary Equation of Comptonization	237
3.4	Solution by the Convolution Method	239
3.5	Double Compton Effect as Source of Low Frequency Photons .	241
3.6	Monte Carlo Calculations of Comptonization Spectra	242
3.7	Bulk Comptonization	243
4	Interaction of X-Rays with Partially Ionized Media	249
4.1	X-Ray Reflection	250
4.2	Scattering of X-Ray Lines on Neutral Hydrogen and Helium ..	256
5	6.4-keV Fluorescent Emission from Molecular Clouds in the Galactic Center	264
5.1	Surface Brightness Distribution of the Neutral and Ionized Iron Line Emission	265
5.2	Sgr B2 Giant Molecular Cloud	266
5.3	X-Ray Archaeology: Activity of Sgr A* in the Recent Past ...	268
6	X-Ray Emission from Supernova 1987A	269
6.1	Analytic Solution of the Problem	270
7	Accretion onto Black Holes and Neutron Stars	272
7.1	Introduction	272
7.2	Efficiency of Accretion onto a Rapidly Rotating Neutron Star	274
7.3	Structure of the Boundary Layer	275
7.4	Time Variability in the Accretion Disk and in the Boundary Layer	277
	References	278
	Index	285

List of Previous Saas-Fee Advanced Courses

- !! 2002 The Cold Universe
A.W. Blain, F. Combes, B.T. Draine
- !! 2000 High-Energy Spectroscopic Astrophysics
S.M. Kahn, P. von Ballmoos, R. Sunyaev
- !! 1999 Physics of Star Formation in Galaxies
F. Palla, H. Zinnecker
- * 1998 Star Clusters
B.W. Carney, W.E. Harris
- * 1997 Computational Methods for Astrophysical Fluid Flow
R.J. LeVeque, D. Mihalas, E.A. Dorfi, E. Müller
- * 1996 Galaxies Interactions and Induced Star Formation
R.C. Kennicutt, F. Schweizer, J.E. Barnes
- * 1995 Stellar Remnants
S.D. Kawaler, I. Novikov, G. Srinivasan
- * 1994 Plasma Astrophysics
J.G. Kirk, D.B. Melrose, E.R. Priest
- * 1993 The Deep Universe
A.R. Sandage, R.G. Kron, M.S. Longair
- * 1992 Interacting Binaries
S.N. Shore, M. Livio, E.J.P. van den Heuvel
- * 1991 The Galactic Interstellar Medium
W.B. Burton, B.G. Elmegreen, R. Genzel
- * 1990 Active Galactic Nuclei
R. Blandford, H. Netzer, L. Woltjer
- ! 1989 The Milky Way as a Galaxy
G. Gilmore, I. King, P. van der Kruit
- ! 1988 Radiation in Moving Gaseous Media
H. Frisch, R.P. Kudritzki, H.W. Yorke
- ! 1987 Large Scale Structures in the Universe
A.C. Fabian, M. Geller, A. Szalay
- ! 1986 Nucleosynthesis and Chemical Evolution
J. Audouze, C. Chiosi, S.E. Woosley
- ! 1985 High Resolution in Astronomy
R.S. Booth, J.W. Brault, A. Labeyrie

XIV List of Previous Saas-Fee Advanced Courses

- ! 1984 Planets, Their Origin, Interior and Atmosphere
D. Gautier, W.B. Hubbard, H. Reeves
- ! 1983 Astrophysical Processes in Upper Main Sequence Stars
A.N. Cox, S. Vauclair, J.P. Zahn
- * 1982 Morphology and Dynamics of Galaxies
J. Binney, J. Kormendy, S.D.M. White
- ! 1981 Activity and Outer Atmospheres of the Sun and Stars
F. Praderie, D.S. Spicer, G.L. Withbroe
- * 1980 Star Formation
J. Appenzeller, J. Lequeux, J. Silk
- * 1979 Extragalactic High Energy Physics
F. Pacini, C. Ryter, P.A. Strittmatter
- * 1978 Observational Cosmology
J.E. Gunn, M.S. Longair, M.J. Rees
- * 1977 Advanced Stages in Stellar Evolution
I. Iben Jr., A. Renzini, D.N. Schramm
- * 1976 Galaxies
K. Freeman, R.C. Larson, B. Tinsley
- * 1975 Atomic and Molecular Processes in Astrophysics
A. Dalgarno, F. Masnou-Seeuws, R.V.P. McWhirter
- * 1974 Magnetohydrodynamics
L. Mestel, N.O. Weiss
- * 1973 Dynamical Structure and Evolution of Stellar Systems
G. Contopoulos, M. H  non, D. Lynden-Bell
- * 1972 Interstellar Matter
N.C. Wickramasinghe, F.D. Kahn, P.G. Metzger
- * 1971 Theory of the Stellar Atmospheres
D. Mihalas, B. Pagel, P. Souffrin

* Out of print

! May be ordered from Geneva Observatory
Saas-Fee Courses
Geneva Observatory
CH-1290 Sauverny
Switzerland

!! May be ordered from Springer



Steven M. Kahn

Soft X-Ray Spectroscopy of Astrophysical Plasmas

S.M. Kahn

Columbia University, New York, USA

1 Introduction

These lectures are intended to provide a review of the basic physics necessary for the interpretation of high resolution soft X-ray spectra of astrophysical sources. While many of the topics I discuss can be found at the requisite level of sophistication in standard textbooks on atomic physics and spectroscopy (e.g. [1]), I have made an attempt to highlight those aspects which are especially important for X-ray transitions, and which are relevant at the characteristic temperatures and densities typically found in various types of X-ray emitting astrophysical plasmas.

My emphasis is on discrete atomic transitions, which dominate the spectra of most cosmic sources in the soft X-ray band ($100\text{ eV} \leq E \leq 10\text{ keV}$). I do not discuss basic continuum processes like bremsstrahlung, synchrotron emission, and inverse Compton emission, as these are covered well in the usual texts used to introduce students to radiative processes in astrophysics (e.g. [2]).

In general, I avoid long derivations, concentrating instead on the key physical ideas that underlie the various formulas, and especially on the definition of terms that appear frequently in the atomic physics literature. The level is intended for advanced undergraduates and beginning graduate students with little or no background in X-ray spectroscopy. While I do assume a rudimentary familiarity with the basics of classical and quantum physics (typical of the preparation one would receive as an undergraduate physics major in an American university), the lectures are self-contained, and were designed to provide a suitable introduction to this field without the need for extensive consultation of other source materials.

The organization is as follows: in the remainder of this initial chapter, I provide a brief introduction to the role of X-ray spectroscopy in astrophysics, and the physical conditions in various types of cosmic X-ray sources. Chapters 1 through 3 cover the essentials of atomic physics: classical and quantum radiation theory, atomic structure, and electron-ion collisional processes, respectively. In Chap. 4, I discuss the various types of equilibria that apply in astrophysical plasmas, and in Chap. 5, I provide a relatively brief review of the most important discrete-line spectral diagnostics that fall in the soft X-ray band. Chapter 6 includes a set of concluding remarks and some thoughts on where this field might be headed in the future.

1.1 The Role of X-Ray Spectroscopy in Astrophysics

X-ray astronomy is not a “new” field of research. Most practitioners date its inception to the serendipitous detection of the very bright binary X-ray source, Scorpius X-1, in June of 1962 [3]. That momentous discovery proved that cosmic systems could be copious X-ray emitters, and that observations in the X-ray band could provide new insights into astrophysical phenomena that could not be gleaned from observations at longer wavelengths. In the ensuing forty years, this field has grown to become one of the major disciplines of observational astrophysics. Hundreds of thousands of discrete sources of X-ray emission have been detected, covering nearly all classes of astrophysical systems.

Until very recently, however, real X-ray *spectra* of astrophysical sources, with sufficient resolution and sensitivity to enable the investigation of individual atomic features, had been largely unavailable. This was principally due to instrumental limitations. Since cosmic X-ray sources are exceedingly faint (typical fluxes for sources of interest are $\sim 10^{-3}$ phot cm $^{-2}$ s $^{-1}$ keV $^{-1}$), early experiments required large area detectors with very high efficiency for photon detection. Gas proportional counters were the instruments of choice. In the soft X-ray band, the spectral resolution achievable with such devices is extremely limited: $E/\Delta E \sim \text{few}$. While the data obtained with those experiments did provide some measure of the overall shapes of cosmic X-ray spectra, they could not be used to derive any real constraints on physical conditions in source emission regions.

The situation improved significantly in the mid 1990’s with the launch of the *ASCA* Observatory. This was the first mission to incorporate charge-coupled device (CCD) detectors at the focus of an X-ray telescope. The energy resolution of CCDs is roughly an order of magnitude better than that achievable with proportional counters. That enabled the detection of broad “humps” in the spectra, which could loosely be identified with complexes of emission lines from particular ions. Yet detailed spectral constraints could still only be derived from model fits to the spectra – even CCD resolution was insufficient to allow for direct interpretation of the intensities of individual features. Hence, the true power of spectroscopy had still not been realized.

Shortly before these lectures were delivered, however, the National Aeronautics and Space Administration launched the *Chandra* X-ray Observatory (June 1999), and the European Space Agency launched the *XMM-Newton* Observatory (December 1999). These two magnificent space missions both incorporate diffraction grating spectrometers, with resolving powers $E/\Delta E \geq 200$ across most of the soft X-ray band. They have collectively provided the first high resolution X-ray spectra of a wealth of astrophysical sources. This has created a revolution in this field, whose significance, even as of this writing two years later, is still continuing to be appreciated. In some cases, the data have provided striking confirmation of existing astrophysical

models. In others, they have presented significant challenges to our basic understanding of the sources involved.

Why is soft X-ray spectroscopy an important tool for astronomy?

There are several unique features of the soft X-ray band that play a role. First, X-ray emitting gas is often the “key” component of the astrophysical system. For many objects (e.g. elliptical galaxies, clusters of galaxies), the virial temperature, $kT \sim GMm_p/R$, lies in the range 10^6 – 10^8 K, where most of the emission comes out at soft X-ray energies. In others (e.g. supernova remnants, binary sources), shocks heat gas into the same temperature regimes. Second, the conventional soft X-ray band (0.1–10 keV) is unusually rich in discrete spectral features. The K-shell transitions of carbon through iron, and the L-shell transitions of silicon through iron fall in this range. In contrast to other wave bands, all charge states are visible in a single X-ray spectrum. This makes the interpretation of the spectrum relatively unambiguous. For example, one can derive relative elemental abundances without invoking any assumptions about the thermal state of the gas. Finally, because of the high radiative decay rates of X-ray transitions, astrophysical emitting plasmas are generally not in local thermodynamic equilibrium. This means that the details of the observed spectra depend on the explicit mechanisms by which the levels are populated. While that can occasionally lead to complications in the interpretation of the data, it also implies that they are quite sensitive to physical conditions in the source. Hence, X-ray spectra have high diagnostic utility.

Astrophysical X-ray spectroscopy can also be of interest as a probe of fundamental physics issues in unusual environments. In particular, cosmic plasmas can achieve extremely low densities, $n_e < 10^{-3} \text{ cm}^{-3}$, orders of magnitude below the densities found in the best vacuum obtainable in a laboratory. At such low densities, radiative decays from very long-lived metastable levels are important. In addition, the time scales for equilibration can be very long in comparison to the lengths of our observations. This means that for some sources, the emitting plasmas appear “frozen” in non-equilibrium states. Finally, given the vast physical scales characteristic of astronomical systems, we can find interesting examples of non-negligible optical depth for exotic absorption and scattering processes.

1.2 Characteristics of Cosmic X-ray Sources

An extensive review of the general science of X-ray astronomy is well beyond the scope of these lectures. However, I believe it is useful, in this introductory chapter, to provide a very brief accounting of physical conditions in the various types of cosmic X-ray sources we are studying with our spectroscopic experiments. More complete discussions of all of these topics can be found in a series of conference proceedings that have appeared within the past year [4].

General introductions to X-ray astronomy, suitable for non-specialists, have been written recently by Schlegel [5] and Tucker & Tucker [6].

Late-Type Stars

X-ray emission from late-type stars (stars of spectral type F, G, K and M) is believed to be produced in coronae, tenuous collections of hot gas confined by magnetic field lines above the stellar photospheres. The best known example, of course, is the solar corona, which was first detected in X-rays by a rocket experiment in 1951. The X-ray luminosity of the quiescent solar corona is $\sim 2 \times 10^{27} \text{ erg s}^{-1}$, which is only of order a part in a million of the total solar luminosity. The characteristic temperature is $\sim 2 \times 10^6 \text{ K}$, and the characteristic electron density is $\sim 10^9 \text{ cm}^{-3}$. However, the Sun turns out to be a rather weak X-ray source. More active late-type stars exhibit X-ray luminosities as high as $10^{32} \text{ erg s}^{-1}$, with temperatures $\sim$ several 10^7 K , and densities that can reach 10^{14} cm^{-3} . Coronal plasmas are optically thin to photoelectric absorption, although line optical depths for the highest oscillator strength lines can be greater than unity. Most active stars exhibit flares, which can increase the luminosities by three to four orders of magnitude on time scales of minutes to hours. There are many issues associated with the formation and energization of stellar coronae that are still poorly understood, making this an active area of research.

Early-Type Stars

Massive early-type stars (spectral types O and B) do not possess the outer convective zones believed to provide the dynamo necessary to generate stellar coronae. On the other hand, these stars possess massive, radiatively driven stellar winds, with mass loss rates $\sim 10^{-6} M_{\odot}$ per yr. X-ray emission from these systems is believed to arise in shocks in the wind, driven by inhomogeneities resulting from both thermal and dynamical instabilities. Typical X-ray luminosities are $\sim 10^{31} \text{ erg s}^{-1}$, with characteristic temperatures $\sim$ several 10^6 K . Since the wind is dense ($n_e \geq 10^{11} \text{ cm}^{-3}$), and far from fully ionized, the overlying photoelectric opacity is significant. However, the shocks are believed to be distributed throughout the wind, so the absorption structure can be quite complex. Emission lines arising from this gas should exhibit velocity broadening with characteristic velocity widths of several 10^3 km s^{-1} .

Supernova Remnants

Supernovae are cataclysmic stellar explosions which drive high temperature blast waves in the surrounding interstellar medium. There are essentially two varieties. Type 2's which result when massive stars exhaust their nuclear fuel and implode, and Type 1a's which result when white dwarf stars accrete material from their binary companions, causing their masses to exceed the critical

“Chandrasekhar limit” ($\sim 1.4 M_{\odot}$) – the maximum allowable for hydrostatic stability. In both cases, $\sim 10^{50}$ – 10^{51} erg of kinetic energy is transferred to the outer layers of the star, which expand into the neighboring environment. Shock waves form in both the stellar ejecta and the surrounding interstellar gas, with initial temperatures $\sim$ few 10^8 K. These shocks radiate brightly at X-ray energies for $\geq 10^4$ yr. As the remnant expands, the temperature drops, roughly as the third power of the radius. The X-ray emitting gas can have a range of densities, from 10^{-2} to 10^1 cm^{-3} . At such low densities, the time scale for ionization balance to be achieved can exceed the age of the remnant, implying that the plasma may be well out of equilibrium. Despite the very large length-scales involved, the density is low enough that the gas is optically thin to both line and continuum radiation.

X-ray Binaries

Nearly half of all stars in the sky are in binaries, i.e. gravitationally bound two-star systems. Stars of higher mass evolve faster, eventually collapsing to form a “compact object” (white dwarf, neutron star, or black hole). Hence, binary systems can form where one member is compact, and the other is relatively normal. If the binary separation is sufficiently short, these systems exhibit mass transfer, wherein the normal star loses mass that is subsequently accreted by the compact companion. Infall into the deep gravitational potential well characteristic of a compact star, shocks the accreting material up to high temperatures, causing these systems to be copious X-ray emitters. For a white dwarf, $\sim 10^{-4}$ of the rest mass energy of the accreting matter may be released in the form of X-radiation. For a neutron star or black hole, the fraction can be much higher, approaching 20%, leading to X-ray luminosities as high as $10^{38} \text{ erg s}^{-1}$.

There are two possible modes of mass transfer. If the companion is an early type star, it may have a significant stellar wind, some of which will be gravitationally captured by the compact object. Such systems are called “high mass X-ray binaries” (HMXBs). On the other hand, if the companion has low mass, but expands to fill the critical equipotential surface that connects to the other star (the so-called *Roche lobe*), material can flow freely through an inner Lagrange point and fall inward to the compact star. This process is called Roche lobe overflow, and the resulting X-ray sources are called “low mass X-ray binaries” (LMXBs). If the compact star is a white dwarf, instead of a neutron star or black hole, the source is called a “cataclysmic variable star” (CV).

The fate of the accreting material is not well understood, and probably varies from source to source. Since it is far easier to dissipate energy than angular momentum, it is thought that the flow should settle into a thin *accretion disk*, with the matter moving in near Keplerian orbits. Some form of viscous interaction between neighboring “rings” allows angular momentum to be transferred out, thereby enabling accretion to proceed, either continuously,

or episodically. Most of the X-radiation is released down near the surface of the compact star (or in the case of a black hole, near the event horizon). Because the material is nearly fully ionized, and the Compton depths are non-negligible, the emergent flux is radiated primarily as a continuum, with characteristic photon energies of order a few keV. However, the transfer of this intense continuum outward through the circumsource medium can generate a wealth of discrete features. The irradiated environment is likely to be severely photoionized, with the energy density in the radiation field nearly four orders of magnitude higher than the thermal kinetic energy of the gas.

Active Galactic Nuclei

The term “active galactic nucleus” (AGN) refers to an intense source of radiation emanating from a compact nuclear region at the center of a galaxy. The first “quasi-stellar objects”, or quasars, were discovered in the early 1960’s. Spectra of these sources indicated significant redshifts, implying large distances, and thus very high luminosities, comparable to that of an entire galaxy. In addition, observed short-term variations in the emission suggested that the emitting regions must be compact, with characteristic dimensions comparable to the size of our solar system.

There is a rich variety of empirical phenomena associated with AGNs, leading to the definition of numerous “classes”, however it is generally believed that most of these can be understood in terms of a grand unified model, wherein a supermassive accreting black hole is surrounded by an obscuring torus of optically thick material, oriented in the equatorial plane. Accretion onto the black hole generates X-radiation, as well as relativistic jets along the spin axis. If our line of sight is oriented above the plane of the torus, we get a direct view of the black hole and the source is bright in X-rays. Such systems are called Seyfert 1 galaxies if they are radio-quiet, or Type 1 quasars if they are radio-loud. If our line of sight is oriented along the plane, the central source is obscured, and the soft X-ray emission we see is mainly reprocessed radiation emanating from the circumsource environment. These are called Seyfert 2 galaxies (radio-quiet), or Type 2 quasars (radio-loud). Finally, if our line of sight is oriented along the jet, the observed emission is greatly enhanced by relativistic beaming. These systems are called BL Lac objects, or more generally, “blazars”.

Our understanding of the accretion process in AGNs is even less well-developed than for X-ray binaries. However, it is believed that similar physical processes must be involved. There is some evidence for the existence of relativistically broadened X-ray emission lines in these systems, which could be produced in the inner most regions of the accretion disk around the black hole. If this interpretation is correct, X-ray spectroscopy of AGNs may provide us with one of our best observational handles on the physics of ultra-strong gravitational fields. For the Seyfert 2 systems, the obscuration

of the central source affords a relatively “clean” view of the surrounding photoionized gas. Soft X-ray spectra of these systems are rich in discrete spectral features.

Clusters of Galaxies

Clusters are massive collections of galaxies that have formed relatively recently via gravitational collapse as the universe has expanded. They are gravitationally bound systems, with most of the mass in the form of dark matter that only interacts weakly (if at all) with ordinary baryonic matter. The richest, most evolved clusters contain hundreds of members, centered on a central dominant galaxy. The intracluster medium is filled with hot gas, in rough hydrostatic equilibrium with the dark matter gravitational potential. Characteristic temperatures are in the range 10^7 – 10^8 K, so that the gas radiates mostly at X-ray energies. Electron densities are $\sim 10^{-3}$ cm $^{-3}$, and typical X-ray luminosities lie in the range 10^{43} – 10^{45} erg s $^{-1}$. Even at such low densities, the cooling timescales appropriate to this gas are often significantly less than the age of the system, especially at the cluster core. This leads to the expectation that gas should continually be cooling out of this medium, perhaps eventually forming stars in the central galaxy. Curiously, however, recent X-ray spectra suggest a deficit of low temperature gas predicted by this scenario. The intracluster media should be mostly optically thin to continuum absorption, but can exhibit non-negligible optical depth for scattering of bright emission lines.

2 Classical and Quantum Radiation Theory

2.1 Introduction

In this chapter, I review the essential components of classical and quantum radiation theory. I assume that most of the material will be very familiar to the reader from undergraduate (and perhaps graduate) coursework in electrodynamics and elementary quantum mechanics. Nevertheless, I believe it is useful to offer this quick review so that we have the relevant formulae ready at hand for reference in later discussions. You might expect that classical radiation theory should find very limited application in a discussion of discrete radiation from atoms, but as I will show, it does provide a quick means of deriving order of magnitude estimates for a number of important processes. In addition, I find it pedagogically useful to discuss the classical and quantum formulae in a unified context. This is rarely done in textbooks, which makes it difficult to follow where and when quantum ideas are important.

In this, and all subsequent chapters, I utilize the CGS system of units. Although this is going out of fashion in most fields of physics (where SI units have indeed become standard), it is still common practice in astrophysics.

In addition, the fundamental equations of radiation theory take on a simple and more elegant form in CGS units. In this system, the unit of charge is the esu, defined such that Coulomb's Law for the attraction between two point charges, q_1 and q_2 is:

$$\mathbf{F}(\mathbf{r}) = \frac{q_1 q_2}{r^2} \hat{\mathbf{r}} , \quad (1)$$

where $\mathbf{r}$ is the vector separation between them and $\hat{\mathbf{r}}$ the unit vector pointing in the $\mathbf{r}$ direction. Thus, $1 \text{ (esu)}^2 = 1 \text{ erg cm}$. In this system, the electric and magnetic fields, $\mathbf{E}$ and $\mathbf{B}$, have the same units, usually expressed as gauss.

2.2 Overview of the Classical Equations

We start with the governing equations of electromagnetism, specifically *Maxwell's equations*:

$$\nabla \cdot \mathbf{E} = 4\pi \varrho , \quad \nabla \cdot \mathbf{B} = 0 , \quad (2)$$

$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} , \quad \nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{j} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} . \quad (3)$$

which relate the spacetime derivatives of the electric and magnetic fields to each other and to the charge and current densities of the medium, ϱ and $\mathbf{j}$, respectively. In addition, the *Lorentz force law*:

$$\mathbf{f} = \varrho \mathbf{E} + \frac{1}{c} \mathbf{j} \times \mathbf{B} \quad (4)$$

relates the force density on a charged volume, $\mathbf{f}$, to the fields and the charge and current densities. Due to the conservation of electric charge, ϱ and $\mathbf{j}$ obey a *continuity equation*:

$$\frac{\partial \varrho}{\partial t} + \nabla \cdot \mathbf{j} = 0 . \quad (5)$$

It is useful to define also the scalar and vector potential functions, φ and $\mathbf{A}$, respectively, which are related to the fields by:

$$\mathbf{B} = \nabla \times \mathbf{A} , \quad (6)$$

$$\mathbf{E} = -\nabla \varphi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} . \quad (7)$$

Equations (6) and (7) do not define φ and $\mathbf{A}$ uniquely. To make the definitions unique, we need to further specify a gauge. For radiation theory, it is most convenient to adopt the *Lorentz gauge*:

$$\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \varphi}{\partial t} = 0 \quad (8)$$

Substitution into Maxwell's equations yield:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \varphi = -4\pi \varrho , \quad (9)$$

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{A} = -\frac{4\pi}{c} \mathbf{j} . \quad (10)$$

which relate the potentials to the charge and current densities. These equations have solutions of the form:

$$\varphi(\mathbf{r}, t) = \int d^3\mathbf{r}' dt' \frac{\varrho(\mathbf{r}', t')}{|\mathbf{r} - \mathbf{r}'|} \delta[t' - t_r(\mathbf{r}, t, \mathbf{r}')] , \quad (11)$$

$$\mathbf{A}(\mathbf{r}, t) = \int d^3\mathbf{r}' dt' \frac{\mathbf{j}(\mathbf{r}', t')}{|\mathbf{r} - \mathbf{r}'|} \delta[t' - t_r(\mathbf{r}, t, \mathbf{r}')] \quad (12)$$

where t_r is the retarded time, defined by:

$$t_r \equiv t - \frac{|\mathbf{r} - \mathbf{r}'|}{c} . \quad (13)$$

Differentiation of the right-hand sides of (11), (12) according to (6), (7) yields the electric and magnetic fields associated with arbitrary time-varying charge and current distributions. We return to this shortly.

2.3 Electromagnetic Waves

In charge-free space, Maxwell's equations (2), (3) give rise to wave equations for both the electric and magnetic fields:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{E} = 0 , \quad (14)$$

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{B} = 0 \quad (15)$$

which have plane-wave solutions written in the form:

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} , \quad (16)$$

$$\mathbf{B} = \mathbf{B}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (17)$$

where

$$\omega = kc , \quad (18)$$

$$\mathbf{k} \cdot \mathbf{E}_0 = \mathbf{k} \cdot \mathbf{B}_0 = 0 , \quad (19)$$

$$\mathbf{k} \times \mathbf{E}_0 = \frac{\omega}{c} \mathbf{B}_0 , \quad (20)$$

$$\mathbf{k} \times \mathbf{B}_0 = -\frac{\omega}{c} \mathbf{E}_0 , \quad (21)$$

$$\hat{k} = \hat{E} \times \hat{B} , \quad (22)$$

the waves are transverse, and the fields have equal magnitudes. The first (18) requires that electromagnetic waves travel at the speed of light in the vacuum. The energy flux associated with electromagnetic waves is given by the *Poynting vector*:

$$\mathbf{S} = \frac{c}{4\pi} \mathbf{E} \times \mathbf{B} \quad (23)$$

which has units of $\text{erg cm}^{-2} \text{s}^{-1}$ in the CGS system. For the plane-wave solutions (16), (17), (18)–(22) imply that the real part of the Poynting vector is given by:

$$\mathbf{S}(t) = \hat{k} \frac{c}{4\pi} |E(t)|^2. \quad (24)$$

The plane waves described above are monochromatic. Since the wave equations are linear, however, arbitrary linear combinations of plane-waves also provide solutions. In general, we are interested in the frequency dependence of the radiation, which can be assessed by taking the Fourier transform of the electric field:

$$\tilde{\mathbf{E}}(\omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{E}(t) e^{i\omega t} dt. \quad (25)$$

Parseval's Theorem for Fourier transforms requires:

$$\int_{-\infty}^{\infty} |\mathbf{E}(t)|^2 dt = 2\pi \int_{-\infty}^{\infty} |\tilde{\mathbf{E}}(\omega)|^2 d\omega = 4\pi \int_0^{\infty} |\tilde{\mathbf{E}}(\omega)|^2 d\omega \quad (26)$$

(Since $\mathbf{E}(t)$ is real, $\tilde{\mathbf{E}}(-\omega) = \tilde{\mathbf{E}}(\omega)$). The energy in the radiation per unit area is given by:

$$\frac{dW}{dA} = \int_{-\infty}^{\infty} \mathbf{S}(t) dt = c \int_0^{\infty} |\tilde{\mathbf{E}}(\omega)|^2 d\omega \quad (27)$$

so the energy per unit area per unit frequency is:

$$\frac{dW}{dA d\omega} = c |\tilde{\mathbf{E}}(\omega)|^2. \quad (28)$$

2.4 The Classical Multipole Expansion

As for the electric field, we can take the Fourier transform of the charge, current density and vector potential:

$$\tilde{\varrho}(\mathbf{r}, \omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \varrho(\mathbf{r}, t) e^{i\omega t} dt, \quad (29)$$

$$\tilde{\mathbf{j}}(\mathbf{r}, \omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{j}(\mathbf{r}, t) e^{i\omega t} dt, \quad (30)$$

$$\tilde{\mathbf{A}}(\mathbf{r}, \omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{A}(\mathbf{r}, t) e^{i\omega t} dt. \quad (31)$$

Using (12) and (13) we get:

$$\tilde{\mathbf{A}}(\mathbf{r}, \omega) = \frac{1}{c} \int d^3\mathbf{r}' \frac{\tilde{\mathbf{j}}(\mathbf{r}', \omega) e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \quad (32)$$

where $k = \omega/c$. If we are interested in the character of the radiation far from the charge distribution, then $|\mathbf{r}| \gg |\mathbf{r}'|$, so that $|\mathbf{r}-\mathbf{r}'| \approx r - \hat{n} \cdot \mathbf{r}'$, where $\hat{n}$ is the unit vector pointing in the direction $\mathbf{r}$. We thus obtain:

$$\tilde{\mathbf{A}}(\mathbf{r}, \omega) \approx \frac{e^{ikr}}{rc} \int d^3\mathbf{r}' \tilde{\mathbf{j}}(\mathbf{r}', \omega) e^{-ik(\hat{n} \cdot \mathbf{r}')} . \quad (33)$$

The classical multipole expansion involves a Taylor expansion of the complex exponential inside the integral in (33), assuming that $k(\hat{n} \cdot \mathbf{r}') \ll 1$. To see why this might be valid, note that:

$$k(\hat{n} \cdot \mathbf{r}) \sim R\omega/c \sim v/c \quad (34)$$

where R is the characteristic dimension of the charge distribution, and v is a characteristic velocity of the oscillating charge. Thus the multipole expansion is justified in the limit that the charge motions are non-relativistic. The lowest order term is obtained by setting the complex exponential equal to unity. We are then left with a simple integral of the Fourier transform of the current density which can be rewritten in terms of the dipole moment of the charge distribution, $\tilde{\mathbf{d}}(\omega)$

$$\begin{aligned} \int d^3\mathbf{r}' \tilde{\mathbf{j}}(\mathbf{r}', \omega) &= - \int d^3\mathbf{r}' (\nabla' \cdot \tilde{\mathbf{j}}) \mathbf{r}' = -ikc \int d^3\mathbf{r}' \tilde{\rho}(\mathbf{r}', \omega) \mathbf{r}' \\ &\equiv -ikc \tilde{\mathbf{d}}(\omega) . \end{aligned} \quad (35)$$

We thus refer to this term as the *electric dipole* or (E1) term. Expressions for $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{B}}$ in the electric dipole limit can be found by taking the appropriate derivatives of $\tilde{\mathbf{A}}$. Then, converting back to the time domain, we obtain:

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{c^2 r} [\hat{n} \times (\hat{n} \times \ddot{\mathbf{d}}(t_r))] \quad (36)$$

where $\ddot{\mathbf{d}}(t_r)$ is the second time-derivative of the electric dipole moment evaluated at the retarded time. The Poynting vector is:

$$\mathbf{S} = \frac{c}{4\pi} |\mathbf{E}|^2 \hat{n} = \frac{|\ddot{\mathbf{d}}|^2 \sin^2 \theta}{4\pi r^2 c^3} \hat{n} \quad (37)$$

Integrating over the surface of a sphere of radius r yields the total energy radiated per unit time:

$$\frac{dW}{dt} = \int \mathbf{S} r^2 d\Omega = \frac{2}{3} \frac{|\ddot{\mathbf{d}}|^2}{c^3} . \quad (38)$$

For a single, accelerating point charge, this reduces to the well-known *Larmor formula*:

$$\frac{dW}{dt} = \frac{2}{3} \frac{q^2 a^2}{c^3} \quad (39)$$

where a is the acceleration of the charge.

For non-relativistic motions, the electric dipole term will dominate whenever it is non-zero. If it is zero, however, the next highest term will be important. In that case, the relevant integral in (33) is:

$$\int d^3 \mathbf{r}' \tilde{\mathbf{j}}(\tilde{\mathbf{r}}', \omega)(\hat{n} \cdot \tilde{\mathbf{r}}') \quad (40)$$

which can be “broken” into two parts:

$$\tilde{\mathbf{j}}(\mathbf{r}', \omega)(\hat{n} \cdot \mathbf{r}') = \frac{1}{2} \left[\tilde{\mathbf{j}}(\hat{n} \cdot \mathbf{r}') - \mathbf{r}'(\tilde{\mathbf{j}} \cdot \hat{n}) \right] + \frac{1}{2} \left[\tilde{\mathbf{j}}(\hat{n} \cdot \mathbf{r}') + \mathbf{r}'(\tilde{\mathbf{j}} \cdot \hat{n}) \right] \quad (41)$$

The integral of the first term on the right-hand side of (41) can be shown to be related to the magnetic dipole moment of the current distribution:

$$\tilde{\boldsymbol{\mu}}(\omega) \equiv \frac{1}{2c} \int d^3 \mathbf{r}' \mathbf{r}' \times \tilde{\mathbf{j}}(\mathbf{r}', \omega) \quad (42)$$

while the integral of the second term is related to the electric quadrupole tensor of the charge distribution:

$$\tilde{Q}(\omega) \equiv \int d^3 \mathbf{r}' \left[3\mathbf{r}'\mathbf{r}' - r'^2 I \right] \tilde{\varrho}(\mathbf{r}', \omega) \quad (43)$$

where I is the identity tensor. The radiated power for the magnetic dipole term is:

$$\frac{dW}{dt} = \frac{2}{3} \frac{|\dot{\boldsymbol{\mu}}|^2}{c^3} . \quad (44)$$

For the electric quadrupole term, the integral over solid angle depends on the explicit form of the quadrupole tensor, but the radiated power is proportional to $|\ddot{Q}|^3 / c^5$. Note that for an oscillating charge distribution:

$$|\ddot{\mathbf{d}}| \sim qR\omega^2, |\ddot{\boldsymbol{\mu}}| \sim R \left(\frac{qv}{c} \right) \omega^2 \quad (45)$$

and

$$|\ddot{Q}| \sim qR^2\omega^3 . \quad (46)$$

Taking $R\omega \sim v$, we find:

$$\frac{dW}{dt} \sim \frac{q^2\omega}{R} \left(\frac{v}{c} \right)^3 \quad (E1) , \quad (47)$$

$$\frac{dW}{dt} \sim \frac{q^2\omega}{R} \left(\frac{v}{c} \right)^5 \quad (M1) , \quad (48)$$

$$\frac{dW}{dt} \sim \frac{q^2\omega}{R} \left(\frac{v}{c} \right)^5 \quad (E2) . \quad (49)$$

So the (M1) and (E2) terms are of the same order and are both down from the (E1) term by a factor $\sim(v/c)^2$, where v is a characteristic velocity of the charges.

2.5 The Classical Oscillator

An important application of classical radiation theory, and one that proves useful in understanding radiation from atoms, is the *classical harmonic oscillator*, in which the acceleration of the charge is given by:

$$\ddot{\mathbf{x}} = -\omega_0^2 \mathbf{x} \quad (50)$$

where $\mathbf{x}$ is the position of the oscillating charge, and ω_0 is the oscillation frequency. However, since an oscillating charge radiates energy, there must be a damping force associated with the radiation, which gradually reduces the amplitude of the oscillation to zero. This is called the “radiation reaction”. We can approximate it by noting that the power dissipated by the drag force must agree with the Larmor formula for the radiated energy. That yields:

$$\mathbf{F}_{drag} \approx \frac{2}{3} \frac{q^2}{c^3} \ddot{\mathbf{x}} \approx \frac{2}{3} \frac{q^2}{c^3} \omega_0^2 \dot{\mathbf{x}} \quad (51)$$

so the equation of motion becomes:

$$\ddot{\mathbf{x}} + \Gamma \dot{\mathbf{x}} + \omega_0^2 \mathbf{x} = 0 \quad (52)$$

where

$$\Gamma = \frac{2}{3} \frac{q^2 \omega_0^2}{mc^3}, \quad (53)$$

and m is the mass of the charge. The solution has the form:

$$\mathbf{x}(t) = \mathbf{x}_0 e^{-\Gamma t/2} \cos(\omega_0 t + \varphi), \quad (54)$$

and thus the Fourier transform of the electric dipole moment ($\mathbf{d}(t) = q\mathbf{x}(t)$), becomes:

$$|\tilde{\mathbf{d}}(\omega)|^2 = q^2 \left(\frac{x_0}{4\pi} \right)^2 \frac{1}{(\omega - \omega_0)^2 + (\Gamma/2)^2}. \quad (55)$$

The radiated spectrum in the (E1) limit is given by:

$$\frac{dW}{d\omega} = \frac{8\pi}{3} \frac{\omega^4}{c^3} |\tilde{\mathbf{d}}(\omega)|^2 \approx \left(\frac{1}{2} k |\mathbf{x}_0|^2 \right) \left[\frac{\Gamma/2\pi}{(\omega - \omega_0)^2 + (\Gamma/2)^2} \right]. \quad (56)$$

Here $1/2 k |\mathbf{x}_0|^2$ ($k \equiv m\omega_0^2$ is the “spring constant” of the oscillator) is the initial energy of the oscillation, and the term in brackets describes a Lorentzian line profile, centered at ω_0 , with width equal to Γ . Note that:

$$\Delta\nu = \frac{1}{2\pi} \Delta\omega = \frac{4\pi}{3} \frac{q^2 \nu_0^2}{mc^3} \quad (57)$$

and

$$\Delta\lambda = \frac{c}{\nu_0^2} \Delta\nu = \frac{4\pi}{3} \frac{q^2}{mc^2} \quad (58)$$

is a constant, independent of frequency or wavelength. This is the classical *natural line width* for electric dipole transitions. For an electron, $\Delta\lambda \approx 1.2 \cdot 10^{-4} \text{ \AA}$. For the soft X-ray transitions we are concerned with in this lecture, $\lambda \approx 1 - 100 \text{ \AA}$, so the natural width is nearly always a very small component of the line broadening. The time-averaged radiated power of the classical oscillator is:

$$\frac{dW}{dt} = \frac{1}{3} \frac{q^2 \omega_0^4 |\mathbf{x}_0|^2}{c^3} . \quad (59)$$

Since the initial energy, $W_0 = 1/2 k |\mathbf{x}_0|^2 = 1/2 m \omega_0^2 |\mathbf{x}_0|^2$, the *classical radiative decay rate* is given by:

$$A_{cl} \equiv \frac{(dW/dt)}{W_0} = \frac{2}{3} \frac{q^2 \omega_0^2}{mc^3} \quad (60)$$

which turns out to be equal to the damping constant, Γ . In terms of the linear frequency:

$$A_{cl} = \frac{8\pi^2}{3} \frac{q^2}{mc^3} \nu^2 \approx 2.5 \cdot 10^{-22} \nu^2 \text{ s}^{-1} \quad (61)$$

for an electron, where ν is in Hz. Note that in the X-ray band, where $\nu \approx 10^{16} - 10^{18} \text{ Hz}$, radiative decay rates are extremely fast, $\sim 10^{10} - 10^{14} \text{ s}^{-1}$. This has an important effect on level populations for X-ray emitting plasmas, as we will see later.

The discussion above pertains to spontaneous emission. To model induced processes, like photoexcitation, we must consider driven oscillations, where there is an applied force due to an incoming wave, given by: $\mathbf{F}_{appl} = q\mathbf{E}_0 e^{i\omega t}$. The equation of motion is now that of a damped, driven harmonic oscillator. The time-averaged radiated power for this case becomes:

$$\frac{dW}{dt} = \frac{q^4 |\mathbf{E}_0|^2}{3m^2 c^3} \frac{\omega^4}{(\omega^2 - \omega_0^2)^2 + (\Gamma\omega_0)^2} . \quad (62)$$

Since the time-averaged incident energy flux in the wave is $\langle S \rangle = c/8\pi |\mathbf{E}_0|^2$, the cross-section for scattering is:

$$\sigma(\omega) = \frac{dW/dt}{\langle S \rangle} = \frac{8\pi}{3} \frac{q^4}{m^2 c^4} \frac{\omega^4}{(\omega - \omega_0)^2 + (\Gamma/2)^2} . \quad (63)$$

In the vicinity of line center: $\omega^2 - \omega_0^2 \approx 2\omega_0(\omega - \omega_0)$, so this becomes:

$$\sigma(\omega) = \frac{2\pi^2 q^2}{mc} \frac{\Gamma/2\pi}{(\omega^2 - \omega_0^2)^2 + (\Gamma\omega_0)^2} \quad (64)$$

where we have used our earlier expression for Γ . The scattering cross-section again has the Lorentzian line profile with width in angular frequency equal to Γ . Integrating over frequency yields:

$$\int_0^\infty \sigma(\omega) d\omega = 2\pi \int_0^\infty \sigma(\nu) d\nu = \frac{2\pi^2 q^2}{mc} \quad (65)$$

so

$$\sigma(\nu) = \frac{\pi q^2}{mc} \varphi(\nu) . \quad (66)$$

where $\varphi(\nu)$ is the normalized line shape (it may have other components associated with Doppler broadening, etc.). Note that the coefficient is independent of frequency. For an electron, it has the value: $\pi e^2/mc = 2.7 \cdot 10^{-2} \text{ cm}^2 \text{ Hz}$.

2.6 Quantum Radiation Theory – Overview

We now turn to the quantum theory. There are two fundamental differences between the classical and quantum treatments of the interaction between radiation and matter:

- In quantum mechanics, charge configurations are expressed in terms of quantum “states”. Radiative interactions involve an exchange of energy and momentum, so they are associated with a change of state. The only stationary quantum states are the eigenstates of the Hamiltonian, which is the operator associated with the energy of the system. The rates for various processes therefore involve quantum “matrix elements” of the form $\langle f | H_{rad} | i \rangle$, where f represents the final state, i the initial state, and H_{rad} is the perturbing Hamiltonian associated with the radiation field. In the classical picture, charges radiate when they are accelerated. Acceleration requires an external applied force, which can be identified with the perturbing Hamiltonian.
- In the quantum treatment, the radiation field is described in terms of discrete particles or “photons”. The energy of an individual photon is $E = \hbar\omega = h\nu$, where h is Planck’s constant, and has the value $6.626 \cdot 10^{-27} \text{ erg s}$. The momentum of a photon is given by $\mathbf{p} = \hbar\mathbf{k} = (\hbar\omega/c)\hat{\mathbf{k}}$, directed along the direction of propagation. Photons are spin 1 particles, and therefore the emission or absorption of a photon changes the angular momentum of the system by one unit of $\hbar$.

The key rates and cross-sections for various radiative processes follow from time-dependent perturbation theory. We begin with the time-dependent *Schrodinger equation*:

$$H | \psi \rangle = i\hbar \frac{\partial}{\partial t} | \psi \rangle . \quad (67)$$

The energy eigenstates satisfy:

$$H | \psi_E \rangle = E | \psi_E \rangle \quad (68)$$

and therefore have a time dependence given by:

$$| \psi_E(t) \rangle = | \psi_E \rangle e^{-iEt/\hbar} . \quad (69)$$

Let the total Hamiltonian contain a dominant “unperturbed part” and a small additional “perturbing part”:

$$H = H^0 + H' \quad (70)$$

and let $| n \rangle$ represent a complete set of energy eigenstates of H^0 . An arbitrary state $| \psi(t) \rangle$ can be expanded in terms of these energy eigenstates:

$$| \psi(t) \rangle = \sum_n a_n(t) | n \rangle e^{-iE_n t/\hbar} . \quad (71)$$

Substituting into (67) and taking the scalar product with a specific energy eigenstate $\langle k |$ to both sides, then yields the differential equation:

$$i\hbar \frac{\partial a_k}{\partial t} = \sum_n a_n \langle k | H' | n \rangle e^{i\omega_{kn} t} \quad (72)$$

where $\omega_{kn} \equiv (E_k - E_n)/\hbar$. Here, we have used the fact that the energy eigenstates are orthonormal: $\langle k | n \rangle = \delta_{k,n}$. Suppose the system is initially in state “m”, so that $a_k(0) = \delta_{k,m}$. Then, to lowest order in the perturbing Hamiltonian, the coefficients a_k at some later time are given by:

$$a_k(t) = (i\hbar)^{-1} \int_0^t \langle k | H'(t) | m \rangle e^{i\omega_{km} t} dt . \quad (73)$$

For application to radiation theory, we are interested in perturbations which are oscillatory in time:

$$H'(t) = H' e^{\pm i\omega t} \quad (74)$$

where ω is some angular frequency. Thus:

$$a_k(t) = (i\hbar)^{-1} \langle k | H' | m \rangle \int_0^t e^{i(\omega_{km} \pm \omega)t} dt . \quad (75)$$

The probability at time t that the system has made the transition from “m” to state “k” is given by $| a_k(t) |^2$. The transition rate, R , is thus given by:

$$R = \lim_{t \rightarrow \infty} \frac{| a_k(t) |^2}{t} = \frac{2\pi}{\hbar^2} | \langle k | H' | m \rangle |^2 \delta(\omega_{km} \pm \omega) \quad (76)$$

$$= \frac{2\pi}{\hbar} | \langle k | H' | m \rangle |^2 \delta(E_k - E_m \pm \hbar\omega) . \quad (77)$$

This last expression indicates that the transition is possible only if the change of state is accompanied by the emission or absorption of a single photon with energy equal to the energy difference between the states. Note that this is a first order perturbation result. Multi photon processes occur via higher order terms in the perturbation expansion.

2.7 The Radiation Hamiltonian

The appropriate Hamiltonian to use for the interaction between charged particles and electromagnetic fields is derived from the formalism of classical mechanics. Defining a Lagrangian of the form:

$$L = \frac{1}{2}mv^2 - q\varphi + \frac{q}{c}\mathbf{A} \cdot \mathbf{v} \quad (78)$$

where φ and $\mathbf{A}$ are the classical scalar and vector potentials, respectively, and applying Lagrange's Equation:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\mathbf{r}}} \right) = \frac{\partial L}{\partial \mathbf{r}}, \quad (79)$$

we arrive at the desired Lorentz force law for the electromagnetic force on a single charge:

$$\mathbf{F} \equiv m\dot{\mathbf{v}} = q\mathbf{E} + \frac{q\mathbf{v}}{c} \times \mathbf{B}. \quad (80)$$

The canonical momentum of the particle is defined by:

$$\mathbf{p} \equiv \frac{\partial L}{\partial \dot{\mathbf{r}}} = m\mathbf{v} + \frac{q\mathbf{A}}{c}. \quad (81)$$

The Hamiltonian is then:

$$H \equiv \mathbf{p} \cdot \mathbf{v} - L = \frac{1}{2}mv^2 + q\varphi = \frac{1}{2m} \left(\mathbf{p} - \frac{q}{c}\mathbf{A} \right)^2 + q\varphi. \quad (82)$$

It is the canonical momentum $\mathbf{p}$ that we associate with the quantum mechanical operator $(\hbar/i)\nabla$. Substituting into (82) yields:

$$H = -\frac{\hbar^2}{2m}\nabla^2 + \frac{iq\hbar}{mc}(\nabla \cdot \mathbf{A}) + \frac{iq\hbar}{mc}(\mathbf{A} \cdot \nabla) + \frac{q^2}{2mc^2}A^2 + q\varphi. \quad (83)$$

For an electromagnetic wave, $\varphi = 0$, and therefore, in the Lorentz gauge, $\nabla \cdot \mathbf{A} = 0$. The term involving A^2 is small compared to the first order terms in $\mathbf{A}$, so we ignore it. In addition, there may be a non-radiation potential $V(\mathbf{r})$, e.g. the binding potential of the atom. In that case:

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) + \frac{iq\hbar}{mc}(\mathbf{A} \cdot \nabla) \quad (84)$$

The first two terms on the right hand side are usually taken to be the unperturbed Hamiltonian. The perturbing Hamiltonian associated with the interaction with radiation is given by the third term. For a strictly monochromatic wave, we can write the vector potential in the form:

$$\mathbf{A}(\mathbf{r}, t) = \text{Re} \left[\mathbf{A}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \right] = \frac{1}{2} |\mathbf{A}_0| \left[\hat{\mathbf{e}} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} + \hat{\mathbf{e}}^* e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \right] \quad (85)$$

where $\mathbf{A} = |\mathbf{A}_0| \hat{\varepsilon}$, and $\hat{\varepsilon}$ is the polarization vector of the wave. From (7) and (24), we find that the time-averaged Poynting vector is given by:

$$\langle \mathbf{S} \rangle = \frac{\omega^2}{8\pi c} |\mathbf{A}_0|^2 \hat{k} \quad (86)$$

Recall that $\langle \mathbf{S} \rangle$ represents the energy flux of the radiation. If we think in terms of discrete photons, the photon flux, $dN/dAdt$, is given by:

$$\frac{dN}{dt dA} = \frac{|\langle \mathbf{S} \rangle|}{\hbar\omega} = \frac{\omega}{8\pi\hbar c} |\mathbf{A}_0|^2. \quad (87)$$

Note from (85) that the perturbing Hamiltonian in (84) has two pieces, one proportional to $e^{-i\omega t}$, and one proportional to $e^{+i\omega t}$. The former leads to the absorption of a photon ($E_k = E_m + \hbar\omega$), while the latter leads to the emission of a photon ($E_k = E_m - \hbar\omega$). For a given set of initial and final states, only one of the two terms can satisfy energy conservation, so we can treat them separately. The expression for the transition rate between initial state i and final state f is thus:

$$R = \frac{2\pi}{\hbar} \frac{q^2 \hbar^2}{4m^2 c^2} |\mathbf{A}_0|^2 \left| \left\langle f \left| e^{\pm i\mathbf{k} \cdot \mathbf{r}} \hat{\varepsilon}^{(*)} \cdot \nabla \right| i \right\rangle \right|^2 \delta(E_f - E_i \mp \hbar\omega) \quad (88)$$

where the top sign corresponds to absorption (with $\hat{\varepsilon}$ in the matrix element) and the bottom sign corresponds to emission (with $\hat{\varepsilon}^*$ in the matrix element).

2.8 Bound-Free Absorption (Photoionization)

Consider first the application to bound-free absorption, where the initial state of an electron is a bound state in an atom, and the final state is that of a free particle. To get the total transition rate, we must integrate over all possible final states. For a free particle, the states are characterized by the momentum vector $\mathbf{p}$. However, the *uncertainty principle* requires that a particle cannot be localized in a 6-dimensional phase space cell smaller than $d^3\mathbf{r}d^3\mathbf{p} = (2\pi\hbar)^3$. Therefore, the density of states for a free particle is given by:

$$\varrho(E)dE = \frac{V d^3\mathbf{p}}{(2\pi\hbar)^3} = \frac{Vm(2mE)^{1/2}dE d\Omega}{(2\pi\hbar)^3} \quad (89)$$

where V is the allowable volume for the free particle (it will drop out of the later expression), $d\Omega$ is a differential element of solid angle, and we have assumed non-relativistic dynamics. The free particle final state of the charge can be represented by:

$$\psi_f(\mathbf{r}) = V^{-1/2} e^{i\mathbf{p}_f \cdot \mathbf{r} / \hbar} \quad (90)$$

where the coefficient has been introduced for normalization, i.e. $\langle \psi_f | \psi_f \rangle = 1$ when the integration is performed over the allowable volume.

Taking $(2mE_f)^{1/2} = mv_f$, and integrating over energy in (88), we obtain:

$$dR = \frac{1}{4} \frac{q^2}{(2\pi\hbar c)^2} v_f |\mathbf{A}_0|^2 \left| \left\langle e^{-i\mathbf{p}_f \cdot \mathbf{r}/\hbar} \mid e^{i\mathbf{k} \cdot \mathbf{r}} \hat{\varepsilon} \cdot \nabla \mid \psi_i \right\rangle \right|^2 d\Omega. \quad (91)$$

The differential cross-section for this process is given by:

$$\frac{d\sigma}{d\Omega} = \frac{dR/d\Omega}{dN/dt dA} = \frac{dR/d\Omega}{\omega |\mathbf{A}_0|^2 / 8\pi\hbar c} \quad (92)$$

$$= \frac{q^2}{2\pi\hbar c} \frac{\nu_f}{\omega} \left| \left\langle e^{-i\mathbf{p}_f \cdot \mathbf{r}/\hbar} \mid e^{i\mathbf{k} \cdot \mathbf{r}} \hat{\varepsilon} \cdot \nabla \mid \psi_i \right\rangle \right|^2. \quad (93)$$

Actually, this expression is an approximation to the real photoionization cross-section because the liberated electron is not really “free” – it still feels the Coulomb attraction to the nucleus. A more accurate treatment would use a true continuum wave-function for the electron subject to the atomic potential. We will come back to this later.

2.9 Bound-Bound Transitions

In the case of bound-bound transitions, which give rise to emission or absorption lines, both the initial and final states are discrete. Equation (88) indicates that if the incoming wave is perfectly monochromatic, then the transition rate will be infinite if $\hbar\omega = |E_f - E_i|$, and zero otherwise. To derive a meaningful cross-section, we must integrate over a finite spectrum of the incident radiation field. This is characterized by a continuum photon flux, $dN/dt dA d\omega$. Setting:

$$|\mathbf{A}_0|^2 = \frac{8\pi\hbar c}{\omega} \frac{dN}{dt dA d\omega} d\omega \quad (94)$$

in (88) and integrating over frequency, yields:

$$R_{i \rightarrow f} = \frac{4\pi^2 q^2 \hbar}{m^2 c \omega_{if}} \frac{dN}{dt dA d\omega}(\omega_{if}) \left| \left\langle f \mid e^{\pm i\mathbf{k} \cdot \mathbf{r}} \hat{\varepsilon}^* \cdot \nabla \mid i \right\rangle \right|^2 \quad (95)$$

where $\omega_{if} \equiv |E_f - E_i|/\hbar$. Here again the (+) sign corresponds to absorption and the (−) sign to emission. The emission case is actually *induced* emission, since the transition rate is proportional to the incident flux. Because the radiation Hamiltonian operator is Hermitian, the rates for emission and absorption are identical (with the appropriate reversal of initial and final states).

Dividing the transition rate in (95) by the continuum flux yields a quantity with units of $\text{cm}^2 \text{ Hz}$, which is the cross-section integrated over frequency:

$$\frac{R_{i \rightarrow f}}{dN/dt dA d\omega} = \int \sigma(\omega) d\omega = 2\pi \int \sigma(\nu) d\nu. \quad (96)$$

This yields:

$$\int \sigma(\nu) d\nu = \left(\frac{\pi q^2}{mc} \right) \frac{2\hbar}{m\omega_{if}} |\langle f | e^{i\mathbf{k} \cdot \mathbf{r}} \hat{\varepsilon} \cdot \nabla | i \rangle|^2 . \quad (97)$$

Notice that the term within parentheses is the classical expression we had earlier (65). The remainder of the right hand side is the “quantum correction” to the classical result, and is called the *oscillator strength*, usually denoted by the symbol f :

$$f_{i \rightarrow f} \equiv \frac{2\hbar}{m\omega_{if}} |\langle f | e^{i\mathbf{k} \cdot \mathbf{r}} \hat{\varepsilon} \cdot \nabla | i \rangle|^2 . \quad (98)$$

2.10 The Quantum Multipole Expansion

The matrix element which appears in (98) involves the complex exponential factor: $e^{i\mathbf{k} \cdot \mathbf{r}}$. This is reminiscent of the classical expression (33) where we found it useful to expand this expression as a Taylor expansion in $\mathbf{k} \cdot \mathbf{r}$. The logic in the quantum calculation is the same: $\mathbf{k} \cdot \mathbf{r} \approx v/c$, where v is the characteristic velocity of oscillating charges in the system. For non-relativistic motions, this is a small parameter. In the lowest order limit, the electric dipole (E1) limit, we set the complex exponential to unity. The matrix element becomes:

$$\langle f | \hat{\varepsilon} \cdot \nabla | i \rangle = \frac{i}{\hbar} \hat{\varepsilon} \cdot \langle f | \mathbf{p} | i \rangle \quad (99)$$

where $\mathbf{p}$ is the momentum operator. Using the commutation relation:

$[p^2, \mathbf{r}] = -2i\hbar\mathbf{p} = 2m [H^0, \mathbf{r}]$, we can rewrite this in the form:

$$\begin{aligned} \langle f | \mathbf{p} | i \rangle &= \frac{mi}{\hbar} \langle f | [H^0, \mathbf{r}] | i \rangle \\ &= \frac{mi}{\hbar} (E_f - E_i) \langle f | \mathbf{r} | i \rangle = im\omega_{if} \langle f | \mathbf{r} | i \rangle . \end{aligned} \quad (100)$$

The (E1) expression for the oscillator strength is therefore:

$$f_{i \rightarrow f} = \frac{2m\omega_{if}}{\hbar} |\hat{\varepsilon} \cdot \langle f | \mathbf{r} | i \rangle|^2 \quad (101)$$

Averaged over polarization directions, this becomes:

$$f_{i \rightarrow f} = \frac{2}{3} \frac{m\omega_{if}}{\hbar} |\langle f | \mathbf{r} | i \rangle|^2 . \quad (102)$$

A simple set of operator manipulations shows that the (E1) oscillator strengths satisfy a sum rule (the *Thomas-Reiche-Kuhn sum rule*):

$$\sum_f f_{i \rightarrow f} = Z \quad (103)$$

where Z is the number of bound electrons in the atom. This provides a useful limit on the oscillator strengths for highly excited transitions, which are numerous and therefore unwieldy to calculate. The next term in the multipole expansion has the form $\langle f | (\mathbf{k} \cdot \mathbf{r})(\hat{\mathbf{e}} \cdot \mathbf{p}) | i \rangle$, which, as in the classical case, can be broken into two pieces:

$$\begin{aligned} & 1/2 \langle f | (\mathbf{k} \cdot \mathbf{r})(\hat{\mathbf{e}} \cdot \mathbf{p}) - (\mathbf{k} \cdot \mathbf{p})(\hat{\mathbf{e}} \cdot \mathbf{r}) | i \rangle \\ & + 1/2 \langle f | (\mathbf{k} \cdot \mathbf{r})(\hat{\mathbf{e}} \cdot \mathbf{p}) + (\mathbf{k} \cdot \mathbf{p})(\hat{\mathbf{e}} \cdot \mathbf{r}) | i \rangle . \end{aligned} \quad (104)$$

The first term can be rewritten as:

$$(\mathbf{k} \times \hat{\mathbf{e}}) \cdot (\mathbf{r} \times \mathbf{p}) \sim \boldsymbol{\mu} \cdot \mathbf{B} \quad (105)$$

where $\boldsymbol{\mu}$ is the magnetic dipole moment of the orbiting electron. This is the magnetic dipole term (M1). For atomic transitions, we need to include both orbital and intrinsic spin contributions to the magnetic dipole moment. The second term above gives rise to electric quadrupole (E2) transitions. Here again, (M1) and (E2) transitions are of the same order in v/c . The (E1) term always dominates unless the matrix element of the position vector vanishes between the initial and final states. Transitions for which this is the case are called “electric dipole forbidden”, or simply “forbidden”. This condition gives rise to certain “selection rules” for (E1) transitions, which we discuss later in the context of atomic structure. Transitions for which the expression in (98) vanishes to all orders in $(\mathbf{k} \cdot \mathbf{r})$ are called “strictly forbidden”. These can only go by two-photon decay.

2.11 Spontaneous Emission

The quantum theory summarized so far only works for induced transitions, where an external electromagnetic field is introduced as a perturbation. This is because the treatment is semi-classical, i.e. the radiation field is still modeled classically even though the radiating system is treated quantum mechanically. Spontaneous emission, in which a system in an excited state decays on its own by emitting a photon, does not occur in this picture because the initial state involves no radiation field, so there is no perturbing Hamiltonian. The correct treatment of this process requires the quantization of the radiation field. That is straightforward, but too time-consuming to review here. However, another form of semi-classical argument can be invoked to derive what turns out to be the correct result. In the (E1) limit, our classical expression for the radiated power is given by (38). For an oscillator at a particular frequency:

$$|\ddot{\mathbf{d}}|^2 = \omega^4 (|\tilde{\mathbf{d}}(\omega)|^2 + |\tilde{\mathbf{d}}(-\omega)|^2) = 2\omega^4 |\tilde{\mathbf{d}}(\omega)|^2 . \quad (106)$$

Using (35), we can write this in terms of the integrated current density:

$$|\tilde{\mathbf{d}}(\omega)|^2 = \frac{1}{\omega^2} |\mathbf{j}_0|^2 \quad (107)$$

where $\mathbf{j}_0 \equiv \int d^3\mathbf{r}' \tilde{\mathbf{j}}(\mathbf{r}', \omega)$ Thus:

$$\frac{dW}{dt} = \frac{4}{3} \frac{\omega^2}{c^3} |\mathbf{j}_0|^2 \quad (108)$$

In quantum mechanics, the charge density for a point charge is $\varrho = q |\psi(\mathbf{r})|^2$. From the continuity equation (5) and the time-dependent Schroedinger equation (67), it can be shown that the current density must be given by:

$$\mathbf{j} = -\frac{iq\hbar}{2m} [\psi^* \nabla \psi - \psi \nabla \psi^*] \quad (109)$$

An appropriate “quantization” of the classical expression (108) can thus be obtained by setting:

$$|\mathbf{j}_0|^2 = \left| \int d^3\mathbf{r}' \left(\frac{-iq\hbar}{2m} \right) [\psi_f^* \nabla \psi_i - (\nabla \psi_f)^* \psi_i] \right|^2 \quad (110)$$

$$= \frac{q^2}{m^2} |\langle f | \mathbf{p} | i \rangle|^2 \quad (111)$$

$$= \frac{q^2 \omega_{if} \hbar}{2m} f_{i \rightarrow f} \quad (112)$$

where $f_{i \rightarrow f}$ is the electric dipole oscillator strength of (102). The resulting expression for the decay rate is then:

$$A_{i \rightarrow f} = \frac{1}{\hbar \omega_{if}} \frac{dW}{dt} = \frac{2}{3} \frac{q^2 \omega_{if}^2}{mc^3} f_{i \rightarrow f} \quad (113)$$

Comparison with (60) shows that this is simply the expression for the radiative decay rate of the classical oscillator multiplied by the absorption oscillator strength.

3 The Structure of Multi-Electron Atoms

3.1 Introduction

This chapter is devoted to the structure of multi-electron atoms. This is a vast and complex subject and time limitations will unfortunately prevent me from going into any real depth on most of the topics I will cover. My main focus will be on defining the relevant terms and outlining the basic principles and approximations which are used in modern atomic physics calculations. I will not discuss computational techniques or the specifics of particular codes.

Once again, I assume that much of this material is familiar to the reader from undergraduate and graduate courses in quantum mechanics.

The physics of atomic structure basically involves the solution of the time-independent Schroedinger equation:

$$H\psi = E\psi \quad (114)$$

where H is the Hamiltonian operator, E the energy and ψ is the wave-function for the electrons in the atom, usually expressed as a function of spatial and spin coordinates. For all but the simplest atoms, this equation is not analytically solvable and various approximation techniques are required. The most common, and most general is time-independent perturbation theory, in which one writes the Hamiltonian in terms of two parts:

$$H = H^0 + H^1 \quad (115)$$

a zeroth-order Hamiltonian H^0 , which is amenable to direct solution and an additional perturbation H^1 which has much smaller amplitude. In first order perturbation theory, the corrections to the energy levels due to the presence of the perturbation are given by:

$$\Delta E_n^{(1)} = \left\langle \psi_n^{(0)} \mid H^1 \mid \psi_n^{(0)} \right\rangle \quad (116)$$

where $\psi_n^{(0)}$ is the zeroth-order wave-function associated with the n -th energy level, E_n , and the corrections to the wave-functions are given by

$$\Delta \psi_n^{(1)} = \sum_{k \neq n} \frac{\left\langle \psi_k^{(0)} \mid H^1 \mid \psi_n^{(0)} \right\rangle}{E_k^{(0)} - E_n^{(0)}} \psi_k^{(0)}. \quad (117)$$

The zeroth-order wave-functions are orthonormal by construction and the perturbed wave-functions remain orthonormal to lowest order in H^1 .

Another approach which is frequently used for more complex atoms is the *Ritz variational method*. Its utility follows from the fact that the expectation value of the Hamiltonian with respect to an arbitrary normalized wave-function ψ , $\langle \psi \mid H \mid \psi \rangle$, is a minimum when ψ is the ground state eigenfunction of H . Even more generally, if the functional $\langle \psi \mid H \mid \psi \rangle$ is *stationary* with respect to perturbations in ψ , then ψ must be an eigenfunction of H . Typically, one uses this method by choosing a form for a trial wave-function characterized by a set of adjustable parameters and then minimizing the expectation value of the Hamiltonian with respect to those parameters.

3.2 Hydrogen-like Ions

We will begin the discussion with a quick review of the structure of hydrogen-like ions or one-electron atoms. Hydrogen-like ions are important for a number of reasons. First, in the non-relativistic limit, the time-independent

Schroedinger equation is exactly solvable so we can get analytic expressions for all important quantities. Second, the “hydrogenic approximation” is often useful for orders of magnitude estimates of rates for important processes and for simple scaling laws with the nuclear charge Z . Finally, hydrogen-like ions are quite important contributors to the soft X-ray emission from astrophysical plasmas. Indeed, the brightest lines are usually Lyman series transitions from hydrogen-like oxygen, neon, silicon and other low- Z elements.

The non-relativistic Hamiltonian for a single electron in an attractive central potential is given by:

$$H = \frac{p^2}{2m_e} - V(r) . \quad (118)$$

Making the usual substitution: $\mathbf{p} = -i\hbar\nabla$ we get the relevant form of (114):

$$\left(-\frac{\hbar^2}{2m_e} \nabla^2 - V(r) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}) . \quad (119)$$

It is convenient to use *atomic units* where the natural unit of length is the *Bohr radius*: $a_0 \equiv \hbar^2/me^2 = 0.529 \cdot 10^{-8}$ cm, and the natural unit of energy is twice the Rydberg constant: $e^2/a_0 \equiv 2Ry = 27.2$ eV = $4.36 \cdot 10^{-11}$ erg. In these units, $e = \hbar = m = 1$.

Equation (119) then takes the form:

$$\left(\frac{1}{2} \nabla^2 + E + V(r) \right) \psi(\mathbf{r}) = 0 . \quad (120)$$

Equation (120) is spherically symmetric, so it is useful to write it in spherical coordinates. A spherically symmetric Hamiltonian commutes with the total angular momentum operator $\mathbf{l} = \mathbf{r} \times \mathbf{p}$, which implies that eigenstates of H are also eigenstates of l^2 and l_z . In spherical coordinates (120) becomes:

$$\left(\frac{1}{2} \left[\frac{1}{r^2} \left(r \frac{\partial}{\partial r} \right) \left(r \frac{\partial}{\partial r} \right) + \frac{1}{r} \frac{\partial}{\partial r} - \frac{l^2}{r^2} \right] + E + V(r) \right) \psi = 0 . \quad (121)$$

The only dependence on the angular coordinates (ϑ, φ) in this expression is the l^2 term. That implies that the equation is separable and ψ can be written as a product of radial and angular parts:

$$\psi(r, \vartheta, \varphi) \equiv \frac{R(r)}{r} Y(\vartheta, \varphi) . \quad (122)$$

The eigenfunctions of l^2 and l_z are called *spherical harmonics* and have the form:

$$Y_{lm}(\vartheta, \varphi) \equiv \left[\frac{(l-|m|)!}{(l+|m|)!} \frac{2l+1}{4\pi} \right]^{1/2} (-1)^{(m+|m|)/2} P_l^{|m|}(\cos\vartheta) e^{im\varphi} \quad (123)$$

where P_l^m is the associated Legendre Polynomial. The spherical harmonics obey the eigenvalue equations:

$$l^2 Y_{lm}(\vartheta, \varphi) = l(l+1) Y_{lm}(\vartheta, \varphi) \quad (124)$$

$$l_z Y_{lm}(\vartheta, \varphi) = m Y_{lm}(\vartheta, \varphi) \quad (125)$$

where l and m are integers, with $-l \leq m \leq l$.

After substitution of (122) into (121), we are left with the *radial equation*:

$$\left(\frac{1}{2} \frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} - \frac{l(l+1)}{2r^2} + E + V(r) \right) \frac{R(r)}{r} = 0. \quad (126)$$

For bound-states, $E < 0$, the solutions are discrete and are characterized by an integer index n called the *principal quantum number*. Bound-state wave-functions are only obtained for $n \geq l+1$, so for a given principal quantum number, the only allowed angular momentum states are $l = 0, 1, 2, \dots, n-1$. The radial eigenfunctions are thus characterized by the two indices n and l .

For the particular case of the Coulomb potential $V(r) = Z/r$, (126) is exactly solvable, and leads to the radial wave-functions:

$$R_{nl}(r) = - \left(\frac{Z(n-l-1)!}{n^2[(n+l)!]^3} \right)^{1/2} e^{-\rho/2} \rho^{l+1} L_{n+l}^{2l+1}(\rho) \quad (127)$$

where $\rho \equiv 2Zr/n$ and $L_{n+l}^{2l+1}(\rho)$ are associated Laguerre polynomials. The energy eigenvalues, in atomic units, have the form:

$$E_n = \frac{-Z^2}{2n^2} \quad (128)$$

and are independent of l . This is a unique property of the Coulomb potential.

The probability density of finding the electron in the radial range $r \rightarrow r + dr$ is given by $R_{nl}^2(r)$. Plots of this function for a few low order orbitals are given in Fig. 1. Several key features of these radial wave-functions are immediately apparent from the plots. First, most of the charge is concentrated in a spherical shell of moderate thickness, whose radius increases with n . This is expected classically, i.e. smaller binding energy is associated with larger orbits. Note that for a given n , the radius of this shell decreases with increasing l . Again, this is in line with classical expectations. For a fixed energy, smaller angular momentum implies an elliptical orbit with higher eccentricity, in which the electron spends most of its time further away from the nucleus. Finally, note that as r goes to zero, the probability density goes to zero for all but the $l = 0$ states. Hence only these states are appreciably affected by nuclear interactions.

Since the energy only depends on n for hydrogen-like ions, there are n degenerate l states for each value of n , and $2l+1$ degenerate m states for each value of l . In addition, the electron is a spin $1/2$ particle, so there are

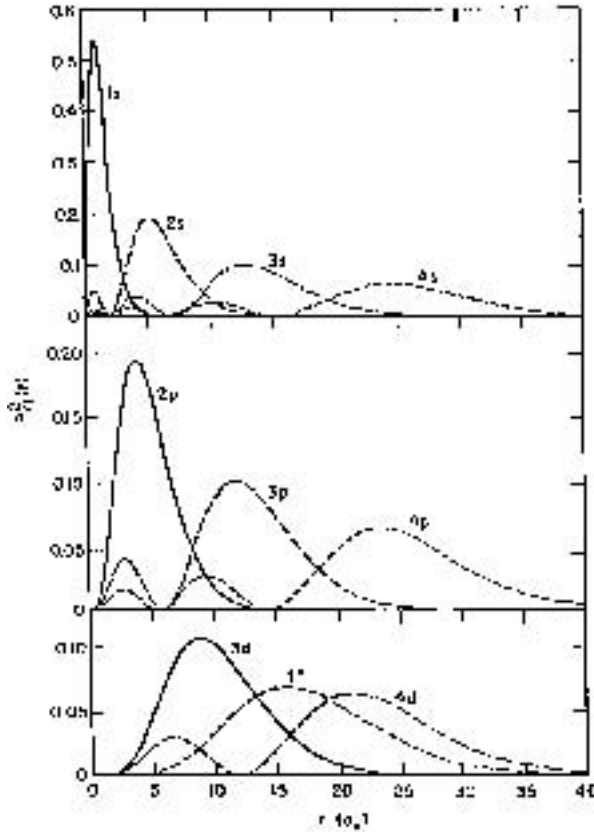


Fig. 1. Probability density to find the electron as a function of r (from Rybicki and Lightman, Fig. 9.1)

two degenerate spin states for each spatial state. The total degeneracy of level n is therefore given by:

$$g_n = 2 \sum_{l=0}^{n-1} (2l+1) = 2n^2 \quad (129)$$

3.3 Scaling with Nuclear Charge

It is useful, at this stage, to look at the scaling of various quantities with the nuclear charge Z . First note that the energy levels scale like Z^2 , which implies that the frequencies of key transitions also scale like Z^2 . The Lyman- α or $n = 2 \rightarrow 1$ transition, specifically, has photon energy given by:

$$\hbar\omega_{K\alpha} = (10.2 \text{ eV})Z^2. \quad (130)$$

Note that this line falls in the soft X-ray band (0.1–10 keV) for $Z = 3\text{--}31$, which includes the abundant elements: C($Z = 6$), N($Z = 7$), O($Z = 8$), Ne($Z = 10$), Si($Z = 14$), S($Z = 16$), Ar($Z = 18$), Ca($Z = 20$) and Fe($Z = 26$). The energy of this line is only slightly affected by the presence of additional electrons. So (130) gives a rough idea of the energies of all K-shell feature transitions down to $n = 1$, for these and other elements.

Transitions down to $n = 2$ are called L-shell transitions. For hydrogen-like ions, the brightest is the Balmer- α transition corresponding to $n = 3 \rightarrow 2$, whose energy is given by:

$$\hbar\omega_{L\alpha} = (1.89 \text{ eV})Z^2 \quad (131)$$

Note that the L-shell transitions for Fe fall close to 1 keV, in the center of the soft X-ray band. These are especially important for diagnostic purposes, as we will review in a subsequent chapter.

Equation (127) implies that the scaling of the radial wave-function is like Z^{-1} . Specifically, the characteristic size of hydrogen-like ions is given roughly by a_0/Z , where a_0 is the Bohr radius we defined earlier. Recall from (102) that the oscillator strength for an E1 transition is proportional to $\omega_{ij} |\langle f | \mathbf{r} | i \rangle|^2$. This scales like $Z^2 Z^{-2}$, and thus is independent of Z . The radiative decay rates for E1 transitions are proportional to $\omega^2 f$, so they scale like Z^4 .

The Coulomb potential for a hydrogen-like atom is proportional to $1/r$ so classically, the electron orbit obeys the Virial theorem, i.e. the kinetic energy is $-1/2$ times the potential energy:

$$\frac{1}{2}mv^2 = \frac{Ze^2}{2r}.$$

For the ground-state:

$$r \simeq \frac{a_0}{Z},$$

and thus:

$$v \simeq \left(\frac{Z^2 e^2}{ma_0} \right)^{1/2} = (Z\alpha)c \quad (132)$$

where $\alpha \equiv e^2/\hbar c \simeq 1/137$ is the fine structure constant. We saw earlier that the expansion parameter for both the classical and quantum multipole expansion $(\mathbf{k} \cdot \mathbf{r}) \sim v/c$, where v is a characteristic velocity of the system. For atomic transitions, we see that this parameter is $\sim Z\alpha$. The magnetic dipole and electric quadrupole terms are thus $\sim (Z\alpha)^2$ times smaller than electric dipole terms, so they scale like Z^6 . For low- Z abundant elements (C, N, O), $(Z\alpha)$ is indeed a small parameter. However for Fe, it is ~ 0.2 , so higher order multipole terms are non-negligible and can often be important in the spectrum.

3.4 Relativistic Corrections

The time independent Schroedinger equation as expressed in (119) assumes non-relativistic dynamics. For relativistic charges, one must use the Dirac equation instead. However, since $v/c \sim Z\alpha$, atomic electrons are only mildly relativistic, even for iron which is the highest Z abundant element. Thus, it is sufficient to use (119) and to treat relativistic corrections as a simple perturbation to the atomic structure.

To lowest order, there are three contributions to the relativistic corrections:

$$H_1^1 = -\frac{1}{8} \frac{p^4}{m_e^3 c^2} \quad (133)$$

which is the lowest order correction to the kinetic energy,

$$H_2^1 = \frac{1}{2m_e^2 c^2} \left(\frac{1}{r} \frac{dV}{dr} \right) \mathbf{l} \cdot \mathbf{s} , \quad (134)$$

the spin-orbit term, which represents the magnetic interaction between the magnetic dipole moment of the electron associated with its intrinsic spin and the magnetic field that it sees as it orbits in the electric field of the nuclear charge, and

$$H_3^1 = \frac{\hbar^2}{4m_e^2 c^2} \left(\frac{dV}{dr} \right) \frac{\partial}{\partial r} , \quad (135)$$

the so-called Darwin term, which is a relativistic correction to the potential energy produced by the non-localizability of the electron associated with its rest mass energy.

For the Coulomb potential in hydrogen-like atoms, a simple first order perturbation theory calculation using zeroth-order wave-functions yields the energy shift:

$$\Delta E_n = +E_n \frac{(Z\alpha)^2}{n^2} \left[\frac{n}{(j+1/2)} - \frac{3}{4} \right] \quad (136)$$

where j is the eigenvalue associated with the *total angular momentum* – specifically $j(j+1)\hbar^2$ is the eigenvalue of j^2 , where $\mathbf{j} = \mathbf{l} + \mathbf{s}$. The fact that the perturbed energies depend on j is a consequence of the spin-orbit term, which is proportional to the operator:

$$\mathbf{l} \cdot \mathbf{s} = \frac{1}{2} (j^2 - l^2 - s^2) . \quad (137)$$

Ignoring the relativistic corrections, eigenfunctions of the Hamiltonian for a one-electron central potential are simultaneous eigenfunctions of H_0 , l^2 , l_z , s^2 and s_z , so the states are characterized by the quantum numbers n , l , m_l , s , m_s . When the spin-orbit term is included however, l_z and s_z no longer commute with the Hamiltonian. The states are then characterized by n , l , s , j , m_j . We will see shortly that this has important consequences for the specification of the states in multi-electron atoms.

3.5 The Central Field Approximation and Quantum Indistinguishability

When there is more than one electron in the atom, the Schroedinger equation acquires an additional term due to the electron-electron repulsion:

$$\left(\frac{1}{2} \sum_j \nabla_j^2 + E + Z \sum_j \frac{1}{r_j} - \sum_{i>j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \psi(\{\mathbf{r}_j\}) = 0 \quad (138)$$

where $\mathbf{r}_j$ is the position coordinate of the j th electron, $\nabla_j \equiv \partial/\partial\mathbf{r}_j$ and the sum is taken over all electrons. As indicated, the wave-function now depends on the set of all electron positions $\{\mathbf{r}_j\}$. Even for the case of just two electrons, (138) is impossible to solve analytically. The main problem is due to the coupling of all of the individual $\mathbf{r}_j$'s.

To make the problem tractable, some simplifying assumptions must be made. The most common is called the *central field approximation*. We partially account for the effects of the electron-electron repulsion by modifying the central potential, and then treat the residual electron-electron repulsion as a perturbation. That is, we define a zeroth order Hamiltonian by:

$$H_0 = -\frac{1}{2} \sum_j \nabla_j^2 + \sum_j V(r_j) \quad (139)$$

and a perturbing Hamiltonian by:

$$H' = \sum_{i>j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_j \left(\frac{Z}{r_j} + V(r_j) \right). \quad (140)$$

Here $V(r)$ takes the form of a screened Coulomb potential. Close to the nucleus,

$$V(r) \rightarrow \frac{-Z}{r} + C$$

where C is a constant. Far from the nucleus

$$V(r) \rightarrow \frac{-(Z - N + 1)}{r}$$

where N is the number of electrons in the atom. The constant C enters in because the outer electrons approximate a uniformly charged sphere where the electron is close to the nucleus, and the potential inside a uniformly charged sphere is constant.

In the central field approximation, the zeroth order Hamiltonian given by (139) is the sum of single particle Hamiltonians, and thus the zeroth order wave-functions can be written as the product of single particle wave-functions:

$$\psi(\{\mathbf{r}_j\}) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\dots\psi_N(\mathbf{r}_N) \quad (141)$$

where the individual $\psi_j(\mathbf{r}_j)$ are solutions to the single electron Schroedinger equation:

$$\left(\frac{1}{2}\nabla_j^2 + E - V(r_j)\right)\psi_j(\mathbf{r}_j) = 0 \quad (142)$$

and are individually characterized by the quantum numbers n , l , m_l , s , m_s . This would be sufficient if it were not for quantum indistinguishability. Because the atomic electrons form a system of *identical particles* and because they are fermions, the total wave-function must be anti-symmetric with respect to particle interchange. We can construct such an anti-symmetric wave-function by forming the following linear combination of product wave-functions:

$$\psi(\{\mathbf{r}_j\}) = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \psi_1(\mathbf{r}_{j_1}) \psi_2(\mathbf{r}_{j_2}) \dots \psi_N(\mathbf{r}_{j_N}). \quad (143)$$

Here, in each term in the sum, the set of single-electron wave-functions is arranged in the same order, but the electron coordinates, $\mathbf{r}_{j_1}, \mathbf{r}_{j_2}, \dots, \mathbf{r}_{j_N}$ have been arranged in a new order which is a permutation of the original set.

The sum is taken over all possible permutations. For each permutation, P represents the number of interchanges. Thus $(-1)^P = +1$ for even permutations and -1 for odd permutations. The wave-function given by (143) is often written in terms of what is called a *Slater determinant*:

$$\psi(\{\mathbf{r}_j\}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \dots & \psi_N(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \dots & \psi_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{r}_N) & \psi_2(\mathbf{r}_N) & \dots & \psi_N(\mathbf{r}_N) \end{vmatrix} \quad (144)$$

and is occasionally referred to as a *determinantal wave-function*. An important consequence of the anti-symmetrization is the *Pauli Exclusion Principle*: “No two electrons can occupy the same individual quantum state”. This can be seen to follow trivially from the Slater determinant. If two of the single particle wave-functions, ψ_i and ψ_j are identical then two columns in the matrix are identical and the determinant vanishes. The Pauli exclusion principle implies that for multi-electron atoms, even the ground state must involve electrons in the individual particle excited states. Recall that for principal quantum number n , there are $2n^2$ distinct spin and angular momentum states. If there are more than two electrons in the atom, at least some must be in an $n = 2$ or higher level. If there are more than ten electrons, some must be in an $n = 3$ or higher state.

The specification of the N individual particle quantum states for the set of N electrons is usually referred to as the *configuration*. The representation of the general wave-function $\psi(\{\mathbf{r}_j\})$ in terms of the Slater determinant is sometimes called the *single configuration approximation*.

3.6 Electron Exchange – Helium-like Atoms

A second important consequence of the anti-symmetrization of the wavefunction is the existence of what are called *electron exchange terms*. These are additional interaction terms which introduce spin dependence in the energy levels even when there is no explicit spin dependence in the Hamiltonian.

The key concepts are most simply illustrated by looking at the detailed level structure of helium-like atoms where there are two orbital electrons. The Hamiltonian for this system is:

$$H = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}} \quad (145)$$

where $r_{12} \equiv |\mathbf{r}_1 - \mathbf{r}_2|$. The Hamiltonian is spin-independent, so the eigenfunctions are functions only of the $\mathbf{r}_1$ and $\mathbf{r}_2$. However, because of the anti-symmetrization, there is a coupling to spin. Specifically, the total wavefunction can be written in only one of the two forms:

$$\psi = \varphi_S(\mathbf{r}_1, \mathbf{r}_2)\chi_A(m_{s1}, m_{s2}) \quad (146)$$

or

$$\psi = \varphi_A(\mathbf{r}_1, \mathbf{r}_2)\chi_S(m_{s1}, m_{s2}) . \quad (147)$$

Here φ denotes the *spatial* component of the wave-function, while χ denotes the *spin* component. The subscripts “S” and “A” indicate the symmetric and anti-symmetric combinations, respectively. Since the total wave-function must be anti-symmetric, one of the two must appear in a symmetric combination while the other must be anti-symmetric.

The symmetric spin-state is the so-called *triplet state*, where the total spin: $\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2$ has eigenvalue $s = 1$. This state has three-fold degeneracy; the degenerate eigenstate can be written in the form:

$$\begin{aligned} &|1/2, 1/2\rangle , \quad m_s = +1 \\ &\frac{1}{\sqrt{2}}(|1/2, -1/2\rangle + |-1/2, 1/2\rangle) , \quad m_s = 0 \\ &|-1/2, -1/2\rangle . \quad m_s = -1 \end{aligned}$$

Here the first index in each case is m_{s1} and the second index is m_{s2} . The anti-symmetric spin state is the *singlet state*, corresponding to $s = 0$. There is no degeneracy in this state. It can be written in the form:

$$\frac{1}{\sqrt{2}}(|1/2, -1/2\rangle - |-1/2, 1/2\rangle) . \quad m_s = 0$$

Invoking the central field approximation, we will treat the electron-electron repulsion term as the perturbation. For simplicity, we will take the central potential to be the simple Coulomb potential of the nuclear charge: $V(r) = -2/r$. In that case the spatial part of the wave-function is the product

wave-function of hydrogen-like eigenfunctions. The symmetric combination is:

$$\frac{1}{\sqrt{2}}(\varphi_1(\mathbf{r}_1)\varphi_2(\mathbf{r}_2) + \varphi_2(\mathbf{r}_1)\varphi_1(\mathbf{r}_2))$$

where φ_1 and φ_2 are each characterized by a particular choice of n , l , m_l . The anti-symmetric combination is:

$$\frac{1}{\sqrt{2}}(\varphi_1(\mathbf{r}_1)\varphi_2(\mathbf{r}_2) - \varphi_2(\mathbf{r}_1)\varphi_1(\mathbf{r}_2)) .$$

Now consider the ground state of the helium atom. Both of the electrons must be in the lowest energy orbital, corresponding to $n = 1$, $l = 0$. Since the two electrons are in the same spatial state, the spatial wave-function must be symmetric. In that case, the spin wave-function is anti-symmetric, so this is a singlet state. In first order perturbation theory, the correction to the energy level is given by:

$$\begin{aligned} \Delta E &= \left\langle \psi \left| \frac{1}{r_{12}} \right| \psi \right\rangle \\ &= \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 |\varphi_{10}(\mathbf{r}_1)|^2 |\varphi_{10}(\mathbf{r}_2)|^2 \frac{1}{r_{12}} . \end{aligned} \quad (148)$$

This expression has a simple classical interpretation: since $|\varphi_{10}(\mathbf{r}_1)|^2$ and $|\varphi_{10}(\mathbf{r}_2)|^2$ represent the probability density of finding the electrons at positions $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively, this is just the weighted average of the electrostatic repulsion energy between them.

Next consider the first excited states. In this case, one of the electrons is in the $n = 1$, $l = 0$ orbital, while the other is in an $n = 2$, $l = 0, 1$ orbital. In this case, there are two possible spatial wave-functions:

$$\frac{1}{\sqrt{2}}(\varphi_{10}(\mathbf{r}_1)\varphi_{20}(\mathbf{r}_2) + \varphi_{20}(\mathbf{r}_1)\varphi_{10}(\mathbf{r}_2))$$

which corresponds to the spin singlet, and

$$\frac{1}{\sqrt{2}}(\varphi_{10}(\mathbf{r}_1)\varphi_{20}(\mathbf{r}_2) - \varphi_{20}(\mathbf{r}_1)\varphi_{10}(\mathbf{r}_2))$$

which corresponds to the spin triplet. The first order perturbation theory correction to the energy level now has two terms:

$$\begin{aligned} \Delta E &= \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 |\varphi_{10}(\mathbf{r}_1)|^2 |\varphi_{20}(\mathbf{r}_2)|^2 \frac{1}{r_{12}} \\ &\quad \pm \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \varphi_{10}^*(\mathbf{r}_1)\varphi_{20}^*(\mathbf{r}_2)\varphi_{20}(\mathbf{r}_1)\varphi_{10}(\mathbf{r}_2) \frac{1}{r_{12}} \end{aligned} \quad (149)$$

where the (+) sign applies to the spin singlet combination and the (−) sign applies to the spin triplet. The first term has the same interpretation that we saw

earlier; it is the weighted average of the electrostatic repulsion energy. However, the second term is new. It appears because of the anti-symmetrization of the wave-function and is generally referred to as the *electron exchange term*. It can be shown that the integral for this term is always positive, so the triplet state has always lower energy. Thus the lowest excited state of helium-like atoms are spin triplet states.

A simple interpretation of the exchange energy is as follows: for a spin triplet combination, the spatial wave-function is anti-symmetric, so the Pauli exclusion principle requires that the electrons stay further apart. In that case, the electrostatic repulsion energy is reduced. For a spin singlet, the electrons are closer together on average and the electrostatic repulsion energy is enhanced.

3.7 Approximation Techniques for Multi-Electron Atoms

For more complicated multi-electron atoms, the electron-electron interaction is a significant perturbation and some form of approximation scheme is required to calculate wave-functions and energy levels. Within the context of the central field approximation, the simplest approach is to assume a central $V(r)$ which suitably accounts for the effects of electron shielding, and then to use this potential to calculate the single electron wave-functions which are the basic ingredients for the Slater determinant wave-function appropriate to the whole atom. Final wave-functions and energy levels are computed using first order perturbation theory, with the perturbation given by (140).

An early candidate functional form for the central potential was the *Thomas-Fermi potential* derived from a statistical treatment of the electron cloud as a gas of free-particle degenerate fermions at zero temperature. The potential is calculated classically from an assumed continuous charge density $\rho(\mathbf{r})$ and the form of $\rho(\mathbf{r})$ is adjusted so as to achieve a minimum in the total (kinetic plus potential) energies. This model yields moderately accurate energy levels for the valence shells of multi-electron near-neutral atoms, where the semi-classical assumptions involved are most reliable.

A more modern, and more accurate approach is to assume a convenient analytic form for the potential such as:

$$V(r) = -\frac{2}{r}((N-1)e^{-\alpha_1 r} + \alpha_2 r e^{-\alpha_2 r} + \dots + \alpha_{N-1} r^k e^{-\alpha_N r} + Z - N + 1) \quad (150)$$

characterized by the adjustable set of parameters: $\alpha_1, \alpha_2, \dots, \alpha_N$. For a given configuration, the values of the α_i 's are determined by minimizing the total energy of the atom. This yields a unique form for the potential for each electron configuration. That is sufficient for calculating energy levels. However, for the calculations of matrix elements (such as oscillator strengths), a common potential must be chosen, or otherwise the wave-functions describing initial and final states are not necessarily orthonormal. The *parametric*

potential method is computationally fast, and has been shown to yield reasonably accurate results, especially for highly charged ions, which are the dominant contributors to astrophysical X-ray spectra.

The most accurate conventional approach however is the *Hartree-Fock* or *self-consistent field* method. Here one takes a direct account of the dependence of the individual electron wave-functions on one another, which is brought about by the electron-electron repulsion term. The governing equations can be derived from the Ritz variational principle, i.e. using total wave-functions, ψ , constructed as Slater determinants of individual electron wave-functions, φ_i , we minimize the quantity $\langle \psi | H | \psi \rangle$ (where H is the total Hamiltonian) subject to the constraint that the individual wave-functions remain orthonormal. This can be accomplished by introducing N Lagrange multipliers ε_i , such that:

$$\delta(\langle \psi | H | \psi \rangle - \sum_i \varepsilon_i \langle \varphi_i | \varphi_i \rangle) = 0. \quad (151)$$

The result is a set of N equations (the Hartree-Fock equations) which look like Schroedinger equations, but with potentials that depends on the wave-function solutions:

$$\left[-\frac{1}{2} \nabla_i^2 - \frac{Z}{r_i} + \sum_{j \neq i} \int d^3 \mathbf{r}_j \frac{|\varphi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \varphi_i(\mathbf{r}_i) - \sum_{j \neq i} \delta(m_{si}, m_{sj}) \\ \times \left[\int d^3 \mathbf{r}_j \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \varphi_j^*(\mathbf{r}_j) \varphi_i(\mathbf{r}_j) \right] \varphi_j(\mathbf{r}_i) = \varepsilon_i \varphi_i \mathbf{r}_i. \quad (152)$$

Here m_{si} and m_{sj} are the eigenvalues of s_z for the i th and j th orbitals in the electron configuration, respectively. The first two terms on the left-hand side of (152) are associated with the single particle Hamiltonian ignoring the electron-electron interaction. The third term comes from the electron-electron repulsion energy. The fourth term is due to the exchange energy. It is zero unless the two orbitals have the same spin ($\delta(m_{si}, m_{sj}) = 1$), so that the spatial part of the wave-function is anti-symmetric.

For a given set of *trial* wave-functions $\varphi_i^{(0)}(\mathbf{r})$, the set of (152) can be solved to yield a new set of wave-functions $\varphi_i^{(1)}(\mathbf{r})$. This is repeated until it converges, i.e. until the resulting set of eigenfunction solutions is “close” to the trial set. The process yields a self-consistent potential for the electron-electron interaction which can then be used to calculate energy levels and matrix elements.

Hartree-Fock calculations are generally time-consuming and unwieldy in comparison to the simpler parametric potential methods discussed earlier. In addition, the self-consistent potential is not always smooth and well-behaved which can complicate the calculation of relativistic corrections (134 and 135) that are important for highly charged ions.

3.8 LS, jj and Intermediate Coupling

The Hamiltonian for the multi-electron atom as incorporated in (138) is rotationally invariant. In addition, it has no explicit spin dependence. This means that H must commute with the operators $\mathbf{J}$, $\mathbf{L}$ and $\mathbf{S}$:

$$[H, \mathbf{J}] = [H, \mathbf{L}] = [H, \mathbf{S}] = 0 \quad (153)$$

where $\mathbf{L}$ is the total orbital angular momentum of all the electrons in the atom: $\mathbf{L} = \sum_i \mathbf{l}_i$, $\mathbf{S}$ is the total spin angular momentum: $\mathbf{S} = \sum_i \mathbf{s}_i$ and $\mathbf{J}$ is the total angular momentum: $\mathbf{J} = \mathbf{L} + \mathbf{S}$. Hence, the eigenstates of H must also be eigenstates of J^2 , J_z , L^2 , L_z , S^2 and S_z and will thus be characterized by definite values of the corresponding eigenvalues: J , M_J , L , M_L , S , M_S , in addition to the energy E .

However, in the central field approximation, we have constructed the eigenfunctions out of single-electron wave-functions, which are themselves eigenfunctions of l^2 , l_z , s^2 , s_z , and are thus characterized by the eigenvalues l , m_l , s , m_s . The simple product wave-functions which comprise the Slater determinant will be characterized by a set of definite eigenvalues $l^{(i)}$, $m_l^{(i)}$, $s^{(i)}$, $m_s^{(i)}$ for each of the electrons in the atom. But L^2 does not commute with the individual $l_z^{(i)}$ operators and S^2 does not commute with the individual $s_z^{(i)}$. Hence these simple products cannot be eigenfunctions of the total Hamiltonian including the electron-electron repulsion.

Product states of definite L , M_L , S , M_S can however be generated by “coupling” individual product wave-functions into suitable superpositions. Here one uses the usual rules of angular momentum addition in quantum mechanics, and the coefficients of the various terms are given by Clebsch-Gordan coefficients. One first couples the spatial wave-functions individually into states of definite L^2 and L_z and the spin wave-functions individually into states of definite S^2 and S_z . One couples their product together to yield states of definite J^2 and J_z . This is called an *LS coupling scheme* or sometimes *Russell-Saunders coupling*.

The anti-symmetrization of the wave-function involves a superposition over permutations of the electron coordinates. Coupling involves a superposition over different values of $m_l^{(i)}$ and $m_s^{(i)}$. In principle, one can anti-symmetrize first and couple afterwards or couple first and anti-symmetrize afterwards. In practice, the latter is usually easier. The calculation of the matrix elements using these anti-symmetrized, coupled wave-functions can be quite complex if carried out by brute force. Fortunately, there is an elegant mathematical formalism known as *Racah algebra* – developed by Racah and Wigner in the 1940’s – which greatly simplifies the angular part of these matrix elements.

The discussion above ignores the relativistic corrections covered in Sect. 3.4. In particular, the spin-orbit term (134) in the single electron Hamiltonian is proportional to the operator $\mathbf{l} \cdot \mathbf{s}$, which does not commute with l_z and s_z , but

does commute with j^2 and j_z . When this term is important, it is convenient to first couple the individual particle wave-functions into states of definite $j^{(i)}$, $m_j^{(i)}$ and then couple these states into states of definite J , M_J . This is known as jj-coupling.

jj-coupling is formally incompatible with LS-coupling because states of definite L^2 , L_z , S^2 , S_z are not characterized by definite values of $j^{(i)}$, $m_j^{(i)}$. In practice, LS-coupling is preferred whenever the electron-electron repulsion term dominates over the spin-orbit terms. This is especially true for low- Z atoms which are not highly ionized. jj-coupling would be preferred for high- Z atoms with only a few electrons. In cases where both electron-electron and spin-orbit terms are important, neither scheme is entirely appropriate. In that case, one chooses one or the other as the basis, and then diagonalizes the “other” perturbing operator in this basis to achieve the appropriate superpositions. This is known as *intermediate coupling*. The final eigenstates are then only characterized by definite values of J and M_J .

3.9 Spectroscopic Notation and Ground-State Configurations

In LS-coupling, a given electron configuration is specified by the quantum numbers $n^{(i)}$, $l^{(i)}$, $s^{(i)}$ for each of the individual electrons and the total quantum numbers L , S , J , M_J for the atom as a whole. In the absence of an external field, the energy levels are degenerate in M_J so this is usually not included. In addition, all electrons have $s = 1/2$, so this too need not be indicated. Over the years, a notational scheme has become standard for designating these configurations. Specifically, for a given nl “shell” the number of electrons in that shell is indicated as an exponent. Recall that there are $2(2l + 1)$ distinct states in such a shell, so the exponent cannot exceed that number. For historical reasons, l is not indicated as an integer, but instead as a letter, with the assignments:

$$\begin{array}{cccccccc} l = & 0 & 1 & 2 & 3 & 4 & 5 & \dots \\ \text{symbol } s & p & d & f & g & h & \dots \end{array}$$

Thus the notation $3d^24f$ indicates two electrons with principal quantum number $n = 3$ and angular momentum $l = 2$ and one electron with $n = 4$ and $l = 3$.

For the total quantum numbers, the standard notation has the form

$$^{2S+1}L_J.$$

Here again a letter is used in place of a number for L and the convention is the same as that used for the individual l 's only with upper case letters instead of lower case. Thus the designation $^2D_{3/2}$ indicates a state with $S = 1/2$, $L = 2$ and $J = 3/2$.

For X-ray emitting astrophysical plasmas, we are mainly concerned with few electron atoms, specifically K- and L-shell ions, isoelectronic with the

neutral elements hydrogen through neon. Only a few key ideas are required to understand the ground configuration of such ions.

1. For a Coulomb potential, we have seen that the energy levels only depend on n not l . This is not true of the screened Coulomb potential appropriate to multi-electron atoms. The lower the angular momentum, the higher the probability that the electron is close to the nucleus where it “sees” less screening of the nuclear charge and hence the lower the energy. The energy therefore increases strongly with n and/or l .
2. Because of the strong dependence on n and l , as electrons are added to an ion, they continue to fill n, l “shells” until they are closed. A shell is closed when all of its magnetic spatial and spin orbitals are filled. A closed shell therefore has J, L and S all equal to zero.
3. For a partially open shell, the state of highest S will have the lowest energy. This is a consequence of the exchange energy, as we saw earlier. If S is maximal, the spin wave-function must be symmetric, which means that the spatial wave-function is anti-symmetric, and the electrons are on average further apart, thereby lowering their repulsion energy.
4. If the partially open shell is less than half-full, the lowest energy state will have the lowest possible value of J . This is a consequence of the spin-orbit interaction, which contributes positive energy that increases with J .
5. If the open shell is more than half-full, it is easier to think in terms of the electron “holes” rather than the electrons. These behave like positive electrons. Their spin-orbit contribution then has opposite sign. As a result, the lowest energy state has the highest possible J .

Using these rules, one can understand now the ground-states of hydrogen-like through neon-like ions have the following configurations:

H: $1s$	$^2S_{1/2}$
He: $1s^2$	$^1S_0$
Li: $1s^22s$	$^2S_{1/2}$
Be: $1s^22s^2$	$^1S_0$
B: $1s^22s^22p$	$^2P_{1/2}$
C: $1s^22s^22p^2$	$^3P_0$
N: $1s^22s^22p^3$	$^4S_{3/2}$
O: $1s^22s^22p^4$	$^3P_2$
F: $1s^22s^22p^5$	$^2P_{3/2}$
Ne: $1s^22s^22p^6$	$^1S_0$

In cases of intermediate coupling, which is important for highly charged ions, it is sometimes useful to also indicate the j -values of the individual electrons. This is done by adding a subscript to the individual shell terms indicating the value of j . Since the spin-orbit interaction for an individual electron has the lowest energy for the lowest values of j , the lower j states are filled first. Thus, in this notation, the ground configuration of oxygen-like ions is represented by $1s^22s^22p_{1/2}^22p_{3/2}^2$. Of course, for intermediate coupling, the L and S values

are not precisely defined. Typically, one lists the notation for the leading term in the LS expansion.

3.10 Configuration Interaction

In Sect. 3.5 we introduced the central field approximation and the associated single configuration approximation, where the total wave-function is written as an anti-symmetrized product of single-electron wave-functions. It should be emphasized that this is an approximation – it is by no means clear that the exact multi-electron eigenfunction of the total Hamiltonian is close to a single configuration wave-function, i.e. to a single Slater determinant. When this is not true, we need to allow for configuration mixing, by forming multi-configuration superpositions derived from matrix elements of the Hamiltonian. Codes which include these effects are called *multi-configuration* calculations.

It is impractical of course to include a large number of configurations in constructing the basis set. However, some guidance comes from the structure of the Hamiltonian. In LS-coupling, only configurations of common L , S , J and parity need be included. In addition, since the Hamiltonian only contains terms involving one or two electrons, interactions can only occur between configurations that differ in at most two orbitals.

Configuration interaction tends to be strong between configurations which are close in energy. For the highly charged ions important in X-ray emitting plasmas, the energy levels are more weakly dependent on l . Thus significant mixing can occur between configurations like $3s^23p^k$ and $3p^{k+2}$. In such cases, the identification of a particular transition with a set of upper and lower configurations is not very meaningful.

3.11 Selection Rules for Radiative Transitions

The matrix elements which appear in the various terms in the multipole expansion for radiative transitions can vanish for particular choices of initial and final states. This gives rise to what are called *selection rules* for the various multipole transitions. Transitions which violate the selection rules are called *forbidden*, while those consistent with the selection rules are *allowed*.

First, consider *electric dipole* transitions. Here the matrix element is $\langle f | \mathbf{r} | i \rangle$, where $\mathbf{r} = \sum_i \mathbf{r}_i$. Since $\mathbf{r}$ is a sum of single electron operators, this matrix element will vanish if the initial and final configurations differ by more than one electron orbital. Hence, only single electron transitions are allowed. Second, note that $\mathbf{r}$ has odd parity. Thus initial and final states must have opposite parity. Finally, since in spherical coordinates $\mathbf{r}_i$ can be written as a superposition of the spherical harmonics with $l = 1$, it is easy to show that this matrix element also vanishes unless $\Delta l = \pm 1$ for the change in the single electron orbital. The essential selection rules are $\Delta l = \pm 1$, $\Delta s = 0$, $\Delta L = 0, \pm 1$, $\Delta S = 0$, $\Delta J = 0, \pm 1$, with $J = 0 \rightarrow 0$ strictly forbidden.

Second, for *magnetic dipole* transitions, the matrix element is $\langle f | \boldsymbol{\mu} | i \rangle$, where $\boldsymbol{\mu}$ is the magnetic dipole moment. Including spin contributions, $\boldsymbol{\mu} \sim \mathbf{L} + 2\mathbf{S} = \mathbf{J} + \mathbf{S}$. Since $\mathbf{J}$ commutes with H , $\langle f | \mathbf{J} | i \rangle = 0$, so we are only left with $\langle f | \mathbf{S} | i \rangle$. This is a pure spin operator, so the net spatial configuration cannot change. Ignoring relativistic terms, $\mathbf{S}$ also commutes with H . However, the spin-orbit interaction introduces some mixing. The selection rules are $\Delta S = 0, \pm 1$ (spin flip), $\Delta J = 0, \pm 1$, no $J = 0 \rightarrow 0$, no parity change, no change in configuration (i.e. $\Delta n = 0$, $\Delta l = 0$ for all electrons).

And third, for *electric quadrupole* transitions, the selection rules are: $\Delta l = 0, \pm 2$, $\Delta L = 0, \pm 1, \pm 2$, $\Delta J = 0, \pm 1, \pm 2$, no $J = 0 \rightarrow 0$, no change in parity.

When configuration interaction is important, these selection rules can appear to be violated because of mixing. That is, even if the dominant configurations in the initial and final states violate the selection rules, there may be small admixtures in each case that do contribute to a non-zero matrix element.

4 Electron-Ion Collisional Processes

4.1 Overview

In the previous two chapters, I have laid out the essential ingredients for the calculation of radiative transitions rates between various energy levels and for the atomic structure effects which give rise to the particular characteristics of those levels. To predict the emergent X-ray spectra of astrophysical plasmas, however, we also need to understand the details of how excited atomic levels are populated. For the most part, that involves the study of electron-ion collisional processes in plasmas. This is also a rich and diverse field and it will not be possible to do justice to the full complexity of this topic. My emphasis, as in the previous chapter, will be on the explication of key concepts, definition of terms commonly used in the atomic physics literature and presentation of some quick back-of-the-envelope type calculations that enable us to derive rough estimates of the rate coefficients for these processes.

Each electron-ion collisional process is accompanied by a quantum mechanical *inverse*, which can be viewed as the same process *time-reversed*. Not surprisingly, the rates for direct and inverse processes involve common matrix elements, and are therefore related. The easiest way to derive these relations is to resort to *detailed balance* arguments, i.e. to set the rates for direct and inverse processes equal in strict thermodynamic equilibrium. I will defer an extensive discussion of thermodynamic equilibrium to the next chapter, but we will anticipate some important results from that discussion and utilize them here.

There are essentially four key electron-ion collisional processes that are important for X-ray emitting plasmas. These are schematically illustrated

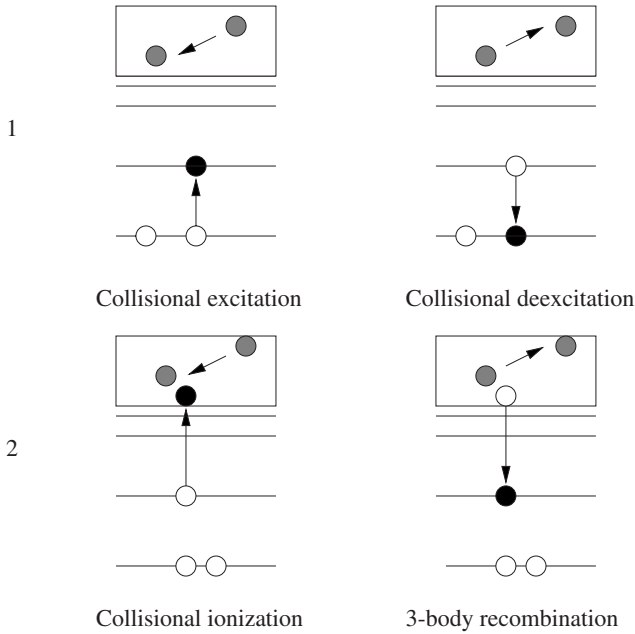


Fig. 2. The first two of the four key electron-ion collisional processes. The “inverse” process is on the right

in Figs. 2 and 3 where the “direct” process is depicted on the left and the “inverse” process on the right.

Collisional Excitation/Deexcitation

In collisional excitation, the interaction between a passing electron in a *continuum state* and a bound electron in a discrete state results in the excitation of the bound electron to a higher energy discrete level. To conserve energy, the colliding electron gives up a fraction of its energy and thus “falls” into a lower continuum state. The inverse process is collisional deexcitation, where a passing electron interacting with an excited atom actually gains energy as a result of the collision.

Collisional Ionization/3-Body Recombination

Collisional ionization is similar to collisional excitation, except that in this case, the final state of the initially bound electron is also a continuum state. The inverse process is 3-body recombination. Here, two, initially free electrons interact with the ion in the same collision. One of the two gets captured into a bound discrete level, while the other carries off the excess energy in a higher continuum state.

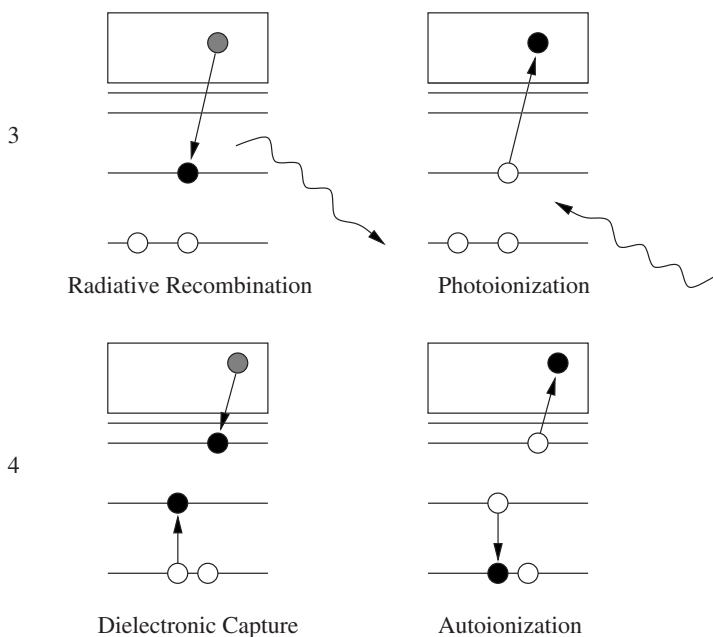


Fig. 3. The last two of the four key electron-ion collisional processes

Radiative Recombination/Photoionization

In radiative recombination a free electron in a continuum state decays into a bound discrete state through the emission of a photon. This is actually a form of spontaneous emission, similar to what we discussed for the radiative decay between two bound levels in Sect. 2.9. The inverse process is photoionization, or bound-free absorption, as discussed in Sect. 2.8.

Dielectronic Capture/Autoionization

Dielectronic capture is a resonant radiationless process in which the decay of an electron from a continuum state to a bound state is accompanied by the elevation of a core electron into an excited state. The resulting atom is doubly excited, and it has a total energy above the ionization potential of the initial ion. The inverse process is autoionization, where a doubly excited atom decays via the emission of a weakly bound outer electron. If the core excitation is associated with a “hole”, in one of the orbitals of an inner shell, this process is usually called *Auger decay*.

In the remainder of this chapter, I will review each of these processes in somewhat more detail.

4.2 Collisional Excitation – Scattering Theory

Collisional excitation is essentially an example of inelastic scattering of an electron off a complex atomic potential, and thus much of the formalism of quantum scattering theory can be applied to this process. Typically, one expresses the continuum wave-function at large distances from the atom as the sum of an incident plane wave and an outgoing spherical wave:

$$\varphi_c(\mathbf{r})_{r \rightarrow \infty} \simeq A \left[e^{i\mathbf{k}_i \cdot \mathbf{r}} + f(\vartheta, \varphi) \frac{e^{i\mathbf{k}_f \cdot \mathbf{r}}}{r} \right] \quad (154)$$

where $\hbar\mathbf{k}_i$ is the initial momentum of the electron, $\hbar^2 k_i^2/2m$ is its initial energy and $\hbar^2 k_f^2/2m$ is its final energy. The flux in the wave is given by:

$$\mathbf{j}(\mathbf{r}) = \frac{\hbar}{2mi} [\varphi^*(\nabla\varphi) - (\nabla\varphi^*)\varphi] \quad (155)$$

(see (109)). For the incident wave, this gives:

$$j_{in} = \frac{\hbar k_i}{m} |A|^2. \quad (156)$$

For the outgoing wave:

$$\begin{aligned} \mathbf{j}_{out} \cdot \mathbf{r} &= \frac{\hbar}{2mi} \left[\varphi^* \frac{\partial \varphi}{\partial r} - \frac{\partial \varphi^*}{\partial r} \varphi \right] \\ &= \frac{\hbar k_f}{m} \frac{|A|^2 |f|^2}{r^2}. \end{aligned} \quad (157)$$

The number of scattered electrons in solid angle element $d\Omega$ is: $(\mathbf{j}_{out} \cdot \mathbf{r})r^2 d\Omega$. Therefore, the differential cross-section for scattering is:

$$\frac{d\vartheta}{d\Omega} = \frac{(\mathbf{j}_{out} \cdot \mathbf{r})r^2}{j_{in}} = \frac{k_f}{k_i} |f(\vartheta, \varphi)|^2, \quad (158)$$

f is called the *scattering amplitude*.

If we limit our consideration to single electron transitions, then the total wave-function can be expressed in terms of product wave-functions for the colliding electron and the bound transitioning electron. These are still identical particles, so the total wave-function must be anti-symmetrized. Due to the exchange terms (see below), we get different answers for the singlet state and the triplet state. Averaging over the four possible spin states, the differential cross-section will then look like:

$$\frac{d\vartheta}{d\Omega} = \frac{k_f}{k_i} \left[\frac{1}{4} |f^+|^2 + \frac{3}{4} |f^-|^2 \right] \quad (159)$$

where the (+) indicates a symmetric spatial wave-function and the (−) indicates an anti-symmetric spatial wave-function.

The calculation of the scattering amplitude proceeds as follows: we write the total wave-function as the sum of anti-symmetrized product wave-functions for the initial and final states:

$$\begin{aligned} \psi = & [\varphi_{c_i}^{\pm}(\mathbf{r}_1)\varphi_{b_i}(\mathbf{r}_2) \pm \varphi_{c_i}^{\pm}(\mathbf{r}_2)\varphi_{b_i}(\mathbf{r}_1)] \\ & + [\varphi_{c_f}^{\pm}(\mathbf{r}_1)\varphi_{b_f}(\mathbf{r}_2) \pm \varphi_{c_f}^{\pm}(\mathbf{r}_2)\varphi_{b_f}(\mathbf{r}_1)] \end{aligned} \quad (160)$$

where $\varphi_{c_{i,f}}^{\pm}$ are the initial and final wave-functions for the colliding electron and $\varphi_{b_{i,f}}$ are the initial and final wave-functions for the bound electron. ψ must satisfy the Schrodinger equation:

$$\left[-\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + V(r_1) + V(r_2) + \frac{1}{r_{12}} \right] \psi = E_{tot}\psi. \quad (161)$$

Therefore, if we take a scalar product with $\varphi_{b_i}^*(\mathbf{r}_2)$ we must get:

$$\int d^3\mathbf{r}_2 \varphi_{b_i}^*(\mathbf{r}_2) \left[-\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + V(r_1) + V(r_2) + \frac{1}{r_{12}} - E \right] \psi = 0. \quad (162)$$

But

$$\left[-\frac{1}{2}\nabla_2^2 + V(r_2) \right] \varphi_{b_i}(\mathbf{r}_2) = E_{b_i} \varphi_{b_i}(\mathbf{r}_2), \quad (163)$$

and

$$E_{tot} = E_{b_i} + \frac{\hbar^2 k_i^2}{2m}. \quad (164)$$

Substitution of (160) into (162) yields

$$\begin{aligned} (\nabla_1^2 + k_i^2) \varphi_{c_i}^{\pm}(\mathbf{r}_1) = & 2 \left[V_{ii}(\mathbf{r}_1) \varphi_{c_i}^{\pm}(\mathbf{r}_1) + V_{if}(\mathbf{r}_1) \varphi_{c_f}^{\pm}(\mathbf{r}_1) \right] \\ \pm 2 \left[\int d^3\mathbf{r}_2 K_{ii}(\mathbf{r}_1, \mathbf{r}_2) \varphi_{c_i}^{\pm}(\mathbf{r}_2) + \int d^3\mathbf{r}_2 K_{if}(\mathbf{r}_1, \mathbf{r}_2) \varphi_{c_f}^{\pm}(\mathbf{r}_2) \right] \end{aligned} \quad (165)$$

where

$$V_{ii} \equiv V(\mathbf{r}_1) + \left\langle \varphi_{b_i} \left| \frac{1}{r_{12}} \right| \varphi_{b_i} \right\rangle \quad (166)$$

$$V_{if} \equiv \left\langle \varphi_{b_i} \left| \frac{1}{r_{12}} \right| \varphi_{b_f} \right\rangle \quad (167)$$

$$K_{ii}(\mathbf{r}_1, \mathbf{r}_2) \equiv \varphi_{b_i}^*(\mathbf{r}_1) \varphi_{b_i}(\mathbf{r}_2) \left[\frac{1}{r_{12}} - E_{tot} - E_{b_i} \right] \quad (168)$$

$$K_{if}(\mathbf{r}_1, \mathbf{r}_2) \equiv \varphi_{b_i}^*(\mathbf{r}_1) \varphi_{b_f}(\mathbf{r}_2) \left[\frac{1}{r_{12}} - E_{tot} - E_{b_i} - E_{b_f} \right] \quad (169)$$

The terms involving the V 's are the direct potential terms, the K 's are the exchange terms. A second similar equation can be obtained (with the i 's and

f's reversed) by taking the scalar product with $\varphi_{b_f}^*$ in place of $\varphi_{b_i}^*$ in (162). The result is a set of two coupled equations which can be solved simultaneously for $\varphi_{c_i}^\pm$ and $\varphi_{c_f}^\pm$ given expressions for φ_{b_i} and φ_{b_f} . They are analogous to the Hartree-Fock equations for a two electron atom. Once the continuum wave-functions are found, the scattering amplitudes can be computed and we obtain the cross-section.

The exchange terms can be important at low collision energies, especially for electric dipole forbidden transitions. At high energies, the continuum wave-functions, φ_{c_i} and φ_{c_f} oscillate strongly in comparison to the slowly varying K-functions and so the integrals on the right-hand side of (165) tend to vanish.

This procedure is still an approximation since we have not allowed the colliding electron to influence the bound-state wave-functions. One approach to correcting this is to include in the trial wave-function (160) other terms allowing for other proper collision channels, involving other sets of bound excited states. That is called a *close coupling calculation* since it couples in other states of the atom. It results in a much larger set of simultaneous equations, depending on how many channels are included.

At energies well above threshold, a much simpler calculation can be performed using the *Born approximation*. Here one assumes plane-wave wave-functions for both the initial and final continuum states. The transition rate can be calculated from time-dependent perturbation theory (see (70)) taking the electron-electron interaction as the perturbing potential:

$$\begin{aligned}
 R &= \frac{2\pi}{\hbar} \left| \left\langle f \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| i \right\rangle \right|^2 \delta(E_f - E_i) \\
 &= \left| \frac{2\pi}{\hbar} \frac{1}{V^2} \int_V d^3\mathbf{r}_1 d^3\mathbf{r}_2 e^{-i\mathbf{k}_f \cdot \mathbf{r}_2} \varphi_{b_f}^*(\mathbf{r}_1) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \varphi_{b_i}^*(\mathbf{r}_1) e^{i\mathbf{k}_i \cdot \mathbf{r}_2} \right|^2 \\
 &\quad \times \delta(E_f - E_i)
 \end{aligned} \tag{170}$$

where we have normalized the plane waves over a finite volume V .

Because exchange effects were found to be small at higher energies, one usually does not need to bother anti-symmetrizing the wave-function. The total rate is found by summing over the trial states of the outgoing electrons so that the δ -function gets replaced by a density of states factor:

$$\rho_f = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right) k_f. \tag{171}$$

The total rate thus scales like $1/V$. However, the incident flux is given by v_i/V in this picture, so the total cross-section is independent of the assumed volume, as expected.

A similar, but somewhat improved calculation can be obtained using continuum wave-functions of the Coulomb potential of the ion in place of the plane-waves. This is called the *Coulomb-Born method*. Even better yet is to

use the continuum wave-functions derived from the effective central potential $V(r)$ of the atom. That is the *distorted wave approach*. Usually distorted wave radial wave-functions are calculated in a *partial wave* expansion, summing over states of definite orbital angular momentum l . Close to threshold, the energy of the outgoing electron is low and only a small number of terms in the partial wave expansion need be kept. The maximum l required can be roughly estimated from classical considerations:

$$L \approx p_f a \Rightarrow l \approx k_f a \quad (172)$$

where a is the characteristic dimension of the atom. At high impact energies, many partial waves are required and the plane-wave Born approach provides a much simpler alternative.

The integral which appears in the plane-wave Born approximation (170), can be simplified using *the Bethe integral*:

$$\int d^3\mathbf{r} \frac{e^{i\mathbf{\Delta} \cdot \mathbf{r}}}{r} = \frac{4\pi}{\Delta^2} \quad (173)$$

which implies that the excitation cross-section is proportional to the square of a matrix element given by:

$$\frac{1}{\Delta^2} \int d^3\mathbf{r} \varphi_{b_f}^*(\mathbf{r}) e^{i\mathbf{\Delta} \cdot \mathbf{r}} \varphi_{b_i}(\mathbf{r}) \quad (174)$$

where $\mathbf{\Delta} \equiv \mathbf{k}_i - \mathbf{k}_f$. Note that the expression in (174) can be approximated by a multipole expansion:

$$e^{i\mathbf{\Delta} \cdot \mathbf{r}} \approx 1 + i(\mathbf{\Delta} \cdot \mathbf{r}) + \dots \quad (175)$$

entirely analogous to the multipole expansion invoked for radiative transitions in Sect. 2.10. Here again, $\mathbf{\Delta} \cdot \mathbf{r} \approx \mathbf{k} \cdot \mathbf{r} \approx v/c$, so for non-relativistic electrons, only the lowest order non-vanishing term usually needs to be considered. We thus obtain selection rules for collisional excitation between bound levels which are identical to the selection rules for radiative transitions between those levels. Therefore, transitions that are electric dipole forbidden also have low cross-section for collisional excitation. The above argument, however, relies on the plane-wave Born calculations, ignoring exchange effects. Generally, exchange terms dominate the cross-section for higher order multipole transitions.

4.3 Collisional Excitation – Classical Estimate

The discussion in Sect. 4.2 provides a sketch of how accurate collisional excitation cross-sections are calculated using sophisticated atomic codes, but is not especially helpful for getting quick quantitative estimates of the magnitude of collisional excitation rates. For this, it is more useful to resort to

simple classical arguments. Imagine a passing electron interacting via the Coulomb force with one of the orbital electrons in the atom. The momentum transfer to the bound electron is approximately:

$$\Delta p \approx \int_0^\infty dt F(t) \approx \frac{e^2}{b^2} \left(\frac{2b}{v} \right) = \frac{2e^2}{bv} \quad (176)$$

where b is the impact parameter of the colliding electron, and $\tau = 2b/v$ is the characteristic duration of the interaction. Thus, the energy transfer to the bound electron is:

$$\Delta E \approx \frac{(\Delta p)^2}{2m} \approx \frac{2e^4}{mb^2 v^2} . \quad (177)$$

The energy transfer must equal the energy of the excitation $\Delta E \approx E_{mn}$, where we are considering a transition from initial state m to final state n . The cross-section at impact parameter b is $\sigma \approx \pi b^2$ so:

$$\sigma_{mn} \approx \frac{2\pi e^4}{mv^2 E_{mn}} = \frac{\pi e^4}{E_e E_{mn}} \quad (178)$$

where E_e is the energy of the colliding electron. In atomic units:

$$\sigma_{mn}(E_e) \approx \frac{4\pi a_0^2}{E_e E_{mn}} \quad (179)$$

where a_0 is the Bohr radius.

It is traditional to express the cross-section in terms of a *collision strength* Ω_{mn} which is specific to the transition, but relatively independent of the electron energy:

$$\sigma_{mn}(E) \equiv \frac{\pi a_0^2}{g_m E_e} \Omega_{mn} \quad (180)$$

where g_m is the degeneracy of the initial state. One thus sees that classically $\Omega_{mn} \approx 4g_m/E_{nm}$. The quantum mechanical treatment (for electric dipole transitions) gives:

$$\frac{\Omega_{mn}}{g_m} = \frac{8\pi}{\sqrt{3}} \frac{f_{mn}\bar{g}}{E_{mn}} \quad (181)$$

where f_{mn} is the dipole absorption oscillator strength for the transition, and $\bar{g}$ is a *Gaunt factor* which is ≈ 1 for $\Delta n = 0$ transitions, and ≈ 0.2 for $\Delta n \neq 0$ transitions.

In thermal plasmas, collisional excitation can be characterized by a rate coefficient $C_{mn}(T)$, which is a function of electron temperature and is specific to the transition. The rate of collisional excitations for transition m to n per unit volume is given by $n_e n_i^m C_{mn}(T)$ where n_e is the free electron density and n_i^m is the density of the relevant ion in state m . In terms of the cross-section:

$$C_{mn}(T) = \int_{v_0}^\infty dv v f(v, T) \sigma_{mn}(v) \quad (182)$$

where $v_0 = (2E_{mn}/m)^{1/2}$ is the threshold velocity for the transition and $f(v, T)$ is the Maxwellian velocity distribution appropriate to a thermal plasma:

$$f(v, T) = 4\pi \left(\frac{m}{2\pi kT} \right)^{3/2} v^2 e^{-mv^2/2kT}. \quad (183)$$

The integration yields:

$$\begin{aligned} C_{mn}(T) &= \left(\frac{\pi a_0^2}{g_m} \right) \left(\frac{2kT}{\pi m_e} \right)^{1/2} \left(\frac{2Ry}{kT} \right) \Omega_{mn} e^{-E_{mn}/kT} \\ &\approx \frac{8.6 \cdot 10^{-6}}{g_m} T^{-1/2} \Omega_{mn} e^{-E_{mn}/kT} \text{ cm}^3/\text{s} \end{aligned} \quad (184)$$

where T is now in K.

The inverse of collisional excitation is collisional deexcitation. The principle of detailed balance asserts that in thermodynamic equilibrium, the rates for a process and its inverse must be equal. The rate for collisional excitation is $n_e n_i^m C_{mn}(T)$. The rate for collisional deexcitation is $n_e n_i^n C_{nm}(T)$. But in thermodynamic equilibrium, the level populations are related by the degeneracies and the Boltzmann factor:

$$\frac{n_i^n}{n_i^m} = \frac{g_n}{g_m} e^{-E_{mn}/kT}. \quad (185)$$

Thus:

$$\begin{aligned} C_{nm}(T) &= C_{mn}(T) \frac{g_m}{g_n} e^{E_{mn}/kT} \\ &= \frac{8.6 \cdot 10^{-6}}{g_n} T^{-1/2} \Omega_{nm} \text{ cm}^3/\text{s} \end{aligned} \quad (186)$$

where $\Omega_{nm} = \Omega_{mn}$. Note that for isoelectronic sequences, Ω_{nm} scales like $E_{nm}^{-1} \sim Z^{-2}$. In contrast, we saw earlier (Sect. 3.3) that radiative decay rates scale like Z^4 . Thus, for X-ray emitting plasmas, whose spectra are dominated by higher Z ions, we need very high electron densities before collisional deexcitation competes with spontaneous radiative decay.

4.4 Collisional Ionization

Collisional ionization is essentially the same process as collisional excitation except that the final state of the initially bound electron is now also a continuum state. The general quantum formalism outlined in Sect. 4.2 can clearly be applied to this case as well. With two continuum states in the final state, the square of the matrix element in (170) is proportional to $1/V^3$ instead of $1/V^2$ but there are now two density of states factors instead of one, so the final expression for the cross-section is still independent of the assumed volume.

As in the case of collisional excitation, there is a simple classical calculation that can be invoked to provide a rough estimate of the cross-section. This is originally due to Thomson and dates back to 1912 (before the discovery of the electron!). Thomson calculated the energy transfer between two same charges, assuming one is initially at rest:

$$\Delta E = \frac{E}{\left[1 + \frac{E^2 b^2}{e^2}\right]} \quad (187)$$

where E is the energy of the colliding electron and b is the impact parameter. Setting $\Delta E \geq \chi$, where χ is the ionization potential of the atom, one finds $b \leq b_c$, where:

$$b_c = \frac{e}{E} \left(\frac{E}{\chi} - 1 \right)^{1/2}. \quad (188)$$

The cross-section is thus given by:

$$\sigma = \pi b_c^2 = \pi e^2 \frac{1}{E^2} \left(\frac{E}{\chi} - 1 \right) = (4\pi a_0^2) \frac{1}{E^2} \left(\frac{E}{\chi} - 1 \right) \quad (189)$$

where the last expression is in atomic units, with E given in Rydbergs. This is a classical ionization cross-section per electron. It must be summed over all the electrons in the atom, using the appropriate χ value for each atomic shell and only including shells for which $E \geq \chi$.

This Thomson exchange cross-section provides a surprisingly good estimate of the true cross-section for $E \gg \chi$, but it gives a significant overestimate near threshold. This is due essentially to two effects:

1. The calculation ignores the initial binding energy of the target electron;
2. It does not allow for the possibility that if too much energy is transferred, the colliding electron itself becomes bound.

Hutchinson [7] suggests a simple modification that partially corrects for these two effects:

$$\sigma = 4\pi a_0^2 \frac{1}{E(E + E_+)} \left(\frac{E}{\chi} - 1 \right) \quad (190)$$

where E_+ is an adjustable parameter which is approximately a few times χ .

The cross-section given in (190) can be integrated analytically over a Maxwellian distribution (as in 182) to yield a rate coefficient. The result involves an exponential integral, but Hutchinson shows that to a good approximation one obtains:

$$\begin{aligned} C(T) &= \langle \sigma v \rangle = 4\pi a_0^2 \left(\frac{8kT}{\pi m} \right)^{1/2} \frac{Ry^2}{\chi(\chi + E_+)} e^{-\chi/kT} \left[1 - e^{-(\chi + E_+)/kT} \right] \\ &\approx (8.5 \cdot 10^{-8}) \left[\frac{Ry^2}{\chi(\chi + E_+)} \right] \left[\frac{kT}{Ry} \right]^{1/2} e^{-\chi/kT} \left[1 - e^{-(\chi + E_+)/kT} \right] \text{ cm}^3 \text{ s}^{-1}. \end{aligned} \quad (191)$$

Similar (but not identical) formulae have been derived empirically from fits to experimental data by Lotz [8] and others. These generally agree with one another to within a factor of two.

The inverse of collisional ionization is 3-body recombination. However, since this process involves the collision of two electrons with the atom in the same interaction, it is usually only important at very high densities ($n_e \geq 10^{19} \text{ cm}^{-3}$), which rarely apply to X-ray emitting astrophysical plasmas.

4.5 Radiative Recombination

Radiative recombination involves the capture of a free electron, accompanied by the emission of a photon with energy given by:

$$\hbar\omega_n = E + \chi_n \quad (192)$$

where E is the initial energy of the electron, and χ_n is the ionization potential of the level into which the electron is captured. Since this is a radiative process, it may be calculated using the techniques outlined in Chap. 2. In particular, we can get a quick semi-quantitative estimate of the cross-section from a classical treatment, where we view radiative recombination as a kind of discrete limit of classical bremsstrahlung, the radiation emitted by an electron as it is accelerated in the Coulomb field of an ion. The energy emitted per unit frequency per unit time per unit volume due to bremsstrahlung by electrons of velocity v is given by:

$$\frac{dW}{d\omega dV dt} = \frac{16\pi e^6}{3\sqrt{3}c^3 m^2 v} n_e n_i Z^2 g \quad (193)$$

where n_e is the electron density, n_i is the ion density, Z is the charge on the ion and g is a Gaunt factor of order unity (see [2]). For radiative recombination, the final state of the electron is discrete, so the energy radiated must all come out at a single frequency given by (192). We may thus write:

$$\frac{dW_n}{dV dt} = \frac{16\pi e^6}{3\sqrt{3}c^3 m^2 v} n_e n_i Z^2 g(\Delta\omega_n) \quad (194)$$

where $(\Delta\omega_n)$ is the frequency difference between two neighboring shells. Adopting an “hydrogenic approximation” for the energy levels:

$$\chi_n \approx \frac{Z^2 Ry}{n^2} \quad (195)$$

$$\Delta\omega_n = \frac{2Z^2 Ry}{\hbar n^3} \approx \frac{2\chi_n}{\hbar n} . \quad (196)$$

We define a cross-section $\sigma_n(v)$ by setting:

$$\frac{dW_n}{dV dt} = n_e n_i v \sigma_n(v) \hbar\omega_n . \quad (197)$$

Plugging in the relevant expressions from (192), (194) and (195) and solving for $\sigma_n v$ yields:

$$\sigma_n(v)v = \frac{16\pi e^6 Z^2 g}{3\sqrt{3}c^3 m^2} \frac{2\chi_n}{n \left(\frac{1}{2}mv^2 + \chi_n\right)} \frac{1}{v}. \quad (198)$$

Finally, averaging over a Maxwellian velocity distribution yields a rate coefficient as a function of temperature:

$$\alpha(T) \equiv \langle \sigma_n(v)v \rangle \approx (5.2 \cdot 10^{-14}) g Z^2 \left(\frac{\chi_n}{kT}\right)^{3/2} e^{\chi_n/kT} E_i\left(\frac{\chi_n}{kT}\right) \text{ cm}^3 \text{ s}^{-1} \quad (199)$$

where $E_i(x)$ is the exponential integral.

To get more accurate estimates from a quantum mechanical calculation, it is usually easier to first calculate the photoionization cross-section and then resort to a detailed balance argument to find the cross-section for radiative recombination. Let $\sigma_{PI}(\omega)$ be the photon cross-section for photoionization at frequency ω and let $\sigma_{RR}(v)$ be the electron cross-section for radiative recombination at electron velocity v . As we have seen, ω and v are related by energy conservation (192) with $E = 1/2 mv^2$. Let n_i be the density of the i th ionic species and n_{i+1} be the density of the one higher ionization state. Then the rate of recombinations per unit volume in the velocity range v to $v + dv$ is given by:

$$dR_{RR}(v) = n_e \sigma_{RR}(v) v f(v) dv n_{i+1} \quad (200)$$

where $f(v)$ is the Maxwellian electron distribution in velocity. The rate of photoionizations per unit volume in the frequency range ω to $\omega + d\omega$ is given by:

$$dR_{PI}(\omega) = \frac{F(\omega)d\omega}{\hbar\omega} \sigma_{PI}(\omega) n_i (1 - e^{-\hbar\omega/kT}) \quad (201)$$

where $F(\omega)$ is the energy flux per unit frequency in the radiation field. In thermodynamic equilibrium, this is given by the expression:

$$F(\omega) = \frac{\hbar\omega^3}{\pi^2 c^2} \frac{1}{(e^{\hbar\omega/kT} - 1)} \quad (202)$$

(see Sect. 5). The last factor which appears in (201) is a correction for stimulated emission – in thermodynamic equilibrium, there are always photon-induced radiative decays in addition to spontaneous radiative decays. Thus (201) gives a *net* photoabsorption rate. Using the expression we had earlier for the Maxwellian distribution (183), and equating the rates in (200) and (201) yields:

$$\frac{\sigma_{PI}(\omega)}{\sigma_{RR}(v)} = \frac{n_e n_{i+1}}{n_i} \left(\frac{m}{2\pi kT}\right)^{3/2} v^3 e^{-(mv^2/2 - \hbar\omega)/kT} \frac{dv}{d\omega}. \quad (203)$$

But $1/2 mv^2 - \hbar\omega = -\chi$ and $dv/d\omega = (d\omega/dv)^{-1} = \hbar/mv$. The ratio of the densities is given by the Saha equation which we will introduce in the next chapter:

$$\frac{n_e n_{i+1}}{n_i} = \frac{2g_{i+1}}{g_i} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2} e^{-\chi/kT} \quad (204)$$

where g_{i+1} is the degeneracy of the final state of the more highly ionized ion, and g_i is the degeneracy of the less ionized ion (see Sect. 5). Collecting terms yields:

$$\frac{\sigma_{PI}(\omega)}{\sigma_{RR}(v)} = \frac{m^2 c^2 v^2}{\hbar^2 \omega^2} \frac{g_{i+1}}{g_i} \quad (205)$$

which is called the *Milne relation*.

The quantum mechanical calculation of photoionization cross-sections was discussed in Sect. 2.8. For hydrogen-like ions, we can obtain an analytical expression. Averaging over l , the cross-section for ionization out of the n th shell is given by:

$$\sigma_n(\omega) = \frac{64\alpha}{3^{3/2}} \frac{Z^4}{n^5} \left(\frac{Ry}{\hbar\omega} \right)^3 \pi a_0^2 g \quad (206)$$

[if $\hbar\omega > Z^2 Ry/n^2$ and is zero otherwise] where g is again a Gaunt factor of order unity. The ω^{-3} dependence is also typical of photoionization cross-sections of more complex atoms. The monochromatic emissivity (energy radiated per unit volume per unit frequency) associated with recombination radiation is given by:

$$\begin{aligned} \frac{dW}{dt d\omega dV} &= n_e n_{i+1} (\hbar\omega) v f(v) \sigma_{RR}(v) \frac{dv}{d\omega} = \left(\frac{2}{\pi} \right)^{1/2} n_e n_i \frac{g_i}{g_{i+1}} \\ &\times \left(\frac{\hbar\omega}{\chi} \right)^3 \left(\frac{\chi^2}{mc^2 kT} \right)^{3/2} \hbar c \sigma_{PI}(\omega) e^{-\hbar\omega/kT} e^{\chi/kT}. \end{aligned} \quad (207)$$

Notice that for $\sigma_{PI}(\omega) \sim \omega^{-3}$, the frequency dependence is essentially exponential above threshold.

4.6 Dielectronic Recombination and Autoionization

Dielectronic capture involves the capture of a free electron into a bound level with the accompanying excitation of a core electron. The resulting recombined atom is doubly excited. It can decay by autoionization, ejecting the captured electron back out into the continuum. In that case, there is no net change in the level of ionization of the atom. However, the doubly excited atom can also decay radiatively, thereby lowering its total energy below the ionization potential of the recombined atom. When this occurs, the recombination is complete and the atom is left in a stable configuration with one extra electron. The complete process – dielectronic capture followed by radiative decay is usually referred to as *dielectronic recombination*. This can be

a very important process in astrophysical plasmas, especially for ions, as we shall see shortly.

Let's first consider the inverse process, autoionization. Its rate (from time dependent perturbation theory) is given by:

$$A_a = \frac{2\pi}{\hbar} \left| \left\langle f \left| \frac{e}{|r_i - r_j|} \right| i \right\rangle \right|^2 \quad (208)$$

where f and i represent the appropriate product wave-functions for the two electrons involved in the interaction in the initial and final states. Note that in the final state, one of the electrons is in a continuum state. Since the continuum states have wave-functions which are normalized to a delta-function in energy, this wave-function has units of energy^{-1/2}. Therefore, the square of the matrix element has units of energy, not energy-squared, as one would otherwise expect. When divided by $\hbar$, it gives a finite rate.

The matrix element which appears in the autoionization decay rate (208) is the same matrix element one would use to calculate the configuration interaction between the doubly bound level and the continuum level with equal energy. In some sense, autoionization is a consequence of configuration interaction. The diagonalized eigenstate of the perturbation is then a superposition of the initial discrete state and a range of continuum states:

$$\psi = a\psi_{discrete} + \int dE b(E)\psi_{continuum}(E) \quad (209)$$

with the coefficient a and $b(E)$ determined by the configuration interaction matrix-element. It can be shown (see [1] pp. 526–535) that the width of the function $b(E)$ is given roughly by $A_a\hbar$, as one would expect based on the energy-time uncertainty principle. The autoionization process by assigning a finite lifetime to the doubly excited level, broadens this level into a narrow continuum whose width is inversely related to that lifetime.

The presence of the configuration interaction also gives rise to characteristic absorption line profiles for photoionization in the vicinity of autoionizing resonances. The continuum state can, of course, be reached by photoexcitation of a core electron. If there were no configuration interaction, these two processes would be distinct and the photoabsorption spectrum would consist of a discrete absorption line on a photoionization continuum, as shown in Fig. 4, left panel. However, with configuration interaction, the final state wave-function is as given in (209), and we get interference between the two channels. The photoabsorption spectrum in this case looks like Fig. 4, right panel, which is called a Beutler-Fano absorption profile. Such features are expected in the extreme ultraviolet spectra of nearby white dwarf stars due to photoabsorption by neutral helium in the intervening interstellar medium [9]. The features so far observed have been associated with autoionizing resonances of neutral helium.

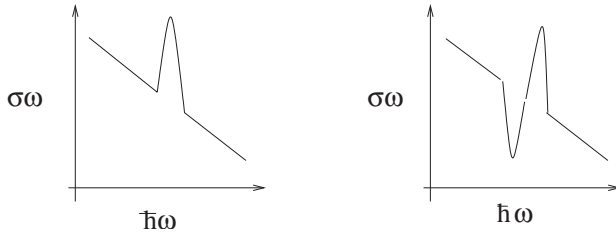


Fig. 4. Spectra without configuration interaction (*left*) and Beutler-Fano profile (*right*)

Note that using the simple Z -scaling arguments we invoked earlier, autoionization decay rates are roughly independent of Z for isoelectronic sequences. This is because the outgoing continuum wave-function is proportional to $E^{-1/2} \sim Z^{-1}$, while the perturbation Hamiltonian $\sim r^{-1} \sim Z^{+1}$. Thus, the matrix element is $\sim Z^0$. This means that autoionization is extremely important for low Z ions, but becomes less and less important in comparison to radiative decay for high Z ions. We will return to this shortly.

We can derive a rate coefficient for dielectronic capture by resorting to detailed balance arguments. The process is resonant, so the cross-section is actually infinite at the velocity which satisfies energy conservation:

$$\frac{1}{2}mv_c^2 = E_i^{**} - E_{i+1} \quad (210)$$

where E_i^{**} is the energy of the doubly excited recombined ion, and E_{i+1} is the energy of the ground-state of the initial ion. That is:

$$\sigma_{dc}(v) = \alpha_{dc} \delta(v - v_c) \quad (211)$$

where α_{dc} has units of cm^3s^{-1} . If n_{i+1} is the density of $i + 1$ ions in the ground state, then the rate of dielectronic captures per unit volume per unit time is given by:

$$R_{dc} = \int d^3v n_e n_{i+1} v \sigma_{dc}(v) f(v) = 4\pi n_e n_{i+1} \alpha_{dc} v_c^3 e^{-mv_c^2/kT} \left(\frac{m}{2\pi kT} \right)^{3/2} \quad (212)$$

where $f(v)$ is the Maxwellian distribution given in (183). If n_i^{**} is the density of i ions in the doubly excited state, then the autoionization rate per unit volume is:

$$R_{auto} = n_i^{**} A_a \quad (213)$$

These rates must be equal in thermodynamic equilibrium. But, in thermodynamic equilibrium, the level populations are given by:

$$\frac{n_i^{**}}{n_i} = \frac{g_i^{**}}{g_i} e^{-(E_i^{**} - E_i)/kT} \quad (214)$$

where n_i , g_i and E_i are the density, degeneracy and energy of the i th ion in the ground state (see Sect. 5), and the ionization structure $n_e n_{i+1}/n_i$ is given by the Saha equation:

$$\frac{n_e n_{i+1}}{n_i} = \frac{2g_{i+1}}{g_i} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2} e^{-\chi/kT} \quad (215)$$

where $\chi = E_{i+1} - E_i$ is the ionization potential for the i th ion. Collecting terms gives:

$$\alpha_{dc} = \frac{g_i^{**}}{2g_{i+1}} 2\pi^2 \left(\frac{\hbar}{mv_c} \right)^3 A_a. \quad (216)$$

Not surprisingly, the temperature drops out since dielectronic capture and autoionization must be related by fundamental constants.

The dielectronic capture rate is obtained by plugging (216) back into (212):

$$R_{dc} = n_e n_{i+1} \frac{g_i^{**}}{2g_{i+1}} A_a \left(\frac{h^2}{2\pi mkT} \right)^{3/2} e^{-mv_c^2/2kT}. \quad (217)$$

To get the *dielectronic recombination rate*, as opposed to dielectronic capture rate, we must multiply the expression in (217) by the probability that the doubly excited atom stabilizes radiatively. Quite generally, this probability is given by the ratio of the sum of all radiative decay rates from the excited state to the sum of all radiative plus autoionizing decay rates:

$$\text{Probability of stabilization} = \frac{\sum A_r}{\sum (A_r + A_a)}. \quad (218)$$

Usually, however, there is only one dominant decay channel in each case, which involves the decay of the core excitation. Thus, the dielectronic rate coefficient becomes:

$$R_{dr} \approx n_e n_{i+1} \frac{g_i^{**}}{2g_{i+1}} \left(\frac{h^2}{2\pi mkT} \right)^{3/2} e^{-mv_c^2/2kT} \left(\frac{A_a A_r}{A_a + A_r} \right). \quad (219)$$

The factor in parenthesis has a maximum when $A_a = A_r$. Hence, dielectronic recombination is efficient when the rates for autoionization decay and radiative decay of the core excitation are approximately equal. Since $A_a \sim Z^0$ and $A_r \sim Z^4$, this is primarily the case for high- Z ions.

We can get a further quantitative feel for how these rates compare by again using a semi-classical treatment. Note that the dielectronic capture process is very similar to collisional excitation, except that the final state of the colliding electron is now a bound state rather than a continuum state. We should therefore be able to get a rough idea of the rate coefficient for this process by extending our earlier classical treatment of collisional excitation to energies below threshold. Recall that our earlier expression for the excitation cross-section was given by (180):

$$\sigma_{mn}(E) \equiv \frac{\pi a_0^2}{g_m E_e} \Omega_{mn} . \quad (220)$$

For capture into principal quantum number n , we can integrate this expression over the velocity range between neighboring Rydberg levels to yield an estimate for α_{dc} associated with this core excitation:

$$\alpha_{ij,n}^{dc} \approx \sigma_{ij}(\delta v) \approx \sigma_{ij} \frac{2Z^2 Ryd}{n^3 m v} = \frac{\pi e^4}{g_i} \Omega_{ij} \frac{Z^2 Ryd}{n^3 m^2 v_c^3} . \quad (221)$$

But $\Omega_{ij}/g_i = 2\pi/\sqrt{3}f_{ij}\bar{g}/E_{ij}$ (181) and

$$f_{ij} = \frac{3}{2} \frac{mc^3}{e^2} \frac{\hbar^2}{E_{ij}^2} A_{ij}^r \quad (222)$$

(Equation 113). Plugging these expressions in and equating $\alpha_{ij,n}^{dc}$ from (221) to α^{dc} from (216) we obtain:

$$\frac{A_{ij}^a}{A_{ij}^r} = \frac{12}{\sqrt{3}} \frac{g_{i+1}}{g_i^{**}} \bar{g} \frac{Z^2}{n^3} \left(\frac{Ryd}{E_{ij}} \right)^3 \frac{1}{\alpha^3} . \quad (223)$$

Note that since $E_{ij} \sim Z^2$, this ratio scales like Z^{-4} , as expected from our earlier discussion. Taking all other features to be of order unity, with $E_{ij} \sim Z^2 Ryd$, this ratio is found to be $\sim Z^{-4} \alpha^{-3}$. Setting it equal to unity (for maximum dielectronic recombination efficiency) then implies $Z \approx 40$. So we see that dielectronic recombination becomes important only for the higher- Z elements, most notably iron.

5 Types of Equilibria

In most astrophysical settings, some form of equilibrium applies, in which there is a balance between competing processes, e.g. heating and cooling, ionization and recombination, excitation and deexcitation, etc. The nature of the equilibrium has a very important effect on the emergent spectrum.

There are three “systems” which may or may not equilibrate with one another:

- the kinetic distributions of the electrons and ions;
- the atomic level populations;
- the radiation field.

We say that we have strict thermodynamic equilibrium when all three systems are characterized by statistical distributions at the same temperature T . In particular, for this case, the radiation field is characterized by the blackbody distribution, so the spectrum is especially simple. For absolute equilibrium, the temperature T , must also be independent of spatial position within the

gas. However, as long as the scale length for temperature variations: $T/|\nabla T|$ is long compared to all relevant mean free paths for particle and photon interactions, it is appropriate to talk about strict local thermodynamic equilibrium, where $T = T(\mathbf{r})$.

The more common term, *local thermodynamic equilibrium* (LTE) usually applies to the situation where the particle distributions and level populations are in equilibrium, but the radiation field is not, i.e. the scale lengths of the system are not sufficient to trap emitted photons and enforce thermalization.

5.1 Properties of LTE

In LTE, the population of a given energy level is proportional to the degeneracy in that level and a Maxwell-Boltzmann factor $e^{-E/kT}$. This gives rise to:

The Maxwellian velocity distribution for free particles

$$\frac{n(v)dv}{n} = 4\pi v^2 \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-\frac{mv^2}{2kT}} dv, \quad (224)$$

The Maxwell-Boltzmann distribution for level populations

$$\frac{n_j^z}{n^z} = \frac{g_j^z}{U^z(T)} e^{-\frac{E_j^z - E_0^z}{kT}}, \quad (225)$$

where $U^z(T)$ is the partition function:

$$U^z(T) = \sum_j g_j^z e^{-\frac{E_j^z - E_0^z}{kT}} \quad (226)$$

and *The Saha equation for the ionization balance*

$$\frac{n_e n^{z+1}}{n^z} = \frac{2U^{z+1}(T)}{U^z(T)} \frac{(2\pi m kT)^{3/2}}{h^3} e^{-\frac{\chi^z}{kT}}. \quad (227)$$

The definition of $U^z(T)$ can be problematic. For example, for H-like atoms

$$g_n = 2n^2 \quad (228)$$

$$e^{-\frac{E_n - E_0}{kT}} = e^{-\frac{z^2 R y}{kT} (1 - \frac{1}{n^2})} \quad (229)$$

$$\Rightarrow U^z(T) \rightarrow \infty; \quad (230)$$

we must truncate the expansion at some high Rydberg level. This is usually a function of the particle density, due to the effects of neighboring charges.

In LTE, the prediction of the emergent spectrum requires the solution of the radiative transfer equations

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu \quad (231)$$

$$S_\nu = \frac{j_\nu}{k_\nu} \quad (232)$$

$$d\tau_\nu = k_\nu ds \quad (233)$$

Here, I_ν is the specific intensity of the radiation field, j_ν is the emissivity of the gas, and k_ν is the opacity, all of which are functions of the position along the path of propagation s . S_ν is called the source function.

For discrete lines:

$$j_{nm} = \frac{h\nu_{nm}}{4\pi} g_m n_n A^{nm} \varphi(\nu) \quad (234)$$

$$k_{nm} = g_m n_m \sigma_{mn}(\nu) - g_n n_n \sigma_{mn}(\nu) \quad (235)$$

But from radiation theory, we found:

$$A_{nm} = \frac{8\pi^2}{3} \frac{e^2}{mc^3} f_{nm} \nu^2 \quad (236)$$

$$\sigma_{mn}(\nu) = \sigma_{nm}(\nu) = \frac{1}{3} \frac{\pi e^2}{mc} f_{nm} \quad (237)$$

and, relating the level populations using the Maxwell-Boltzmann distribution (225), we get:

$$S_{nm} = \frac{j_{nm}}{k_{mn}} = \frac{2h\nu_{nm}^3}{c^2} \frac{1}{(e^{h\nu_{nm}/kT} - 1)} = B_{\nu_{nm}}(T) \quad (238)$$

which is the blackbody function evaluated at the frequency of the transition ν_{nm} !

Looking inward to an optically thick medium at constant temperature, (231) implies:

$$I_\nu(\tau_\nu) = B_\nu(T)(1 - e^{-\tau_\nu}) \quad (239)$$

The line intensities are “limited” to the blackbody intensity evaluated at the local temperature.

For the approximation of LTE to hold, we need the rates for collisional deexcitation of discrete levels to be comparable to the rates for spontaneous radiative decay:

$$n_e C_{nm}(T) \sim A_{nm} \quad (240)$$

$$\Rightarrow n_e \sim 9 \cdot 10^{19} T_K^{1/2} (\delta E)_{keV}^3 cm^{-3} \quad (241)$$

In astrophysical settings, such high densities are only reached in the atmospheres of compact objects like white dwarfs and neutron stars.

When the assumption of LTE is invalid, the calculation of the emergent spectrum can be much more complicated. In general, we have to explicitly

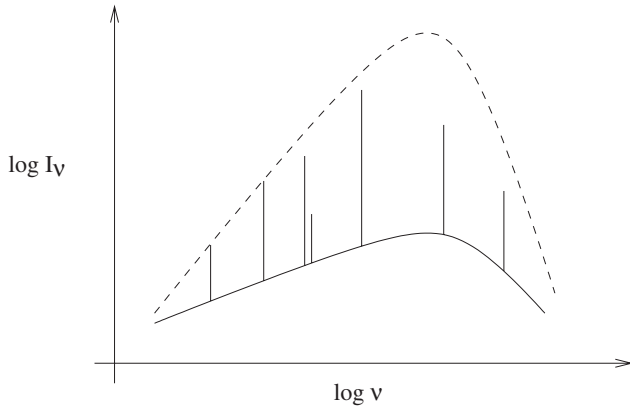


Fig. 5. An illustration of the limitation of line intensities to the blackbody intensity for cases where LTE holds

account for all microphysical processes that feed and deplete the individual quantum levels. The most general, time-dependent equations are of the form:

$$\frac{dn_i^z}{dt} = -n_i^z \sum_j R_{ij} + \sum_{z',k} n_k^{z'} R_{ki} \quad (242)$$

where the R 's represent the rates for collisional and photon interactions coupling levels within the same charge state and in neighboring charge states.

5.2 Coronal Equilibrium

Equation (242) is difficult to solve because of the requirement for inclusion of such a large array of diverse processes. Therefore, it is useful to adopt some approximations, applicable to particular cases. One of the most important sets of approximations applies to the case of *coronal equilibrium*, sometimes also referred to as *collisional ionization equilibrium*.

There are three basic assumptions underlying this limit:

- Excitation and ionization are dominated by electron-ion collisions. Deexcitation is dominated by spontaneous radiative decay.
- Densities are low enough so that atoms are always in their ground states.
- The radiation field has a negligible effect on the atomic populations, and the plasma is optically thin, so photoabsorption and scattering can be ignored.

Sources of applicability for these assumptions include: stellar coronae, the shocked gas of older supernova remnants, and the intracluster media of galaxy clusters.

The charge state distribution in coronal equilibrium is determined by a balance of collisional ionization and radiative and dielectronic recombination:

$$\frac{dn_z}{dt} = -n_e n_z (C_z + \alpha_z) + n_e n_{z+1} \alpha_{z+1} + n_e n_{z-1} C_{z-1} \quad (243)$$

Here C_z represents the rate coefficient for collisional ionization (see Sect. 5.4), and α_z represents the combined RR + DR rate coefficient for recombination (Sects. 4.5 and 4.6, respectively). Note that the characteristic timescales for equilibrium to be established are $\sim(n_e C)^{-1}$ or $\sim(n_e \alpha)^{-1}$. These can be larger than 10^3 yr for $n_e \leq 1 \text{ cm}^{-3}$, as found in young supernova remnants. Since this age exceeds the age of the remnant (for the most recent supernovae), the shocked gas that we observe for these cases may still be ionizing, and the charge balance may be far from equilibrium. A similar situation can be found during weak flares in stellar coronae. Here the electron density is closer to $n_e \sim 10^{10} \text{ cm}^{-3}$, so the equilibration time is of order a few seconds, comparable in some cases to the duration of the flare.

However, if equilibrium is established, so that the left-hand side of (243) vanishes, the electron density n_e , drops out of the equation, and the resulting steady-state ionization structure becomes a function only of temperature. This turns out to be also true of the discrete spectrum. Specifically, since we are assuming that the atoms are “always” in the ground state, the populations of upper levels are given by the ratio of collisional excitation rates from the ground level, to the spontaneous radiative decay rates back down:

$$n_2 = \frac{n_e n_1 \gamma_{12}(T)}{A_{21}}, \quad (244)$$

and the line emissivities become:

$$\epsilon_{21} = n_e n_1 \gamma_{21}(T) E_{12}, \quad (245)$$

where $\gamma_{12}(T)$ is the collisional excitation coefficient (Sect. 5.3), and E_{12} is the energy of the transition. The density of the ion in the ground state is given by $n_1 = A_{elem} f_Z(T) n_H$, where A_{elem} is the abundance of the element relative to hydrogen, and $f_Z(T)$ is the steady-state ion fraction, as discussed above. It is useful to define a *line power* for the transition: $P_{21} = \epsilon_{21}/n_e^2$. We thus get:

$$P_{21}(T) = \left(\frac{n_H}{n_e} \right) A_{elem} f_Z(T) \gamma_{12}(T) E_{12} \quad (246)$$

which is typically expressed in units of $\text{erg cm}^3 \text{s}^{-1}$.

Actually, the “two-state” model discussed above is too simple, since important contributions to upper level populations can also come from ground-state excitations to higher levels, which then radiatively decay to intermediate states. However, even these more complicated “channels” can still be incorporated via the definition of more general, effective excitation rate coefficients that include these terms. A number of coronal equilibrium “spectral synthesis” codes have been developed over the years to provide these line power calculations, and some are in widespread use in the community. The largest

residual uncertainties in these codes generally involve the treatment of the DR rates, and the completeness of the line lists.

For an intermediate charge state, the ion fraction, f_Z , peaks in temperature at some particular value. The excitation rate coefficient, γ , generally increases across the range of temperatures where the ion exists in appreciable abundance. Therefore, the line power, P , exhibits a peak at a temperature often called the *temperature of formation*, T_f . The presence of a particular line in the spectrum implies the existence of plasma at or near the temperature of formation for that line. The modulation of line powers by the temperature dependence of the ion fraction thus gives us a crude temperature diagnostic.

The measured line flux for a collisional plasma is given by:

$$F_{21} = \frac{e^{-N_H \sigma(E_{21})}}{4\pi d^2} \int dV dT n_e^2(T, V) P_{21}(T) \quad (247)$$

$$\sim \frac{e^{-N_H \sigma(E_{21})}}{4\pi d^2} P_{21}(T_f) \int dV n_e^2(T_f) \quad (248)$$

where $e^{-N_H \sigma E_{21}}$ is the attenuation factor through the interstellar and circumsource media, and d is the distance to the source. The integral that remains in (247) is called the *volume emission measure*, $VEM(T_f)$. As indicated, it is a function of temperature. For an assumed set of abundances, and a given column density, N_H , the shape of the emergent spectrum for a coronal plasma is given completely by the shape of the volume emission measure distribution.

5.3 X-Ray Photoionization Equilibrium

A quite different set of approximations applies to the case of *photoionization equilibrium*, where the presence of an intense continuum radiation field has a significant effect on the ionization and thermal structure of the surrounding gas. The electrons are generally too cool to excite prominent X-ray lines in this case, and excited levels are instead populated by direct recombination, by radiative cascades following recombination onto higher levels, and by direct photoexcitation from the continuum.

These conditions are typically found in the circumsource media of accretion-powered sources, such as X-ray binaries and active galactic nuclei. For example, in the accreting gas surrounding an X-ray binary, the energy density in the continuum radiation field is given by:

$$U_\gamma \sim \frac{L}{4\pi R^2 c} \sim 3.7 \cdot 10^4 \text{ erg cm}^{-3} \quad (249)$$

where we have taken $L \sim 10^{38} \text{ erg s}^{-1}$, and $R \sim 10^{11} \text{ cm}$. In contrast, the thermal energy density in the electron distribution is given by:

$$U_e \sim \frac{3}{2} n_e kT \sim 2.4 \text{ erg cm}^{-3} \quad (250)$$

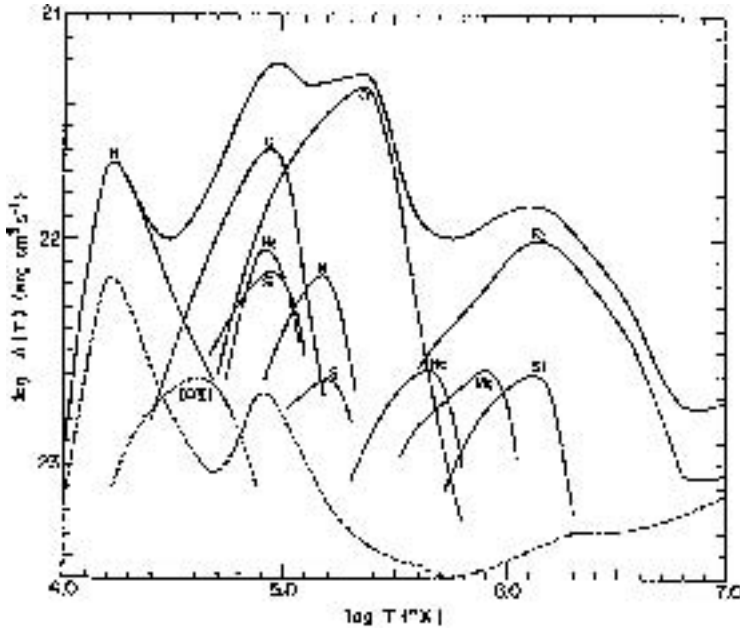


Fig. 6. The power radiated (ϵ/n_e^2) of a cosmic abundance plasma as a function of temperature in coronal equilibrium. The contributions of the individual elements are indicated. Line radiation dominates at temperatures below $10^7 K$

for typical values of the electron density and temperature, $n_e \sim 10^{12} \text{ cm}^{-3}$, $kT \sim 10 \text{ eV}$.

In photoionization equilibrium, the ionization structure is determined by the balance between photoionization and recombination.

$$n_z \int_0^\infty dE \frac{F_E}{E} \sigma_z(E) = n_e n_{z+1} \alpha_{z+1}(T) \quad (251)$$

where F_E is the differential continuum flux, in units of $\text{erg cm}^{-2} \text{ s}^{-1} \text{ keV}^{-1}$, $\sigma_z(E)$ is the photoelectric cross-section as a function of energy (Sect. 3.8), and $\alpha_{z+1}(T)$ is the recombination coefficient, again including both RR and DR contributions. The equilibrium temperature is determined by the solution of the equation of energy balance, where the rate of energy injection is due to photoelectric heating, and the rate of energy loss is due to radiation:

$$\begin{aligned} \sum_{\text{elem}, z} n_{z, \text{elem}} \int_0^\infty dE \frac{F_E}{E} \sigma_{z, \text{elem}}(E) (E - E_{\text{thresh}}^{z, \text{elem}}) \\ = \sum_{\text{elem}, z} n_e n_{z, \text{elem}} \Lambda_{z, \text{elem}}(T) \end{aligned} \quad (252)$$

In the optically thin limit: $F_E = \frac{L}{4\pi R^2} f(E)$, where $f(E)$ is a normalized function containing the details of the spectral shape of the irradiating continuum. In addition, we can write $n_{z,elem} = A_{elem} f_z n_H$, and $n_e = \mu_e n_H$, where μ_e , the mean number of electrons per hydrogen atom, is only a weak function of gas parameters. Therefore “environment specific” factors are all embodied in a single quantity

$$\xi = \frac{L}{nR^2} \quad (253)$$

which is usually referred to as the *ionization parameter*. Given the specification of this ionization parameter, the self-consistent solution of the ionization and energy balance equations yield the $f_z(\xi)$ values for all the elements, and $T(\xi)$. A variety of codes are in widespread use to calculate these quantities.

Plots of the ionization structure of iron as a function of temperature for conditions of coronal equilibrium and photoionization equilibrium are shown in Fig. 7. Two important features are immediately apparent from this figure:

- First, the “dominance of closed shells” is much less obvious in the case of photoionization equilibrium. Given the big jump in ionization potential following the removal of all the electrons in a closed shell, the closed shell charge states (e.g. Ne-like and He-like) dominate over a wide range of temperature for a plasma in coronal equilibrium. However, for a photoionized plasma, photoionization out of inner shells (L-shell and K-shell) plays a significant role for the hard irradiating spectra characteristic of accretion-powered sources. This process is essentially unaffected by the removal of outer valence electrons, eliminating any important distinction between open shell and closed shell charge states.
- Second, the gas is significantly “overionized” relative to the electron temperature in a photoionized plasmas. For example, Ne-like iron (FeXVII) peaks at $kT_e = 10$ eV in the photoionized case, while for the coronal plasma Ne-like iron peaks at $kT_e = 400$ eV.

The significantly different temperatures appropriate to a given charge state for coronal and photoionized plasmas lead to several important characteristic differences in the emergent X-ray spectra. For a coronal plasma, $kT \sim \chi$, the ionization potential of the ion, and δE , the characteristic energies of the line excitations. The lines are formed primarily via collisional excitation from the ground state. The brightest lines are E1 transitions, or those “fed” by E1 transitions. In a photoionized plasma, $kT \ll \chi$ and δE , so the electrons have insufficient energy to collisionally excite X-ray lines. Instead, lines are formed mostly by radiative cascades following recombination. Recombination flux tends to distribute evenly among all the available levels. Hence, the brightest lines tend to come from ions with the fewest states in the upper level configuration (e.g. K-shell ions). In addition, the cascades “rain” into the lowest lying excited levels. Therefore, lines from these levels are usually quite bright. Often, these are higher order multipole transitions, with low collisional coupling strengths to the ground.

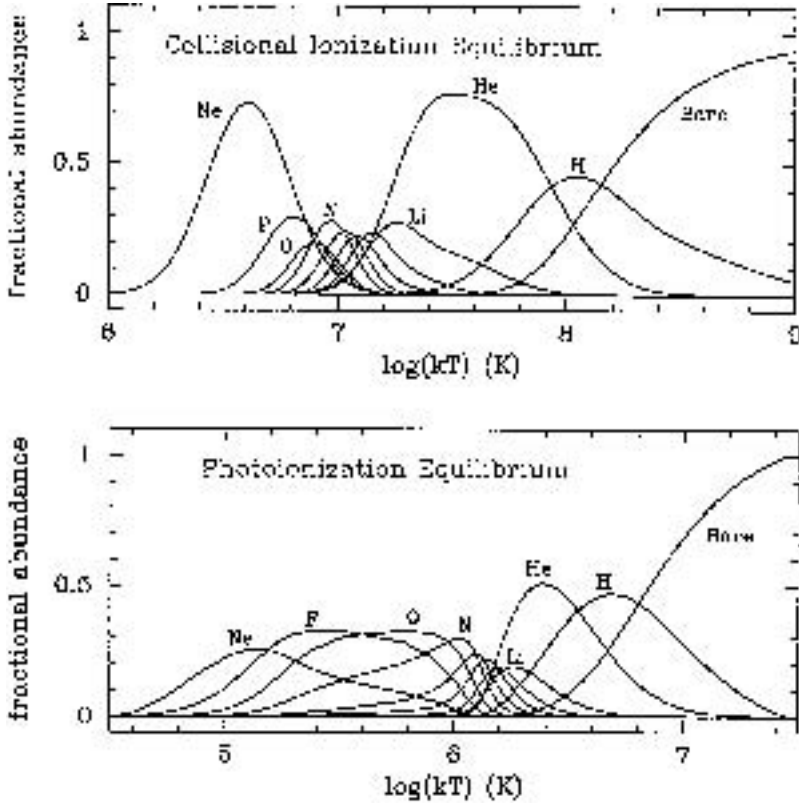


Fig. 7. Plots of the ionization structure of iron as a function of temperature for coronal equilibrium (*top*), and photoionization equilibrium (*bottom*). The element symbols refer to the isoelectronic charge state of iron, e.g. the curve labeled O refers to oxygen-like Fe (figure courtesy of Masao Sako)

However, the most useful spectroscopic diagnostics for distinguishing coronal equilibrium from photoionization equilibrium are the narrow radiative recombination continua (RRC's) expected for the latter case. In Sect. 4.5, we found that RRC's are described by

$$\frac{dW}{dt d\omega dV} \sim \left(\frac{\hbar\omega}{\chi} \right)^3 \sigma_{PI}(\omega) \left(\frac{\chi^2}{mc^2 kT} \right)^{3/2} e^{\chi/kT} e^{-\hbar\omega/kT} \quad (254)$$

For a coronal plasma, $kT \sim \chi \sim \hbar\omega$. The RRC's are broad and do not have high contrast relative to the accompanying bremsstrahlung continuum. On the other hand, in a photoionized plasma, $kT \ll \chi$ and $\hbar\omega$. For this case, the RRCs are strong and fall off steeply with increasing energy. They resemble "lines" at moderate resolution. The relative width of this feature is a good

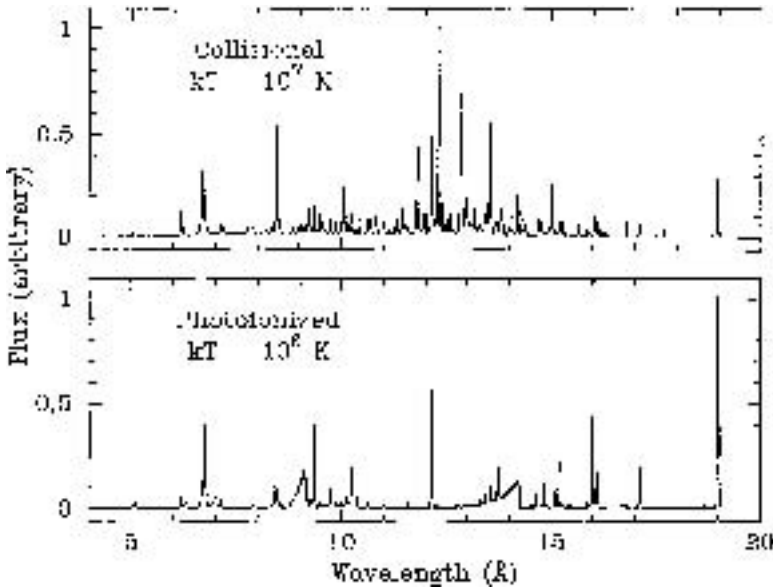


Fig. 8. Plots of characteristic emergent soft X-ray spectra for conditions appropriate to a coronal plasma *top* and an X-ray photoionized plasma *bottom*. Note that the coronal spectrum is more “rich”, due to the greater prominence of the Fe L complex in that case. The photoionized spectrum is dominated by lines from lower-Z K-shell elements, and by low temperature radiative recombination continua (figure courtesy of Masao Sako)

temperature diagnostic, and, if the width is larger than predicted, can signal the presence of extra sources of heating in the gas.

This is illustrated in Fig. 9, which shows the predicted spectrum of neon in a photoionized plasma for electron temperatures of both 10 eV and 50 eV. The former is the expected temperature for these charge states, if photoelectric heating provides the only form of energy injection in the gas. The latter might apply if there are other sources of heating which contribute. As can be seen, the discrete line spectra look very similar for the two cases. However, the RRC (near 9 Å) is much broader and less pronounced at the higher temperature.

With the launches of the grating spectrometers on the *Chandra* and *XMM-Newton* observatories, we now have clear detections of these features in many sources. A particular dramatic case is illustrated in Fig. 10, which shows the spectrum of the bright Seyfert 2 galaxy NGC 1068, as obtained with the reflection grating spectrometer on *XMM-Newton* [10]. As can be seen, the spectrum is rich in emission lines, especially H-like and He-like lines of carbon, nitrogen, oxygen, and neon. The RRC’s from most of these species are labeled in the figure. They are narrow, indicating a low electron temperature of a few eV, characteristic of a photoionized plasma. In NGC 1068, the soft

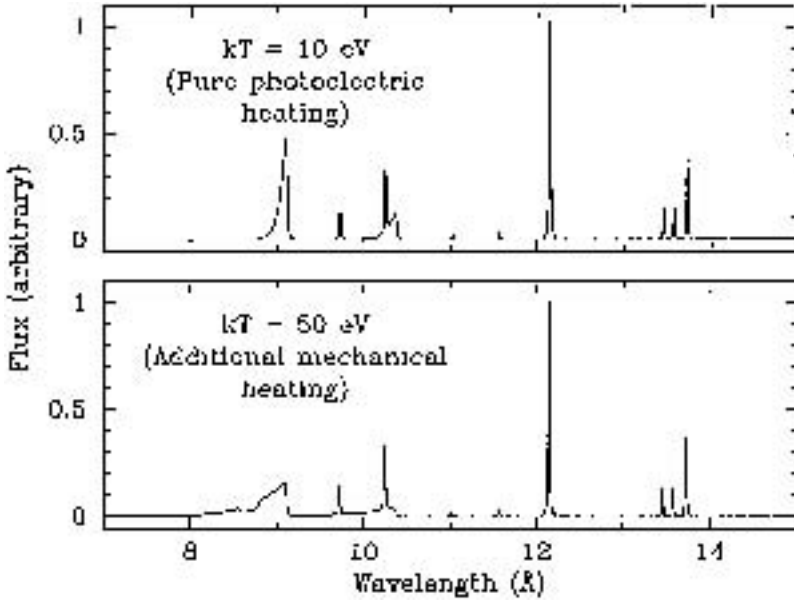


Fig. 9. Plots of the expected spectra of H-like and He-like neon in photoionized plasmas with electron temperatures of 10 eV *top*, and 50 eV *bottom*, but with similar ion fractions. Note the differences in the RRC's for the two cases (figure courtesy of Masao Sako)

X-ray spectrum is produced in an ionization cone, which is irradiated by an intense X-ray continuum emanating from a central obscured nucleus.

5.4 Thermal Instability in Photoionized Plasmas

It has been known for many years that X-ray photoionized plasmas can be thermally unstable in certain regions of ionization parameter space. Typically, this is represented by means of an “S-curve”, a plot of the temperature, derived by solving the equation of energy balance (252), versus an ionization parameter $\Xi = F/n_e T \sim \xi/T$. An example is shown in Fig. 11. On the curve itself, the heating rate is equal to the cooling rate, so the gas is in thermal balance. To the right, heating dominates over cooling, as indicated, while to the left, cooling dominates over heating. On branches of the curve which have positive slope in this figure, the gas is thermally stable. Small perturbations upward in temperature increase the cooling, while small perturbations downward in temperature increase the heating. However, on the branches which have negative slope, the gas is thermally unstable. A small perturbation upward in temperature increases the heating, causing further temperature rise, while a small perturbation downward increases the cooling. Many different calculations of these effects exist in the literature, and

negative during the fall. The net thermal stability is determined by the sum of the contributions from all of these atomic shells. The situation can be quite complex, because the stable and unstable lobes contributed by the different elements occur at different temperatures. One finds that there are “near cancellations”, which makes the total stability quite sensitive to details related to the elemental abundances and the shape of the ionizing spectrum. This can be beneficial, because we can exploit this sensitivity to derive strong constraints on physical conditions in the gas, if the signatures of thermal instability are visible in the spectra.

6 Discrete Line Diagnostics

The relative prominence of various emission line features in cosmic X-ray spectra is determined principally by the abundances of the different elements, and the locations of the K- and L-shell complexes associated with these elements within the X-ray band. Scaling from the H-like isoelectronic sequence, the energies of the K-shell features are given roughly by:

$$E_K \sim (10 \text{ eV})Z^2, \quad (259)$$

while the energies of the L-shell features are approximately:

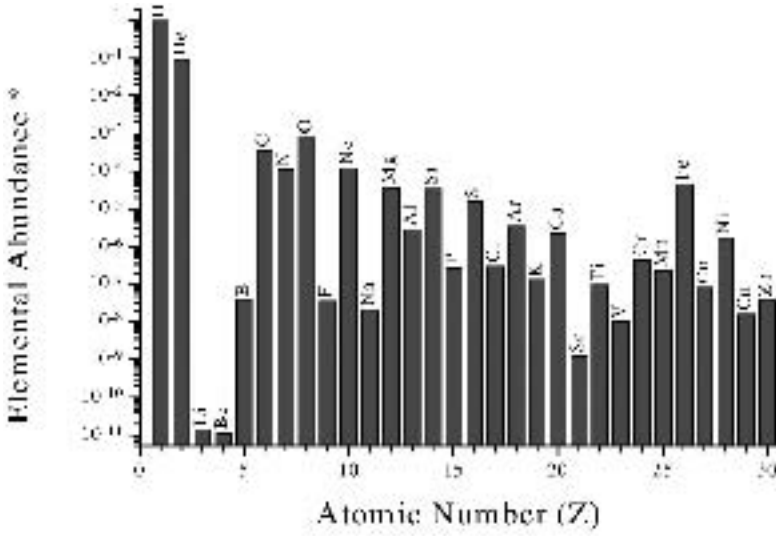
$$E_L \sim (1.5 \text{ eV})Z^2. \quad (260)$$

If we define the conventional soft X-ray band to cover the range $100 \text{ eV} \leq E \leq 10 \text{ keV}$, we see that it includes the K-shell features of beryllium ($Z = 4$) through gallium ($Z = 31$), and the L-shell features of oxygen ($Z = 8$) through thallium ($Z = 81$).

A plot of standard cosmic abundances as a function of atomic number appears in Fig. 12. Several features should be noted:

- The abundances drop precipitously with increasing Z above carbon ($Z = 6$). The abundances of lithium, beryllium, and boron ($Z = 3, 4$, and 5 , respectively) are especially low.
- In general, elements with even values of Z have considerably higher abundances than elements with odd values of Z . This is a consequence of the importance of α -chain reactions, in the production of the heavier elements during the late stages of stellar evolution.
- There is a very prominent abundance peak at iron ($Z = 26$) in the higher Z -range. This is a consequence of nuclear stability. ^{56}Fe has the highest binding energy per nucleon of any nucleus. Fusion reactions that produce lower Z elements are exothermic, while above iron, fusion reactions become endothermic.

Given these considerations, the most significant K-shell complexes in cosmic X-ray spectra are due to C, N, O, Ne, Mg, Si, S, Ar, Ca, Fe, and Ni, while the



* Taken from Anders & Grevesse, *Gesamtesium der Chemischen Erde* 53, 197 (1989).

Fig. 12. A plot of the standard cosmic abundance of the elements as a function of atomic number Z (figure courtesy of Masao Sako)

most significant L-shell complexes are associated with Si, S, Ar, Ca, Fe, and Ni. It is one of the major strengths of cosmic X-ray spectroscopy that such a wide range of elements and charge states is measured in a single wavelength band.

6.1 Lyman Series Transitions in H-like Ions

At the characteristic temperatures of X-ray emitting plasmas, the low- Z abundant elements are often found in their H-like charge states. The most prominent emission lines are the Lyman series transitions:

$$\begin{aligned}
 \text{Ly } \alpha_1: 1s-2p \ ^2P_{3/2}; \text{ Ly } \alpha_2: 1s-2p \ ^2P_{1/2}; \\
 \text{Ly } \beta_1: 1s-3p \ ^2P_{3/2}; \text{ Ly } \beta_2: 1s-3p \ ^2P_{1/2}; \\
 \text{Ly } \gamma_1: 1s-4p \ ^2P_{3/2}; \text{ Ly } \gamma_2: 1s-4p \ ^2P_{1/2} \\
 \dots
 \end{aligned}$$

The ratio of the line intensities for the two transitions in each case is given roughly by the degeneracy factors, e.g.:

$$\text{Ly } \alpha_1 / \text{Ly } \alpha_2 \sim \frac{2(3/2 + 1)}{2(1/2) + 1} = 2.$$

Recall that the splitting is:

(261), the upper levels for the different transitions in the series are close in energy, so the Boltzmann factor in the excitation rates varies only slightly from transition to transition in the temperature range where the H-like ion is the dominant species (see Fig. 14). At the very low temperatures characteristic

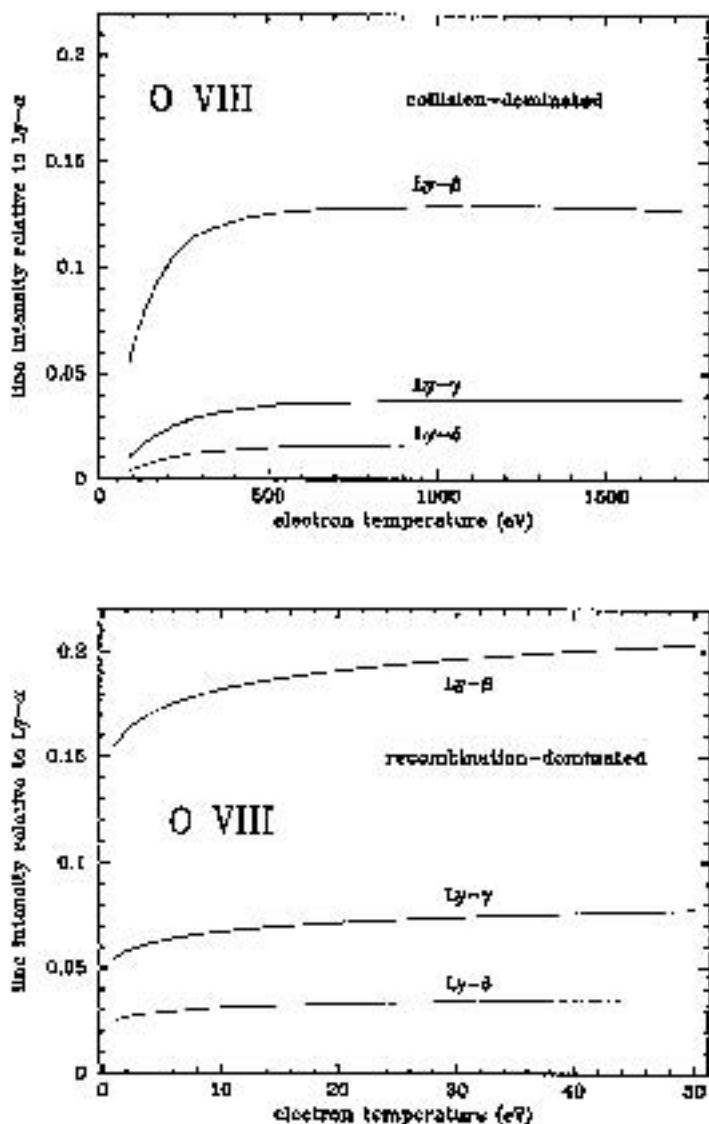


Fig. 14. Plots of the ratio of higher series Lyman line intensities to the Lyman α line intensity as a function of temperature in O VIII, for both coronal plasmas (*top*), and photoionized plasmas (*bottom*)

of photoionized plasmas, Lyman series lines are formed by radiative cascades associated with radiative recombination. The line ratios produced by these processes are somewhat different than those associated with collisional excitation in collisional plasmas. This is apparent from Fig. 14, where it can be seen that the Ly β to Ly α ratio for O VIII is ~ 0.11 for a coronal plasma, and ~ 0.14 for a photoionized plasma. Similar enhancements are found for the higher series line ratios as well.

6.2 He-like Transitions

He-like K-shell lines are among the most important of all in the soft X-ray band. Since the He-like charge state is a tight “closed shell”, this is the dominant ion species over a wide range in temperature, particularly in coronal plasmas. In addition, as explained below, these lines exhibit strong sensitivity to electron density, temperature, and ionization conditions in the emitting plasma.

The most important K-shell He-like transitions are as follows:

$$\begin{aligned} W: & 1s^2 \ ^1S_0 - 1s2p \ ^1P_1 \\ X: & 1s^2 \ ^1S_0 - 1s2p \ ^3P_2 \\ Y: & 1s^2 \ ^1S_0 - 1s2p \ ^3P_1 \\ Z: & 1s^2 \ ^1S_0 - 1s2p \ ^3S_1 \end{aligned}$$

W is an electric dipole transition, also called the *resonance* transition, and is sometimes designated with the symbol r . X and Y are the so-called *intercombination* lines. These are usually blended (especially for the lower- Z elements), and are collectively designated with the symbol i . Z is the *forbidden* line, often designated by the symbol f . It is a relativistic magnetic dipole transition, with a very low radiative decay rate.

The temperature sensitivity of these lines arises as follows [14–16]: Since W is an electric dipole transition, the collision strength for collisional excitation of this line includes important contributions from higher order terms in the partial wave expansion, and thus continues to increase with energy above threshold. By contrast, X and Z are electric dipole forbidden. The dominant term in the excitation collision strength for these transitions involves electron exchange. Therefore, their excitation collision strengths drop off strongly with energy above threshold, whereas Y remains relatively constant. As a result, the line ratio: $G = (X + Y + Z)/W$ is a decreasing function of electron temperature.

The density sensitivity comes from the fact that the $^3S_1$ level can be collisionally excited to the ^3P levels. At high electron density, that process successfully competes with radiative decay of the forbidden line. Therefore, the ratio $R = Z/(X + Y)$ drops off above a critical density, n_c . The critical density depends strongly on Z . For C V, $n_c \sim 10^9 \text{ cm}^{-3}$, while for Si XIII, $n_c \sim 10^{13} \text{ cm}^{-3}$.

However, the R -ratio can also be affected by the presence of a significant ultraviolet radiation field [14]. In particular, the $^3S_1$ level can be photoexcited to the ^3P levels, prior to radiative decay, if there is sufficient ultraviolet intensity at the energy of the relevant transitions. That leads to suppression of the forbidden line and enhancement of the intercombination lines, mimicking the effects of high electron density.

These dependences are illustrated in Figs. 16 and 15, which shows the He-like spectra of oxygen, nitrogen, and carbon for two stellar coronal sources, Procyon and Capella, as measured with the *Chandra* low energy transmission grating spectrometer [17]. The corona of Procyon is cooler than that of Capella. As can be seen, the resonance lines are consequently less intense for Procyon, in comparison to both the intercombination and forbidden lines. Note that the forbidden line of carbon is also comparatively suppressed for Procyon in relation to the intercombination line. While this looks like a density effect, it is actually due to the ultraviolet radiation field from this star. Procyon is an F star, with a relatively high UV flux.

In photoionized plasmas, the excited levels for He-like ions are fed directly by recombination and also by radiative cascades following recombination onto higher levels. The forbidden line is most intense, since most of the cascades from high- n , high- l (high- J) levels land on the lowest lying $1s2s(J=1)$ level, which produces the forbidden line. This is illustrated in Fig. 17, and can also be seen in the spectrum of NGC 1068 shown in Fig. 10 for both the He-like oxygen lines near 22 Å, and the He-like nitrogen lines near 29 Å.

6.3 Iron L-Shell Transitions

Since iron is the most abundant high- Z element, its L-shell spectrum plays a crucial role in astrophysical X-ray spectroscopy. As a result of their higher ionization potentials, the iron L-shell ions contribute significant line emission even when the lower- Z elements are full stripped. For collisionally ionized plasmas, this complex samples a wide range in temperature (0.2–2 keV). In addition, the L-shell spectrum is very “rich”, and there is significant diagnostic sensitivity.

The brightest iron L-shell lines are of the form:

$$\begin{aligned} 2s^2 2p^k - 2s^2 2p^{k-1} 3d \\ 2s^2 2p^k - 2s^2 2p^{k-1} 3s \\ 2s^2 2p^k - 2s 2p^k 3p \end{aligned}$$

The $2p - 3d$ lines generally have the highest oscillator strength. The line positions are a strong function of charge state. Thus, the ionization structure is easily discernible, which provides a simple, abundance-independent constraint on the temperature distribution.

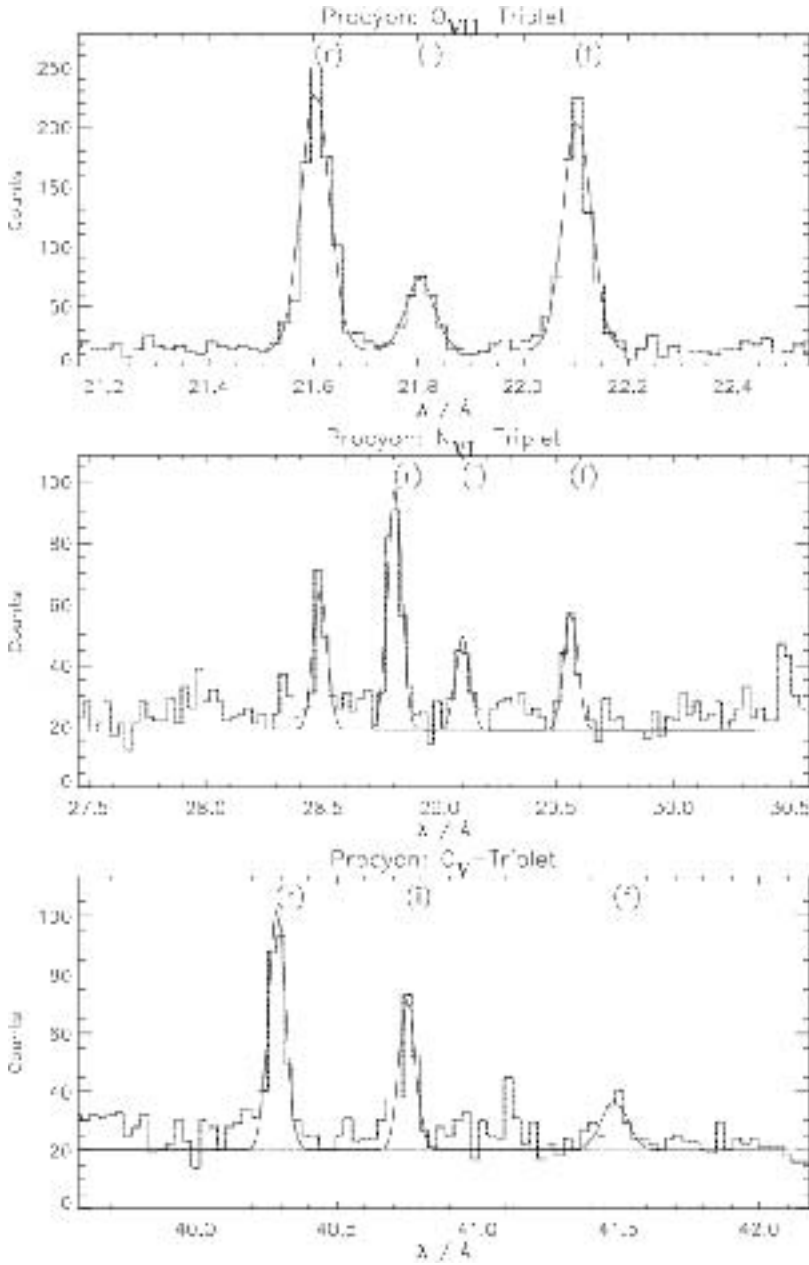


Fig. 15. He-like complexes for O, N, and C from the coronal star Procyon, as measured with the *Chandra* low energy transmission grating spectrometer (From [17])

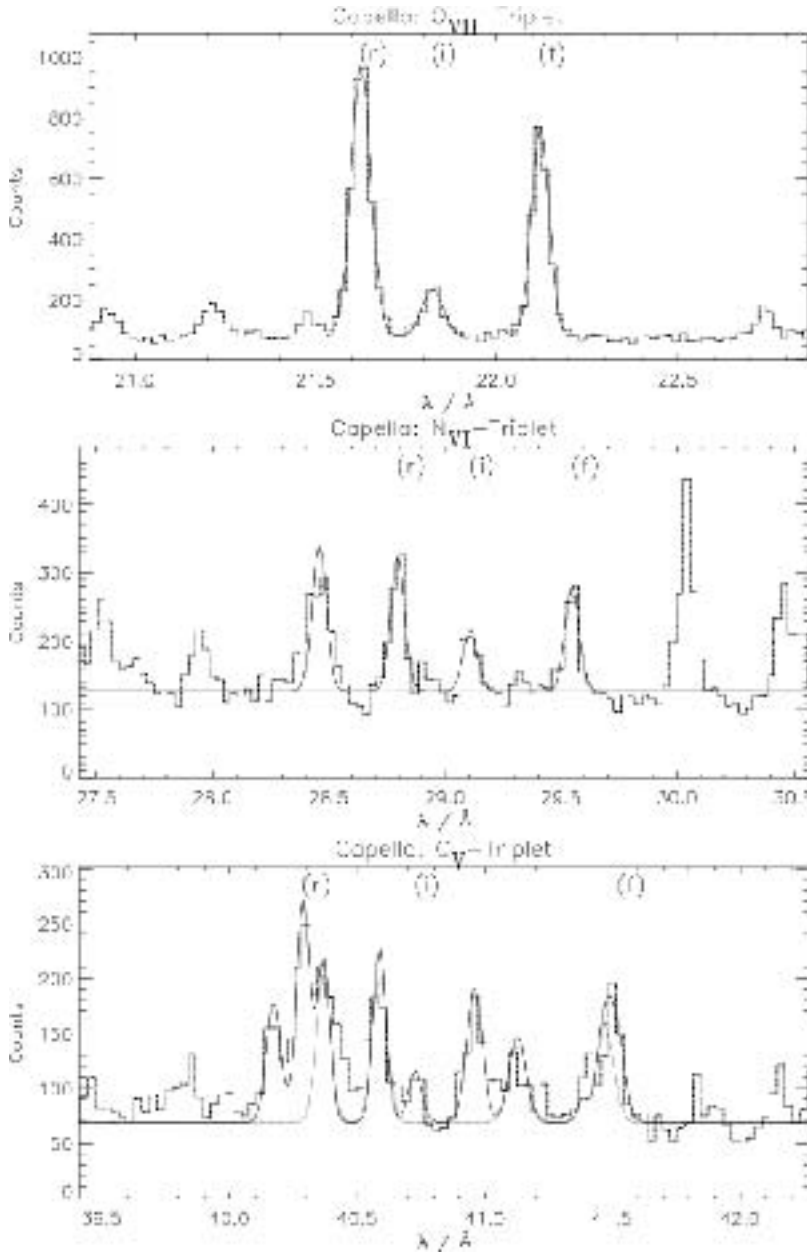


Fig. 16. He-like complexes for O, N, and C from the coronal star Capella, as measured with the *Chandra* low energy transmission grating spectrometer (From [17])

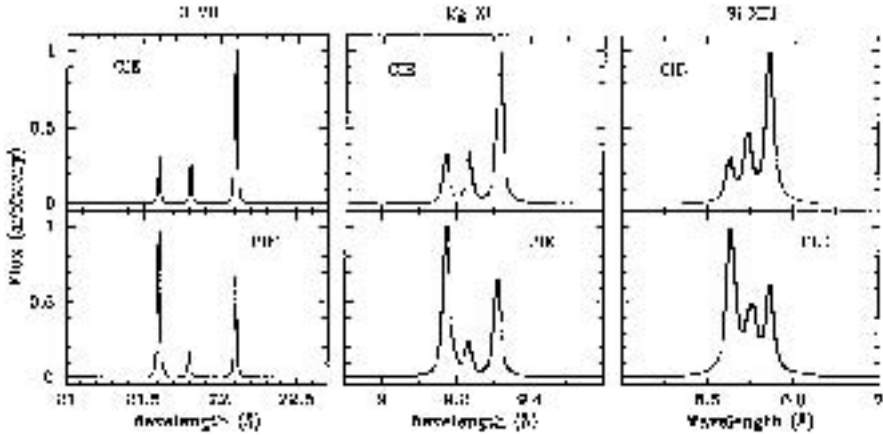


Fig. 17. Calculated He-like emission line spectra of oxygen, magnesium, and silicon for photoionization equilibrium *top* and coronal equilibrium *bottom* plasmas. Note the prominence of the forbidden lines in the case of the photoionized plasmas (figure courtesy of Masao Sako)

This is illustrated in Fig. 18, which shows the iron L spectrum of Capella, as observed with the *Chandra* high energy transmission grating spectrometer. Plotted below the measured data are the calculated contributions from each of the individual charge states, ranging from Na-like iron (Fe XVI) to Be-like iron (Fe XXII). Note the relatively clean separation between the L-shell complexes from each of these ions, allowing for relatively easy decomposition of the spectrum, even with only moderate resolution. The density sensitivity of the iron L complex arises from the fact that the intermediate iron L charge states (e.g. N-like and C-like) possess a number of low lying metastable levels associated with $n = 2 \rightarrow n' = 2$ excitations. These can be populated collisionally, leading to new “seed” states for $2 \rightarrow 3$ excitations, followed by $3 \rightarrow 2$ radiative decays. Such density diagnostics turn on at electron densities $\sim 10^{13} \text{ cm}^{-3}$.

6.4 The Iron K-Shell Complex

The iron K complex is relatively isolated in the spectrum at energies $\sim 6 - 7 \text{ keV}$, where even non-dispersive detectors have moderate spectral resolution. Thus, iron K lines were the first discrete atomic features unambiguously detected for cosmic X-ray sources.

An important contributor to iron K emission, especially for accretion-powered sources, is due to fluorescence from cold material in the vicinity of a bright X-ray continuum. Fluorescence involves a radiative decay following inner shell photoionization, i.e. a transition of the form $1s^2 2s^2 2p^{k-1} nl - 1s 2s^2 2p^k nl$. The excited level, in this case, can also decay via autoionization

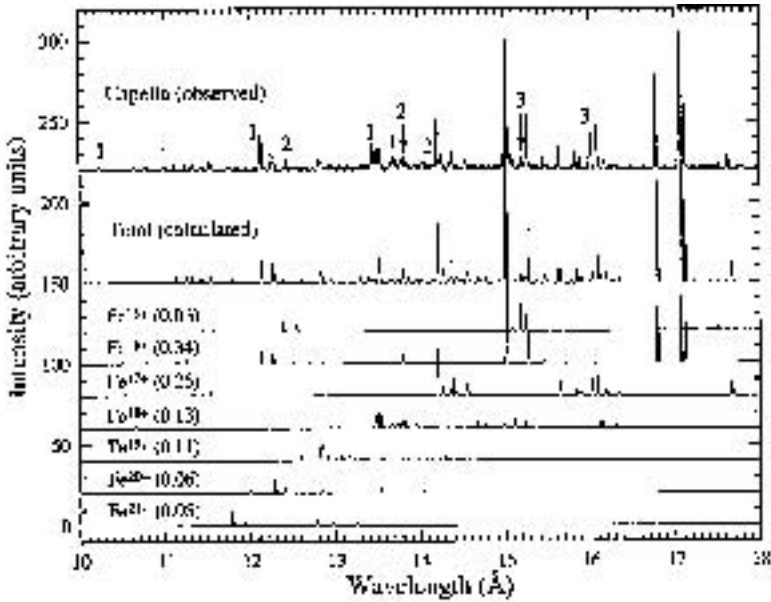


Fig. 18. The spectrum of Capella obtained with *Chandra* high energy transmission grating spectrum, compared with a calculated spectrum showing the separate contributions of each of the iron L charge states (From [19])

by ejecting one of the outer electrons in the valence shell. This latter process dominates for low- Z elements. However, since radiative decay rates scale like Z^4 , and autoionization decay rates scale like Z^0 , the fluorescence yield becomes appreciable for a high- Z element like iron. The near-neutral iron K fluorescence line falls at 6.4 keV, easily distinguishable from the He-like lines near 6.7 keV, and the Lyman α line at 7.1 keV.

The iron K complex also exhibits new features due to the relative importance of dielectronic recombination. DR leads to Li-like “satellites” to He-like K-lines: $1s2pnl - 1s^2nl$. These satellites are shifted down in energy. Higher n implies a smaller shift, and is associated with a higher energy of the recombining electron. Therefore, the satellite spectrum is temperature sensitive (cf. [20]).

At astrophysical densities, all atoms are in the ground state. Most of the satellite lines cannot be produced by collisional excitation of Li-like iron (e.g. $1s2p^2 - 1s^22p$). They come purely from DR on He-like atoms. However, other lines terminate in the ground configuration of the Li-like ion (e.g. $1s2s2p - 1s^22s$). These can be produced by both collisional excitation of Li-like atoms, and DR on He-like atoms. Hence, the line ratios for these various transitions provide an independent measure of the charge balance. Analysis of the Fe K He-like spectrum thus provides independent constraints on the

electron temperature and the level of ionization, and is ideal for investigating departures from ionization equilibrium.

7 Concluding Remarks

As a field, astrophysical X-ray spectroscopy is still in its infancy. While the grating spectrometers on *Chandra* and *XMM-Newton* have already showered us with fascinating results on a wide variety of diverse sources, most of the data have not been completely reduced, and many sources bright enough to provide reasonable spectra have still not yet been observed. A much larger population of interesting sources are too faint for these instruments, but should be amenable for study with the more sensitive experiments planned for future missions such as *Constellation-X* and *XEUS*.

The complete analysis of all of these observations will require a greater level of spectroscopic sophistication than most X-ray astronomers are accustomed to. In the past, we have had the luxury of fitting relatively simple “canned” spectral models to low resolution, low statistics data. As the quality of our spectra improves, these more familiar techniques no longer suffice. Some would prefer to ignore the complications, and continue to work only on the faintest sources where the paucity of photons precludes worrying about spectral details. I have even heard some argue that we should not attempt to build higher resolution spectroscopic instruments, because the data they will acquire will be too difficult to interpret. I find this view to be very unscientific. We will always benefit by better instruments and better data.

In these lectures, I have tried to provide a synopsis of the kinds of issues X-ray astronomers must consider in analyzing their spectroscopic data. But this is by no means a “user manual”. There are no simple codes that will take proper account of all relevant processes, and provide a neat set of “results” at the push of a button. We will all have to continue to learn as we go along. The first data sets we have obtained have already pointed to holes in our existing atomic databases, and in our understanding of particular excitation processes. To make progress, we must complement our data analysis activities with direct involvement in laboratory astrophysics experimentation, and atomic calculation. Astronomers must become spectroscopists, and spectroscopists must become astronomers. This is how real progress will emerge.

Acknowledgments

I am indebted to a number of key individuals for helping me to finally make these lecture notes available for publication. First, I would like to thank Pascal Favre of the Integral Science Data Centre, for his tremendous assistance with the preparation of the manuscript. Second, I would like to thank my students and colleagues at Columbia: Ehud Behar, Jean Cottam, Mingfeng Gu, Ali Kinkhabwala, Maurice Leutenegger, Frits Paerels, John Peterson, Masao

Sako, and Daniel Savin for help with the figures, editing the text, and for contributing many of the ideas that are contained within. I have also benefited from numerous conversations with current and previous collaborators, most notably Peter Beiersdorfer and Duane Liedahl at the Lawrence Livermore National Laboratory, and Bert Brinkman, Jelle Kaastra, and Rolf Mewe of SRON, Utrecht. Finally, I would like to thank my hosts for the Saas Fee program: Manuel Güdel and Roland Walter, for inviting me to Les Diablerets and allowing me to participate in this distinguished lecture series.

References

1. Cowan, R., 1981, The Theory of Atomic Structure and Spectra, Los Alamos series in Basic and Applied Science, University of California Press, Berkeley, CA
2. Rybicki, G. B., and Lightman, A. P., 1979, Radiative Processes in Astrophysics, Wiley, New York, 1979
3. Giacconi, R., Gursky, H., Paolini, F., et al., 1962, *Phys. Rev. Lett.*, **9**, 439
4. Blandford, R., Fabian, A., Pounds, K., 2003, X-Ray Astronomy in the New Millennium, Cambridge University Press
5. Schlegel, E. M., 2002, The Restless Universe: Understanding X-Ray Astronomy in the Age of Chandra and Newton. Oxford University Press
6. Tucker, W., Tucker, K., 2001, Revealing the Universe: the Making of the Chandra X-ray Observatory, Harvard University Press, Cambridge, MA
7. Hutchinson, I. H. 1987, Principles of plasma diagnostics, Cambridge University Press
8. Lotz, W. 1967, *ApJS*, **14**, 207
9. Rumph, T., Bowyer, S., and Vennes, S., 1994, *AJ*, **107**, 2108
10. Kinkhabwala, A., Sako, M., Behar, E., et al., 2002, *ApJ*, **575**, 732
11. Hess, C. J., Kahn, S. M., & Paerels, F. B. S., 1997, *ApJ*, **478**, 94
12. Rasmussen, A. P., Behar, E., Kahn, S. M., et al., 2001, *A&A*, **365**, 231
13. Blair, W.P., Morse, J. A., Raymond, J. C., et al., 2000, *ApJ*, **537**, 667
14. Gabriel, A. H., and Jordan, C., 1969, *MNRAS*, **145**, 241
15. Pradhan, A. K., 1982, *ApJ*, **263**, 477
16. Porquet, D., Mewe, R., Dubau, J., et al., 2001, *A&A*, **376**, 1113
17. Ness, J.-U., Mewe, R., Schmitt, J. H. M. M., et al., 2001, *A&A*, **367**, 282
18. Kahn, S. M., Leutenegger, M. A., Cottam, J., et al., 2001, *A&A*, **365**, 312
19. Behar, E., Cottam, J., and Kahn, S., 2001, *ApJ*, **548**, 966
20. Dubau, J., Volonte, S., 1980, Reports on Progress in Physics, vol. 43, 199



Peter von Ballmoos

Instruments for Nuclear Astrophysics

P. von Ballmoos

1 Introduction

On April 9, 1900, at the session of the *Académie des Sciences*, Paul Villard of the École Normale in Paris, presented a paper “Sur la réflexion et la réfraction des rayons cathodiques et des rayons déviables du radium” [1]. Villard describes a series of experiments with a small radium source, leading to the discovery of a radiation, not deflected by a magnetic field, which was later to be called gamma-rays (the first mention of the term “gamma-ray” is probably from Rutherford in 1903 [2]). Villard’s experiments naturally utilized the first instrument for the detection of gamma rays – a photographic plate wrapped in light-tight black paper and shielded from α and β radiation by a lead foil:

“I think that this effect is due to the presence of non-deviable rays, which are less absorbable than the ones [α rays] that have been described by Mr. Curie. . . . It follows from the facts presented above that the non-deviable rays emitted by radium contain some very penetrating radiations, capable of traversing metal foils and affecting a photographic plate.”

A few weeks later, Villard suggests [3] that the extremely penetrating rays discovered by him were in fact a kind of X-rays, and went on to identify all three components of radium rays (α, β, γ), concluding that “on retrouverait ainsi les trois rayonnements des tubes de Crookes”, i.e., one finds the three kinds of radiation (ions, electrons and X rays) known from experiments with cathode-ray tubes [4].

Whilst High-Energy Astrophysics still is considered a young science, its photon messenger was celebrating his centennial anniversary by the end of the 30th Saas Fee Advanced Course on “High-Energy Spectroscopic Astrophysics”: Happy Birthday, Gamma-Ray!

What made progress so slow? On the threshold to the 21st century, astrophysics has in fact just started to take advantage of the unique insights nuclear gamma-rays can provide: Only today, one century after Villard’s discovery, can we say that the sky has been surveyed for the first time at gamma-ray energies.

The reason for this slow pace is an intricate compound of experimental difficulties that the discipline has to face. The instrumental problems are a major component of this text and will be introduced in Sect. 1.2. First of all, high-energy astronomy had to wait – and still has to wait – for the rare space missions. Unlike the instruments used for research in optical and radio wavelengths, Gamma-ray observations can be done exclusively from space. Even the penetrating MeV photons interact within the top of the atmosphere; as a consequence, gamma-ray telescopes must be carried at altitudes of at least 35 to 40 km in order to observe unscattered photons. Although stratospheric balloons have opened the way, systematic operation of instruments above the atmosphere became practicable only with the era of space exploration, starting in the second half of the 20th century.

1.1 The Instrumental Development of Gamma-Ray Astrophysics

Two major questions scientifically motivated the search for cosmic gamma rays: the origin of cosmic rays, and the quest for a deeper insight into the processes of nucleosynthesis. Accordingly, gamma ray astronomy began to evolve along two lines. The study of high-energy gamma-rays, at energies above say 30 MeV, was tied to cosmic ray research because of their common physics (charged particle collisions and cascades, electromagnetic cascades, cosmic ray acceleration). At lower energies, in the energy range of the nuclear transitions – from about 100 keV to several tens of MeV – gamma-ray astronomy naturally developed with the methods and scopes of nuclear physics (excited nuclei/radioactivity, e^+e^- annihilation). With the breakthrough of X-ray astronomy in the sixties, compact galactic and extragalactic objects gained interest at low and medium gamma-ray energies and had consequential influence on instrument design.

Although the primary scope of this work is spectroscopy in the energy range of the nuclear transitions, the development of high-energy gamma instrumentation will be also summarized below.

The Discovery of Celestial Gamma-Rays

Early efforts to detect a cosmic gamma-ray component had developed at the end of the second world war, with the opportunity to reach high altitude by means of ballistic rocket flights. The first attempts to detect primary photons beyond the Pfotzer maximum were made by Perlow and Kissinger [5,6]. Their two detector systems (0.1–15 MeV and 3.4–90 MeV, respectively), consisted of Geiger–Müller tubes, lead and copper converters; both of them were equipped with a anticoincidence logic for reduction of charged background. The instruments were launched for the first time on a V2 rocket from White Sands, New Mexico on January 28, 1948 and reached an altitude of 61 km. During the 77 seconds considered “above the atmosphere”, an integrated celestial gamma-ray flux of 0.09 ± 0.05 counts per second above 3.4 MeV was

deduced. Perlow and Kissinger regarded the measurement as marginal and did not exclude a null result (the rate is actually more than an order of magnitude higher than what would be expected based on current knowledge of the cosmic diffuse gamma ray intensity). Yet, the authors also recognize that their measurement indicates a cosmic gamma-ray intensity more than three orders of magnitude lower than the total cosmic ray intensity. This fact plagued the newborn discipline and remains one of the major challenges today.

During the difficult pioneer years that follow, the background produced by cosmic rays in the upper atmosphere and in the early passive collimators did not lead to positive detection. What these early attempts to measure gamma-rays did show was that the source fluxes had to be extremely low – orders of magnitudes lower than the predictions made in Morrison’s often cited paper presented at the Vatican conference in 1957 [7].

Ten years after Perlow and Kissinger’s V2 experiment, and nearly six decades after Villard’s discovery, nuclear gamma-ray photons were finally observed unequivocally for the first time. The first significant detection of MeV gamma-rays of extraterrestrial origin was made during a solar flare on March 20, 1958 by a balloon instrument flying above Cuba [8]. A burst of gamma-rays in two detectors, an ion chamber and a Geiger counter, coincided with an unusually strong solar radio flare observed at wavelengths of 3 cm and 27 cm.

The first – still meager – evidence for extrasolar MeV gamma-ray emission came in the early sixties from detectors on two Ranger spacecraft flying towards the Moon where they were to explore the lunar surface [9, 10]. The omnidirectional CsI scintillator detectors could be extended on a 1.8 meter long boom in order to evaluate the spacecraft induced background component. Solid angle considerations indicated a remaining gamma-ray flux of undetermined cosmic origin, that we (still) call the *cosmic diffuse gamma-ray background*.

In 1967, a major discovery was made at MeV gamma-ray energies. While the superpowers of the cold war negotiated treaties to ban nuclear tests, the US Air Force had started to prepare for their verification. Between 1963 and 1969, six pairs of Vela satellites, equipped with X-ray, gamma-ray and neutron detectors, built at Los Alamos and Sandia, were launched as a means of verifying the conditions of the Nuclear Test Ban Treaty of 1963 [11], prohibiting tests in the atmosphere and in space. On July 2, 1967, the Cesium Iodide scintillators of Vela 4 a and b measured an extraordinary enhancement in the count rate lasting six seconds – this was to become the first *gamma-ray burst* observed. The new phenomenon was made public only in 1973 by Klebesadel et al. [12]. It took 25 years more until these enigmatic events finally were observed at other wavelengths. In 1997, the afterglow of a gamma-ray burst was observed by the X-ray satellite Beppo-SAX [13], and subsequently by optical telescopes. Today, host galaxies of gamma-ray bursts have been measured to

have redshifts up to $z = 3.4$ [14], implying energy conversions of $10^{43} - 10^{47}$ J, while variability arguments limit the source regions to less than 100 km. Most models for these cosmic fireballs involve gravitational collapse or accretion of one or several compact objects (hypernova, mergers).

Excited Nuclei and Neutron Capture

In 1972, OSO-7 brought first direct evidence for *gamma-ray lines* in solar flares [15]: Besides the strong e^+e^- annihilation line at 511 keV, the neutron capture at 2.223 MeV resulting from the reaction $^1\text{H}(n, \gamma)^2\text{H}$ was clearly detected. Nuclear excitation lines from carbon and oxygen (^{12}C , ^{16}O – at 4.4 MeV and 6.1 MeV, respectively), although less significant in the OSO-7 data, have since been confirmed and studied extensively by the Solar Maximum Mission SMM [16] along with other excited nuclei from the active sun (^{56}Fe , ^{24}Mg , ^{20}Ne , ^{28}Si).

Apart from a still unconfirmed detection of a neutron capture line at 2.2 MeV [17] from an unidentified source, no evidence for excited nuclei has yet been established for sources beyond the sun. (The possible neutron capture source was found in the generally featureless COMPTEL map of the sky at 2.2 MeV. The point-like feature near $l = 300^\circ$, $b = -30^\circ$, is significant at the 3.7 sigma level. RE J0317-853, one of the hottest known white dwarfs with a strong magnetic field has been discussed as a possible origin of this emission).

e^+e^- Annihilation

Since Anderson's discovery of the positron on August 2 1932 [18], the question on the existence of antimatter in the Universe has puzzled astrophysicists. Besides the production of positrons in the laboratory and by cosmic rays in our atmosphere, it was supposed that they might be produced in a multitude of astrophysical environments (nucleosynthesis, neutron stars, pair plasma etc.). Line emission at 511 keV from the galactic center region has been observed since the early seventies with balloon and satellite experiments. In two balloon flights from Argentina, Haymes' group at Rice University first measured a gamma-ray line at 476 ± 26 keV [19]. Later it was suggested that the line detected was actually the annihilation line, but that the shifted peak could have resulted from the convolution of the broad energy response of the NaI scintillators with the galactic center spectrum consisting of a narrow 511 keV line and the accompanying orthopositronium continuum. In 1977, high resolution Germanium (Ge) semiconductors were flying for the first time on balloons, establishing the detection of a narrow annihilation line at 511 keV (CESR Toulouse [20], Bell-Sandia [21]). The eighties were marked by ups and downs in the measured 511 keV flux in a series of observations performed by the balloon-borne Germanium detectors (principally the telescopes of Bell-Sandia and GSFC). The variable results were interpreted as the signature

of a compact source of annihilation radiation at the galactic center (see e.g. Leventhal, 1991 [22]. Yet in 1990, neither the eight years of SMM data [23], nor the revisited data of the HEAO-3 Ge detectors [24], showed evidence for variability in the 511 keV flux. In the nineties, CGRO-OSSE measured steady fluxes from a galactic bulge and disk component (see Table 2) and rough skymaps [25] are now available based on data from OSSE, SMM and TGRS. A possible third component at positive galactic latitude which was attributed to a annihilation fountain in the galactic center [26], has undergone lively discussions and certainly will have to be confirmed by the next generation of gamma-ray telescopes, particularly SPI-INTEGRAL (see Sect. 4.1).

In fall 1990, the imaging SIGMA telescope detected a strong spectral feature in the spectrum of 1E 1740.7-2942, a source located close to the galactic center [27]. This emission appeared and vanished within days in the energy interval 300–700 keV. Stimulated by this observation, Mirabel et al. [28] performed several radio observations of 1E 1740.7-2942 with the Very Large Array (VLA) revealing two radio jets emanating from the central compact object. Since this discovery of the first galactic “microquasar”, several similar sources have been detected in the inner Galaxy. The spectral and temporal behavior of 1E 1740.7-2942 earned this source the surname “great annihilator” – the data could in fact be explained by pair plasma in the vicinity of a compact object. However, no narrow annihilation line was observed in the center region during the first four years of SIGMA observations [29].

A review of pre-CGRO/GRANAT e^+e^- observations is found in [30], a summary of the 511 keV question during the CGRO/GRANAT era in [31]. *Cosmic radioactivity* was first detected in 1979, by the germanium detector on board the HEAO 3 spacecraft [32]. The discovery of a narrow gamma ray line radiation at 1809 keV emitted by ^{26}Al has since been confirmed by a number of balloon and satellite instruments: here was direct evidence for ongoing synthesis of intermediate and heavy elements in the universe! In order to identify the nucleosynthesis sites, several attempts have been made to analyze balloon- and satellite-data with respect to the angular extent of the ^{26}Al emission. A galactic origin for the line had already been proposed on the base of the HEAO 3 and SMM [33] data; the first sky map in the light of ^{26}Al (inner Galaxy), established the MPI Compton balloon telescope, indicated the inner Galaxy as the principal source [34]. With the first map of the entire sky at 1809 keV by GRO-COMPTEL [35], understanding the origin of galactic radioactivity in a global galactic picture became possible, indicating that massive stars in our Galaxy are as a matter of fact the origin of the observed ^{26}Al [36]. For a review on the discussion over the radioactive ^{26}Al in the Galaxy – observations versus theory – see Prantzos and Diehl [37].

The brightest supernova to be observed for nearly four hundred years, *SN1987A* in the large Magellanic cloud, provided the first opportunity to measure gamma-ray lines from a individual type II supernova. Gamma-rays are of particular interest as a diagnostic of the various progenitor models and

Table 1. Inventory of observed gamma-ray line sources

Physical Process	Energy [keV]	Source	Flux [$\text{ph cm}^{-2} \text{s}^{-1}$]	Instrument (detector type)	Ref.
Nuclear deexcitation					
$^{56}\text{Fe}(\text{p}, \text{p}', \gamma)$	847	Solar flares	up to ≈ 0.05	SMM (NaI scintillator)	[16]
$^{24}\text{Mg}(\text{p}, \text{p}', \gamma)$	1369	Solar flares	up to ≈ 0.08	SMM (NaI scintillator)	[16]
$^{20}\text{Ne}(\text{p}, \text{p}', \gamma)$	1634	Solar flares	up to ≈ 0.1	SMM (NaI scintillator)	[16]
$^{28}\text{Si}(\text{p}, \text{p}', \gamma)$	1779	Solar flares	up to ≈ 0.09	SMM (NaI scintillator)	[16]
$^{12}\text{C}(\text{p}, \text{p}', \gamma)$	4439	Solar flares	up to ≈ 0.1	SMM (NaI scintillator)	[16]
$^{16}\text{O}(\text{p}, \text{p}', \gamma)$	6129	Solar flares	up to ≈ 0.1	SMM (NaI scintillator)	[16]
Radioactive decay					
$^{56}\text{Co}(\text{EC}, \gamma)^{56}\text{Fe}$	847, 1238, 2598	SN 1987A	$\approx 10^{-3\text{b}}$	various scintillators and Ge detectors	[38]
	847, 1238	SN 1991T	$\approx 10^{-3\text{b}}$	COMPTEL (scintillators)	[46]
$^{57}\text{Co}(\text{EC}, \gamma)^{57}\text{Fe}$	122, 136	SN 1987A	$\approx 10^{-4}$	OSSE (NaI-CsI phoswich)	[31]
$^{44}\text{Ti}(\text{EC})^{44}\text{Sc}(\beta^+, \gamma)$	1157	Cas A SNR	$7 \cdot 10^{-5}$	COMPTEL (scintillators)	[47]
	1157	RX J0852.0-4622 ^c	$3.8 \pm 0.7 \cdot 10^{-5}$	COMPTEL (scintillators)	[48]
$^{26}\text{Al}((\beta^+, \gamma)^{26}\text{Mg})$	1809	structured galactic plane	$4 \cdot 10^{-4}/\text{rad}$	COMPTEL (scintillators)	[32-37]
	1809	Cygnus region	$7.9 \pm 2.4 \cdot 10^{-5}$	COMPTEL (scintillators)	[39]
	1809	Vela region	$1-6 \cdot 10^{-5}$	COMPTEL (scintillators)	[32-37]
$e^+ - e^-$ Annihilation					
	511	galactic bulge	$1.7 \cdot 10^{-3}$	OSSE (NaI-CsI phoswich), Ge detectors	[19-25]
	511	galactic disk	$4.5 \cdot 10^{-4}$	OSSE (NaI-CsI phoswich), Ge detectors	[19-25]
	$480 \pm 120^{\text{a}, \text{c}}$	1E 1740-29	$1.3 \cdot 10^{-2}$	SIGMA/NaI scintillator	[27]
	511	Solar Flares	up to ≈ 0.1	SMM (NaI scintillator)	[15, 16]
	$479 \pm 18^{\text{a}}$	Nova Muscae	$6.3 \cdot 10^{-3}$	SIGMA (NaI scintillator)	[40]
	$400-500^{\text{a}}$	Gamma Ray Bursts ^c	up to ≈ 70	various scintillators	e.g. [41]
	$73 \dots 500^{\text{a}}$	Crab Pulsar ^c	$3 \cdot 10^{-4}$	various scintillators	see [42]
Neutron Capture					
$^1\text{H}(\text{n}, \gamma)^2\text{H}$	2223	Solar flares	up to ≈ 1	SMM (NaI scintillator)	[16]
	2223	White dwarf? RE J0317-853 ^c		COMPTEL (scintillators)	[17]
$^{56}\text{Fe}(\text{n}, \gamma)^{57}\text{Fe}$	5947 ^a	June 10 1974 Transient	$1.5 \cdot 10^{-2}$	balloon borne Ge detector	[43]
Cyclotron Lines					
	10-70	Gamma Ray Bursts ^c	up to ≈ 3	various scintillators	[43]
	20-58	various pulsars (9, eg Her X-1)	$3 \cdot 10^{-3}$	scintillators	[49, 50]
	73-79	Crab Pulsar	$4 \cdot 10^{-3}$	scintillator	[44]

a) Redshifted line

b) Maximum emission

c) single and/or marginal detection, feature has yet to be confirmed by other instruments

Table 2. Principal cornerstones in the development of high energy astronomy

1895	G. Roentgen	discovery of X-rays
1896	H. Becquerel	discovery of radioactivity
1899	E. Rutherford	discovery of atomic nucleus
1900	P. Villard	discovery of gamma-rays
1911	V. Hess	discovery of Cosmic Rays (balloons, growth curves)
1932	C. Anderson	discovery of positron (balloon borne Wilson-chamber)
1948	Hulsizer & Rossi	high energy γ 's $< 1\%$ of CR (counters, balloon/B29)
1948	Perlow & Kissinger	marginal measurement of cosmic γ -rays (counters, V2 rocket)
1958	EXPLORER 1	discovery of radiation belts (J. Van Allen)
1958	Peterson & Winckler	first gamma-rays from solar flare (balloon, counters)
1958	Ph. Morrison	Vatican conference (nouvo cimento): predictions ...
1960's	RANGER 3 & 5	cosmic diffuse flux: $dn(E) \sim E^{-2.2}$
1961	EXPLORER 11	22 cosmic HE γ -rays detected, BG of 22000 CR events
1962	ASE-MIT rocket	first cosmic X-ray source: Sco X-1
1967/68	OSO-3	HE γ -rays from the Galaxy
1967	VELA satellites	discovery of γ -ray bursts (nuclear test ban treaty)
1970	UHURU	first X-ray sky survey
1972 ff	balloons	detection of cosmic 511 keV annihilation line
1972,75	SAS-2, COS-B	HE γ -rays from galactic plane, Vela, Geminga
1979	HEAO-3	discovery of galactic ^{26}Al (Ge spectrometer)
1987	SMM, balloons	SN1987A: ^{56}Co line, SN ν detection
1989-98	GRANAT/SIGMA	variable galactic center sources
1991-99	Compton-GRO	^{26}Al sky map, ^{44}Ti from Cas A, compact source spectra
1997	Beppo-SAX et al.	γ -ray burst afterglow/identification of hosts galaxies

explosion scenarios for supernovae because they allow the direct observation of radioactive isotopes – particularly the $^{56}\text{Ni} \rightarrow ^{56}\text{Co} \rightarrow ^{56}\text{Fe}$ decay chain – that power the observable light curves and spectra. Six months after the explosion, SMM discovered the 847 keV gamma-ray line [38] identifying freshly produced ^{56}Co . A rough “light curve” of the 847 keV line was established by SMM and successive balloon observations. The early appearance of the ^{56}Co line has been interpreted as evidence for enhanced mixing of the supernova products within the envelope. After the launch of CGRO in 1991, SN1987A

was observed by the OSSE spectrometer [45]. The evidence for gamma-ray line (122 keV and 136 keV) and continuum emission from ^{57}Co indicates that the ratio $^{57}\text{Ni}/^{56}\text{Ni}$ produced in the explosion was about 1.5 times the solar system ratio of $^{57}\text{Fe}/^{56}\text{Fe}$.

Soon after the beginning of the CGRO mission, SN1991T, a type Ia supernovae has occurred in the direction of the Virgo cluster. A marginal detection of the 847 keV and 1.238 MeV ^{56}Co lines has been reported by COMPTEL [46]. While the SN1991T optical light curve and brightness suggests that $\sim 1.0 M_{\odot}$ of ^{56}Ni were ejected in the event, the COMPTEL observations imply an ejected ^{56}Ni mass of $\sim 1.3 \pm 0.5 M_{\odot}$ (for a distance of 13 Mpc), just about compatible with theoretical SNe Ia model predictions ($M_{56\text{Ni}} \leq 0.9 M_{\odot}$).

In 1994, GRO-COMPTEL discovered a gamma-ray line at 1157 keV emitted by radioactive ^{44}Ti . The source location is compatible with the young (only ~ 300 years old) supernova remnant Cas A [47]. The relatively short decay time of 87 years of ^{44}Ti is comparable to the average time between galactic supernovae and should result in a spotty appearance of the Milky Way at 1157 keV. Based on its 1.15 MeV sky-survey, COMPTEL has announced the tentative detection of a previously unknown supernova remnant, RX J0852-46 or “Vela Junior” [48], which subsequently has been identified in the ROSAT all sky data. Although more complete COMPTEL data indicate that the detection of RX J0852-46 is marginal, it illustrates nevertheless the potential of gamma-ray line astronomy for detection of supernova remnants in otherwise inaccessible regions.

Cyclotron Lines

Since the historic discovery of a cyclotron line in the spectrum of Her X-1 (Trümper, 1977 [49]), such lines have been observed in nine more pulsars – seven of these with Ginga [50] and recently two more with BeppoSAX [51]. The absorption-like features reflect the geometry and physical conditions near the surface of the neutron star. Electrons in an accreting hot, ionized plasma threaded by the strong magnetic fields of the neutron star undergo transitions between discrete Landau levels. This process produces cyclotron resonant scattering lines in the emission spectrum at the fundamental cyclotron frequency, $E_{\text{cyc}} = 11.6(B/10^{12}\text{G}) \text{ keV}$, and its harmonics. While the energy of the line is a direct measurement of the magnetic field strength, the line profile constrains the spatial distribution of the field, the geometry of the accretion flow, and the temperature and optical depth of the X-ray emitting plasma.

High-Energy Gamma Rays

The study of cosmic-rays has progressed with stratospheric balloons ever since their discovery by Victor Hess in 1911–12. At energies above 1 GeV,

Hulsizer and Rossi [52], using a balloon borne ionization chamber, came to the conclusion that less than 1% of the incoming cosmic ray flux was composed of gamma-rays (and electrons).

The first 22 high energy gamma-ray photons were detected by the Explorer-11 spacecraft in 1961 (Kraushaar et al. [53] and [54]). The signal was measured by a scintillator-Cerenkov counter detector, surrounded by a plastic anticoincidence scintillator who efficiently rejected a background of 22000 events induced by charged particles. An improved version of the detector was flown on the OSO-3 satellite [55]. It confirmed the detection of Explorer-1 and indicated an emission of galactic origin.

From here to the 271 high-energy gamma-ray sources of the third EGRET Catalog [56], considerable effort has gone into the development of sensitive detector system. Several types of imaging detectors for high energy gamma-rays were developed and flown on balloons and satellites: conventional optical spark chambers using cameras and film; spark chambers viewed by vidicon tubes; the sonic spark chamber using microphones to record the position of the spark, the proportional counter; and the multiwire magnetic core, digitized spark chamber (see e.g. [57]). Mayor achievements in High-Energy Gamma-rays were the first skymap of the inner galactic plane by NASA's SAS-2 (launched in 1972, see e.g. [58]) and the map of the entire galactic ridge by the ESA satellite COS-B (launched in 1975, see e.g. [59]). The measurements of these two instruments indicated that the gamma-ray emission is strongly correlated with galactic structural features; these results fed a lively discussion on a possible gradient of cosmic rays in the Galaxy, and whether cosmic ray are of galactic or extragalactic origin.

The mayor steps in the history of high energy astronomy are summarized in Table 2, for more information on the development of gamma-ray astronomy, see the historical reviews by Greisen in 1966 [60], in Chupp's book, 1976 [61], or in Pinkau, 1996 [62].

1.2 From Gamma-Ray Astronomy to Nuclear Astrophysics

The Golden Age of Gamma-Ray Astronomy?

With the large satellite platforms of the nineties, the Compton Gamma Ray Observatory and GRANAT/SIGMA, the gamma-ray sky has now been surveyed on various angular scales and a number of new gamma-ray sources has been discovered. The general gamma-ray point source catalog established by Macomb and Gehrels, in 1999 [63] contains 309 objects in the energy range between 50 keV and 1 TeV, and the fourth BATSE gamma-ray burst catalog alone lists 1637 gamma-ray bursts [64].

One of the principal merits of this generation of high energy instruments was their extremely broad coverage – both in energy and angular extent. Together with the operating X-ray telescopes, a quasi-continuous coverage has opened the possibility for multi-wavelength studies of continuum spectra

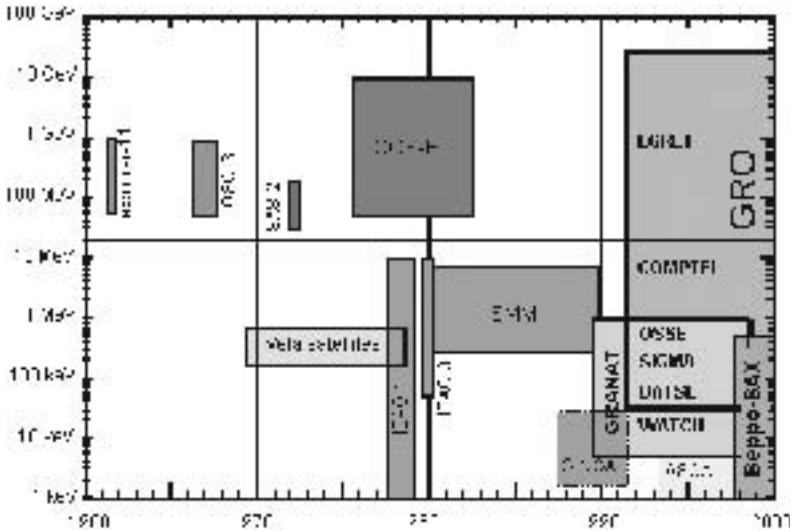


Fig. 1. The “golden age of gamma-ray astronomy”? Never before the high-energy sky has been examined so thoroughly and over such a broad energy range

spanning from the keV- to the GeV-range (Fig. 1). Will the last decade of the 20th century once be called the “golden age of gamma-ray astronomy”?

For many of the high energy sources, multi-wavelength studies may actually be the only way that leads to an understanding of their complex source mechanisms. A model case is the spectrum of the quasar 3C273 that has been observed – partly simultaneously – from radio to gamma-ray energies (see e.g. [65]). Nevertheless, the gamma-ray telescopes on the Compton Gamma Ray Observatory and on GRANAT also have raised new astrophysical questions and highlighted those which remain unanswered. The future goals of gamma-ray astronomy must be defined in this context.

The progress in nuclear astrophysics made during the last decade by SIGMA, BATSE, OSSE and COMPTEL is based primarily on skymaps, excellent timing analysis, and moderate to fair spectral resolution. The observations have revealed specific aspects of the morphology of celestial gamma-ray emitters, yet the physical processes at work are often only poorly understood. Frequently, the observed spectra do not sufficiently constrain the emission mechanisms: explaining a relatively simple, featureless continuum with a complex multiparameter model can be ambiguous, moreover, different components may blend into one another, each of them can depend on various physical parameters in the emitting region.

In many ways, the present situation resembles the situation of optical astronomy in the beginning of the 19th century: Back then, the available observational data mainly consisted in images, starcounts, variabilities, and color indices. Astrophysics was born when G. Kirchhoff and R. Bunsen developed

spectral analysis and explained the Fraunhofer-lines in the spectrum of the sun. The exploration of atomic and molecular lines has since turned out to be the most powerful tool for the study of the physical conditions in celestial sources.

While optical lines reflect structural changes in the electron shell of atoms, caused by collisions with energies of the order of 10^{-3} eV ($T \sim 1000$ K), transition between discrete nuclear energy levels imply MeV energies ($T \sim 10^7$ to 10^9 K), corresponding to the binding energy of nucleons. Collision energies of this order are characteristic of the conditions inside of stars, particles accelerated by electromagnetic fields in solar flares, or interactions of cosmic ray particles with the interstellar medium.

Up to today, little advantage has been taken of the fundamental astrophysical information contained in gamma-ray lines. The reason for this is the modest energy resolution of most of the existing instruments (typically $\Delta E/E \approx 10\%$). Nevertheless, the available elementary spectroscopic measurements (see the inventory in Table 1) already indicate the tremendous potential of gamma-ray lines – here’s a window to nuclear transitions in astrophysical sites – the direct way to study nucleosynthesis and cosmic ray excitation of interstellar matter.

The Challenge of Nuclear Astrophysics

At present, barely three dozen objects are known in the range of nuclear lines [63] (excluding gamma-ray bursts). For comparison, in the soft X-ray domain, more than 60000 sources have been detected during the ROSAT all-sky survey; the ROSAT Bright Source Catalogue [66] alone counts 18811 entries. Based on the databases of ASCA and Beppo-SAX, a rough estimate for the sources known at hard X-ray energies results in several hundred sources above 10 keV, and more than 1000 below this energy. Even in high energy gamma-ray astronomy (> 30 MeV), where sources are typically several orders of magnitude weaker than at MeV energies, 271 sources have been discovered [56] – nearly an order of magnitude more than in the nuclear range.

With all the neighboring domains having come to maturity, why is the MeV range still in its adolescence? Has nature provided this energy band with less sources? Is an intrinsically insurmountable barrier obstructing the view on this range of the gamma-ray sky?

Figure 2 compares the number of sources presently known in the various bands of high energy astrophysics (a) with the relevant physical constraints of the detection process: the mass attenuation coefficient of a typical detector material is shown in Fig. 2(b). The similarities with the source statistics above are striking – here are two ways of expressing the probability for electromagnetic radiation interacting with matter. Besides the minimum of the cross section at MeV energies, telescopes for this domain have to cope with the fact that there is not a single but *three* main interaction processes of gamma-rays with matter.

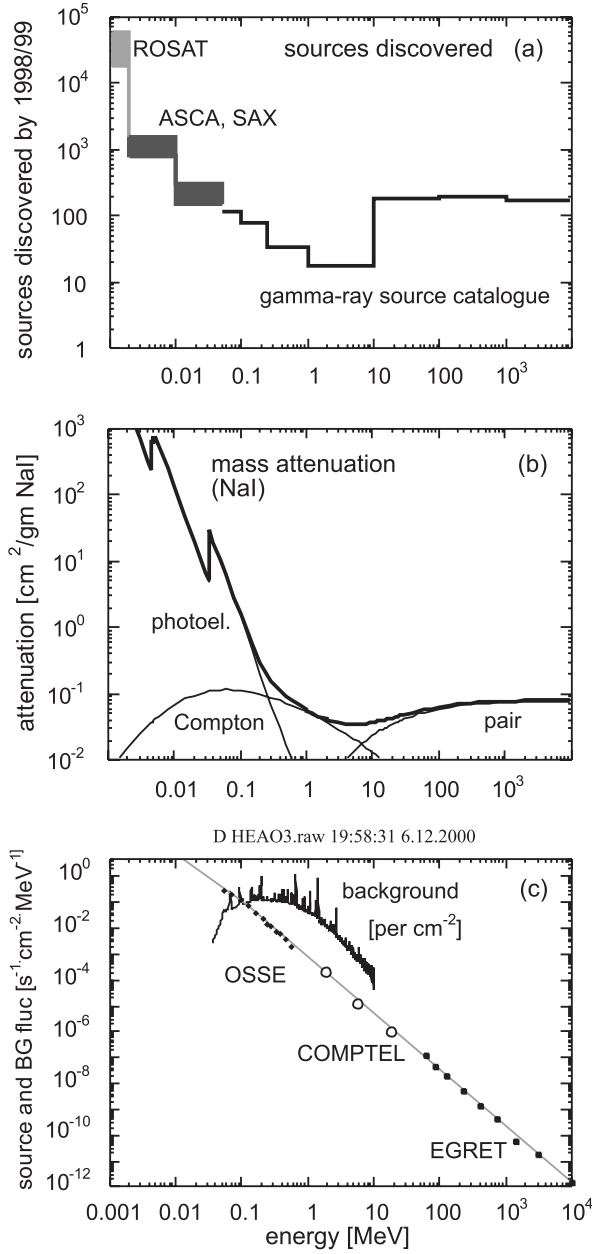


Fig. 2. The discoveries in nuclear astrophysics – confrontation with the realities of detector efficiency, background and source strength (see text)

The bottom panel of Fig. 2 displays the source spectrum of the Crab nebula, the strongest permanent point source at MeV energies as measured by the instruments on CGRO. A typical detector background of a spaceborne gamma-ray spectrometer (HEAO-3) is also shown for comparison. The spectrum shown here is actually an equivalent background flux f_b . It has been obtained by scaling the original HEAO-3 spectrum b [$s^{-1} \cdot cm^{-3} \cdot MeV$] with the photon mean free path μ [cm] in Germanium: $f_b = b \cdot \mu$ [$s^{-1} \cdot cm^{-2} \cdot MeV$]. This quantity not only directly compares with a source flux, it also is the relevant measure for an optimal detector background at a given energy. The background in the nuclear range is maximum not only because of the myriad of physical processes that produce high background rates per unit volume (particularly when exposed to cosmic-ray bombardment in the spacecraft environment outside the atmosphere), but also because the minimum attenuation coefficient (see Fig. 2b) necessitates the thickest detectors, hence very large volumes for background production.

In addition to the difficulties manifest in Fig. 2, the existing telescope systems in MeV astronomy have never used *direct imaging* yet. An important breakthrough for soft X-ray astronomy was in fact direct imaging with high throughput using grazing incidence optics (e.g. EINSTEIN, ROSAT). In high energy gamma-rays, tracking the e^-e^+ pair certainly was the decisive step that brought this domain way ahead of the nuclear range. Tracking makes possible unambiguous backprojection (direct imaging) of every photon, resulting in a tremendous enhancement of the sensibility, since the background in a given source direction is suppressed to virtually zero.

If nature has made the MeV sky almost inaccessible, why should we continue building instruments for nuclear astrophysics? In the first place, there is certainly no evidence for a lack of sources at MeV energies with respect to other energy bands (Fig. 2). Yet, there is physics that could't have been done (e.g. nuclear lines, Sect. 1.2) and discoveries that would never have been made (e.g. gamma-ray bursts) if this window remained closed; and although continuum spectra are steep, the energy flux per decade usually is comparable to neighboring domains. For example, a typical photon spectrum $dN(E_\gamma) \sim E_\gamma^{-2} dE$ implies equal amounts of energy in equal logarithmic energy intervals.

It is certainly not a coincidence that each of the experimental problems represents an exclusive opportunity in the study of astrophysical phenomena: On one side, the low cross section for the interaction of gamma-rays with matter leads to low detector efficiencies, but, on the other side, it makes the universe extremely transparent in this energy range. The struggle dealing with *three different* interaction processes of photons in the detectors – photoeffect, Compton scattering and pair production – is more than matched by the fact that in the most violent astrophysical objects (AGN, gamma-ray bursts), the bulk of the energy transfer occurs in their inverse processes – bremsstrahlung, inverse Compton scattering and matter-antimatter annihilation. Finally, the

numerous background components that experimenters have to contend with: hadronic and electromagnetic cascades from cosmic ray interactions, neutron activation of the spacecraft and telescope materials, elastic neutron scattering, positron annihilation . . . all these processes emphasize the extremely rich physics in the nuclear energy range and most of them correspond to an astrophysical emission mechanism.

1.3 Requirements on Instruments for Gamma-Ray Spectroscopy

Sensitivity is unquestionably the foremost requirement on all future instruments for nuclear astrophysics: spectroscopy will not lead to any physics if the gamma-ray sources are detected just above the sensitivity limit – sufficient statistics are a prime necessity. Furthermore, nuclear astrophysics will not become a full-fledged branch of astronomy unless the number of known sources (Sect. 1.2) is at least equal, and possibly greater than the number of astronomers in the community.

The performance requirements for gamma-ray line spectroscopy missions can be illustrated by comparing measured or anticipated line fluxes with the observed or expected angular scales: Fig. 3 indicates that emissions with a wide range of angular and spectral extent are expected, varying in intensity by several orders of magnitude. The scientific objectives for gamma-ray spectroscopy span through compact sources such as broad class annihilators,

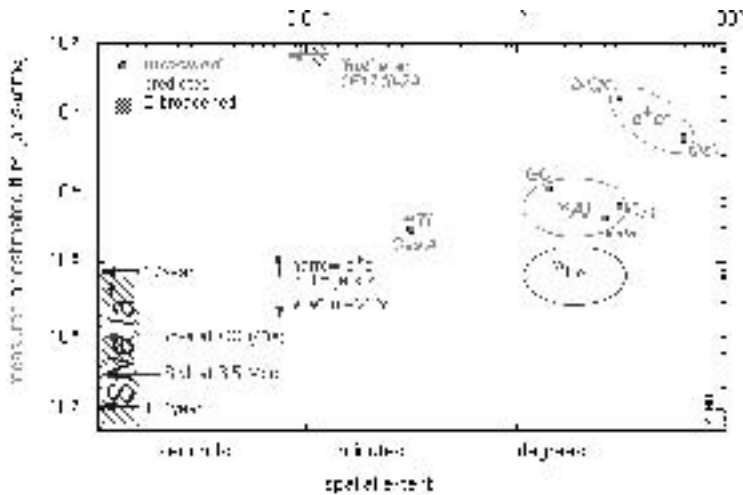


Fig. 3. Future spectroscopy missions have to face emissions with a wide range of angular extent, and with intensities different by several orders of magnitude. The anticipated flux for extragalactic SNe of type 1 has been deduced from the COMPTEL detection of SN1991T [46] and by scaling its ^{56}Co 847 keV gamma-ray flux with the optical peak magnitude of observed SNIa

long-lived galactic radioisotopes with hotspots possibly in the degree-range, to the extremely extended galactic disk and bulge emission of the narrow e^+e^- line.

From the previous generation of instruments (sensitivity $> 10^{-5} \text{ ph} \cdot \text{s}^{-1} \cdot \text{cm}^{-2}$) we have learned that narrow lines generally seem to be emitted from extended distributions, while broad lines tend to be radiated by compact sources. Hence, a natural next objective for gamma-ray line spectroscopy is the mapping of the relatively intense sources (on the upper right of Fig. 3) which are typically emitting $10^{-4} \text{ ph} \cdot \text{cm}^{-2} \cdot \text{s}^{-1}$ to a few $10^{-6} \text{ ph} \cdot \text{cm}^{-2} \cdot \text{s}^{-1}$.

Candidate sources of this intensity are mostly galactic and include the sites of recent nucleosynthesis, regions of e^+e^- annihilation and clouds where nuclear deexcitation by energetic particles takes place. Some of them might appear as extended structures: either because of their apparently diffuse origin – as in the case of narrow 511 keV line – or because they are relatively close by as the nucleosynthesis sites in the local spiral arm (^{26}Al in the Vela and Cygnus region). An instrument that is adequate for this kind of objectives should provide a sensitivity of several $10^{-6} \text{ ph} \cdot \text{cm}^{-2} \cdot \text{s}^{-1}$, a wide field of view and an angular resolution in the degree range. Such a profile corresponds to the performance of the coded mask spectrometer SPI on ESA's INTEGRAL mission (Sect. 4.1).

On a more distant horizon, experimental gamma-ray astronomy has to find ways to further extend the limits of resolution and sensitivity: At energies above $\sim 511 \text{ keV}$, Compton telescopes might achieve line sensitivities of several $10^{-7} \text{ ph} \cdot \text{cm}^{-2} \cdot \text{s}^{-1}$ and provide angular resolutions of fractions of degrees.

Apart from a few exemptions (SN1987A, possibly SN1991T and very few compact galactic objects), the evidence for point-like sources of narrow gamma-ray line emission has been mostly implicit. Yet, in the area at the lower left of Fig. 2 various objects like e.g. galactic novae and extragalactic supernovae are predicted. These sources will have small angular diameters but very low fluxes – mostly because such objects are relatively rare and therefore are more likely to occur at large distances. In order to cover the objectives in this area, experimental gamma-ray astronomy has to find new ways to improve the observational performance.

In the following chapter, the groundwork needed to understand gamma-ray detection will be laid in a summary of the relevant interactions of photons with matter. The various types of detectors for the gamma-rays are discussed in Sect. 3. Finally, three families of telescope systems for gamma-ray astronomy will be discussed: coded aperture systems (Sect. 4.1), Compton telescopes (Sect. 4.2), and focusing instruments (Sect. 4.3).

2 Interaction of High Energy Photons with Matter

How does radiation interact with matter? Instruments for high energy spectroscopic astrophysics must answer several aspects of this question: they not only have to collect photons, but also measure their energy and determine their arrival direction.

A gamma-ray photon has four properties – energy, momentum, spin, and polarization – any interaction will have to satisfy the corresponding laws of conservation. Table 3 summarizes thirteen interaction processes relevant in the energy range of interest for nuclear astrophysics. For the instrumentation in gamma-ray astronomy, three processes are of practical interest:

(I) *photoelectric absorption*: The photon cedes all of its energy to a bound atomic electron. The kinetic energy carried away by the photoelectron is the difference between the photon energy and the binding energy of the electron. The photoelectric effect dominates at low energies (up to several hundred keV).

(II) *scattering by atomic electrons*: The photon is deflected from its original direction, with or without losing energy. If the incident photon energy is sufficiently high compared to the electron binding energy, gamma-rays are scattered by electrons that can be considered free and at rest (Compton effect). While Compton scattering predominates in the MeV region, it represents the high energy limit for the general case of inelastic scattering from bound atomic electrons. Coherent scattering, or Rayleigh scattering (the elastic case) takes place if the electron returns to its original state after the interaction. No loss of energy and phase information takes place, the momentum is transferred to the atom as a whole.

(III) *pair production*: For gamma-ray energies exceeding twice the electron rest mass, the creation of an electron–positron pair becomes possible in the vicinity of a nucleus. While the photon disappears, the particles carry the excess energy above 1.02 MeV. Pair production dominates above 5 to 10 MeV.

For spectroscopic detectors the *energy loss processes* – photoelectric absorption, Compton effect, pair production – are of particular importance. Figure 4 illustrates their relative importance as a function of the atomic number (Z) of the medium. The signature of the primary processes in gamma-ray spectra, and the signature of secondary energy loss processes, will be discussed in Sect. 2.5.

Attenuation Coefficients

Since gamma-ray photons are removed individually from the beam in a single event, the number of photons removed, dI , is proportional to thickness dT of the matter traversed

$$dI = -\mu I_0 dT . \quad (1)$$

Here, I_0 is the number of incident photons, and μ is called the linear attenuation coefficient, it is the probability of an interaction – absorption, scattering

Table 3. Gamma-Ray interaction processes (from C.M. Davisson, 1966 [67])

Process	kind of Interaction	Name	Approximate Energy Range of Maximum Importance	Notation	Approximate Variation (Z)	Remarks
I photoelectric absorption	a with bound atomic e^-	photoelectric effect	dominates at low E (1–500 keV) decreases as E increases	τ	Z^5	
II scattering from electrons	a with bound atomic e^-	Rayleigh scattering coherent or elastic scattering	<1 MeV and greatest scattering angles	$\sigma^{\text{or}}_{\text{R}}$	Z^2 small θ Z^3 large θ	Combines coherently with IVb, IVc, and Vc
coherent	b with free e^-	Thomson scattering	independent of energy	σ_{et}	Z	low frequency limit of Compton scattering
	c with bound atomic e^-		<1 MeV, least at small scattering angles	σ	Z	
incoherent	d with free e^-	Compton scattering	dominates in region of 1 MeV, decreases as energy increases	σ	Z	
III nuclear photoeffect	a with nucleus as a whole	particle production (γ, γ), (γ, n), (γ, p) etc	above threshold has broad maximum between 10–30 MeV			
IV nuclear scattering	a with material as a whole (dep. on nuc E-levels)	Mössbauer effect	important only in very narrow resonance range			Combines coherently with IIa, IVc, and Vc
coherent	b with nucleus as a whole (dep. on nuc E-levels)	nuclear resonance scattering	narrow reson. maxima at low E, broad maxima at 10–30 MeV	$\sigma^{\text{or}}_{\text{r}}$	Z^2/A^2	Combines coherently with IIa, IVc, and Vc
	c with nucleus as a whole (indep. of nuc. E-levels)	nuclear Thomson scattering		$\sigma^{\text{or}}_{\text{r}}$	Z^4/A^2	Combines coherently with IIa, IVb, and Vc
incoherent	d with individual nucleons	nuclear Compton scattering	$\lambda \leq$ nuclear radius, i.e. > 100 MeV			
V interact. w. Coulomb field	a in Coulomb field of nucleus	elastic pair production	threshold ~ 1 MeV, dominant at HE, ϕ increases with E	κ or κ_{pair}	Z^2	
pair production	b in Coulomb field of electron	triplet production	threshold at 2 MeV increases as E increases	κ_{tr} or κ_{triplet}	Z	
Delbrück scattering	c in Coulomb field of nucleus	nuclear potential scattering	real part $>$ imaginary (below 3 MeV) $<$ imaginary (above 15 MeV), real and imaginary both increase as energy increases	$\sigma^{\text{or}}_{\text{D}}$ or $\sigma^{\text{or}}_{\text{D}}$	Z^4	Combines coherently with IIa, IVb, and IVc

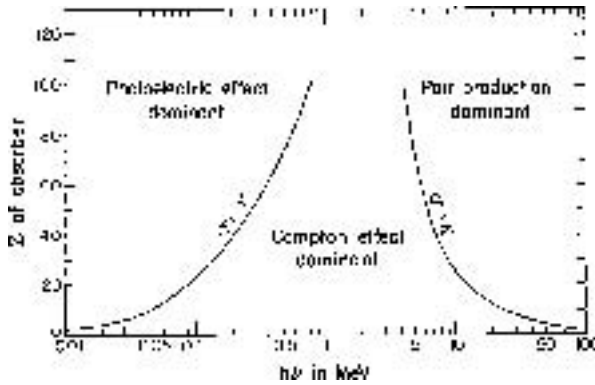


Fig. 4. Z -dependent boundaries of the three principal interactions (from Evans 1955 [68]): The *solid lines* indicate equal interaction probabilities for the photoelectric and Compton effect ($\sigma = \tau$), and for Compton effect and Pair production ($\sigma = \kappa$)

or pair production, occurring per unit path length of the absorber. The total attenuation probability μ is composed of the three independent interaction processes

$$\mu = \tau + \sigma + \kappa \quad (\text{photoelectric, Compton-effect, pair-prod.}) \quad (2)$$

The linear attenuation coefficients, τ , σ , and κ are related to the fundamental cross sections (discussed in Sects. 2.1–2.5) by:

$$\tau = a^\tau \cdot N \quad [\text{cm}^{-1}] \quad \text{photoelectric} \quad , \quad (3)$$

where a^τ is the atomic cross section [cm^2/atom] for photoelectric absorption and N the atomic number density [atoms/cm^3].

$$\sigma = e^\sigma \cdot Z \cdot N [\text{cm}^{-1}] \quad \text{Compton} \quad , \quad (4)$$

where e^σ is the cross section for removing the photon from beam [cm^2/e^-] and Z the number of electrons per atom.

$$\kappa = a^\kappa \cdot N [\text{cm}^{-1}] \quad \text{pair production} \quad , \quad (5)$$

where a^κ is the pair production cross section per nucleus [$\text{cm}^2/\text{nucleus}$].

Instead of using the linear attenuation coefficients which depend on the density and physical state of the absorber, it is more practical to employ the *mass attenuation coefficients*. The total mass attenuation is defined as μ/ρ , with ρ being the density of the material [g/cm^3]. For the individual processes, the mass attenuation coefficients are obtained by dividing the linear coefficients by the density ρ . Tables and graphs generally contain mass attenuation coefficients. As an example relevant to gamma-ray spectroscopy, the total and

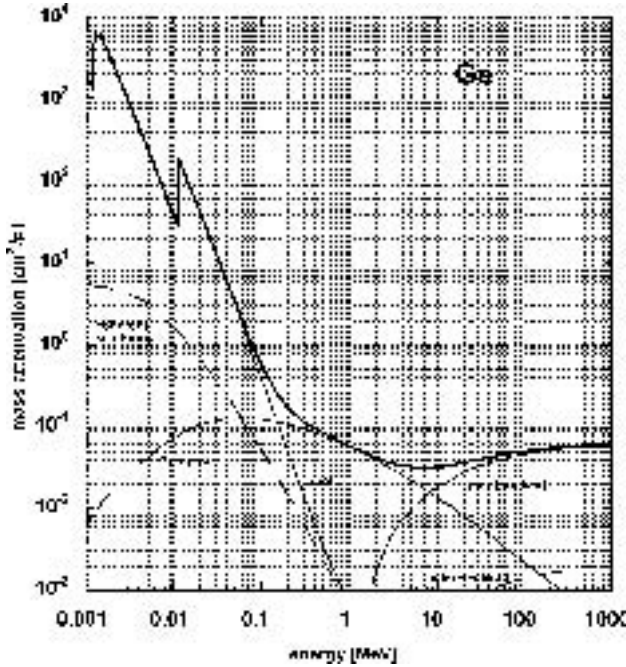


Fig. 5. Mass attenuation coefficients (μ/ρ) for Germanium [69]

individual mass attenuation coefficients of Germanium are shown in Fig. 5. These and other data can be found in the “Photon Cross Sections Database” of the National Institute of Standards and Technology Standard Reference Database XGAM [69].

The number of transmitted photons, I , of a collimated beam traversing the distance t of a medium characterized by μ is attenuated by a factor $e^{-\mu x}$

$$I = I_0 e^{-\mu t} = I_0 e^{-(\mu/\rho) \cdot \rho t} . \quad (6)$$

The product ρt , the *mass thickness* of the absorber, is the relevant parameter used along with mass attenuation coefficients. It is often preferable to express the thickness of an absorber in mass thickness – with respect to the attenuation processes, this quantity has more physical meaning than the geometrical thickness – for example, the essential parameter for a stratospheric balloon observation is the residual atmospheric mass, expressed in ρt [g/cm²].

The mean free path λ is the average distance a photon traverses in a medium before being removed, it is given by

$$\lambda = 1/\mu . \quad (7)$$

For the standard detectors used in gamma-ray spectroscopy (e.g. inorganic scintillators or solid state detectors) the typical mean free path is of the order of a few millimeters up to several centimeters.

2.1 Photoelectric Effect

The photoelectric effect is the complete transfer of the photon energy $h\nu$ to a bound atomic electron. While the incident photon disappears in the interaction with the absorbing atom, the photoelectron carries away the kinetic excess energy E_e that is left after overcoming the binding energy of the electron E_b

$$E_e = h\nu - E_b . \quad (8)$$

Photoelectric absorption cannot take place with free electrons, a third particle is needed for momentum conservation. The interaction is with the entire atom, yet, due to its high mass, the kinetic energy of the recoil atom is usually negligible.

The probability for interaction with an electron of a certain shell is highest for photon energies $h\nu$ slightly greater than E_b . For energies where both are possible, absorption by a K-shell electron is more probable than that by an L shell one, the L shell usually only contributes about 20%. As the photon energy increases, the atomic electrons appear less tightly bound and the absorption cross section drops, approximately according to the power law

$$a^\tau \sim h\nu^{(-7/2)} . \quad (9)$$

The curve of the photoelectric attenuation shows several discontinuities (see e.g. Fig. 5). While a photoelectric interaction with an electron of a certain shell is possible above the binding energy E_b , the photon energy is insufficient just below, causing a sharp drop in the absorption cross section τ . These absorption edges thus mark the binding energies E_b of the various electron shells (K, L, M) of the absorber-materials.

The vacancy in one of the bound shells of the now ionized absorber atom, is filled through capture of a free electron or by rearrangement of the electron shells. This either leads to the emission of an Auger electron or to the emission of one or several X-ray photons.

The photoelectric effect strongly increases with the atomic number Z – as the effect increases with the tightness of the electron binding, it dominates at low energies. Whereas no single analytical formula describes the photoelectric effect over all energies and Z -values, the following expression for $a^\tau(Z, h\nu)$ serves as a rough approximation:

$$a^\tau \sim Z^n h\nu^{(-7/2)} , \quad (10)$$

with n varying from 4 to 5. The case of photoelectric absorption cross section $a^\tau K$ for K shell electrons and photon energies $h\nu \ll m_0 c^2$ can be described [70]:

$$a^\tau K = \sigma_{\text{et}} Z^5 \alpha^4 2^{5/2} (m_0 c^2 / h\nu)^{7/2} , \quad (11)$$

where σ_{et} is the Thomson cross section (see (19) below), r_0 the classical electron radius ($r_0 = e^2 / m_0 c^2$) and α the fine structure constant ($\alpha = 2\pi e^2 / hc$).

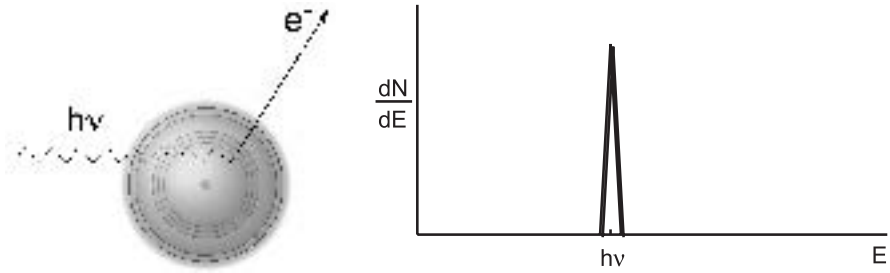


Fig. 6. Monoenergetic gamma-rays detected via the photoelectric effect (most likely with an inner shell electron) and the corresponding energy loss spectrum in a large detector

The photoelectric absorption is the ideal interaction process for spectroscopic detectors (Fig. 6). While the photoelectron carries off most of the gamma-ray energy, it is generally unlikely to escape from a detector due to its short range (typically 1 mm per MeV in moderate density materials). Characteristic X-rays resulting from electron rearrangement have ranges of the same order or shorter; their escape from the large detectors used in gamma-ray astronomy is generally not a significant effect. Hence, monochromatic gamma-rays impinging on a large detector and interacting by photoelectric absorption will result in an energy loss spectrum with a single peak at the energy of the gamma-rays.

2.2 Scattering from Free Electrons

Compton Scattering

For photon energies $h\nu$ largely superior to the electron binding energies, atomic electrons can be considered free. The limiting case of a photon scattered by an electron that is free and at rest is described by the *Compton effect*. The incoming photon transfers a part of its energy and momentum to an electron, the *recoil electron*. The energy $h\nu'$ of the scattered photon can be obtained from the relativistic equations for conservation of energy and momentum:

$$h\nu' = \frac{h\nu}{1 + \alpha(1 - \cos \theta)} . \quad (12)$$

θ is the scatter angle, $\alpha = h\nu/m_0c^2$, with m_0 being the rest mass of the electron.

The intensity I of the scattered photons at the angle θ and distance r from a single scattering electron is

$$I = \frac{I_0}{r^2} \frac{h\nu'}{h\nu} \frac{d_e\sigma}{d\Omega} , \quad (13)$$

where I_0 is the intensity of the incident beam, $d_e\sigma$ is the cross section per electron for the number of photons scattered into the solid angle $d\Omega$ in the direction θ . The differential cross section for Compton scattering has been calculated by Klein and Nishina [71] – for unpolarized radiation they obtain

$$\frac{d_e\sigma}{d\Omega} = \frac{1}{2}r_0^2 \left(\frac{h\nu'}{h\nu} \right)^2 \left(\frac{h\nu}{h\nu'} + \frac{h\nu'}{h\nu} - \sin^2\theta \right), \quad (14)$$

with r_0 the classical electron radius (e^2/m_0c^2). By substituting the ratio $h\nu'/h\nu$ from (12), $e\sigma$ is obtained as a function of θ

$$\frac{d_e\sigma}{d\Omega} = \frac{1}{2}r_0^2 \frac{1}{[1 + \alpha(1 - \cos\theta)]^2} \left[1 + \cos^2\theta + \frac{\alpha^2(1 - \cos\theta)^2}{[1 + \alpha(1 - \cos\theta)]} \right]. \quad (15)$$

A graphical representation of the differential Compton cross section is shown in Fig. 7. Equation (15) describes the cross section per electron, in order to obtain the atomic cross section for a certain element, they have to be multiplied by the atomic number Z : $d_a\sigma = Zd_e\sigma$.

In Fig. 8, the energy loss spectrum of monochromatic gamma-rays interacting through a single Compton scattering is sketched. The energy E_{e-} deposited in the detector is the difference between the incident and scattered gamma-ray energy, $E_{e-} = h\nu - h\nu'$. E_{e-} can take any value in the continuum between zero and $E|_{\theta=\pi}$. This spectrum is called the *Compton continuum*, the maximum energy E_{ce} ($\theta = \pi$) is called the *Compton edge*. In the extreme case of forward scattering, angles $\theta \approx 0$, the scattered gamma ray have

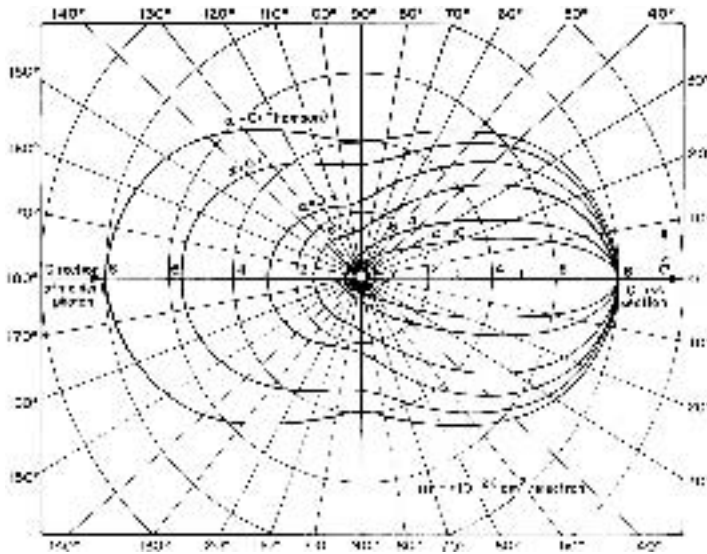


Fig. 7. The differential Compton cross section in polar coordinates $\alpha = h\nu/m_0c^2$ [68]

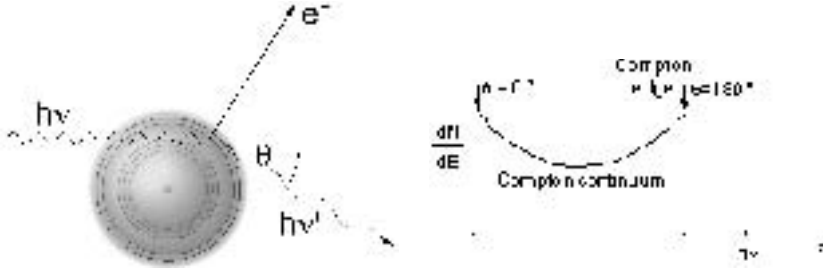


Fig. 8. Monoenergetic gamma-rays interacting through a single Compton scattering (most likely with an outer shell electron, because they are more numerous) and the corresponding energy loss spectrum

energies $h\nu' \approx h\nu$, and the energy E_{e^-} deposited is near zero. For backscattering of the photon at $\theta = 180^\circ$, the maximum energy is transferred to the electron that recoils along the direction of incidence. In this case, the energy of a backscattered gamma-ray becomes

$$h\nu'|_{\theta=\pi} = \frac{h\nu}{1 + \frac{2h\nu}{m_0c^2}} , \quad (16)$$

while the energy loss spectrum shows an event at the Compton edge:

$$E|_{\theta=\pi} = h\nu \frac{2h\nu/m_0c^2}{1 + 2h\nu/m_0c^2} . \quad (17)$$

Thomson Scattering

In the limiting case of low photon energies ($h\nu \ll m_0c^2$) scattering on a free electron, the Klein–Nishina equation (15) reduces to the classical equation for Thomson scattering, also called elastic or coherent scattering.

$$\frac{d\sigma_{\text{et}}}{d\Omega} = \frac{1}{2}r_0^2(1 + \cos^2 \theta) . \quad (18)$$

Here, the electron is considered a harmonic oscillator in the E-field of the incident radiation. The total Thomson cross section is

$$\sigma_{\text{et}} = \frac{8\pi}{3}r_0^2 . \quad (19)$$

2.3 Scattering from Bound Electrons

For atomic electrons, the Compton effect discussed above is the limiting case for high photon energies where electrons can be considered free. A general

description of photons scattering on matter has to include the effects of the binding energies of the electrons, their motion and distribution within the atom. Two cases are possible – *coherent* and *incoherent* scattering from bound electrons of an atom. When scattering coherently ($h\nu' = h\nu$), the electrons return to their original state after the interaction while the entire atom absorbs the momentum (in the case of concrete instrumental applications, the momentum may be transferred to an array of atoms – e.g. a crystal). In order to obtain the intensity of the radiation scattered by the atom, the *amplitudes* of the scattered radiation by each electron *are added* and the sum is squared. In the case of incoherent scattering ($h\nu' < h\nu$) by the electrons of an atom, there is no phase relation between the radiation of the different electrons, the total scattered intensity is obtained by *adding the intensities* scattered from each electron of the atom.

The following discussion of coherent and incoherent scattering from atoms follows the one given in Davisson [67]. An approximate differential cross section per atom $d_a\sigma$ for both incoherent ($d_a\sigma_{\text{in}}$) and coherent ($d_a\sigma_{\text{co}}$) scattering can each be written as the product of two factors:

$$\begin{aligned} d_a\sigma &= d_a\sigma_{\text{in}} + d_a\sigma_{\text{co}} \\ &= Z d_e\sigma \cdot S + d_e\sigma_{\text{et}} \cdot f^2 \end{aligned} \quad (20)$$

The two cases of incoherent scattering and coherent scattering are discussed below. In both cases, the first factor concerns the probability that the photon be deflected by a certain angle, transferring a corresponding amount of momentum to the electron as though the electron were free. In the case of incoherent scattering, the second factor, is the probability that the electron having received this momentum, will absorb a certain amount of energy and thereby become excited or leave the atom. The second factor, in the case of coherent scattering, is the probability that the Z electrons of an atom take up a recoil momentum p without absorbing energy. In both cases, the second factor is a function of the momentum transfer p

$$p = \frac{2h}{\lambda} \sin(\theta/2) \quad \text{for } h\nu(1 - \cos\theta) \ll m_0c^2. \quad (21)$$

Incoherent Scattering from Bound Atomic Electrons

The probability to deflect the photon by a given angle, the first factor, may be taken as the Klein–Nishina cross section $d_e\sigma$. The second factor (per electron) is the incoherent scattering function S . S can be derived from the square of a generalized atomic form factor, summed over all excited states of the atom, integrated over the continuous spectrum and divided by Z . It can be expressed as $1 - Z^{-1} \sum_1^Z f_n^2 - C$, where $\sum_1^Z f_n$ is the atomic structure factor f (see below), and f_n is the electronic structure factor, it gives the amplitude of the radiation scattered coherently by the n^{th} electron, in terms of that scattered

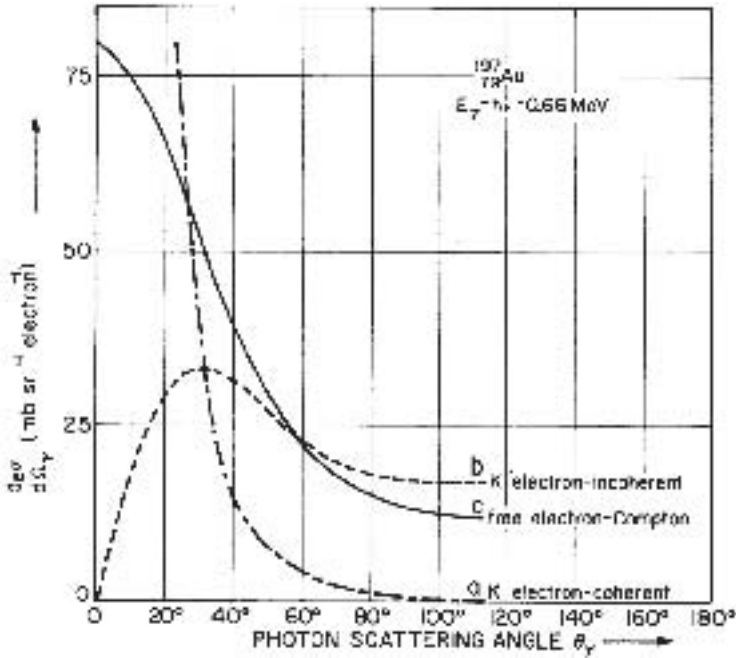


Fig. 9. Differential cross section per unit solid angle for the scattering of 662 keV photons on gold (from Motz and Missioni [72])

- (a) coherent scattering from K electron, calculated from [73]
- (b) incoherent scattering from K electron, experimental
- (c) Compton scattering – free electron (Klein–Nishina equation)

coherently by a free electron. C is a corrective factor taking into account electron transitions forbidden by the exclusion principle. The differential cross section per atom for incoherent scattering can now be written

$$d_a \sigma_{in} = Z d_e \sigma \cdot S. \quad (22)$$

For 662 keV photons scattering on gold, Fig. 9 compares the cross section for incoherent K electron scattering [72], with calculations for the two limiting cases: Compton scattering as calculated by the Klein–Nishina equation, and the calculated cross section for coherent K electron scattering. As expected from theory, the cross section for bound electron scattering approaches zero for small scattering angles (yet, it was found to be greater than expected at large angles). The curves clearly show the predominance of coherent scattering at small scatter angles.

Coherent Scattering from Bound Atomic Electrons

The intensity of the radiation scattered from an atom is obtained by squaring the *sum of the amplitudes* of the scattered radiation by each electron of

the atom. The amplitude of scattering from an atom is called *atomic scattering factor* or *form factor* f . The atomic scattering factor f depends on the structure of the atoms electron envelope – it is the ratio of the amplitude of the radiation scattered by the atom to the amplitude which a single electron would scatter. If all the Z electrons in an atom were concentrated at one point, the amplitude scattered by the atom would simply be Z times the amplitude scattered by a single free electron. Yet, the diffuse cloud of varying electron density causes scattering from one part of the atom to be out of phase with scattering from another part so that the two contributions to the total scattering cancels instead of adding. The atomic scattering factor f will therefore, in general, be less than Z .

The phase difference between the radiation scattered by a charge at $\vec{r}$ and the radiation scattered by the same charge at the center of the atom is

$$\phi = 2\pi/\lambda(\vec{s} - \vec{s}_0)\vec{r}. \quad (23)$$

Here, $\vec{s}_0$ is the unit vector of the incident photon direction, and $\vec{s}$ is the vector of the scattered direction. The amplitude of electric field scattered by the Z electrons of an atom is then given by sum of the electronic structure factors f_n (amplitude of the radiation scattered coherently by the n^{th} electron)

$$f = \sum_{n=1}^Z f_n = \sum_{n=1}^Z \int e^{i(2\pi/\lambda)(\vec{s}-\vec{s}_0)\vec{r}} \rho_n(\vec{r}) d^3r, \quad (24)$$

with ρ_n the electron charge density distributions, the probability of finding the n^{th} electron in the volume element d^3r being $\rho_n \vec{r} d^3r$.

At zero scattering angle θ (this is for $\sin \theta/\lambda \rightarrow 0$), the value for a scattering factor f of a given atom has a value equal to the number of electrons Z in the atom. As $\sin(\theta/\lambda)$ increases, the value of the scattering factor f decreases.

The differential cross section for coherent scattering per atom is now obtained by multiplying the differential Thomson cross section $d_e\sigma_{\text{et}}$ (see above) by the square of the atomic scattering factor f

$$d\sigma_{\text{co}} = d_e\sigma_{\text{et}} \cdot f^2. \quad (25)$$

Scattering in a Crystal

Besides of the scattering factors of its constituting atoms, the intensity of the radiation scattered from a crystal depends on the arrangement of the atoms in its unit cell, the thermal disarrangement of the regular lattice and the mosaic structure of the actual macroscopic crystal.

In a crystal, the difference in path length from the scatter-centers is translated into a phase difference between scattered waves. The scattered intensity is proportional to the square of the Fourier transform of the charge density.

In a unit cell consisting of m atoms at the positions $\vec{r}_j = 1, 2, \dots, m$, the scattered radiation from atom j has the relative amplitude f_j . Its contribution

to the total amplitude of the scattered beam is defined by the phase difference $(2\pi/\lambda)(\vec{s} - \vec{s}_0)\vec{r}_j$, with $\vec{s}$ and $\vec{s}_0$ is the unit vectors of the incident and scattered direction, respectively. The scattered amplitude from all j atoms that make up the unit cell is expressed by the structure factor,

$$F = \sum_{j=1}^m f_j e^{i\phi_j} = \sum_{j=1}^m f_j e^{i(2\pi/\lambda)(\vec{s} - \vec{s}_0)\vec{r}_j} . \quad (26)$$

In the terminology of cristallography, a crystal lattice, defined by the fundamental translation vectors $\vec{a}$, $\vec{b}$, $\vec{c}$, the position of every atom in the unit cell is $\vec{r} = x\vec{a} + y\vec{b} + z\vec{c}$. A crystalline plane is described by the Miller indices (hkl) , defined by the reciprocal intercepts on the three basis axes: na/h , nb/k , nc/l with $n \in \mathbb{N}$. The spacing d of the crystalline planes (hkl) in a cubic unit cell with volume a^3 is given by

$$d = \frac{a}{\sqrt{h^2 + k^2 + l^2}} . \quad (27)$$

The *direction* of maximum intensity of scattered radiation is for constructive interference from the atoms of a given crystalline plane (Fig. 10). The Bragg condition gives the relation between the spacing of atomic planes d (hkl) and the angle of incidence θ_B with respect to this set of planes.

$$2d \sin \theta_B = n\lambda . \quad (28)$$

Since $(\vec{s} - \vec{s}_0)\vec{r}_j = \lambda(hx_j + ky_j + lz_j)$, (26) for the structure factor can now be rewritten using Miller indices for a set of crystalline planes (hkl)

$$F(hkl) = \sum_{j=1}^m \int f_j e^{i(2\pi)(hx_j + ky_j + lz_j)} . \quad (29)$$

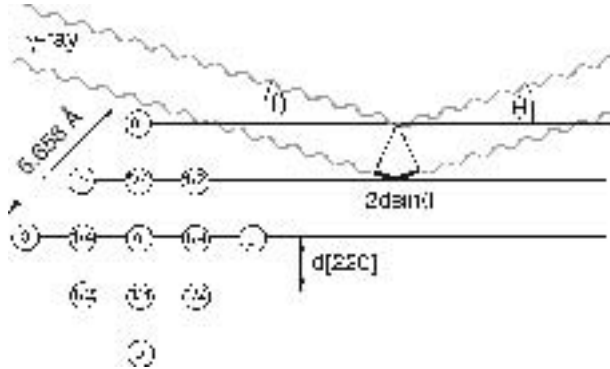


Fig. 10. The Bragg condition for constructive interference from the atoms of a given crystalline plane (e.g. Germanium $[220]$ planes)

For a lattice consisting of identical atoms, the structure factor can be expressed by the *geometric structure factor* S_{hkl} and the atomic scattering factor f : $F(hkl) = fS_{hkl}$, where

$$S_{hkl} = \sum_{j=1}^m = e^{i(2\pi)(hx_j + ky_j + lz_j)} . \quad (30)$$

Because of their relevance in Sect. 4.3, the example of Germanium or Silicon is presented here. Both crystals have diamond structure, with 8 atoms per unit cell at positions $(0, 0, 0)$, $(1/2, 1/2, 0)$, $(1/2, 0, 1/2)$, $(0, 1/2, 1/2)$, $(1/4, 1/4, 1/4)$, $(3/4, 3/4, 1/4)$, $(3/4, 1/4, 3/4)$, and $(1/4, 3/4, 3/4)$. Their geometric structure factor $S_{hkl} = 8$ for crystalline planes (hkl) where $h + k + l = 4n$, with $n \in \mathbb{N}$, and $S_{hkl} = 5.66$ for h, k, l all odd. Every other combination of h, k, l results in $S_{hkl} = 0$.

The absence of reflections ($S_{hkl} = 0$) from certain crystalline planes (hkl) is explained by destructive interference between sets of intervening planes of atoms. For example, the reflection from the (222) plane in Ge is canceled because the atoms at the points 0 and $1/2$ on the fcc (face-centered cubic) lattice produce a phase shift of π with respect to the atoms at $1/4$ and $3/4$ on a similar fcc lattice displaced along the body diagonal by one forth of its length.

The thermal motion of atoms about their equilibrium positions does not broaden the reflection, but leads to a reduction of the scattered intensity. The scattered amplitude is reduced by the Debye–Waller factor, generally written as

$$f_{DW} = e^{-2\langle u^2 \rangle \sin^2 \theta / \lambda} , \quad (31)$$

here, isotropic harmonic vibrations about the equilibrium positions are assumed with a mean quadratic amplitude $\langle u^2 \rangle$. As the temperature T rises, the mean quadratic amplitude $\langle u^2 \rangle$ of the atoms from their rest-position increases and f_{DW} decreases.

A *perfect crystal* reflects monochromatic radiation over an angular range ω_D called the Darwin width, that is

$$\omega_D = \frac{4r_0|F|d^2}{\pi V} \tan \theta_B , \quad (32)$$

with r_0 the classical electron radius, F the structure factor, d the crystalline plane spacing, V the unit cell volume, and θ_B the Bragg angle [74]. At 200 keV, the Darwin width is 0.1 arc seconds for Ge (111), and 0.02 arc seconds for Ge (333). The very narrow angular acceptance of perfect crystals, and, as a consequence (differentiating (28)) the very narrow energy bandpass, has lead to the use of so-called *mosaic crystals*.

The present discussion of *scattering in mosaic crystals* follows the description of Kohnle [75]. In the *Darwin model* [76] for mosaic crystals, the true defect structure of the crystal, which may be due to dislocations, inhomogeneous strains, etc., is described by an agglomerate of perfect crystal blocks.

Each block is in itself an ideal crystal, but adjacent blocks are slightly offset in angle with respect to one another. The relative displacements of the blocks are large compared to the Darwin width, so that the blocks scatter incoherently. Since the block size is microscopic or sub-microscopic, a large number of blocks take part in the scattering process, and the angular distribution of the blocks can be defined as a continuous function. It is assumed that this function is a Gaussian with a FWHM called the *mosaic width* ω . A thorough description of diffraction in mosaic crystals is given in Zachariasen [74].

The intensity of the reflected beam from a mosaic crystal is governed by the diffraction coefficient α (its definition is analog to the absorption coefficient μ). The relative power change α due to diffraction in the layer of thickness dT equals the integrated reflecting power R of a single block times the probability that the block has the “correct” inclination W times the number of single block layers in dT : dT/t_0 where t_0 the block size.

$$\alpha dT = \frac{WR}{t_0} dT. \quad (33)$$

R is given by the zero-absorption thin crystal reflecting power if the so-called primary extinction is negligible, meaning that the attenuation of the incident beam power by the diffraction process inside each mosaic block can be neglected. This is the case if the mean block size is much smaller than the so-called primary extinction depth t_{ext} [77], which for the Laue case (where the diffracted photon passes through the crystal volume):

$$t_0 \ll t_{\text{ext}} = \frac{2V}{\pi r_0 F \lambda} \sin(90^\circ - \theta). \quad (34)$$

In the Darwin model, and for the case of negligible primary extinction, the diffraction coefficient α for a glancing angle θ near the Bragg angle θ_B can be expressed as

$$\alpha(\theta - \theta_B) = \frac{r_0^2 F^2 \lambda^3 f_{\text{DW}}}{V^2 \sin(2\theta_B(E))} \cdot \frac{1}{\sqrt{(2\pi) \omega'}} \cdot e^{-(\theta - \theta_B(E))^2 / 2\omega'^2} \quad (35)$$

Here, ω' is the mosaic width ω times $1/2\sqrt{2 \ln 2}$.

The efficiency ε of scattering by a mosaic crystal can now be defined as the ratio P_H/P_0 : the number of reflected to the number of incident photons. The variation of the power of the incident and diffracted beams as a function of the penetration depth T inside the crystal is described by the two transfer equations [74]

$$dP_0 = -\mu P_0 \frac{dT}{\cos \theta_B} - \alpha P_0 dT + \alpha P_H dT \quad (36)$$

$$dP_H = -\mu P_H \frac{dT}{\cos \theta_B} - \alpha P_H dT + \alpha P_0 dT, \quad (37)$$

with α the diffraction coefficient, P_0 the power in the direct beam, P_H the power in the reflected beam, μ the absorption coefficient for all incoherent processes, and the direction cosine of the Bragg angle θ_B which scales the thickness dT for absorption along an oblique path. The first two terms in (36) and (37) describe the decrease in power due to absorption and diffraction, the third term is the increase of the incident or of the diffracted beam due to reflection of the diffracted and incident beams respectively. The solution of (36) and (37) for the Laue case, i.e. for the boundary conditions $P_H(0) = 0$, leads to the following expression for the diffraction efficiency.

$$\varepsilon_D = \frac{P_H(T)}{P_0(0)} = 0.5 \cdot e^{-(\mu T / \cos \theta_B)} \cdot [1 - e^{-2\alpha T}] . \quad (38)$$

The diffraction efficiency is the product of an absorption term and a diffraction term. Because of multiple reflections and absorption in the crystal, the diffraction efficiency is < 0.5 in the Laue geometry.

2.4 Optical Properties of Materials: Reflection and Refraction

When a stream of photons encounters a medium with a change in the index of refraction some photons are reflected and some are refracted into the medium. The complex index of refraction of a material is written

$$n = 1 - \delta + i\beta , \quad (39)$$

the parameters δ and β are called the refractive index decrement and the absorption index, respectively.

At gamma ray energies the real part of the refractive index is very close to unity, its behavior can be qualitatively understood in an atomistic picture: The electrons of the material are brought to a forced oscillation by incident electromagnetic radiation, the atoms become dipoles, creating a dipole moment P per unit volume. At gamma-ray energies, where frequencies ω are much higher than the last resonance ω_0 , (the hardest absorption in the X-ray band), the atomic electrons behave as if they were free. Their displacement is 180° out of phase with the driving force of the electric field (this is, the dipole is lagging by π). The vectors of E and j (the current density) are $\pi/2$ out of phase, and the acceleration of the electrons is in phase with the electric field vector E . However, the oscillators do not change the wavelength: Although the system takes up power (jE) every second quart of a period, it returns the same energy during the following quarter period. For the case of quasi-free oscillation ($\omega \gg \omega_0$), the dipole-moments of the oscillators are always opposed to the direction of the electric field vector. With the dielectric polarization of a material $P = (\varepsilon - 1)\varepsilon_0 E$, the dielectric constant ε becomes < 1 , and hence the refractive index $n < 1$ (Maxwell relation $n = c/v = \sqrt{\varepsilon}$).

In the high-frequency limit of scattering an electromagnetic wave, the real part of the refractive index $1 - \delta$ can be estimated using the plasma frequency ω_p of the material. The plasma frequency ω_p is a function of the average electron density n_e of the material, which in turn depends on the atomic number Z , the atomic weight A , and the mass density ρ : $\omega_p \sim (\rho Z/uA)^{1/2}$. For a frequency ω , the refractive index decrement is expressed by

$$\delta = \omega_p^2 / \omega^2, \quad (40)$$

$$= \frac{r_0}{2\pi} \frac{\rho}{u} \frac{Z}{A} \left(\frac{hc}{E} \right)^2. \quad (41)$$

Here r_0 is the classical electron radius, u is the mass corresponding to 1 a.m.u., and E is the photon energy. Some representative values for δ and β are shown in Table 4 for 10 keV photons [78]. Note that the real part of n is ever so slightly less than one.

Table 4. Refractive properties of selected materials for 10 keV photons

Element	Z	δ	β	θ_c
C (diamond)	6	$4.6 \cdot 10^{-6}$	$4.5 \cdot 10^{-9}$	0.173°
Si	14	$4.9 \cdot 10^{-6}$	$7.4 \cdot 10^{-8}$	0.180°
Cu	29	$1.6 \cdot 10^{-5}$	$1.9 \cdot 10^{-6}$	0.326°
Au	79	$3.0 \cdot 10^{-5}$	$2.2 \cdot 10^{-6}$	0.443°

Finally it should be mentioned that the macroscopic constants δ and β are related to the dispersive part of the atomic scattering factor $f(0)$ from the microscopic theory (see Sect. 2.3, (24) ff) – the complex atomic scattering factor for the forward scattering direction is $f(0) = f + if'$

$$\delta = \frac{r_0 \rho \lambda^2}{2\pi v A} f, \quad (42)$$

$$\beta = \frac{r_0 \rho \lambda^2}{2\pi v A} f'. \quad (43)$$

Since the value of f equals the number of electrons Z in the atom for forward scattering ($\sin \theta/\lambda \rightarrow 0$), (41) and (42) are identical ($\lambda = hc/E$). The imaginary part β describes the absorption through a layer of thickness t as given by Beer's law $I(t) = I_0 e^{-\mu t}$ for the atomic cross section ((1) ff)

$$\mu = 2r_0 \lambda f'. \quad (44)$$

Total External Reflection

Let us consider a beam incident on a material with index $n < 1$ for gamma ray photons, producing a reflected and a transmitted beam. The incident,

transmitted and reflected beam form angles θ_i , θ_t , θ_r , with respect to the boundary of the material. Here, Snell's law takes the form $\cos \theta_i = n \cos \theta_t$. As the angle θ_i decreases, the transmitted beam approaches tangency with the boundary, and as it does more and more of the flux appears in the reflected beam. For $\theta_t = 0^\circ$, all the incoming energy is reflected back into the incident medium in a process known as total external reflection. This is analog to total internal reflection of visible light in a prism. Total external reflection occurs for angles equal or smaller than the *critical angle* θ_c , Snell's law reduces to

$$\cos \theta_c = n . \quad (45)$$

The critical angle is expressed by the refractive index decrement, with (41) it becomes

$$\theta_c = (2\delta)^{1/2} , \quad (46)$$

$$= \left(\frac{r_0}{\pi} \frac{\rho}{u} \frac{Z}{A} \right)^{1/2} \left(\frac{hc}{E} , \right) \quad (47)$$

Total external reflection takes place at larger incident angles for high Z materials and low photon energies (see Table 4). At incidence angles larger than θ_c , reflectivity drops steeply with increasing angle.

Refraction

Focusing gamma-ray instruments using refractive optics have recently been proposed by Skinner [79] (see Phase Fresnel Lenses, Sect. 4.3). The underlying principle is a combination of diffraction and refraction: a gamma-ray beam going through a certain material thickness experiences a phase shift. The material thickness necessary to produce a phase shift of 2π can be derived from the refractive index decrement

$$t_{2\pi} = \frac{\lambda}{\delta} \approx 6 \left(\frac{\rho}{\text{g cm}^{-3}} \right)^{-1} \left(\frac{E}{1 \text{ MeV}} \right) \text{ mm} . \quad (48)$$

Figure 11 shows the thickness t_π as a function of energy, along with the corresponding absorption loss in a layer of thickness t_π for Nickel and Gold. For any material for which Z is not unnecessarily high, the loss is no more than a few per cent over a wide range of gamma ray energies.

2.5 Pair Production

For photon energies exceeding a certain threshold, the production of an electron-positron pair can take place in the field of a nucleus or of an electron. The incident gamma-ray is annihilated, the excess of the photon energy above the threshold for pair production E_{th} being imparted as kinetic energy to the e^+e^- pair.

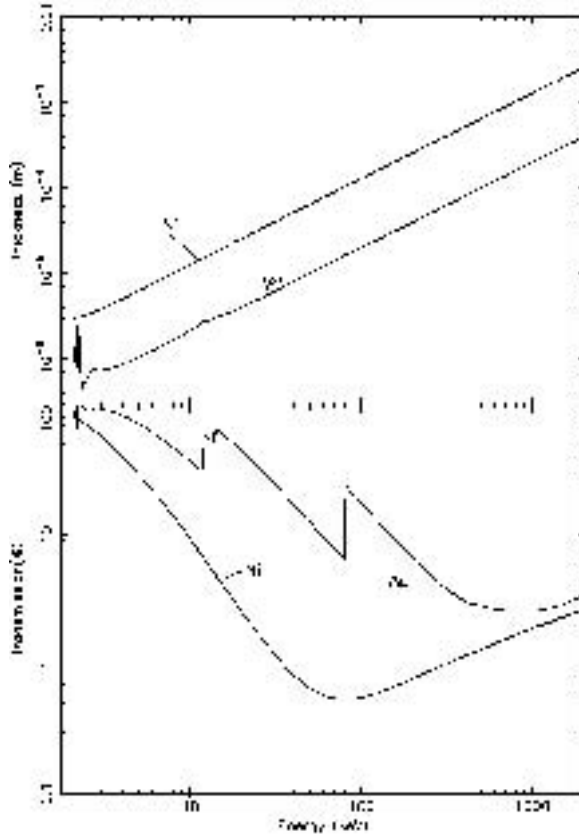


Fig. 11. Thickness t_π producing a phase shift of π in Ni and Au (*top*) and corresponding absorption losses (*bottom*) from [79]

$$E_{\text{th}} = 2m_0c^2(1 + m_0/M) , \quad (49)$$

where m_0 is the rest-mass of the electron, M is the mass of the Coulomb charge. For protons or nuclei, $E_{\text{th}} = 1.022$ MeV; for the weaker field of electrons, pair production is always less probable than for protons and $E_{\text{th}} = 2.044$ MeV.

Pair production cannot occur without the Coulomb field of a charged particle as partner: in the system “photon - e^+e^- pair” alone, conservation of energy and momentum cannot be satisfied simultaneously (the momentum of the photon is $p = E/c$ and always higher than $2m_0c$, this is, the pairs would have to be faster than c).

There is no simple closed expression for the pair cross-section. An order of magnitude figure can be obtained by estimating the momentum transfer to the nucleus with atomic number Z if an electron or a positron is removed

from a distance r with a velocity of $\sim c$. With $F = Ze^2/(4\pi\epsilon_0 r^2)$, the force on the electron or positron, the momentum transfer to the nucleus becomes

$$p = \int_r^\infty \frac{Ze^2}{4\pi\epsilon_0 r^2} dt = \int_r^\infty \frac{Ze^2}{4\pi\epsilon_0 r^2} \frac{dr}{c} = \frac{Ze^2}{4\pi\epsilon_0 rc} . \quad (50)$$

In order to take up a momentum of the order of mc required by the production of the pair, the distance r has to be $r \approx Ze^2/(4\pi\epsilon_0 mc^2)$. This is the classical electron radius r_0 times Z !

For the transformation of the photon into an e^+e^- pair, the probability is about equal to the fine-structure constant α . The cross section for pair production then becomes about

$$a^\kappa \approx \alpha Z^2 r_0^2 . \quad (51)$$

The cross section a^κ increases with increasing energy. For the relatively low energy photons, $m_0 c^2 \ll E \ll 137 m_0 c^2 Z^{-1/3}$, pair-production occurs in the vicinity of the nucleus, screening of its Coulomb field by the electrons is negligible. The cross section in this regime has been given by Bethe and Heitler [80] as

$$a^\kappa = 4\alpha Z^2 r_0^2 \left[\frac{7}{9} \ln \left(\frac{2h\nu}{m_0 c^2} \right) - \frac{109}{54} \right] . \quad (52)$$

At very high energies, $E \gg 137 m_0 c^2 Z^{-1/3}$, the screening of the nuclear Coulomb field by the electrons can be considered total, the cross section becomes

$$a^\kappa = 4\alpha Z^2 r_0^2 \left[\frac{7}{9} \ln \left(\frac{191}{Z^{1/3}} \right) - \frac{1}{54} \right] . \quad (53)$$

The idealized energy loss spectrum for gamma-rays of energy $h\nu$ that interact via pair production is sketched in Fig. 12: it shows a single peak at an energy $h\nu - 2m_0 c^2$ – the double escape peak. While the kinetic energy of the electron and the positron is transferred to the detector medium, the two 511 keV photons resulting from the annihilation of the positron are supposed to escape from the detector.

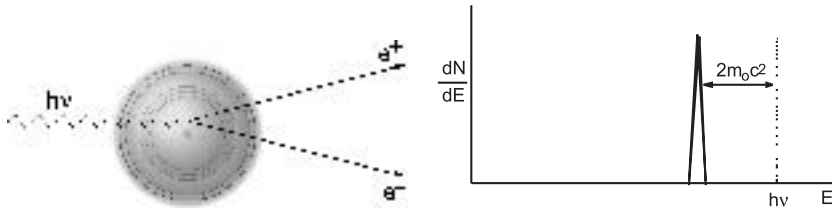


Fig. 12. The energy loss spectrum for monoenergetic gamma-rays interacting via pair production (in the field of the atomic nucleus) when the photons of the subsequent e^-e^+ annihilation escape

2.6 The Spectral Signatures of Energy Loss Processes

Detector Response

The response of a gamma-ray detector to monochromatic radiation is schematically illustrated in Fig. 13a – the principal energy-loss processes (simple photoelectric absorption, Compton scattering, and pair production) are reflected in idealized spectral features (Fig. 13b,c). In addition to the basic energy-loss processes and their combinations, a gamma-ray spectrum may show the signatures of secondary energy-loss processes, such as *annihilation radiation*, *electron escape*, *bremsstrahlung escape*, and *fluorescence radiation*.

For detectors whose size is large with respect to the mean free path of the secondary radiation produced by the incident photons ($h\nu$), no energy escapes from the detector. The sum of the primary and subsequent energy deposits (1, 2, 5) produce a peak at the energy $h\nu$ which will dominate the spectrum; this peak called is the *photopeak*, or (better) the *full energy peak*.

In real detectors, secondary radiation generated by the incident photons may escape and produce characteristic features. If a Compton scattered gamma-ray leaves the detector without further interaction (3), its energy $h\nu'$ (see (12)) is lost and the energy deposit is $E_c-(h\nu, \theta)$. This *Compton continuum* that spans from the low energy threshold up to the Compton edge $E|_{\theta=\pi}$.

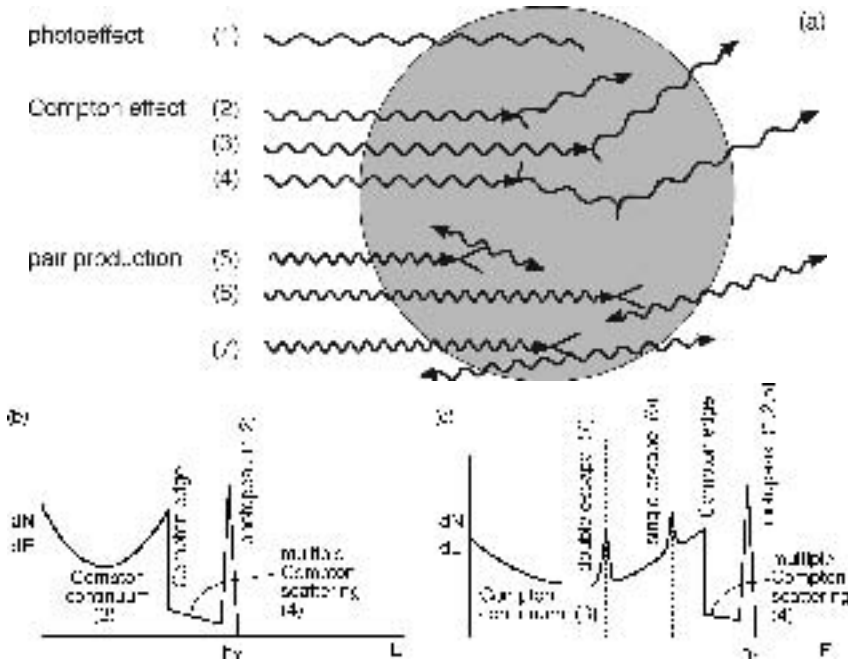


Fig. 13. The spectral signatures of energy loss processes resulting from monochromatic photons ($h\nu$) impinging on a gamma-ray detector (see text)

The interval between the full energy peak and the Compton edge is partially filled up by a continuum produced by *multiple Compton scattering* (4).

For sufficiently high incident gamma-ray energies $h\nu \gg m_0c^2$, pair production becomes possible and the subsequent e^+e^- annihilation produces two 511 keV photons that may escape from the detector (Fig. 13c). A single (6) and a double (7) escape peak appear in the spectrum at energies $E = h\nu - m_0c^2$, and $E = h\nu - 2m_0c^2$, according to whether one or both annihilation photons are lost.

The escape of secondary electrons, also called *electron leakage*, can become important for detector sizes small with respect to the mean range of the secondary electrons. Also, energy is lost close to the surface of the detector by bremsstrahlung from secondary electrons, and by characteristic X-rays resulting from electron rearrangement (*X-ray escape peaks*). As a consequence of all these secondary processes, events are measured at energies $E' < h\nu$ and the full energy peak efficiency (ε_{fep}) decreases.

Passive Material

In an actual instrument, the detector is always surrounded by materials – shields, aperture systems, front-end electronics etc. – that will interact with the incoming gamma-rays. Beyond the background radiation emitted by these materials (in space, background is mostly induced by Cosmic Rays), their secondary radiation produces various types of spectral features. The spectral signatures of the principal energy-loss processes (simple photoelectric absorption, Compton scattering, and pair production) are illustrated in Fig. 14.

Photons of energy $h\nu$ that pass through the instruments aperture and interact with the passive materials surrounding the detector through Compton Scattering (II) can reach the detector with energy $h\nu'$ (16). For a large variety of input spectra and detector geometry (backscatter-angles), these events result in a continuum peaking between 150–250 keV. The so called *backscatter peak* remains between 170 keV ($m_0c^2/3$) and 255 keV ($m_0c^2/2$) for incident photon energies between 511 keV and ∞ , respectively.

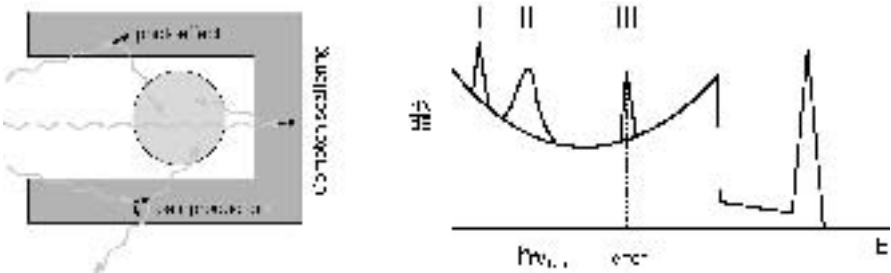


Fig. 14. Spectral signatures of the principal energy-loss processes in the passive material surrounding detector

Escaping X- and gamma-rays, produced by fast electrons and annihilating positrons in the passive materials result in characteristic X-ray lines and the 511 keV annihilation line (I and III).

Besides of efficiently reducing various background components, an *active shield* and anticoincidence electronics help in suppressing the backscatter peak (“anti-Compton” shield). The pair-production component of the omnipresent 511 keV line is also suppressed with an active shield. Characteristic X-ray lines may be avoided by shielding the detector by a *graded shield*: an outer layer of high Z-material efficiently shields against exterior background. A second inner shield of a lower Z-material attenuates the characteristic X-rays produced in the outer layer while only emitting weakly its low energy characteristic X-rays.

2.7 Characterizing the Detector Response

Spectra

The spectral information of the raw data is displayed as a *differential pulse height spectrum* $S_T = dn/dH[\text{counts/channel}]$: the number of counts in a channel versus the channel-number (ADC) corresponding to the pulse amplitude H . For astrophysical interpretation, this raw spectrum of energy losses has to be deconvolved (or model fitted) with the *spectral response matrix* M in order to obtain the *photon spectrum* $S_e = dn/dE$ e.g. $[\text{photons}/\text{cm}^{-2} \text{ s}^{-1} \text{ MeV}^{-1}]$. The spectral response matrix M is defined as

$$S_T = M * S_E . \quad (54)$$

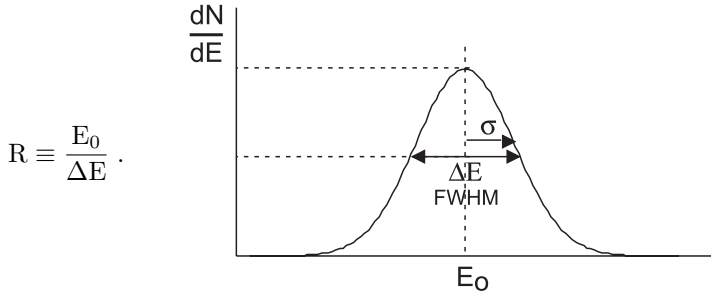
Efficiency

The principal characteristics of a detector for spectroscopy are its efficiency and its energy resolution. Other important features are its spatial resolution, timing resolution, dead time, and possibly polarimetric capabilities. The intrinsic efficiency relates the total number of counts detected with the incident counts; the relevant figure of merit for spectroscopy is the full energy peak efficiency which is defined as the ratio of photons detected within the gamma-ray line divided by the number of monochromatic photons incident on the detector

$$\begin{aligned} \varepsilon_{\text{int}} &= \frac{n_{\text{counts}} \text{ (detected over entire energy range)}}{n_{\text{photons}} \text{ (incident on detector)}} , \\ \varepsilon_{\text{fep}} &= \frac{n_{\text{counts}}(\lambda) \text{ (detected in line)}}{n_{\text{photons}}(\lambda) \text{ (incident on detector)}} . \end{aligned}$$

Energy Resolution

The width of a narrow gamma-ray line ΔE as observed in a detector is usually defined by its full width half maximum (FWHM), assuming a Gaussian centered on the line energy E_0 . The energy resolution R of the detector at photon energy E_0 is then defined as



This is the definition as used in most other domains of spectroscopy – here, better resolution is reflected in a higher R . In gamma-ray astronomy, the resolution is often expressed in percents of the inverse quantity $\Delta E/E_0$. For example, scintillator detectors have resolutions in the range $\Delta E/E_0 = 5\text{--}15\%$, Germanium detectors have resolutions $\Delta E/E_0$ of the order of 0.2% . Informally, experimentalists often speak of the resolution as the width ΔE (in keV) at a typical energy of a calibration source, e.g. at 1.33 MeV (^{60}Co line). In a Germanium detector, for example, ΔE is typically between 1.7 keV and 2.5 keV .

The resolution depends on a variety of fluctuations in the detector response, the most important are *statistical noise* in the formation of the charge carriers, variations in the *efficiency* of the signal collection, and *noise* in the detector and its electronics. Independently of the detector type, statistical noise will always be present and often is the dominant component. If the formation of charge carriers is assumed to be a Poisson process, the standard deviation σ is proportional to the square root of the variance N , with N the number of charge carriers created on the average. In this case, the line-width becomes $\Delta E_P = 2.35E_0/\sqrt{N}$ and the limiting resolution due to statistical fluctuations is

$$R_{PS} = \frac{\sqrt{N}}{2.35}, \quad (55)$$

since the width ΔE (FWHM) corresponds to 2.35σ for a Gaussian shape.

Yet, for certain types of detectors, the *measured energy resolution* is much lower than as calculated by Poisson statistics (in Ge detectors, factors of 3–4 are not uncommon). If the charge carriers are not formed independently, the process *cannot* be described by Poisson statistics (in the extreme case where the incident energy is transformed with constant efficiency into charge carriers which are then *all* collected, the signal would show no statistical fluctuations at all). The ratio between observed variance and the one predicted from

Poisson statistics is called the Fano factor

$$F = \frac{\text{observed variance}}{\text{Poisson predicted variance (N)}} \quad (0 < F \leq 1) .$$

While scintillators have Fano factors close to unity, in semiconductor diode detectors F may be as small as 0.06–0.16. The statistical limit to the resolution now becomes

$$R_S = \frac{1}{2.35} \sqrt{\frac{N}{F}} . \quad (56)$$

3 Detectors

The essence of any detector system is determined by the way gamma-rays transfer energy to the matter they traverse. Unlike charged particles (α -, β -radiations) that continuously interact with the matter through Coulomb force, γ -rays do not exchange energy with matter unless they undergo a “catastrophic interaction” (Fig. 15). Gamma-rays typically have a mean free path of the order of several centimeters; electrons only have a characteristic pathlength of a millimeter in common detector materials.

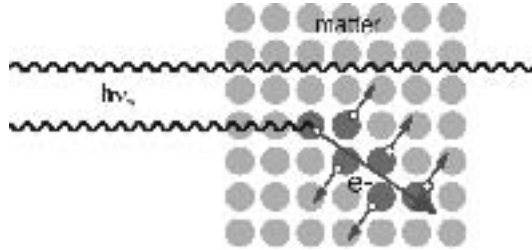


Fig. 15. Gamma-rays and matter

- no interaction
- catastrophic interaction: mainly photo-, Compton-, pair-effect
- e^- interacts continuously (Coulomb force/ionization)

The detection of a gamma-ray photon can be generalized as a three step process 1) *Conversion*: In all cases of practical interest – photoelectric effect, Compton effect, and e^+e^- pair-production – the full or partial energy of the photon is transferred to secondary electrons. 2) *Ionization of detector medium* by secondary electrons creates a large number of charge carriers and excited atoms or molecules along their path. 3) *Signal collection*: While certain types of detectors directly collect the charge carriers created by the fast electrons, others rely on scintillation light from recombination of electrons with ions or on the small temperature increase (or phonons) in the absorber material.

According to the way the signal is collected, different classes of detectors can be identified. In *gas-filled counters* and *semiconductor detectors* (Sects. 3.1 and 3.3), an electric field causes the charge carriers created to migrate and be collected. In *scintillators* (Sect. 3.2), the emission of visible and UV light by deexcitation atoms is favored. The light-signal is reconverted into a electric-current by a photomultiplier tube or photodiode before. In *phonon detectors*, the temperature rise is transformed into a current by a termistor, for example.

Yet, the above detector categories, particularly the division between gas-filled detectors and scintillators, actually reflect the historic development of radiation-detectors rather than on the physical processes of the signal collection. Just as certain gas-filled detectors can be operated as scintillators, there are liquid detectors relying on the drift of ions pairs.

The choice of an optimal detector is driven by the main requirements on gamma-ray spectroscopy (Sect. 1.3) which are – first of all – sensitivity and energy resolution. Complete conversion of the gamma-ray energy within the detector is thus crucial for almost all detectors – this translates into the need for high density high Z materials. An exception is the use of low Z-materials preferred in the scatter-detectors of Compton telescopes; here imaging and spectroscopy rely on two or more partial energy deposits.

In classical spectrometers, relying on total energy absorption, the principal energy range of operation will influence the choice of the detecting medium. Since gamma-ray attenuation scales with $\sim \rho Z^{-7/2}$ in the photo-absorption range, Z is the critical parameter for the stopping power of a material below the Photoelectric/Compton cross over energy. In the Compton dominated region above the cross over, the mass attenuation coefficient (μ/ρ) is nearly independent of the Z of the material, making the density a prime parameter.

3.1 Gas-filled Detectors

Gamma-rays passing through the detector transfer energy to one or several electrons of the fill-gas or the chamber-wall. Along the path of the energetic electron through the gas, ionized molecules and low energy electrons are created. The positive ions and free electrons are called *ion pairs*. An electric field causes the ions to drift towards the cathode (e.g. the cylindrical wall of the gas chamber), the electrons towards the anode.

The detector can be thought as a capacitor into which a charge is deposited. The signal is a voltage drop across the bias resistor V_R . For a time constant of the circuit RC (see Fig. 16) sufficiently long with respect to the charge collection time, a signal pulse is produced with an amplitude proportional to the energy loss within the chamber. As in all detectors, the preamplifier is best located close to the detector in order to reduce capacitance of the leads – this allows for high gain and fast rise time. The pre-amplified

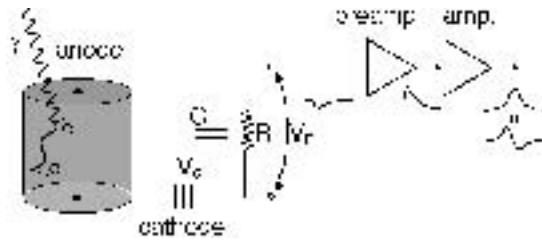


Fig. 16. The components and simplified circuitry of a gas-filled detector. C represents the capacitance of the chamber and any additional capacity. V_R is the output pulse

signal goes through further amplification (Pulse Height Analysis, PHA) before being converted to a number by an Analog-to-Digital Converter (ADC).

The various types of gas-filled detectors – ionization chambers, proportional counters, and Geiger counters – correspond to different operation regions of such detectors, as illustrated in Fig. 17. As long as the applied voltage is very low, the field strength is insufficient to prevent recombination of electrons and ions – the charge collected is thus lower than the sum of the ion pairs created by the fast electron. With increasing electric field, recombination of electrons and ions becomes less likely and is eventually overcome. The region of *ion saturation* is reached when all charge carriers created by direct ionization are collected ($\sim 10^4$ V/m for a typical gas at 1 atm). This domain is

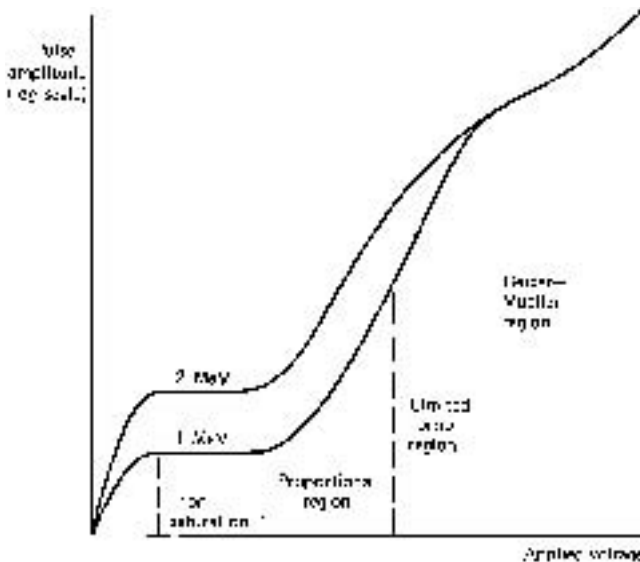


Fig. 17. The regions of operation for gas-filled detectors, explanations see text (from Knoll [81])

the normal mode of operation for *ionization chambers* that will be discussed below.

As the voltage is further increased, *gas multiplication* sets in above a threshold of about 10^6 V/m (typical gas at 1 atm): Between collisions with the gas molecules, the secondary electrons are now sufficiently accelerated to acquire kinetic energies greater than the ionization energy of the gas. Hence, additional ion pairs are created and new electrons are accelerated that result in a cascade amplifying the pulse. In the region of *true proportionality*, the gas multiplication process is linear, this is, the amplified charge is proportional to the directly ionized charge. This is the region of operation for *proportional counters* discussed below.

If the voltage is further increased, nonlinear effects begin to degrade the spectroscopic properties of the detector (*limited proportionality*). The cloud of ions produced in the gas multiplication only drifts slowly towards the cathode, meanwhile, this space charge alters the shape of the field which governs the multiplication process. As a result, the pulse amplitude increases nonlinearly with increasing initial energy deposit.

At even higher voltage, the *Geiger-Müller* region of operation is reached: with the enhanced intensity of each gas multiplication avalanche, the probability for secondary avalanches triggered by UV photons becomes very high. The number of UV photons, emitted by the decay of the excited states produced by electron collisions with the fill gas, is now above “criticality”: the photons either directly ionize gas molecules or strike the cathode wall, liberating additional electrons that quickly produce additional avalanches at sites removed from the original. The multiplication can reach factors of typically 10^6 to 10^8 . Yet, the strong space charge that is created by the ions will eventually reduce the electric field below the threshold for gas multiplication – the process is therefore self-limiting. The output pulse of the Geiger-Müller discharge has the same amplitude regardless of the gamma-ray energy loss.

Below, the types of gas-filled detectors suitable for gamma-ray spectroscopy are discussed: both ion chambers and proportional counters are based on conversion of the photon with the fill medium (whereas in typical Geiger-Müller counters, the gamma rays interact primarily with the wall of the tube) and produce output pulses proportional to the initial energy deposit.

Proportional Counters

In a proportional counter the signal-to-noise ratio is improved with respect to simple ionization chambers because internal gas multiplication amplifies the output signal by a factor of more than 10^3 . Above a threshold of the order of 10^6 V/m for typical gases at atmospheric pressure, the secondary electrons become sufficiently energetic to ionize the gas and create more free electrons – the ensuing cascade is called Townsend avalanche. The increase of free electrons (density n_e) per unit pathlength dx is

$$\frac{dn_e}{n_e} = \alpha dx, \quad (57)$$

where α is the Townsend coefficient which depends on the field strength. For a spatially constant field, the solution of the Townsend (57) predicts an exponential growth of the electron density $n(x) = n_e(0)e^{\alpha x}$, yet, in most proportional counters the geometry is closer to the one shown in Fig. 16. For cylindrical geometries, the electric field at a distance r from the anode wire is

$$E(r) = \frac{V}{r \ln(b/a)}, \quad (58)$$

here V is the potential applied to the anode with respect to the cathode, a is the anode wire radius, and b the radius of the inner wall of the cathode. Usually only very close to the thin anode wire (typically a few wire radii) is the field strong enough to produce an avalanche. Hence, most of the ion pairs are created outside of the very small part of the detector volume where E is above the threshold. Electrons first drift to this region before gas multiplication sets in. The avalanche terminates when all electrons are collected. Each electron therefore undergoes the same amplification and the gas multiplication factor (also termed gas gain) remains constant, independently of the initial interaction site.

The choice of the fill gas is driven by the requirement of efficiently stopping gamma-ray photons within the active volume of the detector and providing the best energy resolution possible. At lower energies, good efficiencies are obtained with modest gas pressures (<5 atm) and with various fill gases – commonly used are noble gases (Ne, Ar, Xe) and hydrocarbon gases such as methane or ethylene. At higher energies, above say 100 keV, heavier fill gases like krypton or xenon and high gas pressures are preferable to achieve reasonable efficiencies. The efficiency of a xenon filled gas detector as a function of energy is shown in Fig. 18a [82].

The collisions in the gas multiplication process not only ionize the gas, but a part of the kinetic energy goes into the excitation of the gas molecules. Consequently the counting statistics are reduced and so is the energy resolution. Furthermore, when these excited molecules decay to the ground state, the emitted photons (visible or UV) can produce additional electrons by ionization of the gas or by photoelectric interaction with the detector housing. In xenon, the W -value (the average energy to form an ion-pair) is 21.5 eV while the ionization energy is only 12 eV – nearly half of the energy goes into excited states of the Xe atoms! To avoid loss of proportionality and spurious pulses caused by these deexcitation photons, most detectors contain a small quantities of a stabilizing gas component called *quench gas*. The complex molecule of the quench gas is selected to have a lower ionization energy than that of the fill gas. Upon collision, the fill gas ion gives up energy to the quench molecule rather than losing its energy by radiative emission. The use of *Penning mixtures* as quench gases can even help to improving the energy

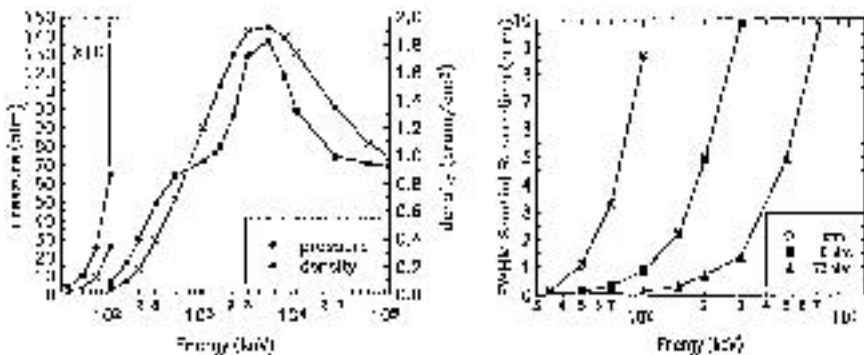


Fig. 18. (a) The pressure (left axis) and density (right axis) needed to give 50% detection efficiency in 10 cm of xenon. (b) Upper limit for the spatial resolution in a xenon gas detector due to the range of fast electrons (from [82])

resolution. In the *Penning effect*, the ionization potential of the quench gas is matched to the metastable energy of the principal fill gas which is resonantly deexcited while the quench gas is ionized, increasing the number of electrons. A comprehensive study of quench gases for xenon detectors is given in [83].

The statistical limit to the energy resolution of a proportional counter is estimated ([81], p. 178)

$$R = \frac{1}{2.35} \sqrt{\frac{E_0}{W(F+f)}} , \quad (59)$$

with F , the Fano factor (≈ 0.17 for Xe), and f the multiplication variance characterizing the avalanche statistics ($f \approx 0.6 - 0.8$ for the multiplication factors/electric field strengths typical in proportional counters). At $E_0 = 100$ keV, a gas detector filled with pure Xenon will therefore have an upper limit to the energy resolution of $R \approx 30$ ($\Delta E/E \approx 3.3\%$). In practice the resolution is limited by preamplifier noise, acoustical noise and certain physical processes that become particularly important at high pressures (that is, particularly in detectors optimized for gamma-ray energies).

Position-sensitive Proportional Counters

Most telescope systems for nuclear astrophysics (Sects. 4.1 and 4.2) require large area detectors with spatial resolution – not only for imaging of the gamma-ray sky, but also as a means to reduce background in order to achieve good sensitivity. *Multi-wire proportional counters* (MWPC) consist of a grid of anode wires between two large flat plates or grids serving as cathodes. The main design characteristics of an MWPC sensitive up to energies above say one hundred keV are schematically sketched in Fig. 19. For optimal energy resolution, recombination effects should be minimal. This requirement

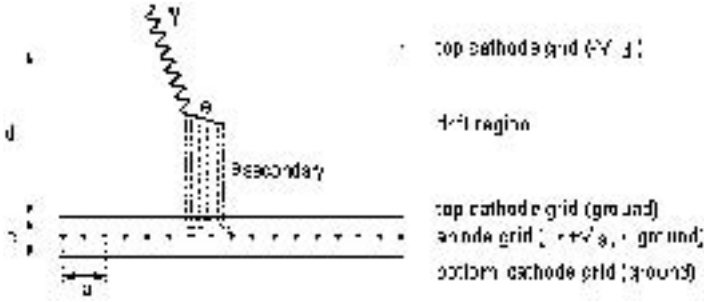


Fig. 19. The principal design characteristics of a multi-wire proportional counter

is best satisfied in a field hyperbolically decreasing with anode distance. The optimal field geometry in the multiplication region around the anode wires is for a $\approx c$: the distance between anodes optimally is about equal to the distance between the cathode planes. Between two cathodes wires, a ground wire prevents warping of the anode grid. The volume between the cathodes is consequently dimensioned by the desired spatial resolution – i.e. typically a few mm. Since such a small detector volume would result in very low detection efficiencies, conversion of the incident photons has to take place in the *drift region*; the secondary electrons are first drifted to the top cathode before entering into the region of gas multiplication. According to the gas pressure, the thickness d of the drift region might be several cm, while field strengths of $V_d \approx 250 [\text{Vbar}^{-1}\text{cm}^{-1}]$ are typical in xenon.

Typical distances between anode wires is of the order of a millimeters, but the spatial resolution is limited by the range of the fast electrons in the fill gas. The interaction is localized, in one dimension, by identifying the anode wire showing the signal. The perpendicular coordinate can either be determined by the charge division method, by the rise time method or by using the image charge on the cathode plane, with cathode wires running perpendicular to the anode wires. A charge division circuit uses two amplifiers on either end of the anode wire that has a significant resistance per unit length; the ratio of the charges collected by two amplifiers is proportional to the position of the interaction. The rise time method is based on the rise time difference between the signals from the preamplifiers placed on either side of the anode wire. For a review of position sensitive MWPC in X- and gamma-ray astronomy see e.g. Ubertini [84].

Microstrip gas counters (MSGC) offer several advantages with respect to MWPCs and also have potential application as focal instruments for concentrating telescopes (Sect. 4.3). Microstrip gas detectors reproduce the field structure of multiwire chambers; they use an electrode structure made of a sequence of alternating thin anode and cathode strips on an insulating or partially insulating support. The classical MSGC is built on a glass support a few hundred μ thick, and the drift volume is defined by a drift cathode

situated at a typical distance of 2–6 mm from the plane of the strips. The typical pitch (the repetition sequence) is 100–200 μ . The anodes and cathodes are deposited on the support using techniques from microelectronics, e.g. planar technology. The benefits of MSGC's are their ease of construction, uniform response, reduced operating voltage for a given gain, reduced charge saturation at high gain, better spatial resolution, better energy resolution, and higher efficiency for the detection of fluorescent pairs.

The detectors of INTEGRAL's JEM-X telescope consist of two identical, high pressure, imaging microstrip gas chambers, each with a collecting area of 500 cm². The gas is a mixture of xenon (90%) and methane (10%) at 1.5 bar pressure. Microstrips are patterned in a 0.15 μ m thick Au layer deposited on a semiconducting substrate. The 27 cm wide pattern and dimensions are shown schematically in Fig. 20. The electrode structure is built on a glass support glued to a titanium frame. While the detector entrance window is made from Beryllium and only 250 μ m thick, the detector box is made of stainless steel with a minimum thickness of 2 mm. This provides good background suppression in the primary energy range below 35 keV and shields the internal electronics from radiation damage. Charged particles can be identified and rejected based on longer pulse rise times, veto signals, or deposition of charge over several strips.

Laboratory measurements with Xe(90%)/CH₄(10%) at 1.5 bar have demonstrated a detector energy resolution of $R = 2.5\sqrt{E[\text{keV}]}$ – i.e. a resolution $\Delta E/E_0$ of 16% at 6 keV and 6.7% at 35 keV.

Ion Chambers

In ion chambers, the simplest type of all gas-filled detectors, all the charges created by direct ionization are collected through the application of an electric field. If the energetic electrons generated by the photon deposit all their kinetic energy within the gas, the number of ion pairs generated is proportional to the incident gamma-ray energy. The number of pairs formed can be estimated by dividing the electrons energy deposit by the average energy to form an ion-pair. This energy, called the W-value, is always higher than the ionization energy of the least bound electron shell for the gases used in detectors. While the ionization energy in such gases 10–20 eV, the W-value is typically 25–40 eV per ion pair (e.g. 21.9 eV/ion pair for xenon, 26.4 eV/ion pair for argon, and 33.8 eV/ion pair for air).

In xenon, a photon losing all its energy $E_0 = 1$ MeV will therefore create about $n_0 = 45000$ ion pairs. With a Fano factor $F = 0.17$ in xenon, and since no gas multiplication takes place, the expression (57) for the upper limit of the energy resolution due to counting statistics reduces to

$$R = \frac{1}{2.35} \sqrt{\frac{E_0}{WF}} \approx 200. \quad (60)$$

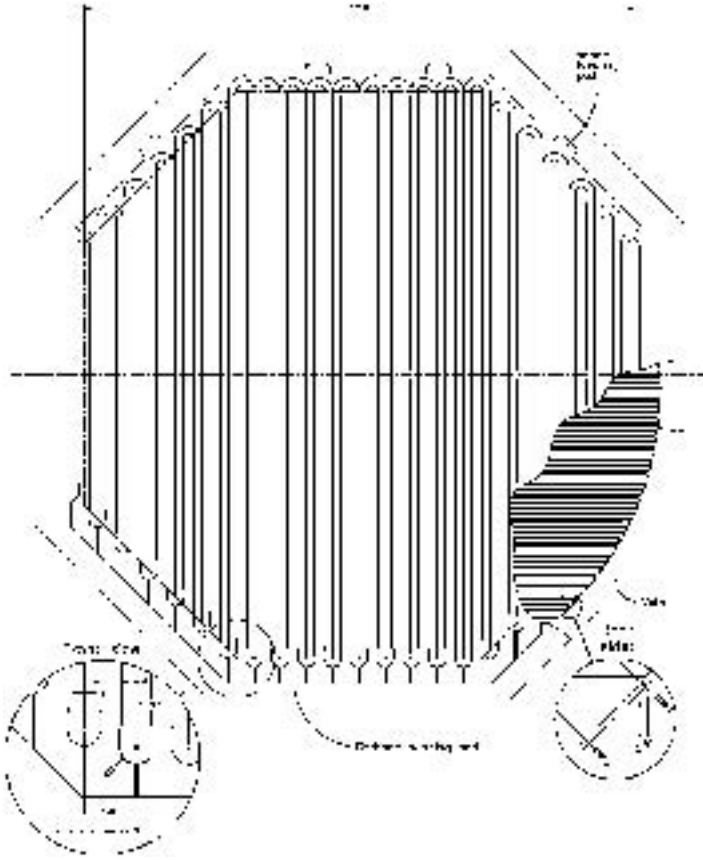


Fig. 20. Microstrips pattern of INTEGRALs X-ray monitor JEM-X [86]

This theoretical limit ($\Delta E/E_0 \approx 0.5\%$) corresponds to the outstanding resolution achieved today only by semiconductor detectors. Various experiments have shown that the performance is limited by the electron transport which is extremely sensitive to impurities in the compressed xenon. Furthermore, given the capacitance C of a typical ion chamber (≈ 100 pF), the maximum pulse amplitude is given by

$$V_{\max} = \frac{n_0 e}{C} = \frac{E_0 \cdot e}{W \cdot C} \approx 5 \cdot 10^{-5} \text{ V} . \quad (61)$$

While detectable, this is a weak signal and susceptible to deterioration by the various sources of noise in the amplification chain.

Despite the experimental challenges, outstanding energy resolutions of the order of $\Delta E/E_0 \approx 2\text{--}3\%$ have been measured in ionization chambers [87–90].

In the laboratory, ionization chambers with classical cylindrical geometry (see Fig. 16), such as e.g. the large volume spectrometer [85] which is filled

with 5 liters of xenon under 35 atm pressure (density of 0.3 g/cm^3), have shown energy resolutions of $\Delta E/E_0 \approx 2\%$ at 662 keV.

A high pressure xenon ionization chamber for the observation of cosmic gamma-ray lines was flown on the MIR station [87]. The 3 liter chamber was filled with 0.6 g/cm^3 density xenon mixed with hydrogen for increasing the drift velocity of electrons. At 1 MeV, the energy resolution without electronics noise was $\Delta E/E_0 \approx 1.3\%$ and the total energy resolution was $\Delta E/E_0 \approx 2.0\%$. During the two years operation with frequent passages through the South Atlantic Anomaly (SAA), no degradation of the performance was observed.

A compact detector system sensitive from 100 keV to over 1 MeV has been built by the Brookhaven Gas Detector Group [88]. In a parallel plate detector a linear drift field of 2 kV/cm is applied. The pressure of the xenon gas is 0.55 g/cm^3 ; particular care has gone into the gas purification and filling system.

Since the detector volume is only 160 cm^3 , the full energy peak efficiency is 30% at 200 keV and 2% at 662 keV. Figure 21 shows the pulse height spectrum of a ^{137}Cs source for an optimally collimated beam: the gamma-ray 662 keV peak has a FWHM of 13.2 keV ($\Delta E/E_0 \approx 2.0\%$); the contribution of electronic noise measured with a pulser is just over 8 keV FWHM. Under more general conditions, the resolution is slightly worse – e.g. because of the finite lifetime ($\sim 5 \text{ ms}$) of the secondary electrons (although this is ~ 100 times longer than the drift time). At energies above 1 MeV, the resolution is further degraded due of the larger range of the fast electrons causing *ballistic deficit* effects (different locations within the detector cause different collection times,

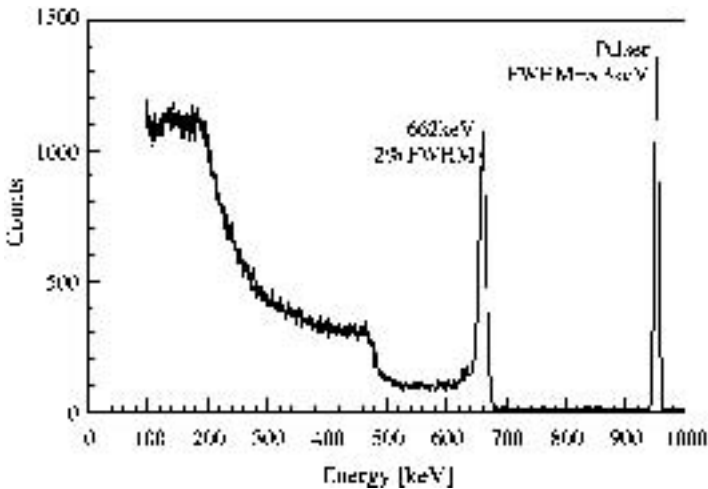


Fig. 21. Anode pulse height spectrum of collimated 662 keV photons entering a parallel plate high pressure xenon detector (0.55 g/cm^3) at right angles to the linear drift field [89]

since the amplifiers shaping time constant is fixed and finite, the amplitude of the shaped pulse might at times be less with respect to the one obtained with an infinite shaping time.)

A comprehensive study of high-pressure xenon detectors for gamma-ray spectroscopy in the energy range between 0.1–2.0 MeV has been undertaken by Bolotnikov and Ramsey [89]. Their measurements of the intrinsic energy resolution (noise subtracted) as a function of the density in cylindrical ionization chambers are shown in Fig. 22. At densities below 0.6 g/cm³ the resolution is determined mainly by electronic noise; the best energy resolutions measured were obtained for rise-time selected events: $\Delta E/E = 2.0\%$ at 662 keV and $\Delta E/E = 2.2\%$ at 511 keV. The sharp deterioration in energy resolution above 0.55 g/cm³ is poorly understood today. According to [90] it can be explained by the appearance of the first exciton band, which is formed inside a cluster of at least 10 atoms due to density fluctuations in dense Xe, introducing an additional energy loss for ionizing electrons.

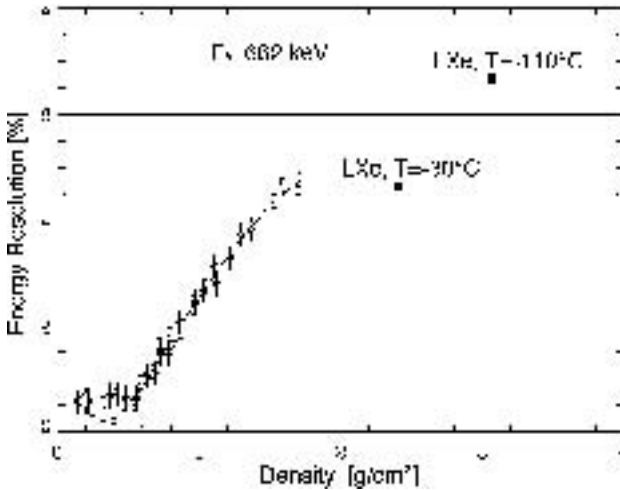


Fig. 22. Density dependencies of the intrinsic (noise subtracted) energy resolution measured for 662 keV gamma-rays in a cylindrical ionization chamber [90]. At a density of 0.5 g/cm³, the total energy resolution $\Delta E/E \approx 2.2\%$ at 662 keV

Time Projection Chambers

A promising perspective for imaging gamma-ray spectrometers – particularly advanced Compton telescopes (see Sect. 4.2) – are ionization chambers that allow localizing the position of the conversion or even track the fast electrons. A time projection chamber (TPC) measures the energy and all three spatial coordinates of every ionizing interaction in the sensitive volume.

Aprile et al. [91] propose to combine the spectroscopic properties of xenon filled ion chambers with three-dimensional localization, having demonstrated the power of event imaging with a *liquid xenon time projection chamber* of LXeGRIT [92, 93]. The balloon-borne LXeGRIT, conceived to be operated over an energy range from 200 keV to 20 MeV, contains high purity liquid xenon at a temperature of about -100 C° . In a time projection chamber, both the ionization and scintillation signals are detected in order to measure the energy and 3D position of an interaction. The fast ($<5\text{ ns}$) Xe scintillation light, detected by photomultiplier tubes, provides an event trigger. The drift of free electrons in a uniform electric field of typically 1 kV/cm , induces charge signals on a pair of orthogonal planes of parallel wires with a 3 mm pitch, before collection on four independent anodes (see Fig. 23). The X-Y coordinate information is obtained from the pattern of hits on the wires, while the energy is obtained from the amplitude of the anode signals. The Z-coordinate is determined from the drift time measurement referred to the light trigger. The drift time is also used to improve the spectral performance. After removing the dependence of the signal amplitude on the distance from the anode, an energy resolution of $\Delta E/E_0 = 10\%$ is obtained at 1 MeV scaling with $E^{-1/2}$ (the noise subtracted value is 8.8% FWHM at 1 MeV). The angular resolution is composed of two contributions. A fixed angular uncertainty of $\sim 2^\circ$, due to the spatial resolution of the interactions within the detector.

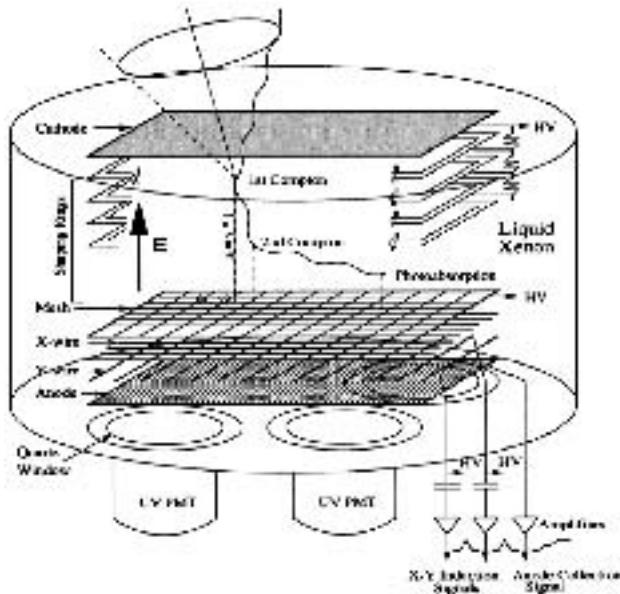


Fig. 23. Schematic of the Liquid Xenon Time Projection Chamber LXeGRIT from Aprile [81]. The sensitive area is $20 \times 20\text{ cm}^2$ and the maximum drift length is 7 cm – explanation see text

A second contribution from the uncertainty in the Compton scatter angle θ (12) which depends on the energy resolution; for small scatter angles, this contribution is about 3° , increasing for larger scatter angles ($\sim 5^\circ$ at $\theta \approx 50^\circ$).

Since the energy resolution measured in high-pressure Xe is superior to the mediocre LXeGRIT values (attributed mostly to density fluctuations associated with the formation of clusters of atoms in dense Xe – equivalent to densities higher than 0.6 g/cm^3 in high pressure Xe detectors, see Fig. 22), Aprile et al. [92] expect much better performance for gas filled time projection chambers. Besides the energy resolution, which approaches the statistical limit set by the number of charge carriers produced (see ion chambers above), the angular resolution consequently improves. Both qualities will also enhance the sensitivity of the instrument. An even more dramatic increase in sensitivity would be achieved if the Compton recoil electrons could be tracked in low-medium pressure xenon.

3.2 Scintillators

Fast electrons passing through a scintillator transfer a part of their energy to excited atomic or molecular states that quickly decay through the emission of visible or ultraviolet light. This *prompt fluorescence* is collected by photomultiplier tubes or photodiodes that convert the signal into an electric pulse.

The fundamental properties characterizing a scintillator are its scintillation efficiency (fraction of fast electron energy converted into scintillation light), the decay time of the induced luminescence, and of course its stopping power (which is related to its Z value – see Table 3). In addition, a scintillator detector should offer linear conversion of the deposited energy into scintillation light, and the medium, which must of course be transparent to the scintillation light, preferably has a refraction index close to that of glass ($n \approx 1.5$) in order to favor light collection by the photomultiplier tubes (PMT).

Two classes of materials partly fulfill of the above requirements – inorganic crystals, with their high scintillation efficiency and high Z-value, and organic liquids and plastics with their short light decay times.

Organic Scintillators

In an organic scintillator, fluorescence originates from transitions in the energy levels of single molecules, consequently they are independent on the physical state of the molecule and organic scintillators take many different forms. Certain organics, such as anthracene ($\text{C}_{14}\text{H}_{10}$), are used in solid polycrystalline detectors, as a vapor, or as compound in a solution (see Table 5).

Organic scintillators contain aromatic compounds consisting of planar molecules made of benzenoid rings. Some of the energy deposited in the

Table 5. Properties of certain organic scintillators

	Scintillator	Density [g/cm ³]	Refractive Index n	Relative Light Output % Anthracene	Decay Time [ns]	λ of Max Emission [nm]
Crystal	Anthracene	1.25	1.62	100	30	447
	Stylbene	1.16	1.626	50	4.5	410
Plastic	NE-102	1.03	1.581	65	2.4	423
	NE-110	1.03	1.58	60	3.3	434
Liquid	NE-213	0.874	1.508	78	3.7	425
	NE-226	1.61	1.38	20	3.3	430

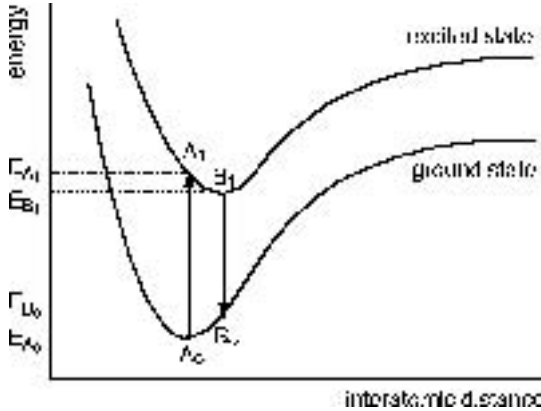


Fig. 24. Simplified energy-level diagram in an organic scintillator

detector by the fast electron will be absorbed by elevating the electron configuration into one of the numerous excited states. The scintillation process is schematically represented in the energy-level diagram of Fig. 24, showing the potential energy of a molecule as a function of interatomic distance. The lower curve represents the potential energy for all electrons in the ground state, the upper curve shows an excited state. The Franck–Condon principle (electronic transitions in the molecule occur very fast with respect to the readjustment time of the interatomic distance) states that the energy deposited raises the molecule from A_0 to A_1 ($E_e = E_{A_1} - E_{A_0}$) in a time (~ 0.1 ps) short compared to the vibration time. Since a state with excess vibrational energy is no longer in thermal equilibrium with its neighbors, vibrational energy is quickly lost moving the molecule to B_1 . After a time (~ 10 ns) long compared to the vibrational time the excited state decays to ground level B_0 , the excess energy ($E_p = E_{B_1} - E_{B_0}$) being carried away by a photon. This fluorescent emission produces of the order of 1 photon per 100 eV of energy deposited. It should be noted that the energy required to excite a state E_e , exceeds the energy carried away by a photon E_p . $E_e \neq E_p$ is important since it signifies

different emission-and absorption-spectra; this translates into negligible re-absorption, making the scintillator transparent to the scintillation photons.

One of the main advantages of organic scintillators is their short decay times of the induced luminescence so that fast signal pulses are generated (Table 5). Together with their low Z-value, organic scintillators are well matched to the requirements for the upper detectors (D1 – see Sect. 4.2) in Compton telescopes, where excellent timing is required for pulse shape discrimination and time of flight measurement, and low Z-values are welcome in order to favor the Compton effect (see Fig. 4). The D1-detectors of GRO-COMPTEL used seven cells filled with the liquid-scintillator NE-213A [94]. With each of the 28 cm diameter, 8.5 cm thick cells viewed by eight photomultiplier tubes, the scintillator detectors offer event localization (Anger-camera), the average 1σ spatial resolution is 2.3 cm. Figure 25 shows the spectrum of Compton scatter events depositing 468 keV in one of COMPTEL’s D1 detectors. It has been obtained by measuring backscattered events from a ^{137}Cs source (662 keV) – i.e. events that produce a coincidence in an auxiliary detector placed to allow only for scatter angles of 180° . The energy resolution is $\Delta E/E_0 = 13\%$ at 1 MeV scaling with $E^{-0.43}$.

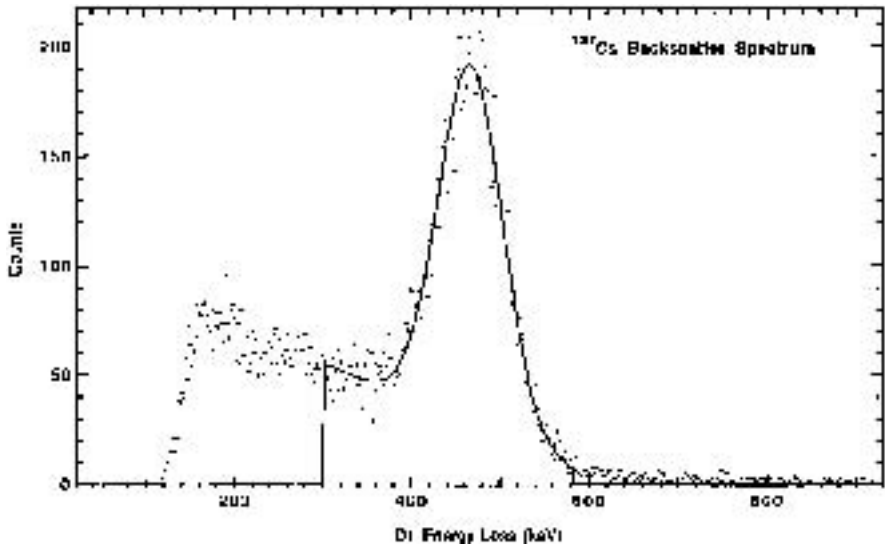


Fig. 25. Compton scatter events depositing 468 keV in one of COMPTEL’s D1 liquid scintillator detectors (see text) [94]

While gamma-ray telescopes often use organic scintillators as charged particle anticoincidence shields, they are rarely used as spectrometers: their lower scintillation efficiency with respect to inorganic scintillators, typically 18% of the light yield of NaI(Tl), results in inferior energy resolution, and

the low Z -values make them poor gamma-ray absorbers. With sensitivity (stopping power) and energy resolution (light yield) as principal requirement for nuclear spectroscopy, inorganic scintillators have been most widely used in gamma-ray telescopes.

Inorganic Scintillators

The scintillation process in inorganic crystals relies on the energy states in a solid insulator where the band theory is applicable. In a pure crystal, electrons can only occupy two discrete energy levels – the valence band (electrons that are bound at lattice sites) and the – usually empty – conduction band (only electrons with sufficient energy to migrate through the crystal). In the intermediate forbidden band of energies, called the band gap, free electrons cannot exist in a pure alkali halid crystals.

The ionization energy produced by fast electrons moving through a crystal, causes electrons to move from the valence band up to the conduction band, producing a vacancy in the valence band that is called a hole. Yet, the “direct” return of the electron to the valence band with emission of a photon is an inefficient process, and the band gap energy corresponds to UV photons with short absorption lengths. In 1948, Robert Hofstadter [95] first described the very high light output obtained from *activated* sodium iodide crystals i.e. with a trace amount of thallium impurities. The role of the impurity is to generate meta-states between the pure crystal valence and conduction bands. Electrons in the conduction band can drop in one of these meta-states and deexcite from it to the valence band. This process not only is more efficient (with respect to the deexcitation over the entire band gap), but it also leads to the emission of visible light photons.

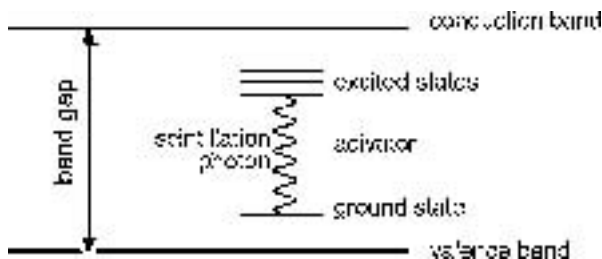


Fig. 26. Band structure of a crystal with activators

The scintillation mechanism and production of a signal-pulse in an activated inorganic scintillator can be summarized as follows: The fast electron(s) produced by the conversion of the gamma-ray generates a large number of e^- /hole pairs – the electrons are raised from the valence-band to the conduction-band, the holes quickly drift to an activator site and ionize it

(the ionization energy of impurities is lower than the ionization energy of a typical lattice site). The electrons in the conduction band are free, until they encounter an ionized activator, creating a neutral, excited atom. For appropriate activators, there are allowed transitions from the excited state to the ground state that are very rapid, emitting a photon in the visible domain. Since typical decay times for the excited states are of the order of $\tau_{1/2} \approx 10^{-7}$ s, much longer than the time for which electrons migrate, the excited states form essentially at once. The scintillation light emission is therefore characterized by the decay times of the excited states.

Inorganic scintillators can be divided into three main groups:

- (a) *impurity activated* inorganic scintillators, the activator sites are produced by adding impurities to the crystal, examples are NaI(Tl), Thallium activated Sodium Iodide, CsI(Tl), Thallium activated Cesium Iodide, and $\text{Gd}_2\text{SiO}_5(\text{Ce})$ Cerium activated Gadolinium Orthosilicate.
- (b) *self activated* – here, a stoichiometric excess of one of the constituents of the solid produces the activator sites, examples are BGO, Bismuth Germanate ($\text{Bi}_4\text{Ge}_3\text{O}_{12}$), or CdS Cadmium Sulfide with excess Cd.
- (c) pure crystals – activator sites are produced by imperfections in the crystal lattice – an example is Diamond.

NaI(Tl): The most extensively used inorganic scintillator is sodium iodide with about 1‰ thallium activator content. NaI(Tl) has an unusually large light yield corresponding to an absolute scintillation efficiency of about 13 percent. The material exhibits no significant self-absorption of the scintillation light. Its dominant decay time is 230 ns, slower than organic scintillators but fast enough for gamma-ray telescopes, including solar flare studies. The emission spectrum of NaI(Tl) is peaked at a wavelength corresponding to the blue region of the electromagnetic spectrum and is well matched to the spectral response of photomultiplier tubes. The principal deficiencies of NaI(Tl) are its mechanical fragility, the need for a hermetic sealed enclosure (NaI(Tl) is hygroscopic) and its extreme toxicity (Thallium). NaI(Tl) is susceptible to radiation damage, i.e. prolonged exposure to intense radiation degrades the scintillation performance. Radiation damage has been observed above levels of 1 Gray (100 rad).

While the popularity of NaI scintillator is based on its good spectroscopic properties, it is also widely used for event location. The NaI Anger camera was invented in 1957 [96] for medical imaging: in a larger flat single crystal, the interaction location is determined by comparing the relative amplitudes of the several photomultipliers viewing the crystal.

The SIGMA telescope [97] which operated between 1989 and 1997 on the Granat platform used a thin NaI(Tl) Anger camera as position sensitive detector. The scintillation crystal was 1.25 cm thick and had a geometric area of 784 cm^2 , it was viewed by 61 hexagonal photomultiplier tubes. As well as measuring the energy deposited, the on-board electronics directly provided Cartesian coordinates of the interaction location in the detection

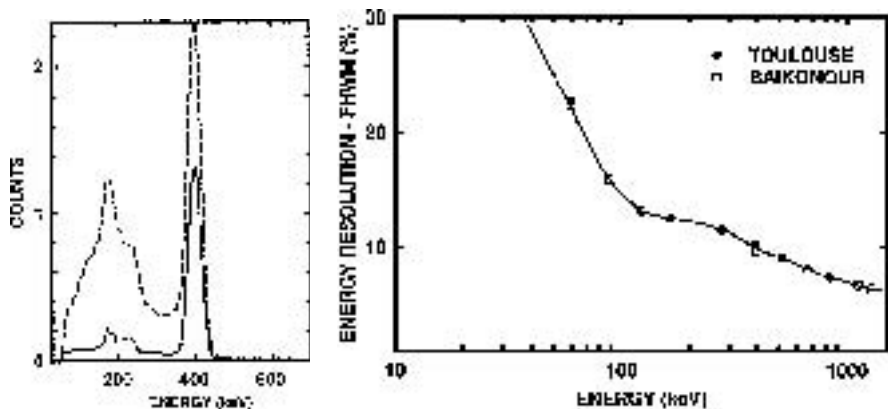


Fig. 27. (a) Spectra of the SIGMA ground calibration with a ^{113}Sn (391 keV) source: integrated counts over the total detector area (*dashed line*) and in imaging mode (*solid line*) (b) the energy resolution of SIGMA as a function of gamma-ray energy [97]

plane (Fig. 27). NaI Anger cameras have been used for a number of other coded mask telescopes (Sect. 4.1), and also for CGRO COMPTEL (Sect. 4.2).

CsI(Tl): Thallium-activated cesium iodide [CsI(Tl)] also produces excellent light yield but has two relatively long decay components with decay times of 0.68 and 3.3 microseconds. Its emission spectrum is shifted toward the longer-wavelength end of the visible spectrum, well matched to the spectral response of photodiodes.

The lower detector level on INTEGRAL's imager IBIS is called PICsIT (Pixelized Imaging CsI Telescope) – a gamma camera consisting of 4096 small CsI(Tl) detector bars. The detector bars have a front surface of $8.55 \times 8.55 \text{ mm}^2$ and a height of 30 mm; the spacing between pixels is only 0.55 mm. Each CsI(Tl) bar is optically bonded to a custom made low leakage silicon PIN photodiode (see photodiodes below). The PICsIT detector layer is divided in eight rectangular modules of 512 detector elements; its total geometric surface is 2994 cm. The energy resolution of the individual CsI(Tl) bars of the Laboratory Model is shown in Fig. 28. PICsIT covers the upper part (150 keV–10 MeV) of the IBIS energy range while ISGRI (Sect. 3.3, CdTe) covers the low energy domain.

BGO: Since its introduction in the 1980s, BGO ($\text{Bi}_4\text{Ge}_3\text{O}_{12}$), a self activated crystal scintillator, has come into wide use. BGO is mechanically and chemically stable (non-hygroscopic) and has a very high density and high Z . Compared to NaI(Tl) it provides a total absorption cross section 2.5 times higher at 1 MeV, permitting compact detector designs. Its disadvantages for spectroscopy are its relatively low light output (20% of NaI(Tl)) and high refractive index resulting in a moderate energy resolution. size is shown in Fig. 29.

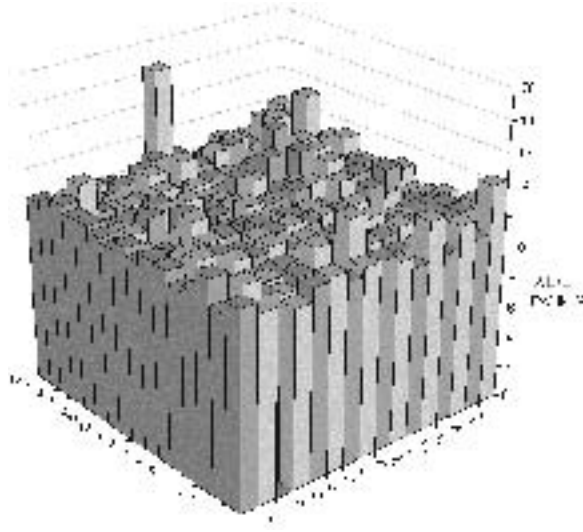


Fig. 28. PICsIT Laboratory Model (PLM): individual energy resolution (FWHM in % at 662 keV) for all pixels. The distribution is quasi-Gaussian centered around 12% [98]

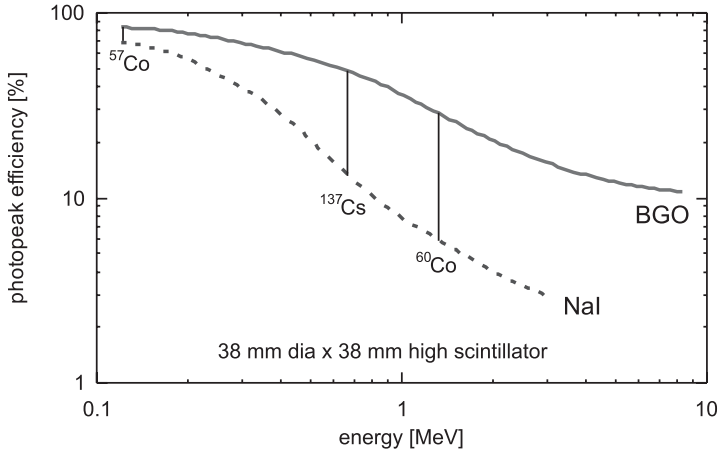


Fig. 29. Photopeak efficiencies for a BGO and NaI(Tl) detector with identical sizes [100]

INTEGRAL SPI's Anti-Coincidence veto System (ACS), consists of 91 BGO blocks in combination with 191 photomultiplier tubes. A description of the SPI ACS is given in Sect. 4.1, a comparison between the photopeak efficiencies of a BGO and a NaI(Tl) detector of identical

PWO: Lead tungstate (PWO or PbWO_4) was selected as the most appropriate scintillator material for future high energy calorimeter projects CMS at CERN's Large Hadron Collider (LHC). PWO has very high absorption power, yet its low light output has limited its use as scintillator to very high energies. The applicability at energies far below 1 GeV was investigated [99] showing $\Delta E/E \approx 15\%$ at 50 MeV. PWO emission spectrum consists of two emission components, the blue one peaking at $\sim 420\text{ nm}$ and a green one peaking around 480–520 nm. The total yield of full size PbWO_4 crystals, integrated in a 100 ns gate, is up to 10 p.e./MeV, as measured by a PMT with a bialkali photocathode, corresponding to a light yield of ~ 100 photons/MeV (assuming an emission weighted quantum efficiency of 10%, see PMTs below). The decay time of the scintillation light from PbWO_4 can be parameterized by three components: one fast ($< 10\text{ ns}$), one slow (20 to 200 ns), and one very slow (500 ns to a few μs).

Interestingly, lead tungstate is also used in ultra-low temperature detectors, achieving energy resolutions of $R > 500$ by detecting ballistic phonons.

Table 6. Properties of inorganic scintillators

	Light Yield ph/keV	Scint. Yield [%NaI]	$\Delta E/E$ at 662	Decay Time [ns]	After- Glow	λ_{peak} [nm]	n Refr.	Hygro	ρ [g/cm ³]
NaI(Tl)*	38	100	7.5%	250	5%	415	1.85	yes	3.67
CsI(Na)*	41	85	9%	630	5%	420	1.84	yes	4.51
CsI(Tl)*	54	45	9%	1005	5%	550	1.79	low	4.51
CaF ₂ (Eu)*	19	50		940		435	1.47	no	3.18
BaF ₂ * fc	1.9	3	$\sim 10\%$	.6–.8	–	225	1.54	low	4.88
sc	10	16		630	–	310	1.50		4.88
BGO*	8–10	20	13%	300	0.1%	480	2.15	no	7.13
CdWO ₄ *	12–15	30–50		14000		475	~ 2.3	no	7.9
PWO [†]	~ 0.1	0.3–1.3		10, 20, 500 ⁽³⁾	(3)	420 500	2.16	no	8.28

Data is derived primarily from *Bircon/Saint-Gobain [100], [†]Zhu et al. 1996 [101]. Light yield values are from measurements with a photodiode with broad spectral response, except for PWO which is measured with a bialkali photocathode PMT. Slow components are measured by the afterglow after 3 ms, for BaF₂ the fast (fc) and slow (SC) is listed separately, for PWO parameters⁽³⁾ see text.

Detecting Scintillation Light

The conversion of the weak light pulse emitted by a scintillator into an electric current requires sensitive detectors for optical/UV photons. Commonly, this conversion is performed by *photomultiplier tubes* (PMT); alternatively, the requirements of certain types of instrument may be satisfied by *photodiodes*.

Photomultipliers

A photomultiplier is a vacuum tube, consisting of a photocathode (conversion of photons into electrons), a multiplier chain (amplification of the signal), and an anode, collecting the resulting current (Fig. 30).

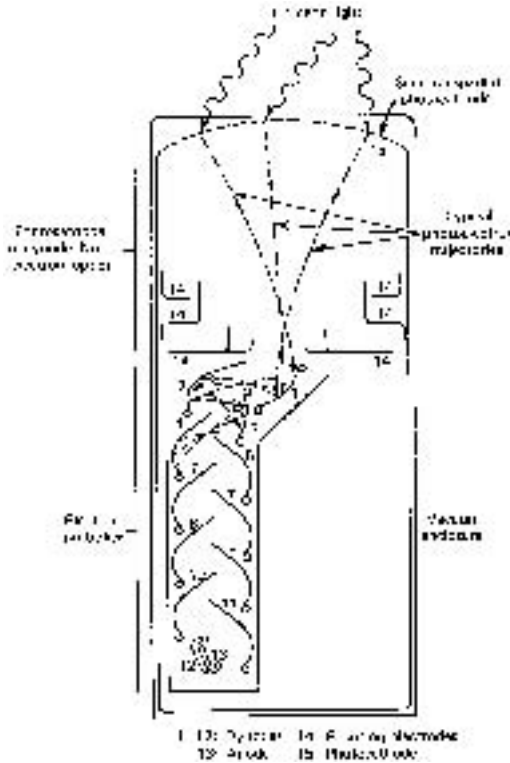


Fig. 30. The elements of a photomultiplier tube [102]

Photocathodes often consist of alkali alloys (such as cesium-antimony, Cs-Sb, or potassium-cesium-antimony, K-Cs-Sb), evaporated as a semi-transparent film onto the entrance window. The conversion of a visible photon (blue) of $h\nu \approx 3 \text{ eV}$ is governed by the photoelectric effect, $h\nu = E_e + W$, where E_e is the kinetic energy of the photoelectrons and W the workfunction,

this is, the potential barrier the electron has to overcome to escape from the photocathode. Also, during the migration of the electrons to the surface of the photocathode, kinetic energy is lost through electron–electron collisions. The efficiency of the entire conversion process strongly depends on the material of the cathode and the entrance window, and on the wavelength of the incident light; it is described by the quantum efficiency, defined by

$$\text{QE} = \frac{\text{number of photoelectrons emitted}}{\text{number of incident photons}}. \quad (62)$$

Practical photocathodes show quantum efficiencies of 10–30% over narrow spectral ranges. Quantum efficiencies of up to 50% have been achieved using GaAsP(Cs) and GaAs(Cs) semi-transparent photocathodes, but these PMTs need moderate cooling to reduce their dark current.

The electron multiplication departs with the photoelectrons produced in the photocathode being accelerated towards the first dynode by focusing electrodes. The dynode chain utilizes the phenomena of secondary emission to multiply the number of primary photoelectrons. Electrons leaving the photocathode have kinetic energies of the order of an eV or less. The number of secondary electrons depends on the coating material and the operating voltage: as the creation of a secondary electron on a dynode requires at least the bandgap energy (2–3 eV), an incident electron may generate about 30 electrons for 100 V accelerating voltage. However, only a small fraction of the exited electrons will contribute to the secondary electron yield – exited electrons may not reach the dynode surface before deexcitation or, if they do reach it, will have lost too much energy to escape from the dynode. As various surface coatings on the dynodes produce 1.5 to 50 secondary electrons for every primary electron that strikes them, net amplification by as much as 10^7 – 10^9 can be achieved in PMT's.

In space experiments, the varying magnetic fields may require careful magnetic shielding of the PMT's: e.g. with the low photoelectron energy at emission ($\sim$ eV), a field of only 1 Gauss can reduce sensitivity by 50%.

Photodiodes

An alternative way to detect the scintillation light from a crystal is the use of a silicon photodiode. In a photodiode, the scintillation photons produce electron–hole pairs that are collected at respectively the anode and the cathode of the diode. Most frequently, reverse biased PIN photodiodes having a low capacitance and leakage current are used. The quantum efficiency of silicon photodiodes is typically 70% between 500 and 900 nm, the wavelength band being well matched to the scintillation light of CsI(Tl) crystals. The lower detector layer of the IBIS telescope on INTEGRAL (PICsIT) is composed of such a combination (see *CsITl detectors* above).

Contrary to PMTs, photodiodes do not require a high voltage power supply but only a bias voltage of about 30 V. Due to the small signal generated

by the photodiode (there is no inherent signal amplification in the photodiode), it is necessary to employ a high quality charge preamplifier in order to keep the noise level as low as possible. Noise is a problem intrinsic to standard photodiodes. The substantial capacitance of the device ($40\text{--}50\text{ pF cm}^{-2}$) for 200 and 300 mm wafer devices) is mainly responsible for the noise which determines for a large part the energy resolution of the detector. Also the dark current of PIN photodiodes may contribute significantly to the noise, especially at larger shaping times. The dark current increases with increasing surface area as well as with increasing temperature.

The low level noise limit can be overcome by using so-called Avalanche PhotoDiodes (APDs), which reach quantum efficiencies as high as 90%. In an avalanche photodiode, an incoming photon creates an electron–hole pair. A large reverse field of up to 2 kV causes electrons to accelerate through the doped silicon toward the device’s cathode, producing an avalanche of electrons by collisional ionization. Each initial photoelectron typically results in several hundred electrons reaching the cathode. Drawbacks of APDs are the poor gain stability (the amplification is a strong function of temperature) and the high room temperature leakage current (requiring cooling).

3.3 Semiconductor Detectors

Semiconductor detectors directly collect the charge carriers that are produced by the incident photon. Along the track(s) of the secondary electron(s) which are created by the gamma-ray interaction with the detector material, electrons are raised from the valence band to the conduction band, leaving an equal number of positive holes in the valence band. The number of electron–hole pairs generated is proportional to the energy loss of the secondary electron. A strong electric field applied across the detector separates the pairs before they recombine. Electrons drift towards the anode, holes to the cathode – the charge collected by the electrodes produces a current pulse whose integral equals the total charge generated by the incident particle; and hence is proportional to the energy deposited in the detector.

Even in the absence of ionizing radiation, the strong electric field required to efficiently collect the charges will induce a leakage current i_l in the semiconductor which has a finite conductivity. The fluctuations of the leakage current are a source of noise over which a charge pulse must be distinguished and increase with the leakage current itself ($\sim i_l^{1/2}$). The leakage current in a semiconductor is due to carrier generation by thermal excitation over the semiconductor band gap E_g . The probability P that electron–hole pair is thermally generated is

$$P(T) \sim T^{3/2} e^{(-E_g/2kT)} . \quad (63)$$

The number of generated pairs is a strong function of the detector temperature T , it also depends critically on the bandgap of the semiconductor

Table 7. Properties of semiconductor materials

Composition	Density [g/cm ³]	Mean Z	Bandgap [eV]	Energy per e ⁻ -hole pair [eV]
Ge	5.32	32	0.74	2.98
Si	2.33	14	1.12	3.61
CdTe cadmium telluride	6.2	50	1.6	4.43
Cd(Zn)Te	6.0	48	1.6	
HgI ₂ mercuric iodide	6.36	62	2.15	4.22

material. The energy ε required to generate an electron–hole pair can be expressed by $\varepsilon = (14/5)E_g + c$, where $0.5 \leq c \leq 1$ eV [103]. Dependent on the size of the bandgap two categories are distinguished: narrow and wide bandgap materials (see Table 7). Below, the two classes will be illustrated by the example of Ge and CdTe detectors.

Semiconductor Junctions

Since the conductivity of even the highest purity semiconductors is not negligible, the reduction of leakage current is the crucial design consideration for such detectors. For example, for high-purity Si values around 50 000 Ω -cm are common. This means that a 1 mm thick slab with a 1 cm² surface area would possess a resistance of 5000 Ω . An applied bias of 500 V would therefore lead to the flow of a leakage current, 0.1 A in magnitude. A pulse of 10^5 charge carriers generated by the passage of a photon of a few 100 keV through such a detector would generate a peak current of around 10^{-6} A – five orders of magnitude inferior than the leakage current!

In the semiconductor junctions used as radiation detectors, leakage current is dramatically reduced by using non-injecting or blocking electrodes as electrical contacts. By doping the surface of e.g. a p-type crystal with acceptor impurities, an n-type contact is created. In the depletion region formed near the junction between the n- and p-type the material, charge carrier diffusion takes place, the electric field generated across the junction makes the contact a diode with very high resistivity. The most typical form of blocking electrode is the PN junction, reverse biased in order to provide sufficient electric field to collect the charge carriers. In order to avoid losses in charge collection, detectors are generally overbiased (typically at 1000 V cm⁻¹).

Narrow Bandgap Semiconductor: Germanium

In narrow bandgap materials such as Germanium thermal excitation at room temperature populates the valence band, leading to an important leakage current that degrades the spectral resolution of the detector. Consequently, narrow bandgap detectors have to be operated below room temperature.

Because of the small energy gap of 0.74 eV, Germanium spectrometers must be cooled to temperatures below 130 K to reduce the leakage currents below 1 nA, which gives comparatively negligible current-generated noise (≈ 50 nA in 0.1 μ s for $E \approx 100$ keV). The most widespread way to cool Ge detectors in the laboratory is by using liquid nitrogen, which keeps the detector at 77 K. In space, passive cooling (radiators) or mechanical coolers (e.g. Stirling machines) are preferable for a longer lifetime of a mission.

An advantage of the modern high purity (HP) Ge detectors, is the high resistivity of the detector which allows depletion depths of several centimeters by applying potentials of several thousand volts. The increase of the achievable depletion depths allows the construction of large high-energy resolution detectors (up to 8 cm in diameter). The impurity concentration must be low enough that the electrons and holes produced by gamma rays are not significantly trapped by impurity levels in the band gap. Present techniques for the production of high-purity germanium provide materials with impurity concentrations of less than 10^9 cm^{-3} (i.e., 1 impurity per 10^{13} atoms).

Gamma-ray spectrometers in space environment are subject to irradiation by cosmic rays, as well as by secondary particles generated in the interaction with the surrounding materials. This affects the crystal structure by increasing the amount of hole trapping within the active volume [104, 105] producing a loss in the charge collection, which depends on the position of the interaction. The resulting effect is a tailing toward the low energy side in the gamma-ray peaks.

Studies performed in the past have shown that protons are more damaging than neutrons by about a factor of 60 [106] and p-type detectors are about 28 times more sensitive than n-type to the irradiation. Studies of the dependence of the radiation damage on the temperature and bias voltage are presented in [107]. Complete recovery from the line broadening can be achieved by annealing the detector at a temperature of about 150°C for several hours [107].

Energy Resolution

The main advantage of semiconductors over scintillators is their excellent energy resolution. A comparison of 1 MeV photons interacting in either one of the two detector types is schematically presented in Fig. 31: In a scintillation detector, only about 120 keV will be transformed into scintillation light ($\sim 12\%$ scintillation efficiency). The number of scintillation photons ($E \approx 3$ eV) generated will be of the order of 40 000, yet only about half of them are assumed to be detected on the photocathode. With a PMT photocathode efficiency of 20%, 4000 photoelectrons are created for a 1 MeV gamma quantum. Statistical fluctuations in this number limit the theoretically achievable energy resolution (see (56)) to $R = 0.42\sqrt{N_{\text{sci}}/F_{\text{sci}}} \approx 25$, corresponding to a line width of 40 keV (FWHM). Non-uniform light collection from the scintillator and the variation in quantum efficiency over the

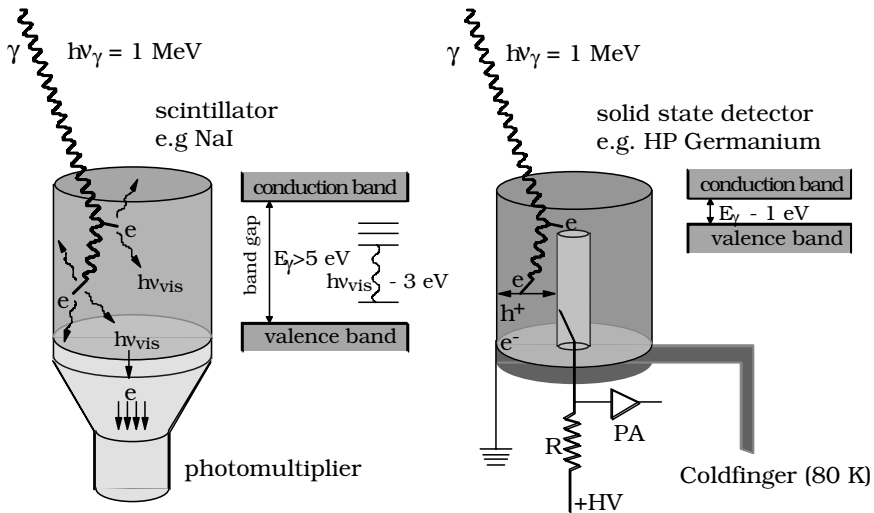


Fig. 31. Scintillator vs. semiconductor detector: comparison of 1 MeV photons interacting in either one of the two detector types

area of the photocathode further degrade this resolution, making scintillators to rather poor spectrometers.

In a semiconductor detector, the small band gap energy increases significantly the number of information carriers per pulse providing better statistics. The energy for creating an electron-hole pair in Germanium is 3 eV (eight times less than the energy per photon for a NaI scintillator counter), resulting in $N_{\text{sem}} \approx 10^6 / 3 \text{ eV} \approx 300\,000$ charge carriers. As the statistical fluctuations in the charge carrier number are lower than expected if the electron-hole pair formation process followed a Poisson distribution, the measured variance is given by the relationship $\Delta N = N_{\text{sem}} F_{\text{sem}}$, where F_{sem} is the Fano factor. For Germanium, Fano factors vary from 0.06 to 0.14 according to different measurements (Knoll [81]), resulting in an energy resolution of the order of $R = 0.42 \sqrt{N_{\text{sem}} / F_{\text{sem}}} \approx 500$. In practice, incomplete charge collection and electronic noise invariably increase this value to around 2 keV. Still, the improvement in energy resolution using Germanium detectors over scintillator counters is about a factor of 30 at 1 MeV.

INTEGRAL-SPI

The detector assembly of INTEGRALs spectrometer SPI consists of an array of 19 n-type Germanium detectors, with a total geometric detection area of 500 cm^2 . Each detector has a hexagonal shape, 3.2 cm on a side, 7 cm deep, and a center-to-center distance of 6 cm, and is mounted inside a tight Aluminum capsule. The preamplified signal of each detector is fed into an analog front-end electronics (AFEE) system where it is amplified, filtered and

and those of scintillators, CdTe can be used well in the low energy domain (down to ~ 20 keV).

Whenever excess charge is generated in a semiconductor, thermal equilibrium is disturbed. The semiconductor returns to equilibrium via recombination, which occurs at trapping sites. The number of free charges decays exponentially, with lifetimes τ_e and τ_h for the electrons and holes respectively. In Si and Ge, there are few trapping sites and the carrier lifetimes are several milliseconds long. A negligible fraction of the charges will be trapped during the charge collection process which typically lasts for 100 ns.

For compound semiconductors such as CdTe or CdZnTe, even the best modern crystal growth practices lead to a much higher density of traps and hence shorter lifetimes. For CdZnTe, typical lifetimes might be $\tau_e = 3 \cdot 10^{-6}$ s and $\tau_h = 5 \cdot 10^{-8}$ s. Because the hole lifetime is much shorter than the hole transit time, a substantial fraction of the hole charge will be lost, leading to a reduced pulse height. Since the charge loss depends on the hole transit time, it depends upon the distance between the gamma-ray interaction and the cathode.

This property can be used to correct for charge loss. Firstly, if the cathode is oriented towards the source signal, low-energy gamma-rays (which preferentially interact at the surface of the detector) will create electron-hole pairs near the cathode, leading to short hole transit times, hence to little charge loss. Secondly, if the interaction depth of the gamma ray can be measured, the charge loss can be estimated from this depth, and a charge loss correction can be applied that reduces the pulse height variation. A reasonable measure of the interaction depth is given by the rise time of the current pulse, and pulse shape rise time measuring electronics are successfully employed for charge loss correction. Thirdly, the electron signal can be measured without a contribution from the hole signal by using a Frisch grid (Frisch grids are a classic solution to incomplete charge collection of ions in gas detectors.)

With their small area, the CdTe detectors are ideally suited to build a pixel-lated imager with good spatial resolution. Outstanding energy resolutions are being achieved with thin detectors e.g. 810 eV FWHM at 59.5 keV in a 1 mm thick CdTe diode [109]. As this type of detector is manufactured in large arrays (eg. 1024 pixels, 38.4×38.4 mm²) they are of particular interest for imaging systems using coded masks or as Compton telescopes. Up to now the use of CdTe was restricted to the low energy domain (e.g. 50% efficiency at 150 keV) due the small thickness necessary to achieve good energy resolution). This is certainly going change over the next years.

INTEGRAL-ISGRI

The upper detector layer of INTEGRALs imager IBIS is an assembly of 16 384 CdTe detectors operating at room temperature, representing a total sensitive area of 2621 cm². The pixels are 4×4 mm large and 2 mm thick; they are spaced by only 600 microns and are organized in 4×4 assemblies called

polycells. Eight identical Modular Detection Units (MDUs) each accommodate 128 (16×8) polycells. An MDU contains 2048 pixels which are read out by 512 Application Specific Integrated Circuits (ASIC). The ASICs have low noise charge-sensitive preamplifiers featuring pulse rise-time measurement in addition to the standard pulse height measurement. This permits a charge loss correction to be computed based on the charge drift-time. The MDUs are connected independently to a Detector Bias Box and to a Module Control Electronics which performs the A/D conversion and provides other on-board processing such as event filtering and active pixel monitoring.

ISGRI covers the lower part (15 keV–1 MeV) of the IBIS energy range while PICsIT covers the high energy domain (see Sect. 3.2, CsI). Since, the energy ranges covered by ISGRI and PICsIT overlap considerably (150 keV–1 MeV), the two cameras can work in coincidence providing a Compton telescope mode which ensures a good background reduction above 200 keV. The dedicated electronics measure simultaneously the rise time and standard pulse-height. This allows the computation of charge loss and ballistic deficit correction (Fig. 33). After application of this correction, a spectral resolution around 7.5% at 122 keV is obtained with the ASICs.

4 The Instruments for Nuclear Astronomy

The instrumental categories which can be identified in the energy range of nuclear astrophysics reflect our current perception of the phenomenon of electromagnetic radiation. Geometrical optics is the base of coded aperture systems; focusing telescopes and Compton telescopes are based on wave and quantum optics respectively (Fig. 34).

A telescope – as we will use the term in this review – is a system characterized by its aperture and its detector. While the aperture defines the method of collecting photons (imaging, and if applicable, concentration) the detector measures their properties (energy). In this chapter, the three families of telescope systems relevant to nuclear astrophysics will be discussed: coded aperture systems (Sect. 4.1), Compton telescopes (Sect. 4.2), and focusing instruments (Sect. 4.3). The three instrumental types have in common the way spectroscopy is performed: the non-dispersive measurement of the photon energy in a detector (Sect. 3). Rather than conducting a comprehensive survey of all existing projects, only a small sample of missions will be presented in order to illustrate the three telescope principles.

4.1 Geometric Optics: Modulating Aperture Systems

Our present understanding of the low energy (<1 MeV) gamma-ray sky has been acquired mainly by modulating aperture systems. The underlying concept of this class of instruments is geometrical optics, that is, the source photons are considered as traveling on rectilinear paths only.

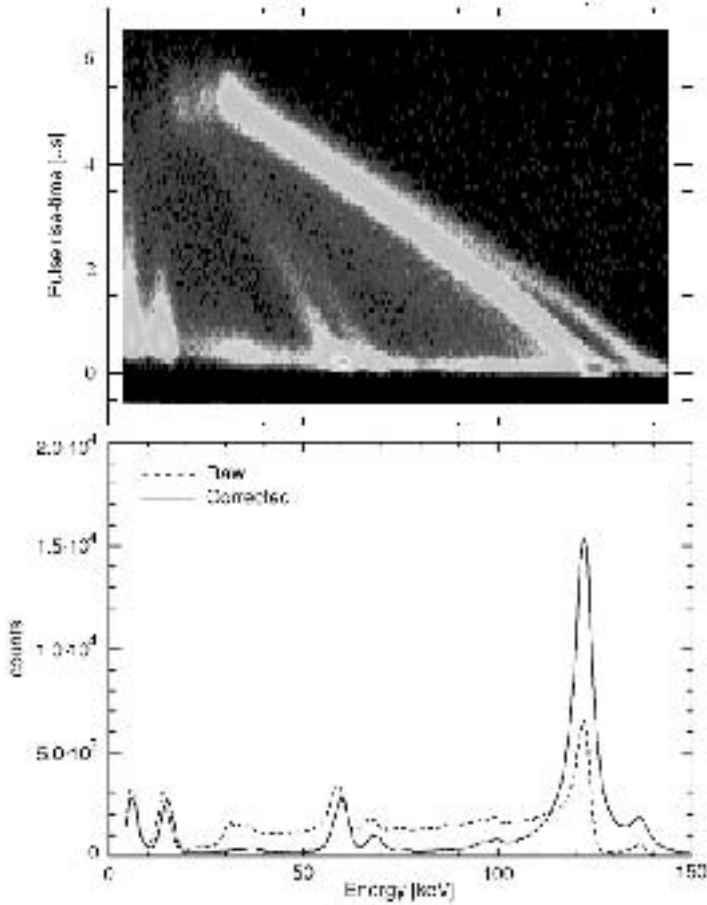


Fig. 33. ISGRI CdTe spectrum of a ^{57}Co source. *above:* a biparametric diagram (pulse height vs. pulse rise time) shows the effect of charge trapping. The strong diagonal ridge represents 122 keV photons – as the distance between interaction site and the cathode increases, the travel time of the charges becomes longer, and charge trapping will reduce the measured pulse height. *below:* *dashed line:* raw energy spectrum, *solid line:* energy spectrum with ballistic deficit correction [110]

The incident radiation is passing through open elements of an otherwise opaque aperture. The aperture system – consisting of masks or collimators – modulates the signal which then reaches the detection plane designed to discern shadow patterns of some kind. The best definition of the photon path is achieved for photoelectric interactions both in the mask and the detector. These systems are well adapted to the hard-X/low energy gamma-ray channel where the photoelectric effect is the predominant mode of gamma-ray interaction in medium- and high-Z materials. Pinhole cameras, rastering

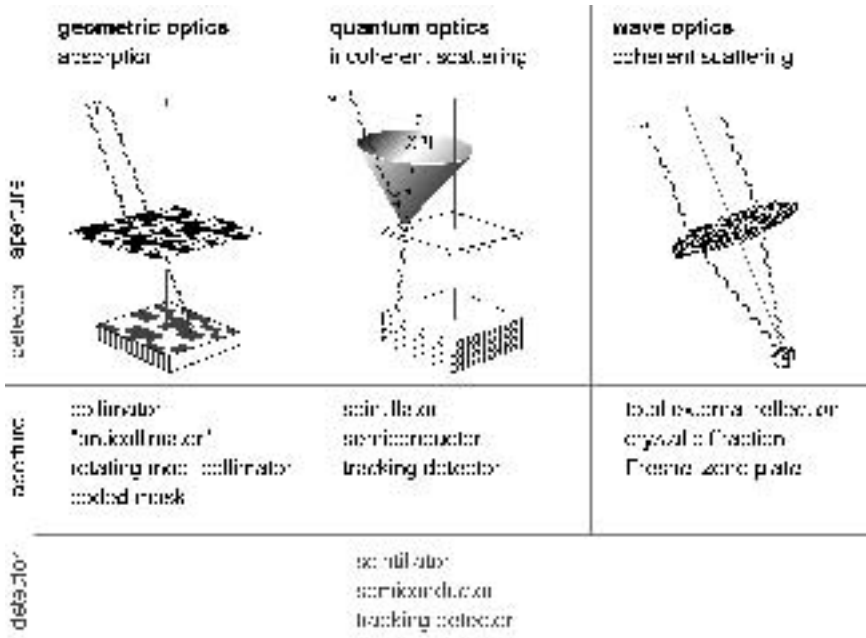


Fig. 34. The three instrumental principles in nuclear gamma-ray astronomy

collimators, coded masks and modulation collimators belong to this category of instruments.

Two main classes of modulating aperture systems can be identified (Fig. 35), according to whether the signal is encoded by *temporal modulation* (e.g. rotating modulation collimators) or by *spatial modulation* (e.g. coded mask telescopes). These two types stand for a whole spectrum of devices mixing the basic concepts of spatial and temporal modulation. Modulating aperture systems of both classes can be used to produce images of the sky.

Over the last decade, large satellite telescopes based on the principles of geometrical optics have prevailed in low- and medium energy gamma-ray astronomy: GRANAT-SIGMA, GRO-OSSE and BATSE. While the detection plane of each of these instruments is based on scintillators ($\Delta E/E \approx 10$), different aperture systems are used. OSSE and BATSE can be considered as modulation collimators (in the case of BATSE, the earth plays the role of an "anticollimator"), SIGMA was a multiplexing device using spatial modulation with a coded mask.

With SPI, IBIS and JEM-X on ESA's INTEGRAL platform, with NASA's HESSI and SWIFT missions, modulating aperture telescopes have come to maturity and will again dominate experimental nuclear astrophysics over the next decade. The coded mask telescope SPI and the rotating modulation collimator HESSI perform high resolution ($R \approx 500$) spectroscopy for sources

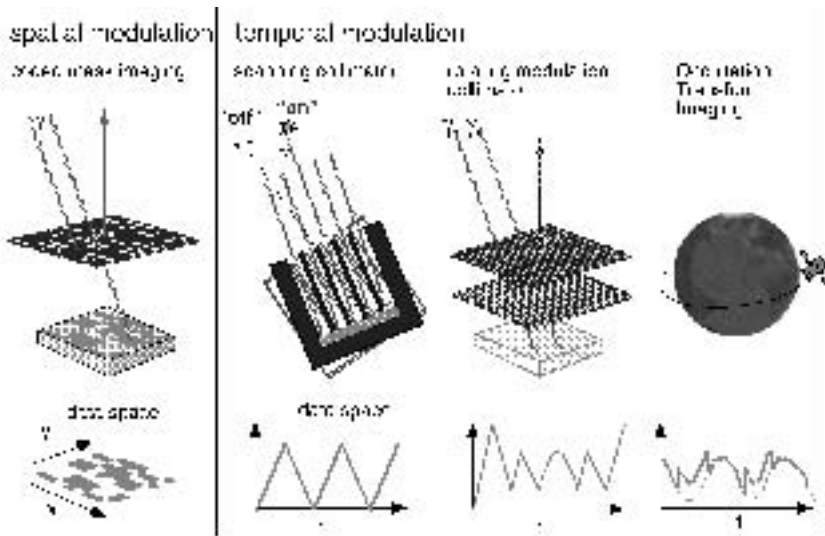


Fig. 35. The categories of modulating aperture systems – spatial and temporal modulation

predominantly in the galactic plane (SPI) and for active regions of the sun (HESSI). These and other instruments will be reviewed in the sections below.

Temporal Modulation – Scanning Collimators

The class of modulating aperture instruments has evolved from plain scanning collimator instruments. Scanning collimators have actually started off the discipline of observational gamma-ray astronomy – until the nineties, most discoveries are owed to this simple type of instrument: from the first observations of the galactic e^+e^- annihilation line with balloon borne “on-off” collimators (see Sect. 1.1 and [19–22]), to the discovery of the first radioactive isotope in the interstellar medium- ^{26}Al - by HEAO-3 [32]. The modulation of the source signal by a collimator system is typically measured by a single detector as a function of time as the entire instrument or parts of it are moved across the sky. Without position sensitivity in the detection plane, the lack of the spatial information in the shadow-pattern is partly compensated by the temporal information of the modulated source flux. For a typical collimator, the variation of the count rate detected from a point source as a function of the scan-angle has a triangular shape. The position of the maximum of the triangle is set by the position of the source along the scanning direction and the height of the triangle is proportional to the flux of the source. Further scans along other directions may then be necessary if the source is to be localized in two dimensions, and particularly for the “imaging” of several

sources or extended emission. Often the collimator has a slat construction defining an aperture with a “long” and a “short” dimension.

HEAO-3

The gamma ray spectroscopy experiment on HEAO-3, which scanned the Milky Way in fall 1979 and spring 1980, consisted of four p-type, high purity germanium detectors, each with a volume of $\sim 100\text{ cm}^3$. The detectors were surrounded by a large CsI shield in electronic anti-coincidence, which was segmented in order to provide crude directionality – the collimator had a 30° (FWHM) field of view. The detectors had an energy range of 50 keV–10 MeV, their initial energy resolution was 3 keV at 1.46 MeV.

OSSE

A prominent representative of the class of scanning collimator instruments was CGRO’s Oriented Scintillation Spectrometer Experiment, OSSE [111], in orbit and operational from 1991 to 2000. With its 2000 cm^2 of effective area (fep) at 511 keV, OSSE performed the first rough mapping of the galactic e^+e^- annihilation line [25]. Four identical detector systems, each one consisting of a large area scintillator and a tungsten collimator, were able to independently scan across the sky to carry out simultaneous source and background pointings. The four detector system were composed of a 330-mm diameter phoswich, consisting of a 102-mm thick NaI(Tl) crystal optically coupled to a 76-mm thick CsI(Na) crystal. Each phoswich was viewed from the CsI face by seven photomultiplier tubes, providing an energy resolution of 8% at 0.661 MeV. Utilizing the differing scintillation decay time constants of NaI(Tl) and CsI(Na), the detector event processing electronics incorporated pulse-shape analysis for the discrimination of events occurring in the NaI crystal from those occurring in the CsI, allowing the CsI portion of the phoswich to act as anticoincidence shielding for the NaI portion. A tungsten alloy passive slat collimator, located directly above the NaI portion of each phoswich, defined the gamma-ray aperture of the phoswich detector, providing a $3.8^\circ \times 11.4^\circ$ FWHM rectangular field-of-view throughout the 0.1–10 MeV energy range. Since the background in LEO is modulated by the spacecraft’s orbital period of about 90 minutes, alternate source and background pointings were executed every 2 minutes by motion scans of the four units.

Temporal Modulation – Occultation Transform Imaging

An unshielded gamma-ray detector orbiting the earth will measure step-like occultation features in its counting rate every time a gamma ray point source crosses the earth’s limb. The occultation features produced by the rising and setting of source can be used to locate and monitor astrophysical sources. In

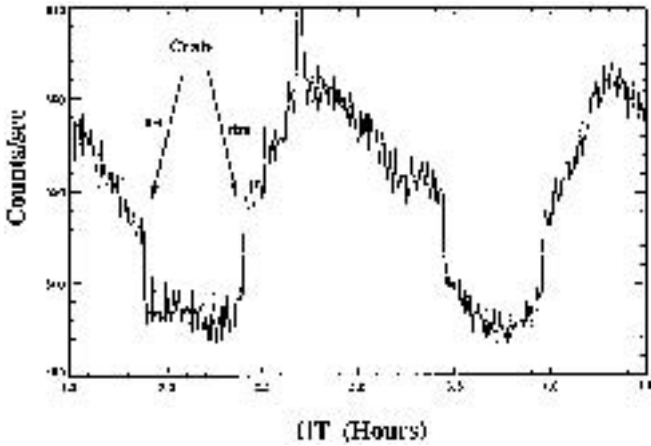


Fig. 36. Example of earth occultation technique with BATSE: As a source (here the Crab nebula) sets below or rises above the earth’s limb, the count rate history shows clearly distinguishable “occultation features” [114]

this approach, called Occultation Transform Imaging, the earth takes the role of an “anticollimator”. The observed change in count rate in several energy bands provides a measurement of the source intensity and spectrum without sophisticated background models.

BATSE

The Burst and Transient Source Experiment, BATSE [112], on CGRO has served as an all-sky monitor using occultation transform imaging. The instrument includes eight Large Area NaI scintillator Detectors (LADs), each 50.8 cm diameter and 1.25 cm thick (2025 cm² geometrical area) operating in the energy range 0.02 to 2 MeV. The eight LADs look out from the corners of the spacecraft such that their surfaces are in the faces of a regular octahedron. Since CGRO orbited the earth at an altitude of about 450 km, about 33% of the sky, as viewed with BATSE, were covered by the earth at any given time. The entire sky was subject to earth occultation for some portion of CGRO’s 52 day precession period. As an example, the “occultation features” produced by the crab nebula are shown in Fig. 36. Using the earth occultation technique, BATSE was able to locate new sources and, for a catalog of moderately strong sources, monitored the photon spectra averaged over weeks and months, and observed light curves in the 35–200 keV band with one day resolution [113].

Temporal Modulation – Bigrig Collimator

Scanning modulation collimators (also *bigrig* or *Oda-collimators*) typically use two or more sets of grids in order to time modulate the intensity at the

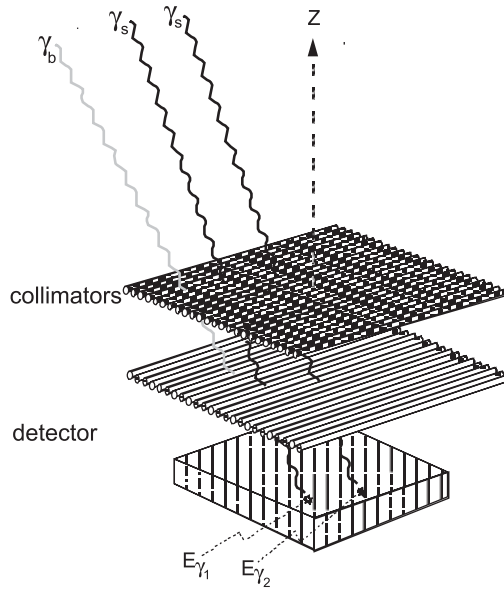


Fig. 37. Temporal modulation – the principle of a bigrid collimator telescope

detection plane. In its simplest version [115] a pair of similar absorbing grids is mounted in front of the detector (Fig. 37). The transmission function of a scanning modulation collimator is determined by the ratio of the pitch of the grid wires (typically twice the wire diameter d) and the distance between the two grids. During the scan of a single point source the count rate at the detector is modulated by the transmission function, typically it is a pattern of periodic windows with opening angles $\Delta = d/D[\text{rad}]$ FWHM. The encoding principle is schematized in the upper part of Fig. 38).

Temporal Modulation – Rotating Modulation Collimators

Following an idea of Mertz [116], suggesting that the collimator be rotated about its axis rather than scanning the axis along a particular straight line, Schnopper et al. [117] proposed a rotating modulation collimator (RMC). The design of instruments belonging to this subclass is virtually identical to the scanning modulation collimator (Fig. 37). The rotation of the collimator results in a cyclic modulation pattern in which the number of cycles per rotation depends on radial position r of the source. The azimuth angle θ of the source determines the phase of the cyclic pattern with an ambiguity of 180° (lower part of Fig. 38). The ambiguity can be avoided by offsetting the axis of rotation during the observation of a source or by e.g. a $1/4$ period shift of one grid.

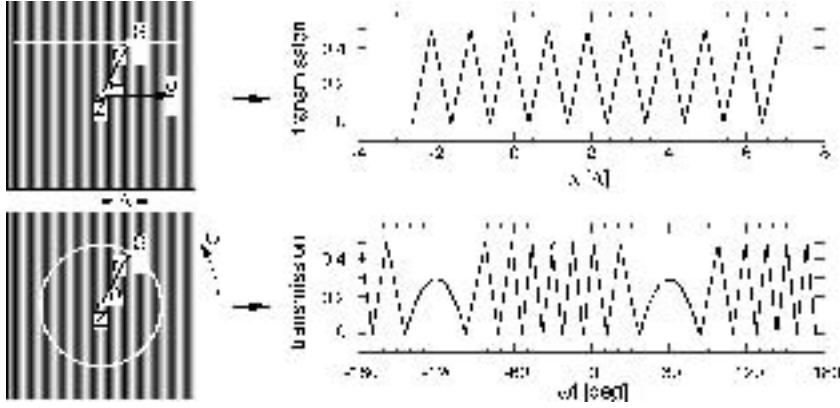


Fig. 38. Temporal modulation – the encoding of a signal by a bigrid collimator telescope (*above*) and a Rotating Modulation Collimator (*below*)

Image Reconstruction – Encoding

For a modulation collimator the expected count rate $N'(t)$ at the detector is of the form

$$N'(t) = \sum_i s_i \cdot \varepsilon \cdot f_i(t) B, \quad (64)$$

where s_i is the flux from the i th source, ε the detection efficiency over the energy band, f_i the transmission function for the i th source at time t , and B is the background count rate. The transmission f_i for a point source located at the position r, θ from the instruments z -axis (Fig. 38) can be written

$$f_i = \left| \frac{1}{2} - (|g_i - \text{int}(g_i)|) \right|, \quad (65)$$

with g_i depending on the type of collimator movement and where $\text{int}(g_i)$ is the integer part of g_i

$$g_i(\alpha) = \frac{r \cos(\theta) - \alpha}{\Delta} \quad \text{scanning modulator}, \quad (66)$$

$$g_i(t) = \frac{r \cos(\theta - \omega t)}{\Delta} \quad \text{rotating modulator}. \quad (67)$$

Image Reconstruction – Decoding

Sources in the field of view can be found by cross-correlating the measured data $N(t)$ with the transmission functions f_p of various trial positions. The cross-correlation function C_p for a trial position p is written

$$C_p = \int_{t_1}^{t_2} N(t) f_p dt . \quad (68)$$

The cross correlation function shows maxima at the position of a point source. However, the correlation map contains artifacts consisting of concentric ring patterns centered on each source and oscillating with a radial periodicity given approximately by Δ . At a position which is symmetric with respect to the rotation axis, the map is enhanced by a ghost mirror image of a source. To find weaker sources that may be masked by such a pattern, Schnopper et al. [117] proposed to subtract the ring pattern of a strong source that has been located. The removal of ambiguous ghost images is improved by offsetting the direction of the instrument z-axis.

Design Considerations for Modulation Collimators

In order to reduce source confusion and the frequency of ghost rings that are due to the different transmission windows additional grids may be added to the collimator system. These systems are called *multi-layer modulation collimators*. Their design and methods for image reconstruction are discussed by Oda et al. [118]. If another grid is inserted in the middle between two grids, every other transmission window is eliminated, while the width of the individual window (and thus the angular resolution of the collimator system) remains constant. However, with increasing number of grid layers the detection efficiency decreases – the system becomes more and more a slat collimator and loses its multiplexing advantage. A concept to overcome the dilemma of the multi-layer modulation collimator is the *multi-pitch modulation collimator (MPMC)* that has been introduced by Makishima et al. [119]. The idea consists in having M separated modules of bigrid modulators, each one having transmission windows with different opening angles. If the largest band spacing is denoted by Δ , the other windows have opening angles $\Delta/2$, $\Delta/3, \dots \Delta/m, \dots$ and Δ/M . When scanning an extended source (of angular size $< \Delta$) each subcollimator detects the corresponding Fourier component of the source profile. With the observed amplitude and phase of all the fundamental Fourier components, the source function can be synthesized through an inverse Fourier transform. This procedure is analogous to aperture synthesis techniques in radio astronomy.

Advantages – Disadvantages

The fact that modulation collimators do not require a position sensitive detector is their principal strength and makes these telescopes technically rather simple. In spite of their simplicity, modulation collimators can survey multiple sources in a wide field of view (multiplexing advantage). On the other hand, the temporal modulation is difficult to apply to variable sources, at least if their variability period is of the order of the modulation period. The range of spatial frequencies of a collimator being limited – mostly only one

frequency is used – the range of angular scales that can be imaged is necessarily limited. Imaging of extended sources is difficult due to the reduced modulation contrast.

HESSI

On February 5, 2002, the High Energy Solar Spectroscopic Imager (HESSI) was launched by a Pegasus XL rocket into a 600 km-altitude orbit. HESSI’s imaging system is made up of nine rotating modulation collimators, each consisting of a pair of 1.5 m separated grids mounted on a sun-pointed spacecraft rotating 15 times per minutes [120]. The grid pitches range from 34 μm to 2.75 mm in steps of the square root of 3 resulting in angular resolutions that are spaced logarithmically from 2.3 arcsec to $\gtrsim 3$ arcmin, allowing sources to be imaged over a wide range of angular scales. Diffuse sources larger than 3 arcmin are not imaged but full spectroscopic information is still obtained. The spectrometer has nine segmented Germanium detectors, one behind each RMC, to detect photons from 3 keV to 20 MeV. The (n-type) coaxial Ge detectors (7.1-cm diameter $\times$ 8.5-cm long) are cooled to $\lesssim 75$ K by a mechanical cryocooler. The inner electrode is segmented into three contacts that collect charge from three electrically independent detector segments, defined by the electric field pattern. This provides the equivalent of a ~ 1 -cm thick planar GeD in front of a thick ~ 7 -cm coaxial GeD, plus a bottom 0.5-cm “guard-ring”. The spectral resolution is ~ 1 keV (FWHM) in the front segment up to ~ 100 keV, ~ 3 keV in the rear segment up to ~ 1 MeV increasing to ~ 5 keV at 20 MeV. Pointing information is provided by a solar aspect system and roll angle system. The spacecraft rotation rate of 15 rpm provides a complete image with the maximum number of Fourier components in 2 seconds, however spatial information from fewer Fourier components is still available on time scales down to tens of ms, provided the count rates are sufficiently high.

The primary scientific objective of HESSI is to understand particle acceleration and explosive energy release in the magnetized plasmas at the Sun, processes which also occur at many other sites in the universe.

Utilizing HESSI’s technology for extrasolar astrophysical targets, the CYCLONE mission is in the proposal stage for NASA’s SMEX program. The high angular resolution that can be achieved with HESSI type RMC’s, associated with the energy resolution of the Germanium detectors, make CYCLONE an attractive alternative for the energy range of 3–200 keV [121]. Besides of performing sub-arcminute mapping of galactic supernova remnants in ^{44}Ti emission, CYCLONE would study the cyclotron lines in accreting neutron stars, the crowded galactic fields of compact objects, and active galactic nuclei.

Spatial Modulation – Coded Mask Imaging

A coded mask telescopes typically consists of a planar array of opaque and transparent elements located in front of a position sensitive detection plane.

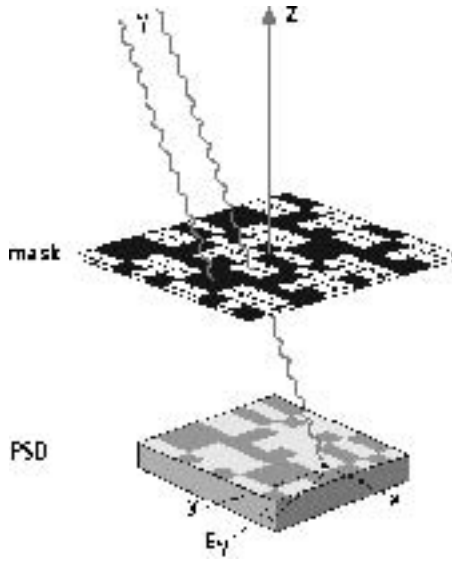


Fig. 39. The principle of spatial modulation by a coded mask telescope

A point source above the instrument projects a shadow of the mask onto the detection plane (Fig. 39). For every gamma-ray event interacting on the detector, the energy E_γ (or an energy interval), the arrival time t (or a time interval), and interaction location (x,y) on the detector are measured – the distribution of interaction locations is called *shadowgram*. The position of the source can be reconstructed by measuring the angular offset (in orthogonal angles expressed by x - and y -coordinates) of the shadowgram relative to Z , the optical axis of the telescope.

The ongoing discussion on the ancestry of coded aperture masks should be extended to the 4th century BC when Aristotle noticed that specks of sunlight under a large tree always are rounded (Aristotle, “problemata physica”). In problem XV, 6 [122], Aristotle describes a setup conspicuously resembling a coded mask that produces multiple images of the sun. His explanation of the phenomenon invokes the principles of geometric optics, although comprehension is rendered difficult through the multiple translations of terminology with over two millennia. Aristotle also describes and correctly interprets the observation of the crescent-shaped images of the sun during an eclipses produced by different “masks” [123].

Nonetheless, many authors regard the *Camera obscura* (Mo Ti, Al-Haitham, R. Bacon) as predecessor of coded mask instruments. In its basic form it consists of a darkened box in which images of external objects, received through a small aperture, are projected on a screen. This type of device – in photography it is known as pinhole camera – is applicable to observations at any wavelength, as long as diffraction at the aperture (d) is

negligible. While diffraction does not limit application in gamma-ray imaging ($2f\lambda/d \ll d$), it is the complementarity of throughput and angular resolution that make the pinhole camera impracticable in gamma-ray astronomy: As the angular resolution linearly improves with decreasing diameter of the “pinhole”, the observed countrate from a source decreases quadratically – with its intrinsically weak source fluxes, gamma-ray astronomy has therefore not used pinhole cameras. It is noteworthy, however, that one of the early X-ray satellites, ARIEL V (launched in 1974) was equipped with the all sky monitor ASM [124] which produced the true images in the keV range by using a 1 cm^2 “pinhole”.

Interestingly, when indirect imaging with a coded mask first was proposed in 1961, the underlying idea had not been a pinhole or multiple-pinhole camera, but the novel concept of holography. Mertz and Young [125] recognized that shadowcasting of a large, coarse Fresnel zone plate (Fig. 40; negligible diffraction in X-rays) could mimic a hologram that, after photographic reduction, can be used to reconstruct an optical image by exposing it to a source of coherent light. A hologram has indeed similarities to the intensity distribution recorded by the position sensitive detection plane behind a coded mask: every region of the two dimensional pattern – hologram and shadowgram – contains information of the whole object, including its three dimensional structure.

However, other modulators than Fresnel zone plates can be used for coded mask instrument when digitized methods of numeric deconvolution are used. The concept of the camera obscura inspired Dicke [126] and, independently, Ables [127] who proposed multiple-pinhole masks as modulators. Shadowcasting of a multiple-pinhole mask on a position sensitive device can overcome the conflicting requirements of the single pinhole. The aperture flux is

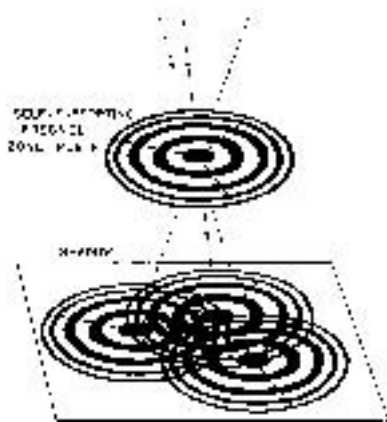


Fig. 40. When Mertz & Young first proposed imaging with a coded mask in 1961 [125] the shadowgram was thought of as a hologram – every region of the projected image contains information of the entire object

multiplied by the number of pinholes without the angular resolution being lost. However, the multiple images projected by this mask now necessitate computer algorithms to reconstruct the emitting object from the intensities measured by the position sensitive detector.

Image Reconstruction – Encoding

Figure 39 schematizes the principle of coded mask imaging: a source within the field of view casts a shadow of the mask on to the detector plane where the two-dimensional intensity pattern of the modulated flux is measured. The position of the shadow pattern allows the location of the source to be found while the size of the projected mask elements determines the source distance. In astronomy, the size of the shadows on the position sensitive detector will equal to the sizes of the mask elements since sources are virtually at infinity. In nuclear medicine and tomography of X-ray emitting plasmas, however, coded mask techniques are also used to extract depth information for volumetric object reconstruction.

The intensity measured by the position sensitive detector can be expressed as a two-dimensional matrix $D_{i,j}$ (the shadowgram) presenting the number of interactions registered in the detector element i, j . The encoding process becomes

$$D = S * A + B, \quad (69)$$

where $S_{i,j}$ is the matrix of the source distribution, $A_{i,j}$ the matrix of the coded mask, and $B_{i,j}$ the background noise matrix representing all contributions not modulated by the aperture. The aperture transmission function $A_{i,j}$ is 1 for transparent mask elements, and 0 for opaque elements. The signal $D_{k,l}$ in a detector element k, l can be written explicitly as

$$D_{k,l} = \sum_{i,j} S_{i,j} \cdot A_{i+k,j+l} + B_{k,l}. \quad (70)$$

The encoded matrix D will have no resemblance with the source distribution S . Some of the techniques that can be employed to reconstruct a conventional image are described below.

Image Reconstruction – Decoding

Mertz and Young (see above, [121]) propose a direct optical reconstruction technique for shadowgrams obtained with their Fresnel zone plate mask pattern. The photographic shadowgram is reduced in size in order to bring the focal length of an individual zone plate to a dimension convenient for visible light. The source distribution is then reconstructed by diffraction of coherent visible light at this reduced shadowgram (or hologram), with a monochromatic point source acting as reference beam.

Figure 41 shows Mertz and Young's demonstration of the principle using visible light. A number of illuminated pinholes simulate the n stars (upper

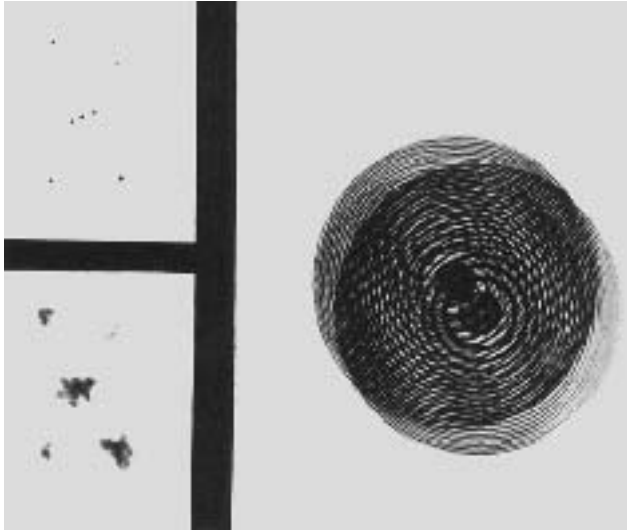


Fig. 41. “Illustrative sample of optical Fresnel transformation”, Mertz & Young, 1961 [125]

left), a Fresnel zone plate as shown in Fig. 40 casts n distinct shadows (right) – this is the hologram. The reconstructed image (lower left) is obtained by diffraction from a reduced copy of this hologram.

Today’s detector and data acquisition systems produce digital information that favor computer algorithms for data analysis. Usually, deconvolution techniques convolve or correlate the encoded matrix D with a decoding array G (also called postprocessing array). The reconstructed source distribution S' can be expressed as,

$$S' = D * G , \quad (71)$$

or, in terms of direct array elements

$$S'_{i,j} = \sum_{k,l} D_{k,l} \cdot G_{i+k,j+l} . \quad (72)$$

Substituting the encoded matrix D (69) results in

$$S' = (S * A) * G + B * G . \quad (73)$$

In order to preserve the object features within the resolution of the system, the choice of the decoding matrix G should be such that $A * G$ is as close as possible to a delta function. For

$$A * G \equiv \delta , \quad (74)$$

Equation (73) reduces to

$$S' = S + B * G . \quad (75)$$

The source is thus reconstructed with the exception of a background term. For a further discussion of coding and decoding coded mask telescopes see e.g [128, 129].

Design Considerations for Coded Mask Systems

Optimal performance of a coded-mask camera requires that every sky position is encoded on the detector in a unique way, and as different from each other as possible. Optimal designs of a wide variety of mask patterns are discussed in the literature (see eg [130]). Most coded aperture telescope designs incorporate so called uniformly redundant array (URA) mask patterns for their apertures. URA mask patterns have autocorrelation functions (ACF) which are δ -functions so that the off-axis image response is constant, thus minimizing imaging systematic noise (sidelobes) due to unequal representations of spatial frequencies. With the high background conditions prevailing in gamma-ray astronomy, maximum sensitivity is achieved if half of the mask elements are selected transparent, and half are set opaque.

The unequivocal determination of source locations, the recognition of multiple sources in the field of view, and the effects of spatial variations in detector background can be dealt with by post-processing techniques, which can be regarded as an intrinsic part of the instrument, and which are usually carried out on the ground (see [128]).

According to ratio of available independent detector elements/mask elements, unambiguous imaging may be aided by dithering the telescope pointing direction – hence introducing additional temporal modulation.

The global imaging characteristics of a coded mask instrument are summarized by simple laws determining the field of view and the angular resolution and of such instruments (Fig. 42).

The *fully coded field of view* α is defined as the angle for which the mask shadow covers the entire detector plane. It is given by

$$\alpha = 2 \arctan \frac{d - 2a}{2b} \quad \text{fully coded FOV} . \quad (76)$$

The properties of a URA mask pattern cited above only apply to sources that lie within this fully coded field of view. In general, the mask dimension d is therefore selected superior to the detector dimension a .

However, sources outside this field of view may also contribute photons to the detector if they are situated within the *partially coded field of view*

$$\beta = 2 \arctan \frac{a + d}{2b} \quad \text{partially coded FOV} . \quad (77)$$

Here, only part of the mask shadow is projected on the detector plane. Finally, the angular resolution $\Delta\theta$ is characterized by the angle subtended by one mask element at the detector,

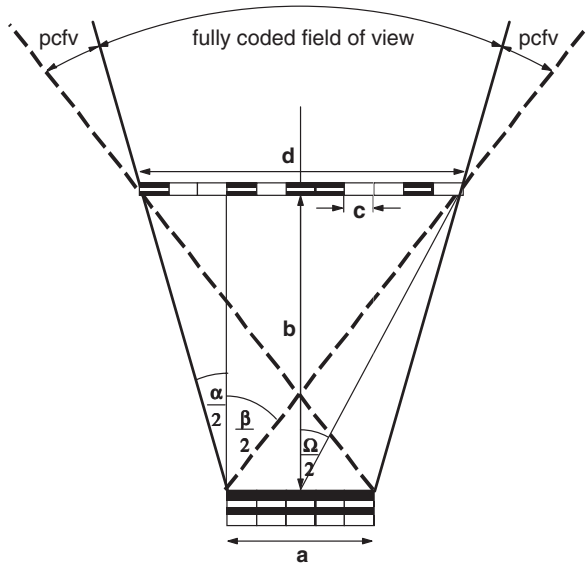


Fig. 42. Schematic side view of a coded mask telescope (pcfv: partially coded field of view)

$$\Delta\theta = r \arctan \frac{c}{b} \quad \text{angular resolution} . \quad (78)$$

Note that this relation only applies, however, if the positional resolution of the detector becomes small compared to the mask elements size c . In many real cases, where the detector resolution is matched to the mask element dimension, the angular resolution is worse by a factor $\sim\sqrt{2}$.

In the case of thick masks used e.g. for MeV observations, the “open” elements of the mask might actually show a reduced area of the sky, due to the thickness of the mask. This effect is called “vignetting”. It is inexistent for a source on the telescope axis and increases for source directions towards the edge of the field of view.

Advantages – Disadvantages

A principal advantage of coded mask telescopes over instruments using temporal modulation is the fact that they observe source and background *simultaneously*. This is not only the key for observing compact galactic and extragalactic sources that all show strong variability on a vast palette of timescales. The simultaneous observation of source *and* background is also the foremost requirement for correct background monitoring and subtraction in space environment, where the background generally varies rapidly.

A coded mask instrument uses a limited range of spatial frequencies to encode the signal from the sky. Its is therefore not surprising that such a

telescope performs well only over limited range of angular scales. Sources containing spatial structures smaller than the angular resolution will not be resolved while larger extended objects will be observed at reduced sensitivity.

As in all other telescopes using geometric optics, a major drawback of coded mask telescopes is that the source photons are spread over the entire detection plane, hence the entire detector volume contributes to the instrumental background noise.

SIGMA

On December first 1989 the first satellite borne coded mask telescope SIGMA was launched on board of the Soviet GRANAT spacecraft in a highly eccentric orbit. SIGMA operated flawlessly from February 1990 and continued its in-orbit activities until October 1997 [97]. SIGMA is the result of a French collaboration between the CESR at Toulouse and CEA at Saclay. Its URA type coded mask consisted of 53×49 elements each $9.4 \text{ cm} \times 9.4 \text{ cm}$ in size, with an underlying basic pattern of 31×29 elements. Located 250 cm above the detector, its opaque elements were 1.5 cm thick blocks of tungsten. Its detection plane consisted of a thin NaI(Tl) Anger camera (1.25 cm thick, detection area 784 cm^2) as used in nuclear medicine, viewed by 61 hexagonal photomultiplier tubes (see Sect. 3.2, Fig. 27). Besides the energy deposit, the on-board electronics directly provided Cartesian coordinates of the interaction location in the detection plane. Beyond the area of the totally coded field of view of $4.7^\circ \times 4.3^\circ$ the half-sensitivity boundary in the partially coded field was a rectangle of $11.5^\circ \times 10.9^\circ$. “Spectral images” and “fine images” were simultaneously recorded by the instrument: “fine images” used precise localization (pixel size 1.6 arcmin) in four contiguous energy bands, whereas the “spectral images” have a two times larger pixel size however in 95 energy channels from 35 keV to 1.3 MeV. The large CsI(Tl) anticoincidence shield (19200 cm^2) was also used for the detection of gamma-ray bursts.

The imaging performance of SIGMA turned out to correspond almost exactly to the intrinsic properties of the telescope: the localization accuracy was of ~ 2 arcmin, while the angular resolution was ~ 13 arcmin.

SIGMA provided a data base of galactic sources in the energy range of a few tens of keV up to several hundred keV. Among the many beautiful results of SIGMA, possibly one of the most relevant scientific consequences of SIGMA is the fact that *all* the compact high energy sources observed are time variable.

INTEGRAL-SPI

The “fine spectroscopy/coarse imaging” concept of SPI is particularly appropriate for nuclear astrophysics since phenomena such as the diffusion of radioactive isotopes into the interstellar medium often lead to narrow lines emitted on a broad angular scale, whereas gamma-ray line emissions from violent compact objects are more likely to be spectrally broadened. SPI will

also further increase our understanding of compact objects – galactic and extragalactic – for example through the observation of spectral features such as cyclotron lines.

The *Germanium detectors* and *Cryostat* of SPI are described in the section on semiconductor detectors (Sect. 3.3).

The aperture system providing the imaging capabilities of the instrument is a coded mask 171 cm above the detector. The mask pattern is a hexagonal uniformly redundant array (HURA) with 127 elements. The 63 opaque mask elements are 3 cm thick tungsten hexagons, 6 cm center-to-center. They assure a signal modulation of more than 90% over SPI's energy range. In the 64 transparent elements the honeycomb mounting plate absorbs less than 10% of the signal at 50 keV.

The anticoincidence subsystem shields the detector assembly and defines a hexagonal aperture of about 24° FWHM. It consists of a large hexagonal container protecting the detector and two hexagonal collimator rings, both made from bismuth germanate scintillators. A total of 500 kg of BGO scintillators form the veto system...shielding 18 kg of Germanium detectors. The BGO thickness is equivalent to 5 cm in all shielded directions. It has been optimized in order to minimize the detector background: whereas photons are rejected more efficiently with a thicker shield, the internally produced background, mainly due to $n\beta$ activations, is enhanced in a more massive shield. A total of 191 photomultiplier tubes are optically coupled to the BGO blocks to detect their scintillation light. A sophisticated read-out electronics of the total shield counting rate allows the measurement of the arrival time of a gamma-ray burst with a 50 ms time resolution. A thin plastic scintillator placed below the tungsten mask reduces the background produced in the mask, particularly in the 511 keV line.

The detection of narrow lines is SPI's main scientific objective and is made possible by the excellent energy resolution of the Ge detectors: 2.35 keV FWHM at 1.33 MeV. The 511 keV line sensitivity for an on-axis point source in $T_{\text{obs}} = 10^6$ sec is $2.8 \cdot 10^5 \text{ ph} \cdot \text{cm}^{-2} \cdot \text{s}^{-1}$. During the galactic plane survey, a point source brighter than $2.2 \cdot 10^5 \text{ ph} \cdot \text{cm}^{-2} \cdot \text{s}^{-1}$ in the inner galaxy will be detected at more than 3 standard deviations. The performance estimates are based on a background model that has been verified by accelerator tests and balloon borne spectrometer data e.g. [131]. A cutaway view of the SPI telescope is shown in Fig. 43; more detailed descriptions of the instrument can be found in [132] and [133].

The coded mask, together with the detector plane, define an angular resolution of about 2.8° within a hexagonal fully coded field of view of $16^\circ \times 16^\circ$ (corner to corner). The partially coded field of view is $34^\circ \times 34^\circ$ (corner to corner) while the anticoincidence shield defines a hexagonal aperture of 25.7° FWHM (corner to corner). The point source location accuracy is 0.5° (90% confidence for 5σ source), it improves with source intensity and exposure

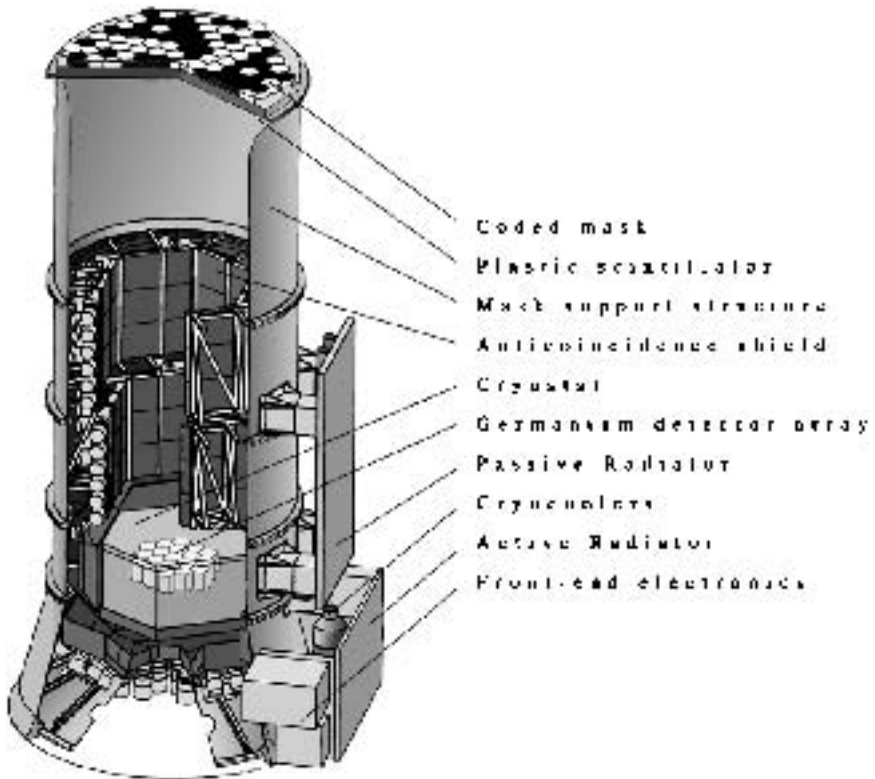


Fig. 43. Cutaway view of SPI and its subsystems

time. Dithering of the telescope pointing axis improves the imaging performance of SPI with its relatively small number of detector elements.

INTEGRAL-IBIS

The “fine imaging/coarse spectroscopy” concept of IBIS is particularly appropriate for the observation of point-like continuum sources: IBIS will study a wide variety of celestial objects ranging from the most compact galactic systems to extragalactic objects, with powerful diagnostic capabilities of fine imaging, source identification and spectral sensitivity in both continuum and lines. It will be able to localize weak sources at low energy to better than a few arcminutes accuracy, covering the entire energy range from 20 keV to 10 MeV. A cutaway view of the IBIS detector system is shown in Fig. 44; for a detailed description see Ubertini et al. 1997 [134].

The detector plane of IBIS features two layers, ISGRI and PICsIT: ISGRI the first is made of Cadmium-Telluride solid-state detectors (see description in Sect. 3.3) and the second of Cesium-Iodide scintillator crystals (see description in Sect. 3.2). The double-layer discrete-element design of IBIS allows the

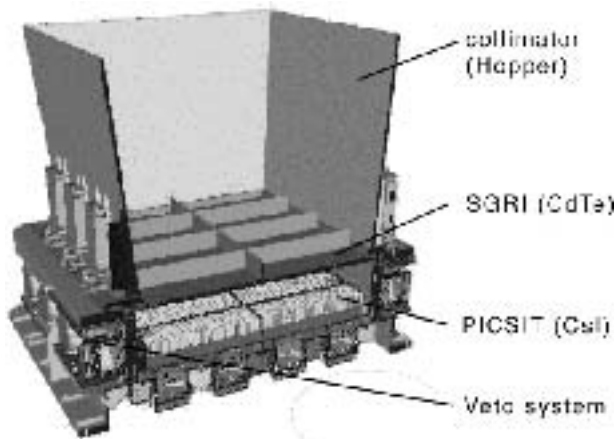


Fig. 44. Cutaway view of the IBIS detectors

paths of interacting photons to be tracked in 3D if the event involves detection in both ISGRI and PICSIT. The application of Compton reconstruction algorithms to these types of events (between few hundred keV and few MeV) allows an increase in signal to noise ratio by rejecting events which are unlikely to correspond to celestial photons (photons outside the FOV). Also, above a few 100 keV, Compton scatter events provide a principle for gamma-ray polarization studies.

An active Bismuth Germanate (BGO) veto shield surrounds the detector planes in the rear and on the sides up to the ISGRI bottom level. Due to the 20 mm of BGO, the detector background from leakage through the shielding of cosmic diffuse gamma-ray background and gamma-rays produced in the spacecraft is reduced to less than the sum of all other background components. A system of passive collimators (tungsten, lead) between the detector stack and the mask limits the solid angle at low energies.

The tungsten mask is placed at a distance of 3.1 m above the ISGRI detector plane. With a thickness of 16 mm, the mask opacity is always larger than 65% throughout the entire energy-range. The coded pattern is a square, $1064 \times 1064 \text{ mm}^2$ in size. It is made up of 95×95 individual square cells of size $11.2 \times 11.2 \text{ mm}^2$. The cells form a modified uniformly redundant array coded pattern of 53×53 elements. The resulting imaging characteristics are a fully-coded FOV of 9° , a partially coded FOV extending to 30° , and an angular resolution of 12 arcmin.

4.2 Quantum Optics: Compton Telescopes

The total interaction cross section for gamma-rays has its minimum in the MeV domain – the nuclear energy range, from several hundred keV up to at least a few MeV. Consequently, instruments making use of modulating

apertures run into several problems: The efficiency of the signal modulation decreases, at the same time the background noise increases with respect to the signal due the growing importance of shield leakage and/or $n\beta$ activation. It is in this same energy range that interactions are dominated by the Compton effect. The idea to make use of the Compton effect instead of fighting it with thicker shielding and modulators has stimulated several groups and resulted in a distinct class of imaging instruments.

The development of the first Compton telescope for observations of celestial gamma-rays began at the Max-Planck-Institute (MPI) at Garching in the early seventies (Schönfelder et al. [135]). The project of a Liquid Xenon Compton Telescope was presented by Alvarez et al. [136] and Dauber and Smith [137] at around the same time – this type of instrument has remained one of the promises for nuclear astrophysics (Sect. 3.1 on Time Projection Chambers).

Similar instruments for neutron measurement had also been proposed and flown (Pinkau [138], White [139], Preszler et al. [140]) to deduce the energy and scattering angle of incident neutrons that elastically scatter off hydrogen nuclei. Imaging Compton telescopes for gamma-ray astronomy have subsequently been improved at the University of California at Riverside [141], by the MPI group [142] and at the University of New Hampshire [143]; the more recent developments are presented at the end of this chapter. Besides their use in imaging telescopes, Compton kinematics can also be used in modulating aperture systems (see Sect. 4.1) for background reduction. The coincidence signature from different segments or planes in the detector array allows the rejection of events which are unlikely to have entered via the instrument aperture. This technique has been applied in the MISO telescope of the Milano–Southampton collaboration [144] and is an operation mode used with the ISGRI and PICsIT detector planes of INTEGRAL-IBIS [134].

Principle of a Compton Telescope

The principle of measurement in a “classic” Compton telescope is illustrated in Fig. 45: An incident gamma-ray is identified by successive interactions in the two detector layers D_1 and D_2 . Compton scattering in the upper detector D_1 is favored when low Z material is chosen. Total absorption of the scattered photon in the lower detector can be expected when high Z materials are used for D_2 . The quantities measured for each gamma event are:

x_1, y_1	the location of interaction in D_1
x_2, y_2	the location of interaction in D_2
E_1	the energy deposited in D_1
E_2	the energy deposited in D_2

From x_1, y_1 and x_2, y_2 the direction X, Ψ of the scattered gamma-ray is obtained; the total energy deposit of the incident photon (energy E_γ) is

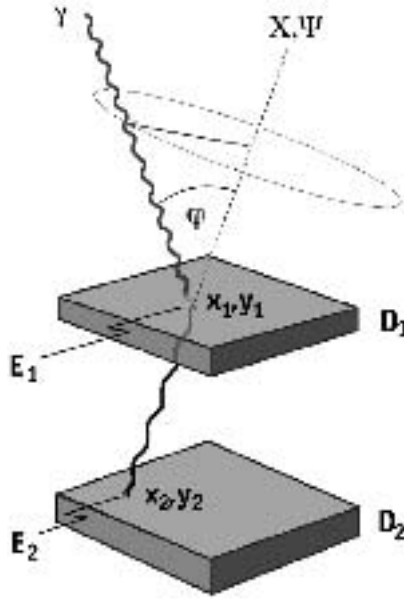


Fig. 45. The principle of a Compton telescope

$$E_{\text{tot}} = E_1 + E_2 . \quad (79)$$

The scatter direction X, Ψ , together with the amounts of energy deposited in the two interactions can be used to reconstruct the arrival direction of the gamma-ray. The Compton equation (80) allows to express the scatter angle φ as a function of the energy-deposits E_1 and E_2 :

$$\cos \bar{\varphi} = 1 - \frac{m_e c^2}{E_2} + \frac{m_e c^2}{E_1 + E_2} , \quad (80)$$

where $m_e c^2$ is the rest energy of the electron; the initial momentum of the electron being neglected here.

If E_1 and E_2 are measured without systematic errors ($E_{\text{tot}} = E_\gamma$), the derived scatter angle $\bar{\varphi}$ equals the true Compton scatter angle φ . The arrival direction of the incident gamma-ray can then be confined to lie on a cone-mantle with axis X, Ψ and opening angle $\bar{\varphi}$ (Fig. 46). The projection of this cone results in a circle on the sky that is generally called the “event circle”. If the direction of the recoil electron is not tracked in the D_1 detector layer (which was the case for GRO-COMPTEL and still is for many state-of-the-art designs), the azimuthal information of the incident photon is lost, and no further information on the circle can be deduced from the measured parameters. Consequently direct imaging is impossible for “classic” Compton telescopes and the image reconstruction process is handicapped by the lack of information on the scatter angle of the incident photon.

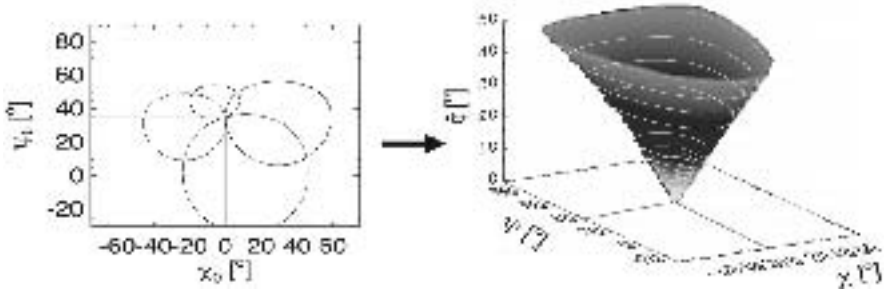


Fig. 46. *Left:* event circles originating from a single point source at X_0, Ψ_0 at $0^\circ, 35^\circ$. *Right:* the three-dimensional data space $(X, \Psi, \bar{\varphi})$ for events from a source at the position X_0, Ψ_0 ; adapted from Oberlack [146]

Data Space and Image Reconstruction

For the analysis in a given energy band (e.g. a gamma-ray line), the data of a Compton telescope are generally arranged in a three-dimensional data-space, spanned by the Compton scatter angle $\bar{\varphi}$ and the scatter direction X, Ψ (Fig. 46). A source distribution I expressed in celestial coordinates (l, b) , emitting gamma-rays of a given energy E_γ can be converted to the expected number of photons in a cell of the data space $(X, \Psi, \bar{\varphi})$:

$$e(X, \Psi, \bar{\varphi}) = b(X, \Psi, \bar{\varphi}) + g(X, \Psi, \bar{\varphi}) \int_l \int_b I(l, b) A(l, b) f(X, \Psi, \bar{\varphi} | l, b) . \quad (81)$$

Here, $A(l, b)$ is the effective exposure of the D_1 detector layer, $f(X, \Psi, \bar{\varphi} | l, b)$ is the instrumental Point Spread Function (PSF) for a hypothetical infinite D_2 layer, $g(X, \Psi, \bar{\varphi})$ is the probability that the trajectory of a photon scattered in D_1 actually intersects D_2 , and $b(X, \Psi, \bar{\varphi})$ is the instrumental and environmental background. The PSF generally depends on the selected energy interval and on the energy of the incident photon.

Further parameters (e.g. pulse shape, time of flight, magnetic cutoff rigidity, aspect angles of telescope versus atmosphere or orbit, etc.) may be necessary to maintain a maximum of information on the background for optimal image reconstruction.

Various methods for the reconstruction of the source distribution from Compton data have been proposed and tested.

Deconvolution Methods – Backprojection

A direct backprojection of an event with the measured parameters $\bar{\varphi}, X, \Psi$ can be realized by the event circle centered on X, Ψ and with a radius of $\bar{\varphi}(E_1, E_2)$ – see Fig. 46. A way to identify a point source within the field of view of a Compton telescope consists of measuring the density of the event circles

for each bins of a skymap. However, in the case of real Compton telescopes a number of effects will severely distort the image.

In the $(\vec{\varphi}, X, \Psi)$ data-space, a point source is represented by a cone centered on the source position. The image reconstruction consists in searching for the source-cones within the data-space. Note that the differential Compton scattering cross-section (15) depends on the polarization of the incoming photons, hence a Compton telescope may be used as polarimeter. In the three-dimensional data-space, a polarized source would then manifest as an asymmetrically populated cone.

Deconvolution Methods – Bayesian Inference

Image reconstruction of the multidimensional dataspace on a two-dimensional skymap I is done by solving (81) for $I(l, b)$. While direct inversion is in principal possible, it is not advisable since the measurement noise propagates uncontrolled into the reconstruction, leading to numerous artefacts and spurious sources in the skymaps. As with other instrument categories that give rise to inverse problems (modulating apertures, but also optical telescopes or radio telescopes), much improved results with respect to direct inversion methods are obtained by means of *Bayesian image reconstruction* which uses Bayes' Theorem

$$P(I|D) \propto P(D|I)P(I) . \quad (82)$$

to derive an expression for the probability $P(I|D)$ of an image I given the measurement D . The first term on the right-hand side, $(PD|I)$, is a goodness-of-fit quantity, measuring the likelihood of the data given a particular image. The second term, $P(I)$, called the “image prior”, expresses the plausibility of a particular image prior to the measurement. The application of different “Bayesian reconstruction procedures” essentially only differ in the choice of the image prior: e.g. the maximum entropy method, or the Richardson–Lucy reconstruction.. For a comprehensive treatment of inverse problems applied to Compton telescopes see Knödlseder [145].

In the maximum entropy method which is widely used with Compton telescopes, the image prior is given by $P(I) \propto \exp(\alpha S)$ where the entropy S measures the deviation of the image I from a default image M ; α is an adjustable parameter that is used to weight the relative importance of the likelihood term $(PD|I)$ and the image prior $P(I)$. For $\alpha \rightarrow \infty$ the probability $(PI|D)$ is dominated by the image entropy, hence the reconstruction tends towards the default image. For $\alpha \rightarrow 0$, the entropy is practically “switched off” and the reconstruction is determined by the goodness-of-fit term $(PD|I)$.

Design Considerations

Detector Coincidence

While Compton scattering in D_1 is favored when a low Z material is chosen, total absorption of the scattered photon in D_2 is most likely when high Z

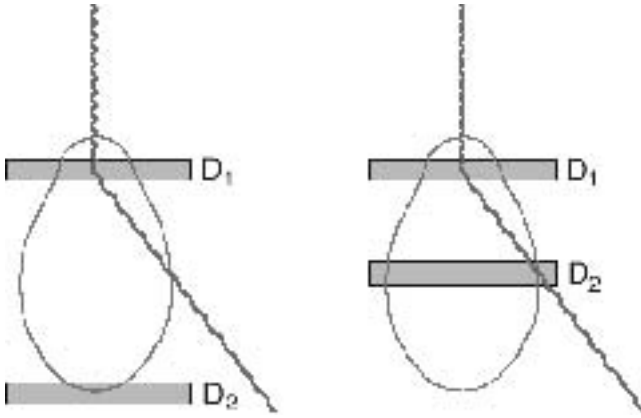


Fig. 47. Efficiency improvement by compact geometry

materials are used (see e.g. Fig. 4). Since the $D_1 \wedge D_2$ coincidence condition discriminates against most of the internal $n\beta$ events, a Compton telescope has an extremely low background. On the other hand, this coincidence condition, at the same time, causes the detection efficiency to be relatively low.

Time of Flight

The residual background can be reduced dramatically if the time-of-flight (TOF) between the two detectors layers is measured. It has been found that for GRO-COMPTEL, the dominant fraction of the instrumental background is due to “upward” scattered events, most likely originating in the massive GRO spacecraft (Fig. 48). Use of the time of flight (measured with an accuracy 1.5 ns in COMPTEL) has proven to efficiently eliminate the photons moving from D_2 to D_1 , reducing the background by 90% to 95% [147]. Compton telescopes designs that are not using TOF (e.g those based on solid state detectors, where rise-times are long with respect to the TOF) may have to use veto shields to suppress “upward” scattered events.

Pulse Shape Discrimination

If the pulse shape of the interaction in D_1 can be measured, it is possible to discriminate between neutron and photon events. The identified neutron interactions can then be used for further background reduction, alternatively, they may be analyzed for the study of the neutron component of the solar wind, for example.

Geometry Options

In spite of their large detector surfaces, the effective area of “classic” Compton telescopes has been rather modest. For example, GRO-COMPTEL’s

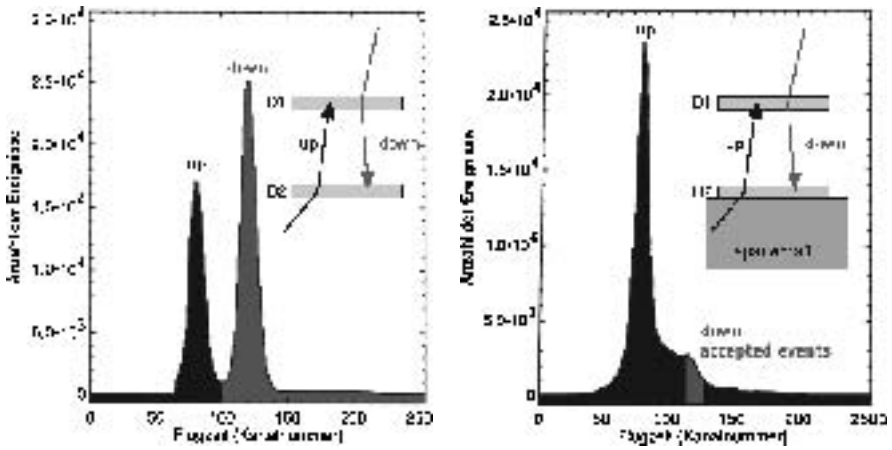


Fig. 48. Two time of flight spectra of GRO COMPTEL. Abscissa: TOF, the channel width is 0.25 ns. Ordinate: number of events detected. The distance D1-D2 being 1.5 m, the time of flight between the detectors is 5 ns. *Left:* ground calibration data, *Right:* flight data – accepted events are those in channels 115–130 [146]

D₁ and D₂ detectors had large geometric surfaces (more than a square meter, combined), however the effective area of the telescope was only a few tens of cm². The main cause for the low efficiency is the “lateral loss” of scattered events, principally due to the unfavorably large distance between D₁ and D₂ with respect the detector dimensions. A more compact geometry will result in an increased detection efficiency: On the one hand more spurious D₁ $\wedge$ D₂ coincidence events will be measured, on the other hand, the field of view of the telescope is larger, leading to a higher exposure time for a certain region of the sky during a survey (multiplexing advantage).

However, when bringing the detectors closer, a corresponding improvement of the spatial resolution of the detectors is required if the angular resolution is to be maintained. Current studies of advanced Compton Telescopes focus on highly segmented solid-state detectors (Si, Ge or CdTe in either stripped or pixelised forms) with sub-mm resolution, resulting in telescope designs with several 10^5 readout channels. A further drawback of a compact configuration is the difficulty of measuring photon flight times between D₁ and D₂, leading to much higher background count rates than in the “classic” configuration (see TOF below). Without shielding against “upward” events, the resulting loss in sensitivity may more than neutralize the gain due to the compactness.

Angular Resolution Through Energy Resolution

As the energy and angular resolution of a Compton telescope are related through (80), the energy resolution affects the angular resolution. The angular resolution is therefore composed of two terms: ΔX , $\Delta \Psi$ the precision

in the measurement of the scatter direction, and $\Delta\bar{\varphi}$, which is related to the errors in the measurements of E_1 and E_2 . An estimate of $\Delta\bar{\varphi}$ is obtained by differentiating (80)

$$\Delta\bar{\varphi}(E_1, E_2) = \sqrt{\frac{\alpha_{\text{tot}}^4(\Delta E_1/m_0c^2)^2 + (\alpha_{\text{tot}}^2 - \alpha_2^2)(\Delta E_2/m_0c^2)^2}{1 - (1 - \alpha_2 + \alpha_{\text{tot}})^2}}. \quad (83)$$

Here $\alpha_2 = m_0c^2/E_2$ and $\alpha_{\text{tot}} = m_0c^2/E_{\text{tot}}$. Achieving high angular resolution is therefore necessarily tied to high energy resolution.

Tracking the Recoil Electron

The possibility of tracking the direction of motion of the recoil electron would restrict the deduced possible arrival directions of an incident gamma ray to within a small arc on the sky, rather than a complete ring. The photons from a particular source would then occupy only a small volume of the data-space, and the signal-to-noise ratio, and hence the sensitivity, would improve. Electron tracking also allows kinematic rejection of various background components, such as events which first interact in D_2 , as well as events which are not completely absorbed by the detector. This would substitute to some extent for the time-of-flight measurement in “classic” Compton telescopes. The use of stripped low-Z semiconductor detectors (such as Si wafers) has recently opened up the possibility of performing this tracking.

COMPTEL

As one of the four experiments on NASA’s Gamma Ray Observatory mission, COMPTEL was in operation from 1991 until 2000, performing the first complete survey of the MeV γ -ray sky. COMPTEL used conventional scintillation detectors, covering energies from 1–30 MeV and a field-of-view of about 1 steradian. A cutaway view of COMPTEL is shown in Fig. 49.

The D_1 detector layer consisted of seven Anger camera cells (NE213), each module being 27.6 cm in diameter, 8.5 cm thick, and viewed by eight photomultiplier tubes. While the sum of the absolute pulse heights gave the energy E_1 , the relative strengths of the pulse heights determined the location of the interaction within the module to within ~ 2.3 cm (1σ). The energy resolution of the D_1 detector modules was 12.5% at 1 MeV; the total area of the upper detector was 4188 cm².

The 14 NaI Anger cameras in the D_2 detector layer were cylindrical NaI (Tl) blocks of 7.5 cm thickness and 28 cm diameter, which were mounted on a supporting baseplate. Each block of NaI was viewed from below by seven photomultiplier tubes. The total geometrical area of the lower detector is 8620 cm². The energy resolution of the D_2 detector modules was 8.3% at 1 MeV, the interaction location is determined with an accuracy of 1.5 cm (1σ).

Each detector layer was entirely surrounded by a thin anticoincidence shield of plastic scintillator which rejects charged particles. The signals from

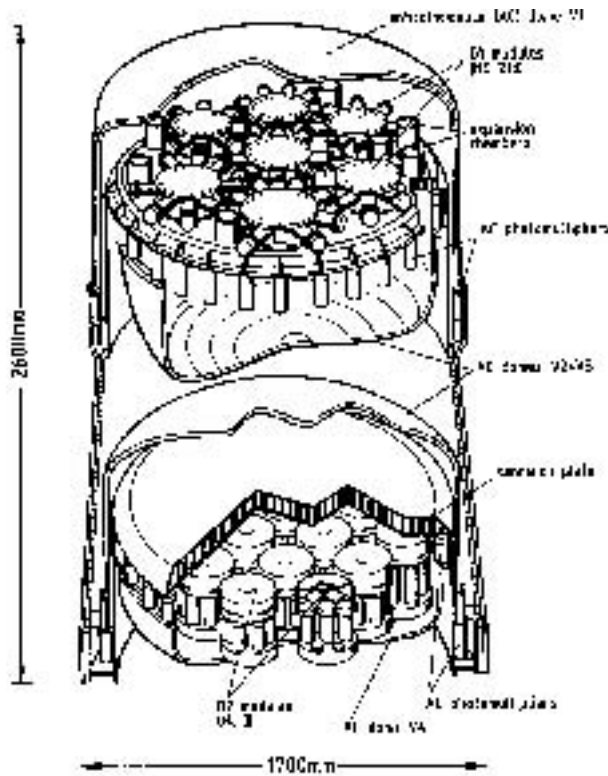


Fig. 49. Cutaway view of GRO-COMPTEL (from [94])

these veto domes, the Time Of Flight (TOF) and the Pulse-Shape Discrimination (PSD) together with the energy, scatter angle and earth-Horizon Angle (EHA) are used to reduce the residual background.

The resolution in φ (83) together with the 1.5 m D₁-D₂ separation, provided an angular resolution of 1°–2°, depending on energy. COMPTEL had an effective area of 10–50 cm² depending on energy an event selection criteria. A complete description of the instrument is given by Schönfelder et al. [94].

Designs for an Advanced Compton Telescope

MEGA

The Medium Energy Gamma-Ray Astronomy telescope is a project for the next generation gamma-ray telescopes for the energy range between 400 keV and 50 MeV [148].

MEGA records and images gamma-rays by completely tracking Compton and pair creation events in a stack of double sided Si-strip track detectors surrounded by a pixelated CsI calorimeter (Fig. 50). The D₁ detector, the

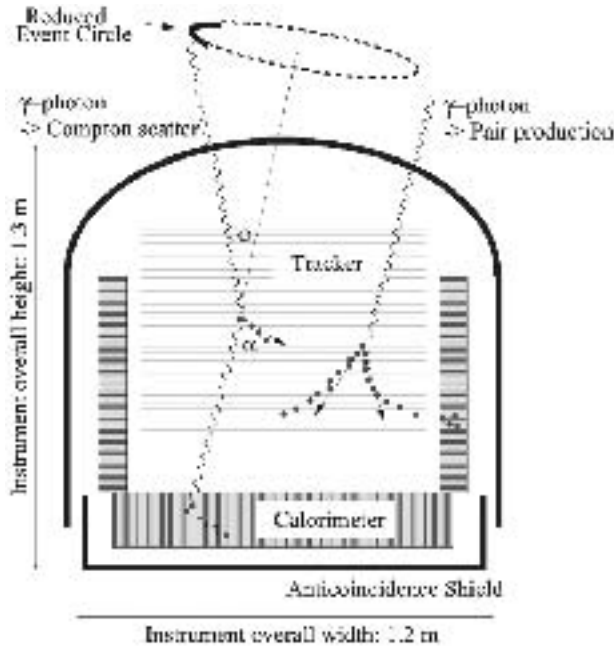


Fig. 50. Schematic design and detection principle of the MEGA telescope [148]

“tracker”, is made from 32 layers of double-sided Si wafers. Each layer is composed of a 3×3 array of $500 \mu\text{m}$ thick silicon wafers, each $6 \times 6 \text{ cm}^2$ in size and fitted with 128 orthogonal p and n strips on opposite sides ($470 \mu\text{m}$ pitch). The biased strips are read out by 128-channel ASICs, creating a total area of $19 \times 19 \text{ cm}^2$ position-sensitive area. For incident energies above about 2 MeV the recoil electron usually receives enough energy to penetrate several Si-layers, allowing it to be tracked. This constrains the incident direction of the photon to a “reduced event circle”, reducing the background.

The D_2 , or “calorimeter”, is a CsI matrix 8 cm deep on the bottom and 4 cm on the side walls. The cross-section of the CsI bars is $5 \times 5 \text{ mm}$, they are read out with Silicon PIN-diodes and low-noise, self-triggering front end electronics. MEGA will have an effective area of $\sim 100 \text{ cm}^2$, a large field of view of about 130° , angular resolution of $\sim 2^\circ$, and energy resolution of $\sim 8\%$ (both FWHM at 2 MeV).

MEGA should operate in a low-inclination LEO (height $\sim 500 \text{ km}$). The telescope with its large field-of-view is best used in a zenith-pointing scan mode to continuously monitor a large fraction of the sky for transient sources and to accumulate exposure for galactic and extragalactic sources. MEGA aims to improve the sensitivity for astronomical sources by at least an order of magnitude with respect to past instruments. Its key science objectives are the investigation of cosmic high-energy accelerators, nucleosynthesis sites

with γ -ray lines, and the mapping of large-scale structures in the Galaxy and beyond.

A prototype of the detectors, tracker and calorimeter, have been integrated on a support structure, which permits the telescope to be tested in beam calibrations and on a balloon payload.

TIGRE

The Tracking and Imaging Gamma Ray Experiment is a mission concept proposing the use of solid state strip detectors to act simultaneously as a Compton telescope and a low energy pair detector. As such, TIGRE will observe with significant sensitivity from 0.3–100 MeV. Its D₁ detector consists of 50 (or more) layers of double sided silicon strip detectors (SSDs). These detect charged particles passing through the detector, and can give the x and y coordinates of the interaction location with a resolution <1 mm. The D₂ layer consists of 5–10 layers of cadmium zinc telluride (CZT) strip detectors. The CZT is arranged to form a five-sided box surrounding D₁. The fine pitch (<1 mm) of the CZT strip detectors allows high spatial resolution to be attained without a large (> 1m) separation between D₁ and D₂. This results in a more compact instrument allowing for more coincidences between D₁ and D₂, improving efficiency in the Compton regime by a factor of 5–10 over “classic” Compton telescope designs. As the dependence of the Klein–Nishina formula on photon polarization is most pronounced for large scatter angles, TIGRE will also be a highly effective gamma ray polarimeter. The use of SSDs gives the possibility of tracking the Compton recoil electron (see “tracking the electron” above).

NCT

The Nuclear Compton Telescope is a Germanium-based prototype design for the Advanced Compton Telescope [149]. The heart of NCT is an array of twelve crossed-strip GeDs with 3-D position resolution. Each of the 15-mm thick planar Ge detectors has an active area 76 mm $\times$ 76 mm. Orthogonal 2-mm electrode strips on the opposite faces, combined with signal timing, provide full 3-D position resolution to 2 mm. Timing techniques for measuring the third dimension (depth) have been verified in the laboratory. The GeDs will be housed in a common cryostat, attached to a liquid nitrogen dewar. The Ge detector array is enclosed by a 5-cm thick active BGO anticoincidence shield. A 10-cm thick CsI front shield collimates the FOV to 40°. NCT has been designed for long duration balloon flights in order to study nuclear line emission and polarization.

ATHENA

The original ATHENA concept [150] is based on D₁ and D₂ layers using Germanium planar strip detectors providing 2–3 keV spectral resolution and

spatial resolution of ~ 2 mm. Such detectors, typically $5\text{ cm} \times 5\text{ cm} \times 1\text{ cm}$, are available today and might be integrated into large panels in the future. The ATHENA concept foresees a 1 m^2 D₁ layer consisting of one panel, and a 1 m^2 D₂ layer of four panels, each panel containing 400 Ge strip detectors. Figure 51 shows a schematic diagram of a solid state Compton telescope for low energy gamma-rays. In the Compton mode ($300\text{ keV} - 10\text{ MeV}$) such an instrument can achieve angular resolutions of $0.2^\circ - 0.3^\circ$ within a field of view of typically one steradian, and a narrow line sensitivity of a few $10^{-7}\text{ ph} \cdot \text{cm}^{-2}\text{s}^{-1}$ above 1 MeV.

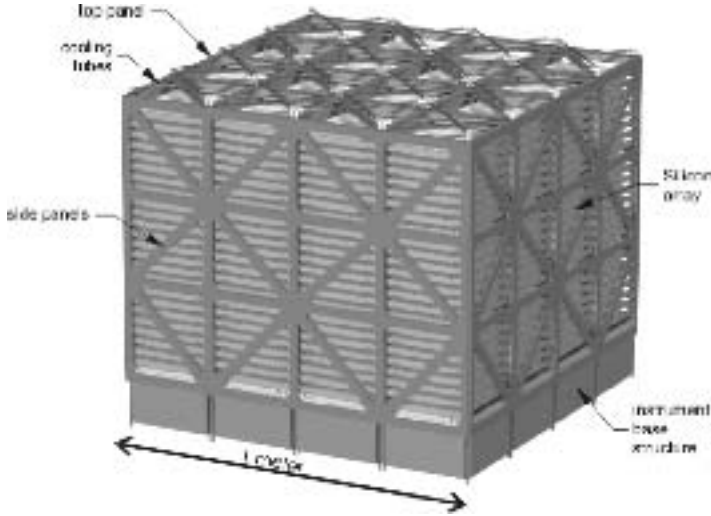


Fig. 51. A possible configuration of an Advanced Compton Telescope – an ATHENA-type Compton telescope equipped with thick lithium drifted silicon detectors [151]

A more recent baseline [151] for an ATHENA-type instrument proposes thick lithium drifted silicon detectors, measuring again roughly $1\text{ m} \times 1\text{ m}$ in frontal area. The individual detectors are $\sim 7\text{ mm}$ thick, and measure $10 \times 10\text{ cm}$ in area using newly emerging technology in crystal growth and lithium drifted silicon (Si(Li)).

LXeGRIT

To demonstrate the operation and performance of a Liquid Xenon Time Projection Chamber (see description in Sect. 3.1) with gamma-rays in the near space environment, the balloon-borne payload LXeGRIT has been flown in a series of stratospheric balloon flights in the period 1999–2001 [152]. The experience with the LXeGRIT prototype have lead to an understanding of

the performance expected from the instrument, they also were useful in identifying weaknesses of the current TPC design and signal readout. The full science potential of a next-generation LXe-based telescope should be tested on future balloon flights.

4.3 Wave Optics: Focusing Telescopes

Since the wavelength of nuclear gamma-ray photons is two to three orders of magnitude shorter than the distance between atoms in solids, astrophysicists have been used to accept that it is “impossible to reflect or refract gamma-rays”. Consequently, present types of telescopes for nuclear astrophysics are based on inelastic interaction processes: most of the instruments are based on geometrical optics (Sect. 4.1) or quantum optics (Sect. 4.2).

Because the collecting area of such systems is equal to the detector area, nuclear astrophysics has come to a mass-sensitivity impasse where “bigger is not necessarily better”. Improving the sensitivity of an instrument can usually be obtained by a *larger* collection area – in the case of classical gamma-ray telescopes this can only be achieved by a *larger* detector surface. Yet, since the background noise is roughly proportional to the volume of a detector, a larger photon collection area is synonymous with higher instrumental background. For such “classic” gamma ray telescopes, the sensitivity is thus increasing at best as the square root of the detector surface.

The ensuing mass/sensitivity dilemma can ultimately only be overcome by concentrating gamma-rays, taking advantage of the phase information of the gamma-ray photons: A gamma-ray optical system is designed to concentrate radiation – by surface reflection, diffraction and/or refraction – collected from a large area into a small focal spot. This allows a modest size, well shielded detector to register a much larger signal than it would have intercepted if it was exposed to the radiation field directly.

Table 8 lists the concepts, main instrumental features, and energy range of various focusing systems for high-energy photons. While the grazing incidence techniques used in X-ray astronomy will be reviewed briefly in the following paragraph, this chapter mainly focuses on the concentration of gamma-rays: diffraction in Fresnel, Bragg- and Laue-lenses.

Grazing Incidence

Total External Reflection

In the gamma-ray domain, the refractive index n of any available material is very close to unity. However, since $n < 1$, efficient reflection is nevertheless possible for very small incidence angles. Total external reflection takes place for angles $\theta < \theta_c$, the *critical grazing angle* (see Sect. 2.4), (47). The critical grazing angle decreases with the square root of the electron density n_e of

Table 8. Focusing systems for high-energy photons

Wolter telescopes	a) total external reflection	$\sim 0.1 - 12$ keV
	b) multilayer mirror interference	$\sim 20 - 100$ keV
Lobster eye telescopes	total external reflection	$0.1 - 3.0 (+)$ keV
Capillary Concentrators	total external reflection	$1 - 60$ keV
Kirkpatrick/Baez optics	total external reflection	
Bragg-lenses	Bragg (surface) diffraction	$10 - 200$ keV
Laue-lenses	Laue (volume) diffraction	200 keV – 2 MeV
Fresnel lenses	refraction/diffraction	1 keV – 10 MeV

a material and with increasing photon energy. For incidence angles larger than θ_c , reflectivity drops steeply with increasing angle.

While telescopes based on total external reflection are widely used in X-ray astronomy, mostly by using nested mirror-arrays of paraboloids and hyperboloids in Wolter-I configuration, the technique becomes much less practical at gamma-ray energies. Whereas at 1 keV the critical angle is of the order of 1 degree for the most commonly used reflecting materials like gold or nickel (high Z materials), grazing angles at gamma-ray energies would be more than two orders of magnitude smaller. As a consequence, the focal length becomes extremely long, and more cumbersome, and the projected effective area of a given mirror surface becomes very small, not to speak of the required surface smoothness which is presently beyond technical feasibility, at least over the large surfaces that would be required.

At present, the highest energy focused by this technique is 45 keV, and has been achieved during a balloon flight of the HERO payload in 2001 using iridium-coated mirrors [153]. Apart from the Wolter-I geometry, which is particularly adapted for imaging and spectroscopy in relatively narrow fields of view, total external reflection is used in Lobster eye geometry (large field of view surveys, [154]), Capillary Concentrators [155], and Kirkpatrick/Baez geometry [156].

Multilayer Mirrors

In order to cover energies up to ~ 100 keV – and maybe even beyond – the above mentioned geometries for grazing incidence telescopes can be used with multilayer coatings as mirror surface. Presently a number of Multilayer Mirrors are under development for use in Wolter telescopes. Although the reflectivity of a single mirror surface at incidence angles greater than the critical angle θ_c is very small, it is not zero, hence a small fraction of the radiation is reflected at reasonably large incidence angles.

Multilayers coatings consist of alternating layers of high and low index n of refraction materials: The reflection by a multilayer mirror is described by the constructive interference of the reflections at all low-high n interfaces

This result in a sizable total reflectivity of the system. Similar to the Bragg-diffraction in crystals (see next section), the reflections have to be added with the correct phase relationship, leading to a boundary condition that relates incidence angle θ_1 , layer thickness d_1 and wavelength λ

$$2d_1 \sin \theta_1 = n\lambda, \quad (84)$$

where n , the order of the reflection is an integer ≥ 1 (multilayers are most commonly used in the first order, $n = 1$). Consequently, the response of so called *Uniform Period Multilayers* results in a narrow energy-bandpass. High reflectivity in a broad energy-bandpass can be achieved with graded multilayer coatings, here the film thickness d is varied over the stack. These *Extremely Broad Band (EBB) Multilayers* with reflectivities over bandpasses of >20 keV are being intensely developed by several groups [157–159]. The materials for the reflector/spacer coatings are selected for their different indices of refraction and for minimum absorption – presently considered material combinations are W/Si, W/C, Ni/C, and Pt/C. A first balloon flight using a 20–40 keV bandwidth mirror utilized at about $\sim 0.2^\circ$ incidence angle has been performed by the InFOCuS project in 2001 [160]. Development work for the hard X-ray telescope on the Constellation-X satellite has indicated potential up to around 200 keV [161] for this technique.

Crystal Diffraction Lenses

Diffraction lenses use the interference between the periodic nature of light and a periodic structure such as the matter in a crystal. The physics of scattering in crystals is discussed in Sect. 2.3 (Coherent scattering from bound electrons). An elementary derivation of the Bragg condition, $2d \sin \theta_B = n\lambda$ see (28 ff) has been given in Fig. 10; it is assumed that the incident waves are reflected by the parallel planes of the atoms in the crystal. (θ_B is the Bragg angle, n is an integer denoting the diffraction order, λ is the wavelength of the gamma-ray being diffracted, and d is the spacing between the crystalline planes used in the diffraction process). There is constructive interference if the optical path difference between neighboring paths is a multiple of the wavelength $n\lambda$.

Bragg- vs. Laue Geometry

The Bragg condition implies that higher incoming photon energies require smaller Bragg angles. At gamma-ray energies, Bragg angles are generally less than one degree. As shown in Fig. 52, reflection can be at the surface (so-called Bragg geometry) or the beam can pass through the crystal volume (so-called Laue geometry). The maximum efficiency for diffraction in the Bragg geometry is close to 100% (assuming no absorption).

A hard-X ray lenses operating in Bragg geometry using mosaic pyrolytic graphite crystals has been proposed [162]. The concentrator consists of 28

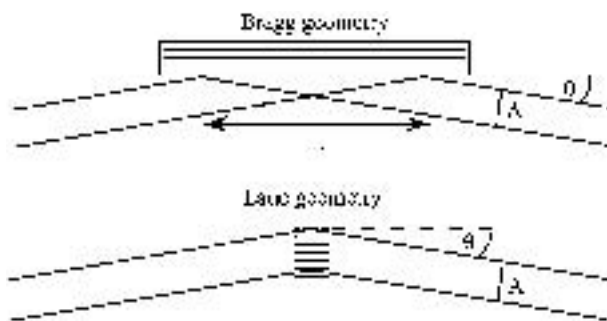


Fig. 52. (a) Bragg geometry (surface reflection) vs. (b) Laue geometry (volume reflection). In Bragg geometry, a crystal would need to have a length $L = A / \sin \theta_B$ to reflect a beam of cross-section A (from [174])

confocal parabolic mirrors. Each mirror is made up of small pieces of mosaic crystal with the diffraction planes parallel to the parabolic surface, which results in a broadband energy response. The outer diameter is 1.3 m, the focal length is 3.8 m. The effective area is 1000 cm^2 at 15 keV decreasing to 35 cm^2 at 100 keV. An angular resolution of a few arc minutes could be achieved. For a discussion of hard-X ray lenses operating in Bragg geometry see eg. [163].

For nuclear energies, Laue geometry is a more appropriate choice: due to the small Bragg angles at high energies, the crystal area in Bragg geometry becomes extremely long. At such energies, the crystal areas needed for Bragg type diffraction would be 100 times the area of crystals used with Laue diffraction: for a 1-cm beam and a Bragg angle of 1 degree, the crystal length $L = A / \sin \theta_B$ would be 57 cm! Laue geometry “only” allows maximum efficiencies of $\leq 50\%$ (assuming no absorption in the crystal). However, the attenuation due to the beam passing through the crystal becomes small at high energies, making Laue geometry possible. In the following, gamma-ray lenses using Laue geometry are discussed.

Laue Geometry Lenses

In a crystal diffraction lens, crystals are usually disposed on concentric rings such that they will diffract the incident radiation of a same energy onto a common focal spot (Fig. 53). A crystal at a distance r_1 from the optical axis is oriented so that the angle between the incident beam and the crystalline planes is the Bragg angle θ_{B1} . Its rotation of around the optical axis results in concentric rings of crystals. With the same crystalline plane $[hkl]$ used over the entire ring, the diffracted narrow energy band is centered on E_1 .

Two subclasses of crystal diffraction lenses can now be identified – *narrow* bandpass Laue lenses and *broad* bandpass Laue lenses.

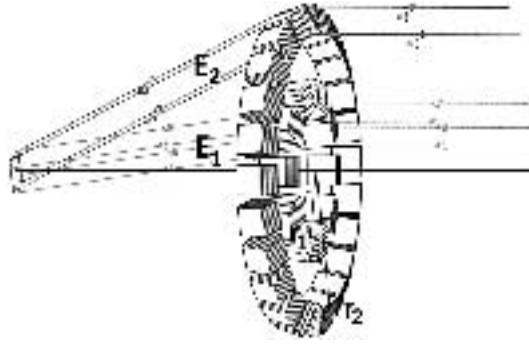


Fig. 53. The basic design of a crystal diffraction lens in Laue geometry

Narrow Bandpass Laue Lenses

Use a different crystalline plane $[hkl]$ for every ring in order to diffract photons in only one energy band centered on an energy $E_1 = E_2$. For a given energy E_1 , a ring with a radius $r_2 > r_1$ must reflect at an angle $\theta_{B2} > \theta_{B1}$ to concentrate the incident beam at a given focal distance. According to the Bragg condition, this is only possible if the crystalline plane spacing d_2 is smaller than d_1 or if a higher order is used. The ring radii are determined by the Miller indices $[hkl]$. For materials with a cubic unit cell (e.g. the face-centered cubic cell of copper, germanium or silicon), the ring radii in small angle approximation are proportional to $\sqrt{h^2 + k^2 + l^2}$. For a given focal distance f of the lens, r_i is the radius of ring “i”,

$$r_i = f \tan[2\theta_{Bi}] = f \tan \left[2 \sin^{-1} \left(\frac{n\lambda}{2d_i} \right) \right], \quad (85)$$

where n is the order of the diffraction process, d_i is the crystalline plane spacing of the “i” ring (see (27)) and λ is the wavelength of the radiation.

As the diffraction efficiency decreases with increasing diffraction order n , a crystal in an exterior rings will add less efficient area to the lens than a crystal on an inner ring. However, since the number of crystals increases with the ring-radius, all rings will usually contribute about the same amount of efficient area to the lens. Using larger and larger Bragg angles with increasing ring radius allows the instrument to be relatively “compact”, featuring a shorter focal length than a broad bandpass Laue lens (see below) with an equivalent amount of efficient area for energy E_1 . This type of instrument has been proposed by B. Smither at Argonne National Laboratories [164], and has been developed for use in nuclear astrophysics by the Toulouse–Argonne collaboration [165, 166]. An example of a narrow bandpass Laue lens, the balloon telescope CLAIRE, will be discussed below.

Broad Bandpass Laue Lenses

Use only one (or very few) set of crystalline planes – typically the lowest order planes e.g. [111], with their optimum diffraction efficiency. Since several concentric rings using the *same set* of planes each focus a slightly different energies because of the varying Bragg angle, a broad energy band can be covered by this type of lens. If the [111] crystals of ring 1 are tuned to diffract photons with energy E_1 onto a certain focal point, the [111] planes of ring 1 are slightly more inclined with respect to the incident beam in order to reflect an energy $E_2 < E_1$ on the same focal spot. Here, the energy E_i diffracted by each ring is proportional to $1/\theta_i$ or $1/r_i$. As a consequence of the small Bragg angles implied by the low order of diffraction, very long focal lengths are required if a large geometrical lens area is required. ((85) above applies e.g. with $i = 1$).

Diffraction lenses with broad energy bandpass have been developed and tested for X-rays since the sixties (e.g. Lindquist and Webber [167]). Today, grazing incidence techniques dominate in X-ray astronomy, either with total external reflection or by using multilayer mirrors. A gamma-ray lens with a very broad continuum coverage has been proposed by N. Lund [168]; here, the wide mosaic structure and the alignment of the crystals placed on an Archimedes' spiral results in a effective area between 350 cm^2 at 300 keV and 25 cm^2 at 1.3 MeV.

The example of a broad bandpass Laue lens for nuclear astrophysics will be discussed below in the context of the projected MAX mission.

Mosaicity

As discussed in Sect. 2.3 (32ff), the acceptance angle of perfect crystals is extremely narrow (fraction of arcseconds for Germanium). The energy bandpass can be increased using so-called mosaic crystals, which are characterized by their mosaic width $\Delta\theta_B$. The mosaic width, or mosaicity, of the crystals governs the flux throughput, the angular resolution and the energy bandpass (see below) of the crystal lens. The diffracted flux from a continuum source increases with increasing mosaic width of the crystal. For a crystal lens telescope, crystals with mosaic widths ranging from a few arc seconds to a few arc minutes are of interest.

Energy Bandwidth

The bandwidth for a source on the axis of the lens is determined by the mosaicity of the individual crystals (see also Sect. 2.3) and the accuracy of the alignment of the crystals. By forming the derivative of the Bragg relation in the small angle approximation (Bragg: $2d\theta_B \approx hc/E$),

$$\Delta\theta_B/\theta_B = \Delta E/E, \quad (86)$$

where $\Delta\theta$ is the mosaic width of the crystal; the energy bandpass ΔE of a reflection becomes

$$\Delta E = \frac{2d \cdot E^2 \cdot \Delta\theta_B}{nhc} . \quad (87)$$

Whereas the energy bandpass of a crystal lens grows with the square of energy, Doppler broadening of astrophysical lines (e.g. in SN ejecta) increases linearly with energy for a given expansion velocity.

Crystal Diffraction Efficiency

As the diffracted photon beam passes through the crystal, photons are diffracted back and forth between the incident beam and the diffracted beam. If the crystal is sufficiently thick, the two beams will emerge from the opposite side of the crystal with equal intensities. Thus the maximum intensity that one can expect in the diffracted beam for the Laue geometry for thick crystals corresponds to 50% of that part of the flux which is not absorbed in the crystal (see Sect. 2.3, (36–38)).

To optimize the intensity in the diffracted beam at a certain energy, one increases the thickness of the crystal until the product of the diffraction efficiency times the transmission through the crystal is maximum. Figure 54 gives an example of the effect for a 10 arcsec mosaicity germanium crystal where the [400] planes are used for the diffraction process [169]. Each curve shows the dependence of the peak diffracted intensity as a function of the thickness of the crystal for a different energy gamma-ray. Each gamma-ray energy has a different thickness for optimum diffracted flux, but, for the higher energies, the maximum is quite broad.

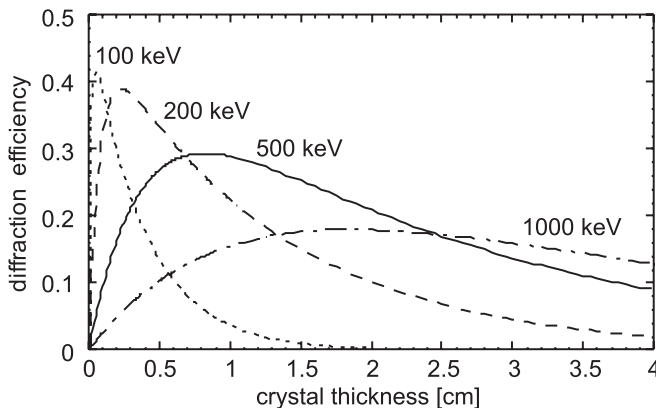


Fig. 54. Diffraction efficiency of a germanium crystal using the [400] diffraction planes, with an acceptance angle of $10''$, as a function of the crystal thickness and for different gamma-ray energies

In order to verify simulations based on the Darwin model for mosaic crystal, the diffraction efficiencies of Ge crystals have been measured at the Advanced Photon Source synchrotron at Argonne National Laboratories [75]. Measured diffraction efficiencies range from 20% to 31% according to energy (200 keV–500 keV) and crystal planes: Ge[111] and [220]. The results (Fig. 55) agree with what is expected from the Darwin model.

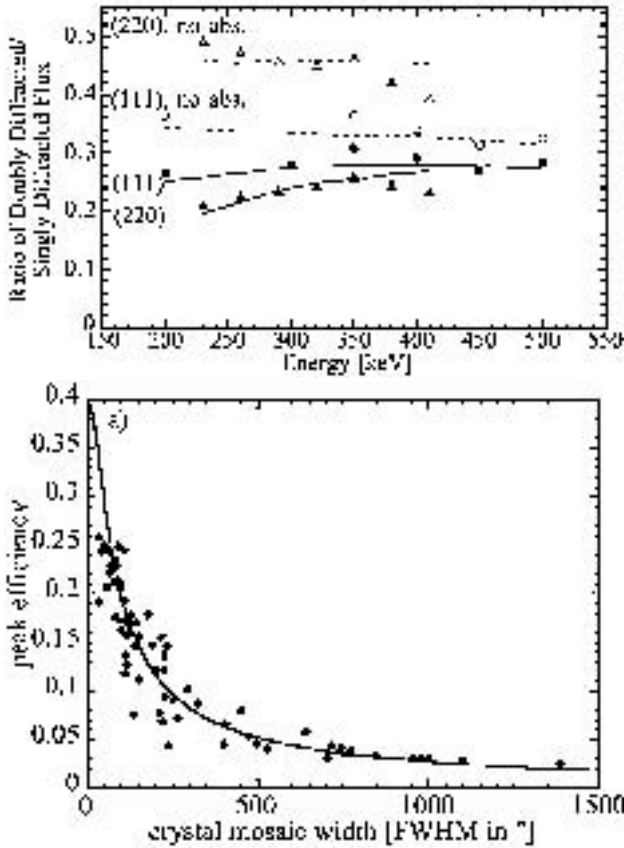


Fig. 55. *above:* Measured diffraction efficiencies (*solid data points*) for a narrow mosaicity (3 arcsec) Ge crystal. The *solid lines* are the results of a simulations using the Darwin model. *below:* The peak efficiency is shown as a function of mosaic width. The data points are from 72 rocking curves evenly spaced over the surface of a 2.46-mm-thick Ge [111] crystal after heating and squeezing the crystal. The measurements were done at 200 keV. The *solid curves* are calculated using the Darwin model [75]

Finite Distance

When tuning/calibrating the telescope in the laboratory, sources with finite distances have to be dealt with. Here the simple lens formula applies:

$$\frac{1}{p'} - \frac{1}{p} = \frac{1}{f},$$

where p is the distance “lens to source”, p' , the distance “lens to focal point”, and f , the focal length. This relationship assumes that $\sin \theta \approx \tan \theta \approx \theta$ (the exact relationship being $\arctan(r/p') - \arctan(r/p) = \arctan(r/f)$).

If a diffracting crystal subtends an angle $\Delta\theta_c$ (as seen from a monoenergetic laboratory source), this may be appreciably larger than the crystal’s mosaicity $\Delta\theta_m$. The fraction of active crystal-volume “seeing” the source is then given by the ratio $\Delta\theta_m/\Delta\theta_c$. The measured efficiency will therefore have to be corrected by a factor $\Delta\theta_c/\Delta\theta_m$ to obtain the diffraction efficiency of the entire crystal. An analogous argument is employed when the radioactive source is replaced by a continuum source (X-ray generator). Here, the energy bandpass corresponding to the mosaicity has to be compared to the energy bandpass defined by the angular extent of the crystal at finite distance – the correction factor still is $\Delta\theta_c/\Delta\theta_m$.

Tunable Crystal Diffraction Lens

Observing in *only one* energy band would clearly be unacceptable for a space instrument using a narrow bandpass Laue lens. In the framework of an R&D project for the French Space Agency CNES, a prototype tunable γ -ray lens (Fig. 56a) has been developed and demonstrated [171]. The capability to observe more than one astrophysical line requires the tuning of two parameters: the Bragg angle θ_B and the focal distance f . While the focal f will have to

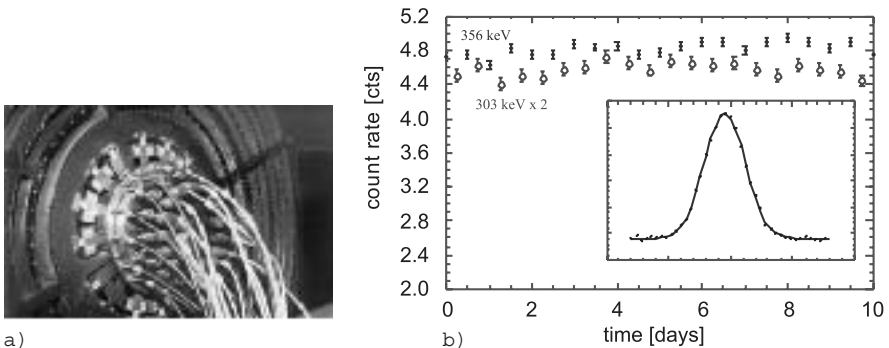


Fig. 56. (a) Prototype tunable lens. (b) The evolution in time of the peak count rate when alternatively focusing 303 keV (circles) and 356 keV (crosses) γ -rays demonstrates the stability and reproducibility of the lens tuning [171]

be controlled to within ~ 1 cm, the precision of the crystal inclination has to be better than the mosaic structure of the crystals. In the setup of Kohnle et al. [171], each crystal is tuned by using piezo-driven actuators to change the crystal inclination, and an eddy-current sensor to determine the current position (Fig. 56a). The resolution of the control-loop permitted an angular resolution of $0.1\text{--}0.4$ arcsec. The stability was found to be better than 0.8 arcsec per day and the reproducibility of a particular tuning better than 5 arcsec (Fig. 56b).

CLAIRE – A Balloon Borne Narrow Bandpass Laue Lens

CLAIRE's objective is to validate the concept of a Laue diffraction lens for nuclear astrophysics. The lens consists of 556 crystals mounted on the eight rings of a 45 cm diameter Titanium frame. In each ring i , the combination of the crystal plane spacing d_i and the Bragg angle θ_{Bi} results in the concentration of 170 keV photon onto a common focal spot of 1.5 cm diameter at 279 cm behind the lens. The geometric area of the lens is 511 cm^2 , its efficiency about 15%, the FOV and the bandpass are $90''$ and ~ 2 keV, respectively. The photons are focused onto a small 3×3 array of high-purity Germanium detectors, housed in a single cylindrical aluminum cryostat. Each of the single Ge bars is an n-type coaxial detector with dimensions of $1.5\text{ cm} \times 1.5\text{ cm} \times 4\text{ cm}$. Focusing onto such a small detector volume results in very low background noise. In order to further reduce the background, the detector matrix is actively shielded by a CsI(Tl) side shield and BGO collimators. The CLAIRE stabilization and pointing system were developed by the balloon division of the French space agency CNES. Two almost independent systems stabilize and point a target close to the sun (the Crab on June 14 and 15!) with a precision better than a few arcseconds: a primary pointing system stabilizes the entire telescope to within 10 arc minutes, while a set of gimbal frames points the gamma-ray lens only. The 3 m telescope structure consists of carbon fiber spars and honeycomb platforms; the entire instrument weighs only 500 kg (the limit for balloon flights in France). CLAIRE was launched by CNES from its base at Gap-Tallard in the French Alps in June 2000 and 2001, the astrophysical target was the Crab nebula. (While the diffraction lens is dedicated to the observation of nuclear lines, a balloon test flight ironically requires observation of a continuum spectrum.) A discussion of the performance of CLAIRE and preliminary analysis of the balloon flights is given by Halloin et al. [172].

MAX – Mission Concept for a Broad Bandpass Laue Lens

Ultimately, the concept of a crystal diffraction telescope should be put to use in space where longer exposures and steady pointing will result in outstanding sensitivities. Ideally, a space borne crystal diffraction telescope will use a gamma-ray lens situated on a stabilized spacecraft, focusing gamma-rays onto a small array of germanium detectors on a small spacecraft flying in formation.

The mission concept MAX [173] proposes simultaneous focusing in two broad energy bands of high astrophysical relevance, using two concentric broad bandpass lenses. As the primary scientific objective of MAX is the study of the $^{56}\text{Ni} \rightarrow ^{56}\text{Co} \rightarrow ^{56}\text{Fe}$ decay chain in type Ia supernovae, the principal energy band is centered on the 847 keV line from ^{56}Co . The corresponding lens is made of copper crystals, each one about 1 cm^3 in size, organized in 10 rings. The crystals of each ring diffract in the [111] plane. While the outermost ring of Cu crystals has a radius of 96 cm and focuses energies of 825 keV, the innermost ring has a radius of 87 cm, concentrating photons of 910 keV. Currently copper crystals can be grown with one arcminute mosaicity, so the energy bandpass is about 70 keV while the peak efficiency reaches 15%. The total effective lens area at 847 keV is 600 cm^2 .

The second energy band of MAX is centered on 500 keV, with the objective of studying electron–positron annihilation emission (X-ray binaries, AGN, spectra of SN Ia ...). The width of the energy band permits the observation of redshifted e^+e^- lines from compact objects (eg. the supermassive black hole in the center of our Galaxy), as well as the study of the 478 keV deexcitation line from ^7Li . The part of the lens concentrating photons in the 500 keV band is made of 14 concentric rings of Germanium crystals on the outside of the Cu one discussed above. The innermost ring has a radius of 97 cm, concentrating photons of 522 keV, the radius of the outermost ring is 110 cm, the diffracted energy being 460 keV. Again, the crystals are each about 1 cm^3 in size and use the [111] diffraction plane. With their 30 arc-second mosaicity, the energy bandpass of every ring is about 20 keV while the peak efficiency reaches 25%. The total effective lens area at 511 keV is 600 cm^2 .

The diffracted photons from both the Germanium and the Copper rings are concentrated onto a 1.5 cm diameter focal spot 133 m from the lens assembly. Here, a small matrix of Ge detectors, shielded by an active BGO shield (thickness 1 cm) performs high resolution spectroscopy. The passively cooled detector matrix is situated on a small spacecraft flying in formation maintaining the focal length to better than $\pm 1\text{ m}$ and by controlling the lateral position to within 1 cm. A high orbit minimizing gravity gradient disturbances allows long uninterrupted viewing, and permits simple passive cooling of the detector to 80–100 K.

The sensitivity of MAX in each energy band is roughly $3 \cdot 10^{-7}\text{ cm}^{-2}\text{ s}^{-1}$ for narrow gamma-ray lines. This estimate has been obtained by completely modeling MAX in the radiation environment conditions encountered outside the magnetosphere. Although a crystal lens telescope is not a direct imaging system, MAX will be able to generate intensity maps, by sweeping the telescope optical axis over a limited target area, or by using its off-axis response for broadened line sources. The angular resolution of a crystal lens telescope is determined by the mosaic width of the crystals, as well as the energy resolution of the detector – here the angular resolution is of the order of

45 arcsec at 511 keV, and about 90 arcseconds at 847 keV. The imaging capabilities of broad bandpass Laue systems have been discussed by Lund [168]. The capability of Laue lenses to resolve possible e^+e^- sources associated with the radiojets of the microquasar 1E1740-29 [28] at 511 keV has been demonstrated by extensive simulations [174].

Fresnel Lenses

Fresnel lenses can focus gamma-rays by using a combination of diffraction and refraction. Because the wavelengths of gamma-ray are so short and the penetrating power high, a phase shift can be achieved in a thickness of material which has a high transparency (see Sect. 2.4).

This type of gamma-ray lens has been proposed by Skinner in 2001 [79, 175, 176] – Fresnel lenses have the potential for revolutionizing gamma-ray astronomy: a telescope based on these principles can have angular resolution better than a micro second of arc – sufficient to resolve the event horizon of black holes in the nuclei of AGNs. At the same time, the sensitivity can be three orders of magnitude better than that of current instrumentation.

Diffraction-limited lenses of several meters in size are feasible and do not require high technology for their manufacture. Focal lengths are long – up to a million kilometers – but developments in formation flying of spacecraft make possible a mission in which the lens and detector are on two separate spacecraft separated by this distance.

Fresnel Zone Plates

In a Fresnel zone plate (Fig. 57) radiation is brought to a focus by blocking parts of the wave front which would arrive at the focal point with an incorrect phase. One can consider a part of the zone plate towards the periphery as

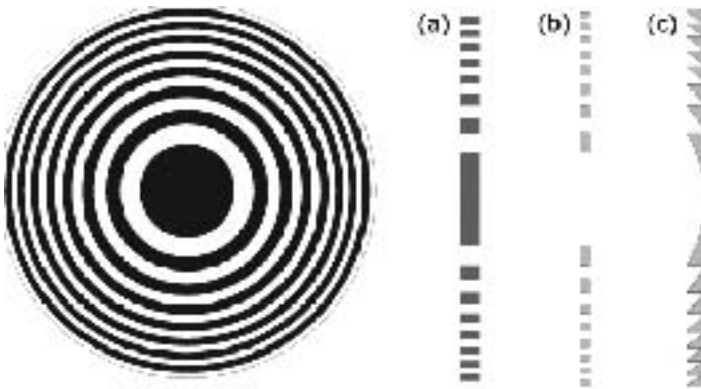


Fig. 57. (a) Fresnel zone plate with absorbing and transmitting zones (b) phase zone plate (c) phase Fresnel lens [79]

a diffraction grating which deviates the radiation towards the focal point. It can then readily be seen that the efficiency for concentrating the radiation into the first order ($k = 1$) focal point cannot exceed π^{-2} , i.e. about 10%, because energy also goes into the zero order ($k = 0$; straight through) and into orders with $k > 1$ and $k < 0$. The energy in these orders is in proportion to the power in the corresponding components in the Fourier transform of a square wave with transmission between zero and one.

Phase Fresnel Lenses

By varying the optical thickness, and hence the phase of the transmitted radiation rather than its amplitude, across the zone plate (Fig. 57c), all of the power can be diffracted into the principle ($k = 1$) focus in a configuration we shall refer to here as a “Phase Fresnel Lens”. The phase shift necessary is, of course, never greater than 2π . The focal length of the lens is a function of the zone widths, characterized by the value $p_{\min}$ at the outer rim where they are finest:

$$f = \frac{d \cdot p_{\min}}{2\lambda} \approx 0.4 \cdot 10^6 \left(\frac{d}{1 \text{ m}} \right) \left(\frac{p_{\min}}{1 \text{ mm}} \right) \left(\frac{E}{1 \text{ MeV}} \right) \text{ km} . \tag{88}$$

Thus very large lens-detector separations are implied. However, with the development of formation flying for space based interferometry, separations of the order of 10^6 km are no more looking ridiculous. Such distances have the benefit of offering a “plate scale” which is convenient for ultra-high angular resolution observations.

FRESNEL – A Conceptual High Angular Resolution Gamma-Ray Mission

Based on the above general arguments for feasibility, a conceptual mission, FRESNEL, using a gamma-ray lens based the principles described here has

Table 9. FRESNEL

nominal γ -ray energy	500 keV 847 keV	2 lenses, selectable by spacecraft rotation
tunable range	325–1200 keV 550–2000 keV	by varying focal length
geometric area	20 m ²	
lens efficiency	> 90%	
focal length	750000 km	at nominal energy
angular resolution	0.7 μ arc seconds	domin. by chromatic aberration
continuum sensitivity	$5 \cdot 10^{-9} \text{ cm}^{-2} \text{ s}^{-1} \text{ keV}^{-1}$	5 σ in 1 d
line sensitivity	$2 \cdot 10^{-9} \text{ cm}^{-2} \text{ s}^{-1}$	5 σ in 10 ⁶ s

been proposed and has been studied by the Integrated Mission Design Center (IMDC) of NASA Goddard Spaceflight Center. The assumed characteristics of the FRESNEL mission are summarized Table 9.

Acknowledgments

Many thanks to my former grad students Pierre Jean, Jürgen Knödseder and Antje Kohnle for letting me use materials of their dissertations. I'm deeply indebted to Gerry Skinner for his careful proofreading and many enlightening discussions. A large part of this manuscript was compiled during a sabbatical semester at IAS Rome. I'm particularly grateful to Pietro Ubertini, Angela Bazzano and the entire gamma-ray astrophysics group at IASR, to whom this work is dedicated.

References

1. Villard, P., 1900, *Comptes rendus*, **130**, 1010–1012
2. Rutherford, E., 1903, *Philosophical Magazine*, **5**, 177–187
3. Villard, P., 1900, *Séances de la Société française de Physique*, 40–46
4. Gerward, L., 1999, *Phys. perspect.* **1**, 367–383
5. Perlow, G.J., and Kissinger, C.W., 1951, *Phys Rev*, **81**, 552
6. Perlow, G.J., and Kissinger, C.W., 1951, *Phys Rev*, **84**, 572
7. Morrison, P., 1958, *Il nuovo Cimento*, Vol. **VII**, N.6, 858
8. Peterson, L., and Winckler, J.R., 1959, *Phys Rev Letters*, **1**, 205
9. Arnold, J.R., 1962, *J.Geophys. Res.* **67**, 4878
10. Metzger, A.E., 1964, *Nature* **204**, 766
11. Treaty Banning Nuclear Weapon Tests in the Atmosphere, in Outer Space and Under Water, 5.8.1963, Moscow
12. Klebesadel, R.W., Strong, I.B., and Olson, R.A., 1973, *Ap.J.* **182**, L85
13. Costa, E., Frontera, F., Heise, J., et al., 1997, *Nature*, **387**, 783
14. Kulkarni, S., et al., 1998, *Nature*, **393**, 35
15. Chupp, E.L., 1973, *Nature* **241**, 333
16. Murphy, et al., 1990, *ApJ*, **358**, 290
17. McConnell, M. et al., 1997, *AIP Conference Proceedings* **410**, 1099
18. Anderson, C.D., 1932, *Phys. Rev* **41**, 405
19. Johnson, W.N., Harnden, F.R., and Haymes, R.C., 1972, *ApJ*, **172**, L1
20. Albernhe, F., et al., 1981, *Astr. Ap.*, **94**, 214
21. Leventhal, M., MacCallum, C.J., and Stang, P.D., 1978, *ApJ*. **225**, L11
22. Leventhal, M., 1991, *Advances in Space Research*, **11**, 8, 157
23. Share, G.H, Leising, M.D, Messina, D.C, Purcell, W.R, 1990, *Ap.J.*, **385**, L45
24. Mahoney, W.A., Ling, J.C., Wheaton, W.A., 1993, *Ap.J.Sup.Ser.*, **92**, 387
25. Purcell, W.R., et al., 1997, *Ap.J.*, **491**, 725
26. Dermer, C.D., Skibo, J.G., 1999, *Ap.J.*, **487**, L57
27. Bouchet, L. et al., 1991, *Ap.J.*, **383**, L45
28. Mirabel et al., 1992, *Nature*, **358**, No 6383

29. Malet, I., 1995, *Ap.J.*, **444**, 222
30. Lingenfelter, R.E., and Ramaty, R., 1989, *Ap.J.*, **343**, 686
31. Kurfess, J.D., *Advances in Space Research*, **25**, 3–4, 631
32. Mahoney, W.A., Ling, J.C., Wheaton, W.A., Jacobson, A.S., 1984, *Ap.J.*, **286**, 578
33. Share, G.H., Kinzer, R.L., Kurfess, J.D., Forrest, D.J., Chupp, E.L., Rieger, E., 1985, *Ap.J.*, **292**, L61
34. von Ballmoos, P., Diehl, R., and Schönfelder, V., 1987, *ApJ.*, **318**, 654
35. Oberlack, U., et al., 1997, *AIP Conference Proc.* **410**, 1109
36. Knödseder, J., et al., 1999, *Astron. Astrophys.* **344**, 68
37. Prantzos, N., Diehl, R., *Physics Reports*, **267**, p. 1–69
38. Matz, S.M., et al., 1988, *Nature* **331**, 416
39. Plüschke, S., et al., 2000, Proc. 5th Compton Symposium, Eds M. McConnell and J. Ryan, p.35
40. Goldwurm, A., et al., 1992, *ApJ*, **389**, L89
41. Mazets, E.P. et al., 1981, *Nature*, **290**, 378
42. Olive, J.-F., 1992, *Ph.D thesis*, Université Paul Sabatier, Toulouse
43. Jacobson, A.S., 1978, NASA GSFC Gamma Ray Spect. Astroph., p 228
44. Ling, J.C., et al., *ApJ Letters*, 1979, **231L**, 896
45. Kurfess, J.D., et al., 1992, *ApJ Letters*, **399**, L137
46. Morris, D.J., et al., 1995, *Ann. New York Acad. Sci.*, **759**, 397
47. Iyudin, A., et al., 1994, *A&A*, **284**, L1
48. Iyudin, A., et al., 1998, *Nature*, **396**, 142
49. Trümper, J., et al., 1977, *Ann. New York Acad. Sci.*, **302**, 538
50. Mihara, T., Makishima, K., Nagase, F., 1995, AAS Meeting, **187**, 104.03
51. Santangelo, A., 1999, *ApJ Letters*, **523L**, 85
52. Hulsizer, R., and Rossi, B.B., 1948, *Phys Rev*, **73**, 1402
53. Kraushaar, W.L., and Clark, G.W., 1962, *Phys Rev Letters*, **8**, 106
54. Kraushaar, W.L., et al., 1965, *ApJ*, **141**, 845
55. Clark, G.W., Gamire, G.P., Kraushaar, W.L., 1968, *ApJ Letters*, **153**, L203
56. Hartman, R.C., et al., 1999, *ApJ Suppl. Series*, **123**, 79–202
57. Fichtel, C.E., and Trombka, J.I., 1997, *Gamma-Ray Astrophysics*, NASA ref. publication 1386
58. Hartman, R.C., et al. 1979, *ApJ*, **250**, 389
59. Mayer-Hasselwander, H.A., et al. 1982, *A&A*, **105**, 164
60. Greisen, K., 1966, in R.E. Marshak (ed.), *Perspectives in Modern Physics*, John Wiley and Sons, New York, p 355
61. Chupp, E.L., 1976, *Gamma-Ray Astronomy*, Reidel, Dordrecht, Holland
62. Pinkau, K., 1996, *Astron. Astrophys. Suppl. Ser.*, **120**, 43
63. Macomb, D.J., and Gehrels, N., *ApJ. Suppl. Ser.*, **120**, 335
64. Paciesas, W.S., et al., 1999, *ApJ. Suppl. Ser.*, **122**, 465.
65. Türler, M., 1999, *A&A Supplement*, **134**, 89
66. Voges, W., et al., 1999, *Astron. Astrophys.*, **349**, 389
67. Davisson, C.M., 1966, in K. Siegbahn (ed.) *Alpha-, Beta-, and Gamma-Ray Spectroscopy*, North-Holland, Amsterdam
68. Evans, R.D., 1955, *The Atomic Nucleus*, Mac Graw Hill Book Company
69. Berger, M. J., Hubbell, J.H., Seltzer, S.M., 1999, *Photon Cross Sections Database*, National Institute of Standards and Technology Standard Reference Database 8, (<http://physics.nist.gov/PhysRefData/Xcom>)

70. Heitler, W., 1954, *The Quantum Theory of Radiation*, Clarendon Press, Oxford
71. Klein, O., and Nishina, Y., 1929, *Z. Physik*, **29**, 853
72. Motz, J.W., and Missioni, G., 1958, *Phys. Rev.*, **124**, 1458
73. Brown, G.E., Mayers, D.F., 1957, *Proc. Roy. Soc. (London)*, **A242**, 89
74. Zachariasen, W.H., 1946, *Theory of X-ray diffraction in Mosaic Crystals*, Wiley & Sons
75. Kohnle, A., 1998, PhD. thesis: *A Gamma-Ray Lens for Nuclear Astrophysics*, Université Paul Sabatier, Toulouse
76. Darwin, C.G., 1914, *Phil. Mag.*, **27**, 315 and 657
77. Schneider, J.R., 1977, *Acta. Cryst.*, **A33**, 235
78. Henke, B.L., Gullikson, E.M., and Davis, J.C., 1993, X-ray interactions: photoabsorption, scattering, transmission, and reflection at $E=50-30000\text{eV}$, $Z=1-92$, *Atomic Data and Nuclear Data Tables*, **54**, (no.2), 181
79. Skinner, G.K., 2001, *Astron. Astrophys.*, **375**, 691
80. Bethe, H.A., and Heitler, W., 1934, *Proc. Roy. Soc.* **A146**, 83
81. Knoll, G.F., 1989, *Radiation Detection and Measurement*, John Wiley and Sons, New York
82. Ramsey, B.D., 1995, *Exp. Astron.* **6**, 119
83. Ramsey, B.D., et al., 1989, *Nucl. Instr. and Meth. in Phys Res.* **A278**, 576
84. Ubertini, P., 1987, *Space Science Rev.* **46**, 1
85. Udin, S.E., et al., 1996, *SPIE Proceeding*, **2806**, 577
86. Kämäräinen, V., et al., 1997, *Proc. 2nd INTEGRAL workshop*, ESA SP-382, 655
87. Dmitrenko, V.V., et al., 1992, *SPIE Proceeding*, **1734**, 90
88. Mahler, G.J., et al., 1998, *IEEE Trans. Nucl. Sci.* **NS-45**, 1024
89. Bolotnikov, A., Ramsey, B.D., 1997, *IEEE Trans. Nucl. Sci.*, **NS-44**, 1006
90. Egorov, E., Ermilova, V. and Rodionov, B., Preprint P.N.Lebedev Physics Institute (USSR), **166**, 1982.
91. Aprile, E., et al., 2000, *Proc. Astronomy with radioactivities*, Schloss Ringberg, Kreuth, Germany, Sept/Oct. 1999, MPE Report 274
92. Aprile, E., et al., 1998, *Nucl. Instr. and Meth. in Phys Res. A* **412**, 425
93. Aprile, E., et al., 2000, *Proc 5th Compton Symp.*, *AIP*, **510**, 799
94. Schönfelder, V., et al., 1993, *ApJ Supp. Series*, **86**, 657
95. Hofstadter, R., 1948, *Phys. Rev.*, **74**, 100
96. Anger, H.O., 1958, *Rev. Sci Instr.*, **29**, 27-33
97. Bouchet, L., et al., 2001, *ApJ*, **548**, 990
98. PICSiT Team ITESRE, 2000, report IN-IM-TES-RP-0038
99. Novotny, R., et al., 1998, *Nucl. Physics B*, **61B**, 613
100. BICRON Saint-Gobain Industrial Ceramics Inc. Catalogue, 3101 (02-2000)
101. Zhu, R.Y., et al., 1996, *NIM A* **376**, 319
102. RCA Photomultiplier Manual, PT-61, RCA Solid State Division, Lancaster, PA, 1970
103. Klein, C.A., 1968, *J. Appl. Phys.*, **39**, 2029
104. Kraner, H.W., Chasman, C., and Jones, K.W., 1968, *Nuclear Instr. And Meth. Sect. A*, **62**, 173
105. Kraner, H.W., Pehl, R.H., and Haller, E.E., 1975, *IEEE Trans. Nucl. Sci.* **22**, 149
106. Pehl, R.H., Varnell, L.S. and Metzger, A.E., 1978, *IEEE Trans. Nucl. Sci.* **25**, 409

107. Koenen, M., Brückner, J., Körfer, M. and Wänke, H., 1995, *IEEE Trans. Nucl. Sci.* **42**, 653
108. Paul, Ph., 2002, PhD. thesis, Université Paul Sabatier, Toulouse
109. Takahashi, T., et al., 2002, *IEEE Trans. Nucl. Sci.*, vol. **49**, No. 3, pp. 1297
110. Limousin, O., 2003, *NIM A*, **504**, 24-37
111. Johnson, W.N., et al., 1993, *ApJ Suppl. Series*, **86**, 693
112. Fishman, G.J., et al., 1989, *Proc. Gamma Ray Observatory Science Workshop*, ed. W. Johnson (Greenbelt: GSFC), 2
113. Harmon, B.A., et al., 2002, *ApJ Suppl. Series*, **138**, 149
114. Ling, J.C., et al., 2000, *ApJ Suppl. Series*, **127**, 79
115. Oda, M., 1965, *Appl. Opt.* **4**(1), 143
116. Mertz, L., 1967, in *Modern Optics*, (New York: Brooklyn Polytechnic Press), p.787
117. Schnopper, H.W., et al., 1970, *ApJ*, **161**, L161
118. Oda, M., et al., 1976, *Space Sci. Instr.*, **2**, 141
119. Makishima, K., et al., 1978, *Cospar: New Instrumentation for Space Astronomy* (Pergamon Press, Oxford and New York), p.277
120. Lin, R.P., et al., 1998, *SPIE Proceeding* **3442**, p.2-12
121. Boggs, S.E., et al., 2001, *ESA Symp. Proc.*, SP-459, 541
122. Aristotle, *problemata physica* - problem XV,6: "Why is it that when the sun passes through quadrilaterals, as for instance wickerwork, it does not produce a figure rectangular in shape but circular?"
123. Aristotle, *problemata physica* - problem XV,11: "Why is it that in an eclipse of the sun, if one looks at it through a sieve or through leaves, such as a planetree or other broad leaved tree, or if one joins one hand over the fingers of the other, the rays are crescent-shaped where they reach the earth? Is it for the same reason as that when light shines through a rectangular peep-hole, it appears circular in the form of a cone? The reason is that there are two cones, one from the sun to the peephole and the other from the peep-hole to the earth, and the vertices meet..."
124. Holt, S.S., 1976, *Astrophys. Space Sci.* **42**, 123
125. Mertz, L., and Young, N., 1961, in *Proc. of the Internat. Conference on Optical Instruments and Techniques* (Chapman and Hall, London), p.305
126. Dicke, R.H., 1968, *Astrophys. J.* **153**, L101
127. Ables, J.G., 1968, *Proc. Astron. Soc. Australia* **1**, 172
128. Skinner, G.K., 1995, *Exp. Astron.*, **6**, 1
129. Caroli, E., et al., 1987, *Space Sci. Rev.*, **45**, 349
130. Skinner, G.K., and Rideout, R.M., 1995, *Exp. Astron.*, **6**, 177
131. Jean, P., et al., 1997, *Proc. 2nd INTEGRAL workshop*, ESA, SP-382, 635
132. Mandrou, P., et al., 1997, *ESA Symp. Proc.* "The Transparent Univers", SP-382, p. 591
133. Lichti, et al., 1996, *SPIE proc*, Vol. **2806**, p.217
134. Ubertini, P., Di Cocco, G., & Lebrun, F., 1997, *ESA Symp. Proc.* "The Transparent Univers", SP-382, p. 599
135. Schönfelder, V., Hirner, A., and Schneider, K., 1973, *Nucl.Instrum. Meth.*, **107**, 385
136. Alvarez, L.W., et al., 1973, *Space Sciences Laboratory UCB, Series 14, Issue 17*
137. Dauber, Ph.M., and Smith, L.H., 1973, *13th ICRC*, Vol **4**, 2716

138. Pinkau, K., 1966, *Zeitschrift f. Naturf.*, **21a**, 2100
139. White, R.S., 1968, *Bull. Am. Phys. Soc.*, **13**, 714
140. Preszler, A.M., Simnett, G.M., White, R.S., 1972, *Phys. Rev. Lettters*, **28** (15), 982
141. Herzo, D., et al., 1975, *Nucl.Instrum. Meth.*, **123**, 583
142. Graml, F., et al., 1975, *Proc. 14th Int. Cosmic Ray Conf.*, Munich, **9**, 3129
143. Lockwood, J.A., et al., 1979, *ApJ*, **248**, 1194
144. Baker, R.E., et al., 1979, *Nuclear Instr. And Meth.*, **158**, 595
145. Knödseder, J., 1997, PhD. thesis, Université Paul Sabatier, Toulouse
146. Oberlack, U., 1997, PhD. thesis, TU München
147. van Dijk, R., 1996, PhD. thesis, Universiteit van Amsterdam
148. Kanbach, G., et al., 2003, *SPIE Proceedings*, Volume **4851**, 1209
149. Boggs, S.E., et al., 2001, *Proc. "Gamma-Ray 2001 Astrophysics"*, Baltimore
150. Kurfess, J.D., et al., 1994, NASA proposal for new mission concepts in Astrophysics, NRA 94-OSS-15
151. <http://heseweb.nrl.navy.mil/gamma/detector/ACT/ACT.htm>
152. Aprile, E., et al., 2002, *SPIE*, 4851, 1196
153. Ramsey, B.D., Alexander, C.D., Apple, J.A., et al. 2002, *ApJ*, **568**, 432
154. Angel, J.R.P., 1979, *Ap. J.* **233**, 364
155. Kumakhov, M.A., 1990, *Nucl. Instr. Meth.*, B48, 288
156. Kirkpatrick, P., and Baez, A.V., 1948, *J. Optic Soc. of America*, **38**, 766
157. Craig, W.W., et al., 1998, *Proc. SPIE*, **3445**, 112
158. Christensen, F.E., et al., 2000, *SPIE* **4012**, 278
159. Owens, S.M., et al., 2002, *Proc. SPIE* **4496**, 115
160. Tawara, Y., et al., 2002, *Proc. SPIE* **4496**, 109
161. Windt, D.L., et al., 2002, *SPIE Proceedings*, Volume **4851**, 639
162. Frontera, F., and Pareschi, G., 1995, *Exp. Astronomy*, **6**, 25 (1995)
163. De Chiara, P., and Frontera, F., 1992, *Applied Optics-OT*, **31**,10, 1361
164. Smither, R.K., 1982, *Rev. Sci. Instr.* **44**, 131
165. von Ballmoos, P., Smither, R.K., 1994, *Astrophys. J. Suppl.*, **92**, 663
166. Naya, J.E., 1996, *Nuclear Instr. And Meth.. Sect. A*, **373**, 59
167. Lindquist, T.R. and Webber, W.R., 1968, *Can. J. Phys*, **46**, 1103
168. Lund, N., 1992, *Exp. Astron.* **2**, 259
169. Smither, R.K., et al., GRO Science Workshop, GSFC, April 1989, NASA Report, Ed. W. Neil Johnson
170. Kohnle, A., et al., 1998, *Nuclear Instr. And Meth.. Sect. A*, Vol. **416**, 493
171. Kohnle, A., et al., 1998, *Nuclear Instr. And Meth.. Sect. A*, Vol. **408**, 553
172. Halloin, H., et al., 2003, *SPIE Proceedings*, Volume **4851**, 895
173. von Ballmoos, P., et al., 2002, CNES proposal (*astropcest-pub_max*)
174. Kohnle, A., 1998, Phd Thesis, Université Paul Sabatier, Toulouse
175. Skinner, G.K., 2002, *Astron. Astrophys.* **383**, 352
176. Skinner, G.K., et al., 2003, *SPIE Proceedings*, Volume **4851**, 1366



Rashid Sunyaev

Hard X-Ray and Gamma Ray Spectroscopy

R. Sunyaev and S. Sazonov

Max-Planck-Institut für Astrophysik, Garching, Germany

A cosmic plasma with a temperature below 10 keV and normal cosmic abundance forms a lot of different spectral lines and features. At higher temperatures and at high optical depths there appears a new very strong player – Comptonization – which determines the formation of the spectra of hard X-ray and soft gamma-ray sources. Comptonization is the process of change of frequency of photons due to scattering on thermal electrons. At a temperature of 10 keV, the average velocity of electrons is close to one fifth of the velocity of light, and consequently the energy of a photon increases or decreases by $\sim 20\%$ in each successive scattering. If we have 50 keV photons, their energies will decrease on the average by 10% after a single Compton scattering on “cold” electrons with $kT \ll h\nu$ due to Compton recoil. In the general case, both the Doppler shift in frequency and the recoil effect work simultaneously. In objects with a finite optical depth for Thomson scattering, this process makes it very difficult to have any narrow features in the spectrum. It leads to the formation of power-law radiation spectra, and in the extreme case of a very high optical depth to the formation of a Wien spectrum with a pronounced broad maximum. If we take into account induced Compton scattering, we will arrive at a situation where a Planck spectrum is formed as a result of photon production by bremsstrahlung and the double Compton effect amplified by Comptonization.

In this review we will concentrate on objects hosting high temperature, rarified plasmas of finite optical depth for Thomson scattering. The best examples of such objects are the accretion disks around accreting black holes and neutron stars in binary X-ray sources, accretion disks in the vicinity of supermassive black holes in active galactic nuclei and quasars, spreading layers on the surface of accreting neutron stars and boundary layers between neutron stars and accretion disks. The same process is extremely important in the hot primordial plasma in the early stages of expansion of the Universe as well as in the hot gas residing in the deep potential wells of clusters of galaxies. Supernovae heated by radioactive decay of Nickel 56 and Cobalt 56 is another example where Comptonization is responsible for the formation of observed X-ray and gamma-ray spectra and for the transfer of energy from gamma-ray photons to an expanding envelope, producing the optical light that we can observe during the exponential decay of the supernova

brightness. At the initial stage, the optical depth of the envelope is huge and the energies of gamma-ray line photons decrease due to recoil down to 20 keV when photon absorption becomes more important than Compton recoil. As the optical depth decreases during the envelope expansion, we begin to see lines shifted by recoil and finally narrow lines appear.

1 Fundamentals of Compton Scattering

1.1 Photon Frequency Shift upon Scattering from a Free Electron

Assume that a photon of energy $h\nu$ and momentum $(h\nu/c)\mathbf{\Omega}$ is scattered by a free electron of energy $\gamma m_e c^2$ and momentum $\mathbf{p} = \gamma m \mathbf{v}$, where $\gamma = (1 - v^2/c^2)^{-1/2}$. Let $h\nu'$ and $(h\nu'/c)\mathbf{\Omega}'$ denote the energy and momentum of the photon after the scattering event.

By introducing the electron and photon four-momenta $p_4 = (\mathbf{p}, i\gamma m_e c)$, $k_4 = (h\nu\mathbf{\Omega}/c, ih\nu/c)$ prior to the scattering event and $p'_4 = (\mathbf{p}', i\gamma' m_e c)$, $k'_4 = (h\nu'\mathbf{\Omega}'/c, ih\nu'/c)$ afterwards, one can easily find how the frequency of the photon will change when it is scattered (see, e.g. [15]). In fact,

$$p_4 + k_4 = p'_4 + k'_4 . \quad (1)$$

Squaring this relation and noting that $p_4^2 = p'^2_4 = -m_e^2 c^2$ while $k_4^2 = k'^2_4 = 0$ we see that

$$p_4 k_4 = p'_4 k'_4 . \quad (2)$$

On the other hand, if we multiply (1) by k'_4 , we find

$$p'_4 k'_4 = p_4 k'_4 + k_4 k'_4 . \quad (3)$$

Defining $\mu = \mathbf{\Omega v}/v$, $\mu' = \mathbf{\Omega' v}/v$, and the scattering angle $\theta = \arccos \mathbf{\Omega \Omega'}$, we may therefore write

$$\frac{\nu'}{\nu} = \frac{1 - \mu v/c}{1 - \mu' v/c + (h\nu/\gamma m_e c^2)(1 - \cos \theta)} . \quad (4)$$

It is customary to speak about *Thomson scattering* if a photon of low energy ($h\nu \ll m_e c^2$) is scattered by an electron at rest ($v = 0$). In Thomson scattering the incident and scattered photons have the same energy ($\nu' = \nu$), so this scattering is coherent, or elastic.

If the photon energy is non-negligible in comparison with the electron rest energy, quantum effects must be taken into account, and the process is called *Compton scattering*. In this case, the photon frequency will decrease because of the recoil effect:

$$\frac{\nu'}{\nu} = \frac{1}{1 + (h\nu/m_e c^2)(1 - \cos \theta)} , \quad (5)$$

and the photon wavelength will increase accordingly:

$$\lambda' = \lambda + \lambda_C(1 - \cos \theta) , \quad (6)$$

where $\lambda_C = h/m_e c$ is the Compton wavelength.

A further interesting situation arises when the electron is moving – in this case energy can be transferred to the photon, and the process is called *inverse Compton scattering*. If a photon is scattered by a moving electron, the Doppler effect will play a role in changing its frequency. In fact, in a reference frame comoving with the scattering electron, the photon frequency prior to the scattering event is $\nu_0 = \gamma\nu(1 - \mu v/c)$, and if $h\nu_0 \ll m_e c^2$, we may neglect the frequency shift of the scattered photon in the electron rest frame: $\nu'_0 \approx \nu_0$. Reverting to the laboratory frame, we obtain

$$\nu' = \frac{\nu'_0}{\gamma(1 - \mu'v/c)} = \frac{\nu_0}{\gamma(1 - \mu'v/c)} = \nu \frac{1 - \mu v/c}{1 - \mu'v/c} . \quad (7)$$

In this review we shall use the term “Compton scattering” to unify Thomson, Compton and inverse Compton scattering.

1.2 Scattering Cross Section

We shall assume that the incident radiation is unpolarized. In this case the differential cross section for Compton scattering is given by [15]

$$\frac{d\sigma}{d\Omega'} = \frac{r_e^2}{2\gamma^2} \frac{X}{(1 - \mu v/c)^2} \left(\frac{\nu'}{\nu} \right)^2 , \quad (8)$$

where

$$\begin{aligned} X &= \frac{x}{x'} + \frac{x'}{x} + 4 \left(\frac{1}{x} - \frac{1}{x'} \right) + 4 \left(\frac{1}{x} - \frac{1}{x'} \right)^2 , \\ x &= \frac{2h\nu}{m_e c^2} \gamma \left(1 - \mu \frac{v}{c} \right) , \quad x' = \frac{2h\nu'}{m_e c^2} \gamma \left(1 - \mu' \frac{v}{c} \right) , \end{aligned} \quad (9)$$

and $r_e = e^2/m_e c^2 = 2.82 \times 10^{-13}$ cm is the classical electron radius.

The quantum-mechanical formula (8) reduces to a classical expression in the *Thomson limit* $\gamma h\nu \ll m_e c^2$:

$$\frac{d\sigma}{d\Omega'} = \frac{r_e^2}{2} \frac{1}{\gamma^2(1 - \mu'v/c)^2} \left[1 + \left(1 - \frac{1 - \cos \theta}{\gamma^2(1 - \mu v/c)(1 - \mu'v/c)} \right)^2 \right] , \quad (10)$$

and further simplifies for Thomson scattering ($v = 0$, $h\nu \ll m_e c^2$):

$$\frac{d\sigma}{d\Omega'} = \frac{r_e^2}{2} (1 + \cos^2 \theta) . \quad (11)$$

The angular part of this expression is the same as for Rayleigh scattering of low-frequency photons by bound electrons.

If a photon of arbitrary energy is scattered by an electron at rest ($v = 0$), the Klein–Nishina differential cross section applies:

$$\frac{d\sigma}{d\Omega'} = \frac{r_e^2}{2}(1 + \cos^2 \theta) \left[1 + (1 - \cos \theta) \frac{h\nu}{m_e c^2} \right]^{-2} \times \left[1 + \left(\frac{h\nu}{m_e c^2} \right)^2 \left(1 + (1 - \cos \theta) \frac{h\nu}{m_e c^2} \right)^{-1} \frac{(1 - \cos \theta)^2}{1 + \cos^2 \theta} \right]. \quad (12)$$

The general formula for the total scattering cross section is

$$\sigma = \int \frac{d\sigma}{d\Omega'} d\Omega' = \frac{3\sigma_T}{4x} \left[\left(1 - \frac{4}{x} - \frac{8}{x^2} \right) \ln(1+x) + \frac{1}{2} + \frac{8}{x} - \frac{1}{2(1+x)^2} \right], \quad (13)$$

where $\sigma_T = 8\pi r_e^2/3 = 6.65 \times 10^{-25} \text{ cm}^2$ is the Thomson scattering cross section. In particular, in the Thomson limit

$$\sigma = \sigma_T \left(1 - x + \frac{13}{10}x^2 + \dots \right), \quad (14)$$

where we have included the Klein–Nishina corrections of first and second order.

In the ultrarelativistic limit ($x \gg 1$), the cross section rapidly decreases with increasing x :

$$\sigma = \frac{3\sigma_T}{4} x^{-1} \left(\ln x + \frac{1}{2} \right). \quad (15)$$

Scattering by an Ensemble of Hot Electrons

Equation (8) describes the differential cross section for Compton scattering by a single electron. Consider now the propagation of photons through a homogeneous gas of electrons with a given isotropic distribution of velocities $f(v)$ (defined so that $\int f(v) d\mathbf{v} = 1$). The probability for a photon originally moving in the direction $\boldsymbol{\Omega}$ to be scattered within a path of length dl into the direction $\boldsymbol{\Omega}'$ is given by

$$\frac{dP}{dl d\Omega'} = N_e \int \left(1 - \mu \frac{v}{c} \right) \frac{d\sigma}{d\Omega'}(\boldsymbol{\nu}, \mathbf{v}) f(v) d\mathbf{v} \equiv N_e \left(\frac{d\sigma}{d\Omega'} \right)_{\text{ens}}. \quad (16)$$

Here N_e is the electron number density, the factor $(1 - \mu v/c)$ takes into account the relative velocity of the electron and photon before scattering [58, 93], and $d\sigma/d\Omega'(\boldsymbol{\nu}, \mathbf{v})$ is given by (8). On the right-hand side of (16) we introduced a new quantity – the ensemble-averaged differential cross section, $(d\sigma/d\Omega')_{\text{ens}}$.

In the nonrelativistic case ($v \ll c$, $h\nu \ll m_e c^2$), $(d\sigma/d\Omega')_{\text{ens}}$ is just the Thomson differential cross section (11), and scattering is characterized by forward-backward symmetry.

When low-energy photons are scattered by ultrarelativistic electrons ($\gamma \gg 1$) but the Thomson limit takes place ($\gamma h\nu/m_e c^2 \ll 1$), the ensemble-averaged cross section takes on another simple form [139],

$$\left(\frac{d\sigma}{d\Omega'} \right)_{\text{ens}} = \frac{2r_e^2}{3} (1 - \cos \theta) . \quad (17)$$

Therefore, in this case photons preferentially scatter backwards, rather than forwards. This phenomenon results from the joint action of two effects. One is that a photon has a better chance of undergoing a scattering by an electron that is moving towards it rather than away from it (the probability is proportional to $1 - \cos \theta v/c$). The other effect is that photons emerge after scattering collimated in the direction of motion of the relativistic electron. The angular distribution of emergent photons in this case contrasts the forward-oriented Klein-Nishina angular function, which corresponds to the case of scattering of energetic photons by an electron at rest ($h\nu \sim m_e c^2$, $v = 0$).

The backward-scattering behaviour of hot plasma has important astrophysical ramifications. For example, a hot electron-scattering atmosphere, such as an accretion disk corona, will be more reflective than a cold one: the fraction of incident low-energy photons reflected by the atmosphere after a single scattering increases by up to 50% [140]. This will affect the cooling rate of the hot plasma by external radiation as well as emergent Comptonization spectra. Also, the spatial diffusion of photons will proceed more slowly in a hot, optically thick plasma, thereby affecting the formation of spectra through Comptonization. These effects are discussed in detail in [52, 54, 55, 64, 129, 157, 174].

Photon Mean Free Path

Integrating the ensemble-averaged differential cross section over all scattering angles gives the effective total cross section σ_{eff} and the photon mean free path $\bar{\lambda}$:

$$\sigma_{\text{eff}} = \frac{1}{\bar{\lambda}} = N_e \int \frac{d\sigma}{d\Omega'}_{\text{ens}} d\Omega' . \quad (18)$$

Several simple asymptotic relations can be derived [132, 150]. In the case of Maxwellian electrons with $kT \ll m_e c^2$ and photons with $h\nu \ll m_e c^2$,

$$\sigma_{\text{eff}} = \sigma_T N_e \left[1 - 2 \frac{h\nu}{m_e c^2} - 5 \frac{h\nu}{m_e c^2} \frac{kT}{m_e c^2} + \frac{26}{5} \left(\frac{h\nu}{m_e c^2} \right)^2 + \dots \right] . \quad (19)$$

In the limit $h\nu kT \ll (m_e c^2)^2$, $kT \gg m_e c^2$,

$$\sigma_{\text{eff}} = \sigma_{\text{T}} N_{\text{e}} \left(1 - 8 \frac{h\nu}{m_{\text{e}} c^2} \frac{kT}{m_{\text{e}} c^2} + \dots \right). \quad (20)$$

In the ultrarelativistic limit $h\nu \gg m_{\text{e}} c^2$, $kT \gg m_{\text{e}} c^2$,

$$\sigma_{\text{eff}} = \frac{3}{16} \sigma_{\text{T}} N_{\text{e}} \frac{m_{\text{e}} c^2}{h\nu} \frac{m_{\text{e}} c^2}{kT} \left[\ln \left(4 \frac{h\nu}{m_{\text{e}} c^2} \frac{kT}{m_{\text{e}} c^2} \right) - 0.077 + \dots \right]. \quad (21)$$

And finally if $h\nu \gg m_{\text{e}} c^2$ and $kT \ll m_{\text{e}} c^2$,

$$\sigma_{\text{eff}} = \frac{3}{8} \sigma_{\text{T}} N_{\text{e}} \frac{m_{\text{e}} c^2}{h\nu} \left(\ln \frac{2h\nu}{m_{\text{e}} c^2} + \frac{1}{2} + \frac{kT}{m_{\text{e}} c^2} + \dots \right) \left(1 - \frac{3}{2} \frac{kT}{m_{\text{e}} c^2} + \dots \right). \quad (22)$$

The above formulae and Fig. 2 (Fig. 7 in Pozdnyakov) demonstrate that the mean free path lengthens as the photon energy or/and the plasma temperature rise. For a given plasma density the minimum mean free path is achieved in the Thomson limit: $\bar{\lambda} = 1/(\sigma_{\text{T}} N_{\text{e}})$.

1.3 Radiation Force

When a photon is scattered by an electron it will transfer to the electron a momentum

$$\Delta \mathbf{p} = \frac{h\nu}{c} \boldsymbol{\Omega} - \frac{h\nu'}{c} \boldsymbol{\Omega}'. \quad (23)$$

Hence a radiation field of intensity $I_{\nu}(\boldsymbol{\Omega}, \nu)$ will impart to an ensemble of electrons a force (per electron)

$$\mathbf{f} = \int \left(\frac{h\nu}{c} \boldsymbol{\Omega} - \frac{h\nu'}{c} \boldsymbol{\Omega}' \right) \frac{I_{\nu}(\boldsymbol{\Omega}, \nu)}{h\nu} \left(1 - \mu \frac{v}{c} \right) \frac{d\sigma}{d\Omega'} f(\mathbf{v}) d\mathbf{v} d\boldsymbol{\Omega} d\boldsymbol{\Omega}' d\nu. \quad (24)$$

Thomson Limit

Let us first evaluate the pressure exerted by low-frequency radiation ($h\nu \rightarrow 0$) on a collimated stream of electrons moving with velocity $\mathbf{v}$. In this case the differential scattering cross section and the photon frequency change are given by (10) and (7), respectively, and we can derive from (24) the force acting on each electron:

$$\mathbf{f} = \frac{\sigma_{\text{T}}}{c} \int \left[\boldsymbol{\Omega} \left(1 - \frac{\boldsymbol{\Omega} \mathbf{v}}{c} \right) - \gamma^2 \frac{\mathbf{v}}{c} \left(1 - \frac{\boldsymbol{\Omega} \mathbf{v}}{c} \right)^2 \right] I_{\nu}(\boldsymbol{\Omega}, \nu) d\boldsymbol{\Omega} d\nu. \quad (25)$$

Consider several examples. In the case of isotropic radiation,

$$\mathbf{f} = -\frac{4}{3} \sigma_{\text{T}} \Sigma \gamma^2 \frac{\mathbf{v}}{c}, \quad (26)$$

where $\Sigma = 4\pi c \int I_\nu d\nu$ is the total radiation energy density. Thus an isotropic radiation field exerts a braking force on a moving electron.

If the radiation is beamed narrowly along the direction $\boldsymbol{\omega}$, the radiation force will be

$$\mathbf{f} = \frac{\sigma_T \mathbf{q}}{c} \left[\boldsymbol{\omega} \left(1 - \frac{\boldsymbol{\omega} \mathbf{v}}{c} \right) - \gamma^2 \frac{\mathbf{v}}{c} \left(1 - \frac{\boldsymbol{\omega} \mathbf{v}}{c} \right)^2 \right], \quad (27)$$

where $\mathbf{q} = \int \boldsymbol{\Omega} I_\nu(\boldsymbol{\Omega}, \nu) d\boldsymbol{\Omega} d\nu$ is the total radiation flux. Note that the above expression can also be derived in terms of the classical radiative damping force exerted on the electron by a plane electromagnetic wave (see [175]).

In the particular case where the radiation beam is directed opposite to the electron velocity $\mathbf{v}$, we obtain the familiar expression [93]

$$\mathbf{f} = \frac{\sigma_T \mathbf{q}}{c} \frac{1 + v/c}{1 - v/c} = \frac{\sigma_T \mathbf{q}}{c} \left(1 + 2\frac{v}{c} + \dots \right), \quad (28)$$

while in the opposite case $\mathbf{q} \parallel \mathbf{v}$,

$$\mathbf{f} = \frac{\sigma_T \mathbf{q}}{c} \frac{1 - v/c}{1 + v/c} = \frac{\sigma_T \mathbf{q}}{c} \left(1 - 2\frac{v}{c} + \dots \right). \quad (29)$$

We see that the accelerating force in this case will be much weaker than the retarding force in the previous case if $v \rightarrow c$.

Integrating equation (27) over $d\mathbf{v}/v$ gives the force that will be exerted by a low-frequency radiation field with an arbitrary angular distribution (not necessarily collimated) on an ensemble of monoenergetic electrons isotropically distributed in velocity space [116]:

$$\mathbf{f} = \frac{\sigma_T \mathbf{q}}{c} \left[1 + \frac{2}{3} \left(\frac{v}{c} \right)^2 \gamma^2 \right] = \frac{\sigma_T \mathbf{q}}{c} \left[1 + \frac{2}{3} (\gamma^2 - 1) \right]. \quad (30)$$

In particular, for thermal plasma with $kT \ll m_e c^2$ we find that

$$\mathbf{f} = \frac{\sigma_T \mathbf{q}}{c} \left(1 + 2 \frac{kT}{m_e c^2} + \dots \right), \quad (31)$$

since $\overline{v^2} \approx 3kT/m_e$. In the ultrarelativistic case, when $\gamma \gg 1$,

$$\mathbf{f} \approx \frac{8\sigma_T \mathbf{q}}{c} \left(\frac{kT}{m_e c^2} \right)^2, \quad (32)$$

because $\langle \gamma^2 \rangle \approx 12(kT/m_e c^2)^2$.

The radiation force in ultrarelativistic electron plasma will be enormously strengthened (and the Eddington luminosity, considered below, will correspondingly diminish) because electrons will preferentially scatter photons by angles close to π , greatly raising the energy of the photons and giving them a large momentum. This scenario will be realized only if collisional or plasma processes are efficient in maintaining the isotropy of the electron distribution.

Klein–Nishina Limit

In the limit $v = 0$, the differential cross section is described by the Klein–Nishina formula (12). After integration over all scattering angle and with (5), (24) becomes

$$\mathbf{f} = \frac{3\sigma_{\text{T}}}{4c} \int \frac{1+a}{a^3(1+2a)^3} \left[\frac{(1+2a)^3}{2a} (a^2 - 2a - 3) \ln(1+2a) + 3 + 17a + 31a^2 + 17a^3 - \frac{10}{3}a^4 \right] I_{\nu}(\mathbf{\Omega}, \nu) \mathbf{\Omega} d\mathbf{\Omega} d\nu, \quad (33)$$

where $a = h\nu/m_e c^2$. In the limit $a \rightarrow 0$ we find asymptotically

$$\mathbf{f} = \frac{\sigma_{\text{T}}}{c} \int \left(1 - \frac{16}{5} \frac{h\nu}{m_e c^2} + \dots \right) I_{\nu}(\mathbf{\Omega}, \nu) \mathbf{\Omega} d\mathbf{\Omega} d\nu, \quad (34)$$

while in the relativistic Klein–Nishina limit, when $a \gg 1$, the radiation force is much reduced:

$$\mathbf{f} = \frac{3\sigma_{\text{T}}}{8c} \int \frac{m_e c^2}{h\nu} \left[\ln \left(1 + \frac{h\nu}{m_e c^2} \right) - \frac{5}{6} \right] I_{\nu}(\mathbf{\Omega}, \nu) \mathbf{\Omega} d\mathbf{\Omega} d\nu. \quad (35)$$

Eddington Critical Luminosity

Many X-ray sources have a luminosity approaching the critical Eddington value. Suppose that an electron at rest is located at distance R from an object of luminosity L and mass M ; then the radiation will exert on it a force (in the Thomson limit)

$$\mathbf{f} = \frac{\sigma_{\text{T}}}{c} \mathbf{q} = \frac{\sigma_{\text{T}} L}{4\pi R^2 c} \frac{\mathbf{R}}{R}. \quad (36)$$

A proton, on the other hand, will be subject to a gravitational force $\mathbf{f}_{\text{grav}} = -(GMm_p/R^2)\mathbf{R}/R$ (nearly the same force will act on a neutron). One may neglect radiation pressure on the proton, since its scattering cross section

$$\sigma_{\text{p}} = \frac{8\pi}{3} \left(\frac{e^2}{m_p c^2} \right)^2 = \left(\frac{m_e}{m_p} \right)^2 \sigma_{\text{T}} \quad (37)$$

is insignificant; and the attractive force exerted on the electron will also be very small, as its mass is small. The electrons and protons in ionized plasma are bound together by electrostatic forces, and charge separation is practically impossible. Both forces mentioned above fall off as R^{-2} and are oppositely directed. They will become equal if the source shines at the Eddington critical luminosity

$$L_{\text{Edd}} = \frac{4\pi GMmc}{\sigma_{\text{T}}} = 1.25 \times 10^{38} \frac{m}{m_p} \frac{M}{M_{\odot}} \text{ erg s}^{-1}. \quad (38)$$

Here m is the mean mass per electron ($m \approx 1.17m_p$ for plasma of normal cosmic composition), and we assume that complete ionization of helium and heavy elements will yield one electron for every two nucleons.

If $L > L_{\text{Edd}}$, no accretion can occur; radiation pressure will overwhelm the gravitational forces and cause material to flow outward. If $L \ll L_{\text{Edd}}$, the light pressure may be neglected; this allows material to be accreted, and makes possible the existence of stars with internal energy sources and stable atmospheres.

Compared with the case of electron–proton pairs, for electron–positron pairs the radiation force will be twice as great, while the gravitational force will be smaller by a factor $2m_e/m_p$. Hence the critical luminosity for electron–positron plasma will be $m_p/m_e = 1846$ times lower than the Eddington luminosity for electron–proton plasma, given by (38). If $L > 7 \times 10^{34} \text{ erg s}^{-1}$, electron–positron plasma will be swept out of high-temperature zones.

Compton Acceleration and Drag

An electron or positron can be accelerated or decelerated by an external source of radiation. The problem greatly simplifies when the Thomson limit holds and the radiation field is axisymmetric (see [103]). In this case, as follows from (25), a blob of matter moving along the axis of symmetry with speed $v = dr/dt$ is accelerated at the rate

$$\begin{aligned} \frac{dv}{dt} &= \frac{1}{\gamma^3} \frac{f}{m} \\ &= \frac{2\pi\sigma_T}{\gamma mc} \left[\left(1 - \frac{v}{c}\right)^2 \int_{-1}^1 I(\mu) \mu d\mu - \frac{v}{c} \int_{-1}^1 I(\mu) (1 - \mu)^2 d\mu \right]. \end{aligned} \quad (39)$$

Here $\mu = (\mathbf{\Omega v})/v$, and we assumed that the gravitational attraction is negligibly small compared to the radiation pressure. In the case of electron–proton plasma this will be true for a super-Eddington radiation source, while for electron–positron plasma this assumption does not require the source to be super- or near-Eddington.

If we consider (39), the first bracketed term represents the boosting effect due to scattering of photons with small incident angles, while the second term describes the *Compton drag* induced by photons coming from angles $\alpha = \arccos(\mathbf{\Omega v}/v) \gtrsim 1/\gamma$, which due to relativistic aberration are perceived by the scattering particle as moving towards the source. This term vanishes in the case of a point-like source, $I(\mathbf{\Omega}) = F(r)\delta(\mathbf{\Omega} - \mathbf{r}/r)$.

In the case of a finite-size source, the right-hand side of (39) vanishes for $\gamma = \gamma_{\text{eq}}$, or $v = v_{\text{eq}}$. Particles are accelerated away from the source as long as $\gamma < \gamma_{\text{eq}}$. If $\gamma > \gamma_{\text{eq}}$, the force reverses, being now directed inwards. This means that at any distance from the source there exists an upper velocity limit, which is independent of the source luminosity, up to which the particle can be accelerated by the radiation. When the particle achieves this velocity,

the net momentum carried by the incident photons disappears in the electron rest frame.

Near the surface of an extended source, $v_{\text{eq}}/c \sim 0.5\text{--}0.7$ depending on the emission angular diagram [115, 175]. Far ($r \gg R$) from a spherical source of radius R and uniform brightness [115, 175],

$$\gamma_{\text{eq}}(\text{sphere}) \sim 3^{1/4} r/R. \quad (40)$$

In the point-source limit ($R \rightarrow 0$), (39) reduces to

$$\frac{d\gamma}{dr} = \gamma^2 \left(1 - \frac{v}{c}\right)^2 \tilde{l} \frac{R}{r^2}. \quad (41)$$

Here the parameter $\tilde{l}$ is the dimensionless compactness, rescaled by the inertia per scattering charge,

$$\tilde{l} \equiv l \frac{m_e}{m} = \frac{L\sigma_T}{4\pi mc^3 R} = \frac{1}{2} \frac{m_p}{m} \frac{L}{L_{\text{Edd}}} \frac{R_S}{R}, \quad (42)$$

where $R_S = 2GM/c^2$ is the Schwarzschild radius. In the case of an electron–positron plasma ($m = m_e$), $\tilde{l} = 306(3R_S/R)(L/L_{\text{Edd}})$, so for a source with $L \sim L_{\text{Edd}}$ and $R \sim 3R_S$, $\tilde{l} \gg 1$. If a particle starts moving with $\gamma_0 = 1$ at radius r_0 , it will attain at infinity a Lorentz factor

$$\gamma_{\infty}(\text{point}) \sim (3\tilde{l}R/4r_0)^{1/3}. \quad (43)$$

Compton acceleration is by far less efficient in the case of electron–proton plasma because of the much greater inertia per unit cross section.

In the case of a finite-size source, the asymptotic solution (43) will be applicable only if the particle trajectory begins at a distance $r_0 > r_t \sim \tilde{l}^{1/4} R$ from the source, where the radiation drag can be neglected. Within the zone $r < r_t$ near the source, the effect of Compton drag is very important due to the presence of a substantial nonradial component of the radiation field, so that the particle Lorentz factor tends to adjust itself very rapidly to the upper limit (40). As a result, for motions starting at $r \ll r_t$ the terminal Lorentz factor will be of the same order as the equilibrium Lorentz factor at the transition radius: $\gamma_{\infty} \sim \gamma_{\text{eq}}(r_t)$. Hence $\gamma_{\infty} \sim \tilde{l}^{1/4}$ in the strong-source limit ($\tilde{l} \gg 1$) [115]. This means that if electron–positron pairs are created near the source, the emergent ultrarelativistic flow will have a narrow distribution in energies. Larger values of γ_{∞} can be obtained only if the particles are injected with relativistic velocities at $r > r_t$.

Accretion disks around black holes and neutron stars provide an example of extended sources where Compton drag can be particularly strong. The radial surface flux distribution of a standard thin disk is given by [148]

$$Q(R) = \frac{3GM\dot{M}}{8\pi R^3} \left[1 - \left(\frac{3R_S}{R} \right)^{1/2} \right], \quad (44)$$

where M is the mass of the compact object and $\dot{M}$ is the accretion rate. In this case the equilibrium Lorentz factor increases only as $\gamma_{\text{eq}} \sim (r/R_{\text{max}})^{1/4}$ (compared to a linear increase for a spherical source). Accordingly, the terminal Lorentz factor is found to be $\gamma_{\infty} \sim \tilde{l}^7$, where now $\tilde{l} = 3GM\dot{M}\sigma_{\text{T}}/28\pi mc^3 R_{\text{max}}^2$ [88, 103, 125]. As a result, the pressure of radiation from a near-Eddington accretion disk can generate only a mildly relativistic electron-positron flow, with $\gamma_{\infty} \sim 2$ –3.

The luminosity of a standard accretion disk cannot exceed the limiting value L_{Edd} . Paczynski and Wiita [119] have shown that there could exist geometrically thick accretion disks emitting at super-Eddington luminosities. The inner region of such a disk should resemble a funnel down toward the black hole. The large surface area of this funnel allows the disk to radiate away much more energy than is possible in the case of a thin disk; the total luminosity may exceed the Eddington limit by more than an order of magnitude. It has been suggested [80, 152] that the thick disks have the potential to form narrow beams (jets), as the super-Eddington emission of the accretion funnel might accelerate particles to relativistic velocities. However, detailed calculations indicate (see [88, 126, 189] and references therein) that the radiation field deep within the funnel should be nearly isotropic and most of it is reprocessed. As a result, the acceleration is limited. An accurate, self-consistent calculation requires taking into account general relativistic effects, the radiation transfer within the funnel, and the stability conditions for thick disks. This constitutes a formidable problem, which has never been solved in full. Approximate solutions typically give terminal Lorentz factors $\lesssim 5$ for electron-positron plasma. Electron-proton plasma beams can only reach terminal velocities of ~ 0.4 – $0.9c$, even when $L \sim 10L_{\text{Edd}}$.

We finally note that the mechanism of Compton acceleration may become more efficient when scattering occurs in the relativistic Klein-Nishina regime, which is possible near compact gamma-ray sources such as blazars, black hole candidates and gamma-ray bursts [103].

1.4 Energy Exchange Between Plasma and Radiation

The Case $h\nu \rightarrow 0$

As a result of the action of the braking force, an electron moving in an isotropic field of low-frequency radiation will be losing energy at the rate

$$\frac{d\gamma}{dt} = -\frac{f}{m_e c} \gamma^{-1} (\gamma^2 - 1)^{1/2} = -\frac{4}{3} \frac{\sigma_{\text{T}} \Sigma}{m_e c} (\gamma^2 - 1), \quad (45)$$

This equation follows from (26) and has a solution of the form $\gamma = [1 + A(t)]/[1 - A(t)]$, with

$$A(t) = \frac{\gamma_0 - 1}{\gamma_0 + 1} \exp\left(-\frac{8}{3} \frac{\sigma_{\text{T}} \Sigma}{m_e c} t\right) = \frac{\gamma_0 - 1}{\gamma_0 + 1} \exp\left(-\frac{t}{t_c}\right), \quad (46)$$

where γ_0 is the initial Lorentz-factor of the electron and

$$t_c = \frac{3m_e c}{8\sigma_T \Sigma} \quad (47)$$

is a characteristic time scale.

But as the electrons cool, the radiation energy density will rise:

$$\frac{d\Sigma}{dt} = -N_e \frac{\gamma m_e c^2}{dt} = \frac{4}{3} \sigma_T \Sigma N_e c (\gamma^2 - 1). \quad (48)$$

Therefore, if $\gamma(t) = \text{const}$

$$\Sigma = \Sigma_0 \exp \left[\frac{4}{3} \sigma_T N_e c (\gamma^2 - 1) t \right]. \quad (49)$$

Since the photon–electron collision frequency is equal to $\sigma_T N_e c$ and the number of photons is conserved by scattering, the energy of a photon will increase, on the average, by

$$\bar{\nu}' = \nu \left[1 + \frac{4}{3} (\gamma^2 - 1) \right] \quad (50)$$

every time it collides with an electron.

Equation (45) enables us to find the rate at which energy will be withdrawn from plasma by Comptonization of low-frequency radiation, whatever the electron temperature may be. For this purpose (45) has to be averaged over the relativistic Maxwellian distribution

$$dN_e \propto \gamma (\gamma^2 - 1)^{1/2} \exp(-\gamma m_e c^2 / kT) d\gamma. \quad (51)$$

In this manner we find that

$$\frac{d\Sigma}{dt} = -N_e m_e c^2 \frac{d\langle \gamma \rangle}{dt} = \frac{4}{3} \sigma_T \Sigma c \langle \gamma^2 - 1 \rangle. \quad (52)$$

Here

$$\langle \gamma^2 - 1 \rangle = \frac{\int_1^\infty \gamma (\gamma^2 - 1)^{3/2} \exp(-\gamma/\eta) d\gamma}{\int_1^\infty \gamma (\gamma^2 - 1)^{1/2} \exp(-\gamma/\eta) d\gamma} = 3\eta(\eta + \langle \gamma \rangle), \quad (53)$$

$$\langle \gamma \rangle = \frac{\int_1^\infty \gamma^2 (\gamma^2 - 1)^{1/2} \exp(-\gamma/\eta) d\gamma}{\int_1^\infty \gamma (\gamma^2 - 1)^{1/2} \exp(-\gamma/\eta) d\gamma} = \frac{3\eta K_2(1/\eta) + K_1(1/\eta)}{2\eta K_1(1/\eta) + K_0(1/\eta)}, \quad (54)$$

where $\eta = kT/m_e c^2$, and $K_p(x)$ are modified Bessel functions. Equations (53), (54) reduce to the standard relations $\langle \gamma^2 - 1 \rangle = v^2/c^2 = 3\eta$, $\langle \gamma \rangle = 3\eta/2 + 1$ in the nonrelativistic case and $\langle \gamma^2 \rangle = 12\eta^2$, $\langle \gamma \rangle = 3\eta$ in the ultrarelativistic case.

Nonrelativistic Case

If $v \ll c$ ($\gamma \approx 1$), then $f = -m dv/dt$ and we have

$$\frac{d}{dt} \left(\frac{mv^2}{2} \right) = -\frac{8}{3} \frac{\sigma_T \Sigma}{m_e c} \left(\frac{mv^2}{2} \right); \quad (55)$$

thus the energy of the electron will decay exponentially as

$$E_e = E_0 \exp(-t/t_c). \quad (56)$$

In the case of thermal electrons,

$$\frac{dT}{dt} = -\frac{8}{3} \sigma_T c \Sigma \frac{kT}{m_e c^2}, \quad (57)$$

$$T = T_0 \exp(-t/t_c), \quad (58)$$

and

$$\frac{d\Sigma}{dt} = -\frac{3}{2} N_e \frac{d(kT)}{dt} = \frac{4\sigma_T \Sigma N_e kT}{m_e c}, \quad (59)$$

$$\Sigma = \Sigma_0 \exp \left(\frac{4\sigma_T N_e kT}{m_e c} t \right) \text{ if } kT(t) = \text{const} \ll m_e c^2. \quad (60)$$

In each scattering event the photon energy will increase, on the average, by

$$\frac{\overline{\Delta\nu}}{\nu} = 4 \frac{kT}{m_e c^2}. \quad (61)$$

Ultrarelativistic Case

If $\gamma \gg 1$, (45) reduces to the familiar expression [22]:

$$\begin{aligned} \frac{d\gamma}{dt} &= -\frac{4}{3} \frac{\sigma_T \Sigma}{m_e c} \gamma^2, \\ \gamma &= \frac{\gamma_0}{1 + (4\sigma_T \Sigma \gamma_0 / 3m_e c) t} = \frac{\gamma_0}{1 + \gamma_0 t / 2t_c}. \end{aligned} \quad (62)$$

Further,

$$\Sigma = \Sigma_0 \exp \left(\frac{16\sigma_T N_e kT}{m_e c} t \right) \text{ if } kT(t) = \text{const} \gg m_e c^2 \quad (63)$$

and

$$\bar{\nu}' = \frac{4}{3} \gamma^2 \nu. \quad (64)$$

The Case $kT \ll h\nu \ll m_e c^2$

Using the Thomson differential cross section (11) and the expression for the frequency shift due to recoil (5), we obtain in this case

$$\frac{\overline{\Delta\nu}}{\nu} = \frac{1}{\sigma_T} \int \left(\frac{\nu'}{\nu} - 1 \right) \frac{d\sigma}{d\Omega'} d\Omega' = -\frac{h\nu}{m_e c^2}. \quad (65)$$

More general analytic relations for the energy transfer rate in the limit $kT \ll m_e c^2$ for arbitrary $h\nu$ can be found in [112, 150].

The Case $h\nu, kT \ll m_e c^2$

The energy exchange due to the recoil and Doppler effects will be small in this nonrelativistic case: $\Delta\nu/\nu \ll 1$. The two effects to a first approximation combine linearly, so that

$$\frac{\overline{\Delta\nu}}{\nu} = \frac{4kT - h\nu}{m_e c^2}. \quad (66)$$

2 Comptonization in Infinite Homogeneous Media

Since Compton scattering changes the photon energy in accordance with (4), the photons composing a monochromatic spectral line will become distributed in frequency after a single electron scattering. The emergent spectrum will depend on the angle between the direction $\boldsymbol{\Omega}$ from which the photons are supplied and the viewing direction $\boldsymbol{\Omega}'$. This spectrum can be described in terms of the redistribution function $K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}')$, which gives the probability for a photon $(\nu, \boldsymbol{\Omega})$ to scatter within a unit path length into a solid angle $d\Omega'$ about $\boldsymbol{\Omega}'$ with a frequency within $(\nu', \nu' + d\nu')$. In the case of thermal plasma, an integral over a Maxwellian velocity distribution $f_M(\mathbf{v})$ arises:

$$K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}') = N_e \int \frac{d\sigma}{d\Omega'} \left(1 - \frac{\boldsymbol{\Omega} \cdot \mathbf{v}}{c} \right) f_M(v_x, v_y, v_z) \left| \frac{\partial v_x}{\partial \nu'} \right| dv_y dv_z. \quad (67)$$

Here $d\sigma/d\Omega'$ is the differential cross section given by (8). The factor $|\partial v_x / \partial \nu'|$ accounts for the fact that only two of the velocity components are independent, and should be calculated from (4).

If the incident radiation is beamed in the direction $\boldsymbol{\Omega}$, one may be interested in knowing the emergent spectrum resulting from a single scattering, integrated over all outgoing directions:

$$P(\nu \rightarrow \nu') = \int K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}') d\Omega' = 2\pi \int K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}') d\cos\theta, \quad (68)$$

where $\theta = \arccos(\mathbf{\Omega}\mathbf{\Omega}')$. The same spectrum will be observed from an arbitrary direction $\mathbf{\Omega}'$ if the incident radiation is isotropic, because of the isotropy of the Maxwellian velocity distribution.

In a more general context, the function $K(\nu, \mathbf{\Omega} \rightarrow \nu', \mathbf{\Omega}')$ represents the kernel (often called the Compton scattering kernel) of the integral kinetic equation governing the Compton interaction of radiation with thermal plasma,

$$\begin{aligned} & \frac{1}{c} \frac{\partial I_\nu(\nu, \mathbf{\Omega})}{\partial t} + (\mathbf{\Omega}\nabla)I_\nu(\nu, \mathbf{\Omega}) \\ &= - \int I_\nu(\nu, \mathbf{\Omega}) K(\nu, \mathbf{\Omega} \rightarrow \nu', \mathbf{\Omega}') [1 + n(\nu', \mathbf{\Omega}')] d\nu' d\mathbf{\Omega}' \\ &+ \int \frac{\nu}{\nu'} I_\nu(\nu', \mathbf{\Omega}') K(\nu', \mathbf{\Omega}' \rightarrow \nu, \mathbf{\Omega}) [1 + n(\nu, \mathbf{\Omega})] d\nu' d\mathbf{\Omega}'. \end{aligned} \quad (69)$$

Here, $I_\nu(\nu, \mathbf{\Omega})$ is the specific intensity of the radiation and $n = c^2 I_\nu / (2h\nu^3)$ is the occupation number in photon phase space. The first integral on the right-hand side of (69) represents the decrement of $I_\nu(\nu, \mathbf{\Omega})$ due to scattering of photons out of the direction $\mathbf{\Omega}$, while the second integral describes the increment of $I_\nu(\nu, \mathbf{\Omega})$ due to scattering into $\mathbf{\Omega}$ from the other directions. The terms $n(\nu, \mathbf{\Omega})$ and $n(\nu', \mathbf{\Omega}')$ in the square brackets represent the contribution of *induced Compton scattering* (discussed in §2.5 below).

In the case of the interaction of isotropic radiation with an infinite homogeneous medium, the kinetic equation reduces to

$$\begin{aligned} \frac{1}{c} \frac{\partial \Sigma_\nu(\nu)}{\partial t} &= - \int \Sigma_\nu(\nu) P(\nu \rightarrow \nu') [1 + n(\nu')] d\nu' \\ &+ \int \frac{\nu}{\nu'} \Sigma_\nu(\nu') P(\nu' \rightarrow \nu) [1 + n(\nu)] d\nu', \end{aligned} \quad (70)$$

with the kernel $P(\nu' \rightarrow \nu)$ being given by (68). Here, $\Sigma_\nu = 4\pi I_\nu / c = 8\pi h\nu^3 n / c^3$ is the radiation spectral energy density.

2.1 Analytic Approximations for the Compton Scattering Kernel

In 1925 Dirac [40] derived an approximate algebraic expression for the kernel $K(\nu, \mathbf{\Omega} \rightarrow \nu', \mathbf{\Omega}')$. The Doppler shift was taken into account to a first approximation, but Compton recoil was neglected. Dirac's formula therefore describes the Doppler broadening ($\Delta\nu/\nu \sim v/c$) of low-frequency spectral lines due to scattering in a nonrelativistic thermal plasma [$h\nu/m_e c^2 \ll (kT/m_e c^2)^{1/2} \ll 1$]. However, since the lowest order terms in the expression (66) for the Compton energy exchange are proportional to $(v/c)^2 \sim kT/m_e c^2$ and $h\nu/m_e c^2$, it is impossible to describe with the help of Dirac's kernel a number of important astrophysical phenomena such as

- the y - and Bose–Einstein μ -distortions of the Cosmic Microwave Background (CMB) spectrum resulting from energy release in the early universe,
- distortions of the CMB spectrum in the directions of galaxy clusters,
- the formation of hard power-law tails in the emission spectra of X-ray binary systems and active galactic nuclei.

After Dirac there have been numerous attempts to propose a better analytic description of the Compton scattering kernel. As follows from (67), any such calculation must deal with three fundamental formulae: (4) for the photon frequency shift, (8) for the scattering cross section, and (51) for the Maxwellian momentum distribution. As each of them is fairly complex, especially the relation giving the scattering cross section, it proves impossible to write down a single analytic expression that would describe the kernel for any values of kT and $h\nu$. Nonetheless, it has been possible to reduce the calculation of the kernel to numerical computation of a single integral over the electron momentum distribution [1, 83, 112]. Apart from these efforts, the Compton scattering kernel has been studied by numerical methods [74, 101, 111, 127, 131].

In astrophysics we are often encountered with the particular case where the energies of both electrons and photons are not too high – $kT, h\nu \ll m_e c^2$. Babuel-Peyrissac and Rouvillois [3] (see also [198]) derived a formula for the kernel that correctly describes the energy transfer between radiation and electrons in this limit. After some modification [141] their formula takes the appearance

$$K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}') = \frac{3}{32\pi} \sqrt{\frac{2}{\pi}} \sigma_T N_e \left(\frac{kT}{m_e c^2} \right)^{-1/2} \frac{\nu'}{\nu g} (1 + \cos^2 \theta) \\ \times \exp \left\{ -\frac{m_e c^2}{2kT g^2} \left[\nu' - \nu + \frac{h\nu\nu'}{m_e c^2} (1 - \cos \theta) \right]^2 \right\},$$

where $g = |\nu\boldsymbol{\Omega} - \nu'\boldsymbol{\Omega}'| = (\nu^2 - 2\nu\nu' \cos \theta + \nu'^2)^{1/2}$. (71)

It can be readily checked that integration (71) over ν' leads to the Thomson differential cross section (11) and an additional integration $\int K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}') d\nu' d\boldsymbol{\Omega}'$ gives $\sigma_T N_e$. This reflects the fact that (71) represents scattering in the Thomson limit.

Since the Maxwellian distribution is the thermodynamic equilibrium distribution for the electrons, the scattering kernel $K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}')$ must satisfy the detailed balance principle. This means that in thermodynamic equilibrium the number of photons which scatter from $d\nu' d\boldsymbol{\Omega}'$ to $\nu d\boldsymbol{\Omega}$ must equal the number scattered from $\nu d\boldsymbol{\Omega}$ to $d\nu' d\boldsymbol{\Omega}'$, allowing for induced effects. Quantitatively, this condition takes the form

$$\begin{aligned}
K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}') & \left[1 + \frac{c^2 B_\nu(\nu')}{2h\nu'^3} \right] \frac{B_\nu(\nu)}{h\nu} \\
& = K(\nu', \boldsymbol{\Omega}' \rightarrow \nu, \boldsymbol{\Omega}) \left[1 + \frac{c^2 B_\nu(\nu)}{2h\nu^3} \right] \frac{B_\nu(\nu')}{h\nu'} , \quad (72)
\end{aligned}$$

where $B_\nu = (2h\nu^3/c^2)[\exp(h\nu/kT) - 1]^{-1}$ is the Planck distribution, so that

$$K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}') = \left(\frac{\nu'}{\nu} \right)^2 \exp \left[\frac{h(\nu - \nu')}{kT} \right] K(\nu', \boldsymbol{\Omega}' \rightarrow \nu, \boldsymbol{\Omega}) . \quad (73)$$

Relation (71) does satisfy this equation. The result (73) also implies that in the absence of induced effects the equilibrium radiation spectrum for Compton scattering in thermal plasma obeys the Wien law $W_\nu \sim \nu^3 \exp(-h\nu/kT)$, since $K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}')W_\nu(\nu)/h\nu = K(\nu', \boldsymbol{\Omega}' \rightarrow \nu, \boldsymbol{\Omega})W_\nu(\nu')/h\nu'$.

It should be noted that when the recoil frequency shift can be neglected ($h\nu \ll kT \ll m_e c^2$), the scattered line profile depends solely on the combination of parameters $[(1 - \cos \theta)kT/m_e c^2]^{1/2}$. Thus, similar profiles can be obtained by varying either the temperature or the scattering angle.

Kernel for the Isotropic Problem

Consider now the kernel $P(\nu \rightarrow \nu')$ corresponding to the isotropic problem. It can be derived by integration in (68) of $K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}')$ over all scattering angles. This integral can be done analytically for the kernel (71) in the limit $h\nu(h\nu/m_e c^2) \ll kT \ll m_e c^2$, when the characteristic frequency shift due to recoil is small compared to the characteristic Doppler broadening. In this case the exponential in the expression for $K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}')$ can be expanded in a Taylor series, and one obtains [141, 167]¹

$$\begin{aligned}
P(\nu \rightarrow \nu') & = \sqrt{\frac{2}{\pi}} \sigma_T N_e \left(\frac{kT}{m_e c^2} \right)^{-1/2} \nu^{-1} \left[1 + \sqrt{2} \delta \left(1 - \frac{h\nu}{kT} \right) \left(\frac{kT}{m_e c^2} \right)^{1/2} \right] \\
& \quad \times \left[\left(\frac{11}{20} + \frac{4}{5} \delta^2 + \frac{2}{5} \delta^4 \right) F + |\delta| \left(-\frac{3}{2} - 2\delta^2 - \frac{4}{5} \delta^4 \right) G \right], \\
F & = \exp(-\delta^2), \quad G = \int_{|\delta|}^{\infty} \exp(-t^2) dt = 0.5\pi^{1/2} \text{Erfc}(|\delta|), \\
\delta & = \left(\frac{2kT}{m_e c^2} \right)^{-1/2} \frac{\nu' - \nu}{\nu' + \nu} . \quad (74)
\end{aligned}$$

Similarly to the kernel (71), the kernel (74) obeys the detailed balance principle:

$$P(\nu \rightarrow \nu') = \left(\frac{\nu'}{\nu} \right)^2 \exp \left[\frac{h(\nu - \nu')}{kT} \right] P(\nu' \rightarrow \nu) . \quad (75)$$

¹ [141] also derived first-order relativistic corrections to the kernels (71) and (74).

Important information about the $P(\nu \rightarrow \nu')$ kernel is provided by its moments, defined as follows:

$$\langle (\Delta\nu)^n \rangle = \frac{1}{\sigma_T N_e} \int P(\nu \rightarrow \nu') (\nu' - \nu)^n d\nu'. \quad (76)$$

The first two moments of the kernel (74) are

$$\begin{aligned} \langle \Delta\nu \rangle &= \left(4 \frac{kT}{m_e c^2} - \frac{h\nu}{m_e c^2} \right) \nu, \\ \langle (\Delta\nu)^2 \rangle &= 2 \frac{kT}{m_e c^2} \nu^2. \end{aligned} \quad (77)$$

The higher moments prove to be at least of the order of $(kT/m_e c^2)^2$, $(kT/m_e c^2)(h\nu/m_e c^2)$ or $(h\nu/m_e c^2)^2$. Note that (77) is valid for arbitrary values of the $h\nu/kT$ ratio, including the case $kT = 0$, even though the kernel (74) itself is only applicable in the limit $h\nu(h\nu/m_e c^2) \ll kT$.

Let us next consider the limiting case $kT = 0$, $h\nu \ll m_e c^2$, when the line profile resulting from a single scattering will be shaped exclusively by the recoil effect. We shall take advantage of the fact that the scattering angle and the emergent photon frequency are uniquely related to each other via (5)², yielding

$$\cos \theta = 1 - \frac{m_e c^2}{h\nu} \frac{\nu - \nu'}{\nu}, \quad d \cos \theta = \frac{m_e c^2}{h\nu^2} d\nu'. \quad (78)$$

As a consequence, the probability $P(\nu')d\nu'$ that the photon frequency after scattering will fall in an interval $d\nu'$ can be expressed through the probability $P(\cos \theta)d \cos \theta$, i.e. through the Thomson scattering cross section (11). We thus find that the line profile is defined in the frequency range $\nu(1 - 2h\nu/m_e c^2) \leq \nu' \leq \nu$ and is given by [131]

$$P(\nu \rightarrow \nu') = \frac{3}{8} \sigma_T N_e \frac{m_e c^2}{h\nu^2} \left[1 + \left(\frac{m_e c^2}{h\nu^2} \right)^2 (\nu' - \bar{\nu})^2 \right], \quad (79)$$

where $\bar{\nu} = \nu(1 - h\nu/m_e c^2)$ is the average frequency of a scattered photon. The kernel is symmetric about $\bar{\nu}$, the point of minimum intensity.

It follows from (79), that the recoil effect leads to a scatter in the emergent frequencies. One can add the corresponding term to the expression (77) for the kernel's second moment:

$$\langle (\Delta\nu)^2 \rangle = \left[2 \frac{kT}{m_e c^2} + \frac{7}{5} \left(\frac{h\nu}{m_e c^2} \right)^2 \right] \nu^2. \quad (80)$$

The additional term is relatively small; for example, when an iron X-ray line with $h\nu = 6.4$ keV is scattered, the additional line broadening due to recoil can be neglected in comparison with the Doppler broadening if $kT \gtrsim 0.1$ keV.

² hence the kernel $K(\nu, \Omega \rightarrow \nu', \Omega')$ is a δ -function when $kT = 0$

Line profiles exhibit a cusp at $\nu' = \nu$ [in the case $kT = 0$, there is an additional cusp at $\nu' = \nu(1 - 2h\nu/m_e c^2)$], so that the profile bears no resemblance to the customary Gaussian profile of an emission line broadened by thermal or turbulent motions of ions. To demonstrate this point, let us assume that $h\nu \ll kT$. For a given plasma temperature, the emission line profile is given by $N(\nu) \sim \exp[-(\nu' - \nu)^2/\Delta\nu_D^2]$ ³, where $\Delta\nu_D = \nu(2kT/m_e c^2)^{1/2}$, so that the mean (rms) frequency shift, $\sqrt{\langle(\Delta\nu)^2\rangle} = \nu(kT/m_e c^2)^{1/2}$. The corresponding value for the electron-scattered line is larger, $\nu(2kT/m_e c^2)^{1/2}$. On the other hand, the FWHM of the Gaussian profile, $2\nu[2\ln(2)kT/m_e c^2]^{1/2}$, is larger than for $P(\nu \rightarrow \nu') - 2\nu[\ln(2)kT/m_e c^2]^{1/2}$. This reflects the fact that a large fraction of photons emerge in the wings of the Compton scattering kernel.

In the vicinity of the cusp, $|\nu' - \nu| \ll \nu(kT/m_e c^2)^{1/2}$, the single-scattering profile (74) can be expanded in terms of $(\nu' - \nu)$:

$$P(\nu \rightarrow \nu')_{+,-} = \frac{11}{20} \sigma_T N_e \nu^{-1} \sqrt{\frac{2}{\pi}} \left(\frac{kT}{m_e c^2} \right)^{-1/2} \times \left\{ 1 + \frac{\nu' - \nu}{\nu} \left[\mp \frac{15}{22} \sqrt{\frac{\pi}{2}} + \frac{1}{2} \left(1 - \frac{h\nu}{kT} \right) \left(\frac{kT}{m_e c^2} \right)^{1/2} \right] + \dots \right\}, \quad (81)$$

where the indices $+$ and $-$ correspond to the right and left wings, respectively. On either side of the cusp the spectrum can be approximated by a power law, with the slopes

$$\begin{aligned} \alpha_+ &= - \left[\frac{d \ln P}{d \ln (\nu'/\nu)} \right]_{\nu'=\nu+0} = \frac{15}{22} \sqrt{\frac{\pi}{2}} \left(\frac{kT}{m_e c^2} \right)^{-1/2} - \frac{1}{2} + \frac{1}{2} \frac{h\nu}{kT}, \\ \alpha_- &= \left[\frac{d \ln P}{d \ln (\nu'/\nu)} \right]_{\nu'=\nu-0} = \frac{15}{22} \sqrt{\frac{\pi}{2}} \left(\frac{kT}{m_e c^2} \right)^{-1/2} + \frac{1}{2} - \frac{1}{2} \frac{h\nu}{kT}; \\ \alpha_- - \alpha_+ &= 1 - \frac{h\nu}{kT}. \end{aligned} \quad (82)$$

It is interesting that when $h\nu = kT$, the line profile in the vicinity of the cusp is symmetric in logarithmic coordinates about $\nu' = \nu$ ($\alpha_+ = \alpha_-$).

2.2 Kompaneets Equation

The *Comptonization process* – the change in the spectrum of radiation due to multiple scatterings of photons with thermal electrons – is governed by the integral kinetic (70) (we consider here the isotropic problem). This equation can generally be solved by numerical methods provided that the Compton scattering kernel is known. Alternatively, Comptonization problems can be treated using Monte Carlo methods (see [132] for a review).

³ Note that the width adopted here is $(M/m_e)^{1/2} = 43(M/m_p)^{1/2}$ times the actual thermal width of lines of an ion of mass M .

In the limit that typical photon energies $h\nu$ and the plasma temperature kT are small compared to the electron rest energy $m_e c^2$, the variation in intensity at a given frequency is largely determined by transitions in a narrow spectral interval near this frequency. If the radiation spectral distribution is sufficiently smooth, the integral equation (70) can be transformed into the differential Fokker–Planck equation describing the diffusion and flow of photons in frequency space:

$$\frac{\partial n}{\partial t} = \sigma_T N_e c \frac{1}{\nu^2} \frac{\partial}{\partial \nu} \left\{ -\nu^2 n \langle \Delta \nu \rangle (1+n) + \frac{1}{2} \left[-\nu^2 n \langle (\Delta \nu)^2 \rangle \frac{\partial n}{\partial \nu} + (1+n) \frac{\partial}{\partial \nu} \nu^2 n \langle (\Delta \nu)^2 \rangle \right] \right\}. \quad (83)$$

Here, $\langle \Delta \nu \rangle$ and $\langle (\Delta \nu)^2 \rangle$ are the first and second moments of the scattering kernel $P(\nu \rightarrow \nu')$, defined in (76). Substituting the values (77) for these moments into (83), we obtain the Kompaneets [87] equation

$$\frac{\partial n}{\partial t} = \frac{\sigma_T N_e h}{m_e c} \frac{1}{\nu^2} \frac{\partial}{\partial \nu} \nu^4 \left(n + n^2 + \frac{kT}{h} \frac{\partial n}{\partial \nu} \right), \quad (84)$$

which plays a central role in the Comptonization theory. The Kompaneets equation is valid in the nonrelativistic limit ($h\nu, kT \ll m_e c^2$) and is accurate to first order in kT and $h\nu$. The first parenthesized term in (84) describes the downward photon flow along the frequency axis due to Compton recoil. The second term, which is due to recoil too, allows for induced Compton scattering. The last term describes the frequency diffusion of photons due to the Doppler effect. It is convenient to introduce dimensionless frequency $x = h\nu/kT$ and interaction time $y = \int (kT(t)/m_e c^2) \sigma_T N_e c dt$. The latter quantity is often called the *Compton parameter*. The Kompaneets equation then becomes

$$\frac{\partial n}{\partial y} = \frac{1}{x^2} \frac{\partial}{\partial x} x^4 \left(n + n^2 + \frac{\partial n}{\partial x} \right). \quad (85)$$

It is not surprising that the main properties of the Kompaneets equation reflect the similar properties of the Compton scattering kernel (see the preceding §2.1). In Compton scattering, the number of photons is conserved, and indeed the Kompaneets equation implies that

$$\frac{d}{dt} N_\gamma = \frac{d}{dt} \int n \nu^2 d\nu = 0. \quad (86)$$

In a plasma of specified temperature, the processes driving the production and absorption of photons (such as free–free processes) will leave the frequency distribution of photons unaltered only if the radiation has the Planck spectrum $n = (e^x - 1)^{-1}$ corresponding to $T_r = T$. But Compton scattering will not affect the frequency distribution for any spectrum of the form $n = (e^{x+\mu} - 1)^{-1}$ with $T_r = T$ and $\mu > 0$, that is, in the more general case

of a Bose–Einstein (BE) equilibrium distribution, as one can easily see by substituting the BE spectrum into the right-hand side of the Kompaneets equation. The chemical potential μ measures the deficiency in the number of BE photons compared with a blackbody spectrum at the same temperature. In the limit $\mu \gg 1$, the BE distribution reduces to the special case of a Wien spectrum, $n = e^{-(x+\mu)}$, or $\Sigma_\nu = 8\pi e^{-\mu} (h\nu^3/c^3) \exp(-h\nu/kT)$, a law which clearly satisfies the Kompaneets equation without the n^2 term responsible for induced processes. In the Wien distribution, the mean photon energy

$$\overline{h\nu} = kT \frac{\int_0^\infty x^3 e^{-x} dx}{\int_0^\infty x^2 e^{-x} dx} = 3kT. \quad (87)$$

We shall find out in this review that, for a given photon number, Compton scatterings tend to establish a Wien spectrum with $\overline{h\nu} = 3kT$.

Alternative Derivation of the Kompaneets Equation

Our derivation of the Kompaneets equation (84) was based on the exact knowledge (in the nonrelativistic limit) of the first two moments (77) of the Compton scattering kernel. These values, in turn, had been found from a fairly involved calculation of the Compton scattering kernel (74). However, the obvious physical requirements on the final equation impose such strong constraints on the parameters that this equation can be derived without explicitly writing down the cumbersome kernel $P(\nu \rightarrow \nu')$. It is this approach which was originally followed by Kompaneets and his collaborators [87, 199]; we describe it below.

Let us seek the first two moments of the scattering kernel in the form

$$\begin{aligned} \langle \Delta\nu \rangle &= \left(A_1 \frac{h\nu}{m_e c^2} + B_1 \frac{kT}{m_e c^2} \right) \nu, \\ \langle (\Delta\nu)^2 \rangle &= B_2 \frac{kT}{m_e c^2} \nu^2, \end{aligned} \quad (88)$$

which ensures that they will be accurate to the first order in $h\nu/m_e c^2$ and $kT/m_e c^2$. Some *a priori* information has been incorporated into (88). First, there is no term proportional to $(kT/m_e c^2)^{1/2}$ in the expression for $\langle \Delta\nu \rangle$, although the shift in frequency in an individual scattering event $\Delta\nu \sim \nu v/c \sim \nu(kT/m_e c^2)^{1/2}$. This is because such linear Doppler shifts can be both positive and negative and should cancel when the average is taken. Second, there is no term $\sim (h\nu/m_e c^2)$ in the expression for $\langle (\Delta\nu)^2 \rangle$ for the following reason. The frequency shift due to recoil $\Delta\nu \sim h\nu^2/m_e c^2$, thus the contribution of the recoil effect to the second moment should be proportional to $(h\nu/m_e c^2)^2$ and can be neglected.

We thus have three coefficients that need to be found. It turns out that once one of these coefficients is known the other two can be determined from

the general properties of the final equation. To this end, let us plug the moments (88) into the Fokker–Planck equation (83). We obtain

$$\begin{aligned} \frac{\partial n}{\partial t} = & \sigma_T N_e c \frac{1}{\nu^2} \frac{\partial}{\partial \nu} \nu^3 \\ & \times \left[\left(-A_1 \frac{h\nu}{m_e c^2} + (2B_2 - B_1) \frac{kT}{m_e c^2} \right) n(1+n) + \frac{B_2}{2} \frac{kT}{m_e c^2} \nu \frac{\partial n}{\partial \nu} \right]. \end{aligned} \quad (89)$$

One property of the final equation, namely conservation of photon number, is already reflected in the expression above – because of its divergent structure it evidently satisfies (86). Another constraint, that the equilibrium Planck distribution of photons must remain unchanged during Compton interaction, $\partial[(\exp(h\nu/kT) - 1)^{-1}]/\partial t = 0$, combined with (89), leads to the equation

$$-(A + B_2/2)h\nu + (2B_2 - B_1)m_e c^2 = 0. \quad (90)$$

This equality will be satisfied for any ν only if

$$B_1 = -4A, \quad B_2 = -2A. \quad (91)$$

Now, let us recall that the average frequency shift due to recoil is given by (65) (we recall that that relation was derived in a very straightforward manner), which immediately gives us $A_1 = -1$. We then find from (91) that $B_1 = 4$ and $B_2 = -2$. On substituting these values into (89), we rederive the Kompaneets equation.

Extension of the Kompaneets Equation

It is possible to generalize the Kompaneets equation beyond its usual range of applicability ($h\nu, kT \ll m_e c^2$) by adding to the Fokker–Planck expansion series (83) terms of higher order in $\Delta\nu$. Itoh et al. [76] and Challinor and Lasenby [27] have done this self-consistently for the mildly-relativistic regime $h\nu, kT \lesssim 0.1 m_e c^2$ by adding terms proportional to $\langle(\Delta\nu)^3\rangle$ and $\langle(\Delta\nu)^4\rangle$ and also first-order corrections to the leading two moments (77) of the scattering kernel. Before them Ross, Weaver and McCray [134], using (80), wrote down the equation

$$\frac{\partial n}{\partial \tau} = \frac{h}{m_e c^2} \frac{1}{\nu^2} \frac{\partial}{\partial \nu} \nu^4 \left(n + \frac{kT}{h} \frac{\partial n}{\partial \nu} + \frac{7}{10} \frac{h\nu^2}{m_e c^2} \frac{\partial n}{\partial \nu} \right) \quad (92)$$

(where the induced term is ignored). The new, third parenthesized term describes the diffusion of photons in frequency due to the recoil effect. This diffusion becomes of importance when narrow X-ray or gamma-ray lines are scattered on cold electrons. Such a situation takes place, for example, during a supernova explosion.

2.3 Plasma Heating and Cooling

Following Levich and Sunyaev [96], let us multiply the Kompaneets equation by $8\pi h\nu^3/c^3$ and integrate it with respect to frequency. On integrating by parts, we obtain the equation

$$\frac{d\Sigma}{dt} = -\frac{3}{2}N_e \frac{dkT}{dt} = -N_e(W_C^+ - W_C^-), \quad (93)$$

where

$$W_C^- = 4 \frac{kT}{m_e c^2} \sigma_T c \Sigma \quad (94)$$

and

$$W_C^+ = \frac{\sigma_T h}{m_e c} \int_0^\infty \nu \Sigma_\nu d\nu + \frac{\sigma_T c^2}{8\pi m_e} \int_0^\infty \frac{\Sigma_\nu^2}{\nu^2} d\nu. \quad (95)$$

The term W_C^- describes the inverse Compton cooling, and W_C^+ the Compton and induced Compton heating of the electrons.

Setting $d\Sigma/dt = 0$ in (93), we obtain an expression, derived in a different way by Peyraud [124] and Zel'dovich and Levich [196], for the stationary electron temperature in a specified radiation field:

$$kT_{\text{stat}} = \frac{1}{4\Sigma} \left(\int_0^\infty h\nu \Sigma_\nu d\nu + \frac{c^3}{8\pi} \int_0^\infty \frac{\Sigma_\nu^2}{\nu^2} d\nu \right). \quad (96)$$

It follows that $T_{\text{stat}} = T_r$ for blackbody and Bose–Einstein distributions.

2.4 Analytic Results for the Homogeneous Problem

Let us apply the Kompaneets equation to a problem which is of particular interest for cosmology. We shall examine how a given initial radiation spectrum evolves as a result of Comptonization in an unbounded homogeneous medium filled with thermal plasma at some temperature T .

Doppler Broadening and Shift

If in the Kompaneets equation (85) we neglect the first two terms in parentheses, it will describe the inverse Compton scattering in thermal plasma:

$$\frac{\partial n}{\partial y} = \frac{1}{x^2} \frac{\partial}{\partial x} x^4 \frac{\partial n}{\partial x}. \quad (97)$$

In 1969 Zel'dovich and Sunyaev [195] found the solution of this diffusion equation:

$$n(x, y) = \frac{1}{\sqrt{4\pi y}} \int_0^\infty n_0(z) \exp \left[-\frac{(\ln x + 3y - \ln z)^2}{4y} \right] \frac{dz}{z}, \quad (98)$$

which indicates how an arbitrary initial spectrum $n_0(\nu) \equiv n(\nu, 0)$ will have evolved at arbitrary time y . Multiplying (97) by $8\pi h\nu^3/c^3$ and integrating over the frequency, Kompaneets [87] found that $\Sigma = \Sigma_0 \exp(4y)$, which is exactly the result (60) obtained in §1.4.

In the case of an infinitely narrow line $\Sigma_\nu(x, 0) = \delta(x - x_0)$, or equivalently $n(x, t = 0) \sim x_0^{-3} \delta(x - x_0)$, we have the solution

$$\Sigma_\nu(x, y) = \frac{1}{\sqrt{4\pi y}} \exp \left[-\frac{(\ln x_0 - \ln x + 3y)^2}{4y} \right], \quad (99)$$

which is valid when $\tau \equiv \sigma_T N_e c t \gg 1$. This last condition means that several scatterings per photon need to occur in order for the original narrow spectral distribution to get broadened enough that the Fokker–Planck formulation of the problem is justifiable.

The line clearly will broaden with time, its center of gravity meanwhile shifting toward higher frequencies [131]. The frequency of peak intensity will increase with y as

$$x_{\max} = x_0 e^{3y}, \quad (100)$$

and the line width at half maximum will be

$$\text{FWHM} = x_0 [\exp(3y + 2\sqrt{y \ln 2}) - \exp(3y - 2\sqrt{y \ln 2})]. \quad (101)$$

So long as $y \ll 1$, the line broadening will dominate over the line shift.

The right-hand side of (99) may be considered the kernel of the truncated Kompaneets equation (97). If the initial spectral distribution is broad enough, $(\Delta\nu/\nu) \gg (2kT/m_e c^2)^{1/2}$, this kernel as well as the differential equation (97) may be applied to the single-scattering problem. That is the initial spectrum convolved with (99) for $y = (kT/m_e c^2)$ ($\tau = 1$) will nearly coincide with the actual single-scattering line profile, resulting from the convolution of the initial spectrum with the Compton scattering kernel (74). One can take advantage of this property when considering the interaction of the cosmic microwave background radiation with an optically thin, hot plasma in the universe.

In the case of an initial blackbody spectrum $n_0 = [\exp(h\nu/kT_r) - 1]^{-1}$, it is convenient to replace x in (97) with $x_r = h\nu/kT_r$. Zel'dovich and Sunyaev [195] found the first iteration (in the limit $y \ll 1$) of the solution of the resulting equation:

$$\frac{\Delta I_\nu}{I_\nu} = \frac{\Delta n}{n} = y \frac{x_r e^{x_r}}{e^{x_r} - 1} \left[x_r \frac{e^{x_r} + 1}{e^{x_r} - 1} - 4 \right], \quad (102)$$

$$\frac{\Delta T_r}{T_r} = \frac{d \ln I_\nu}{d \ln T_r} \frac{\Delta I_\nu}{I_\nu} = y \left[x_r \frac{e^{x_r} + 1}{e^{x_r} - 1} - 4 \right]. \quad (103)$$

This solution is valid in the limit $T_r \ll T$. Equation (103) describes the variation in the brightness temperature. In the Rayleigh–Jeans region ($x_r \ll 1$) of the spectrum, $\Delta T_{RJ}/T_r = -2y$ (for $y \ll 1$). The general solution (98) leads to the law $T_{RJ} = T_r \exp(-2y)$ for arbitrary y .

Recoil Effect

When the temperature of the blackbody radiation is not small compared to T , the terms associated with the recoil effect in the Kompaneets equation (85) will become important. However, (102) will still correctly describe small spectral deviations if we redefine the variable y [186] as

$$y = \frac{k(T - T_r)}{m_e c^2} \sigma_T N_e c t . \quad (104)$$

If $h\nu \gg 4kT$, the time evolution of the line will be determined not by the Doppler effect but by the recoil that results from electron scattering; see (5). The recoil effect should clearly have a substantial influence on the evolution of the spectrum of an X-ray or gamma-ray line. If in the Kompaneets equation (85) we neglect the last two parenthesized terms (the induced scattering and the Doppler frequency shift due to the scattering), then the equation will describe the evolution of the spectrum evolves in the homogeneous case due to the recoil effect:

$$\frac{\partial n}{\partial t} = \frac{1}{x^2} \frac{\partial(x^4 n)}{\partial x} . \quad (105)$$

Arons [2] and Illarionov and Sunyaev [70] have solved this equation: the quantity $n\nu^4$ will be conserved as motion takes place along the trajectory $d\nu/du = -\nu^2$, where $du = (h/m_e c) \sigma_T N_e dt$. To this approximation, the line will evidently remain monochromatic as it evolves, and it can only shift downwards along the frequency axis. Actually, however, the amplitude of the recoil effect depends on the scattering angle ($0 < \Delta\nu/\nu < 2h\nu/m_e c^2$), so the line should in fact broaden somewhat [70, 74, 134]. We already pointed out this broadening effect during our discussion of the Compton scattering kernel (in §2.1) and the Kompaneets equation (in §2.2).

2.5 Induced Compton Scattering

The nonlinear term proportional to n^2 in the Kompaneets equation (84) represents the contribution of induced, or stimulated Compton scattering, which becomes important when $n = c^3 \Sigma_\nu / 8\pi h \nu^3 > 1$. This process is explained by classical electrodynamics [199], and the final expressions, written in terms of spectral energy density Σ_ν rather than n , do not contain the Planck constant. However, its treatment is more straightforward in the framework of quantum theory of photon scattering.

Spectral Evolution and Bose Condensation

How will a radiation spectrum evolve as a result of induced Compton scatterings in an infinite homogeneous plasma? To answer this question, let us consider the Kompaneets equation with only the quadratic term (n^2) left:

$$\frac{\partial n}{\partial \tau} = \frac{h}{m_e c^2} \frac{1}{\nu^2} \frac{\partial}{\partial \nu} \nu^4 n^2, \quad (106)$$

where $\tau = \sigma_T N_e c t$. If we define $f = h n \nu^2$, then the equation simplifies:

$$\frac{\partial f}{\partial \tau} = \frac{2f}{m_e c^2} \frac{\partial f}{\partial \nu}, \quad (107)$$

and can be solved in terms of characteristics; this means that it can be subjected to the further transformation

$$\frac{df}{d\tau} = 0 \quad \text{along} \quad \frac{d\nu}{d\tau} = -\frac{2f}{m_e c^2}. \quad (108)$$

The implicit solution for $\nu(f, \tau)$ has the form

$$\nu(f, \tau) = \nu_0(f) - \frac{2f}{m_e c^2} \tau. \quad (109)$$

The corresponding evolution of the spectrum is very easy to visualize. Let us specify a spectrum in the f - ν coordinates at the instant $\tau = 0$. Each point of the curve moves to the left with a constant, time-dependent velocity. However, this velocity is different for different points – it is proportional to the ordinate of a point. Thus for each point of the initial curve $f_0(\nu)$, it is easy to determine the instant at which it intersects the vertical axis ($\nu = 0$). Now, of course, there can be no zero-frequency photons. Some mechanisms of genuine absorption are bound to appear as $\nu \rightarrow 0$.

Under certain conditions we can expect a spectrum f which has a bend. In that case, even before the Bose condensation described above the formal treatment of the evolution of the spectrum leads to the formation of a characteristic three-valued structure. This phenomenon is completely analogous to the formation of shock waves in gas dynamics. It is impossible to study the structure and subsequent fate of a shock wave using the Kompaneets differential equation, which was derived under the assumption that the spectrum is smooth. In this case, it is necessary to take into account the thermal motions of the scattering electrons and consider the integral kinetic equation:

$$\begin{aligned} \frac{\partial n(\nu, \boldsymbol{\Omega})}{\partial \tau} &= n(\nu, \boldsymbol{\Omega}) \int n(\nu', \boldsymbol{\Omega}') \\ &\times \left[\left(\frac{\nu'}{\nu} \right)^2 K(\nu', \boldsymbol{\Omega}' \rightarrow \nu, \boldsymbol{\Omega}) - K(\nu, \boldsymbol{\Omega} \rightarrow \nu', \boldsymbol{\Omega}') \right] d\nu' d\boldsymbol{\Omega}' \\ &\equiv n(\nu, \boldsymbol{\Omega}) \int n(\nu', \boldsymbol{\Omega}') K_{\text{ind}}(\nu, \boldsymbol{\Omega}; \nu', \boldsymbol{\Omega}') d\nu' d\boldsymbol{\Omega}', \end{aligned} \quad (110)$$

in which we made allowance for possible angular anisotropy of the radiation. The kernel for induced Compton scattering is given by [142, 198]

$$\begin{aligned}
K_{\text{ind}}(\nu, \boldsymbol{\Omega}; \nu', \boldsymbol{\Omega}') &= \frac{3}{32\pi} \sqrt{\frac{2}{\pi}} \sigma_{\text{T}} \left(\frac{kT}{m_e c^2} \right)^{-3/2} \frac{h\nu'(\nu' - \nu)}{m_e c^2 g\nu} (1 + \cos \theta^2) \\
&\quad \times \exp \left[-\frac{(\nu' - \nu)^2 m_e c^2}{2g^2 kT} \right], \\
g = |\nu\boldsymbol{\Omega} - \nu'\boldsymbol{\Omega}'| &= (\nu^2 - 2\nu\nu' \cos \theta + \nu'^2)^{1/2}.
\end{aligned} \tag{111}$$

The characteristic width of K_{ind} is determined by the Doppler broadening, $\Delta\nu \sim (kT/m_e c^2)^{1/2} \nu$, which has the meaning of a free path length of photons in frequency space. By solving the integral kinetic equation, one finds that instead of a simple smoothing of the shock, an oscillatory structure and quasy lines develop with time in the photon spectrum.

Let us consider several astrophysical applications where induced Compton scattering may play a key role.

Plasma Heating

Astronomical radio and infrared sources often exhibit a very high radiation brightness temperature $kT_{\text{b}} = nh\nu \gg m_e c^2$ at low frequencies. Since the brightness temperature usually greatly decreases toward short wavelengths, the radiation flux proves to be extremely small compared with blackbody radiation of temperature T_{r} equal to the T_{b} at low frequencies. In the case of Compton interaction with radio or infrared radiation, however, the electrons “feel” the brightness temperature of the long-wavelength part of the spectrum to a greater extent than the total energy of radiation or the average photon energy. This is due to the high probability (proportional to $n + 1$) of induced interaction of electrons with low-frequency radiation. Though the energy of each photon is quite small, the collision of an electron with a photon is so highly probable that the induced Compton interaction results in electrons taking up considerable energy from the radiation field. As a result, the steady state electron temperature may considerably exceed the average energy of photons. Moreover, the electron temperature tends to approach the brightness temperature of the low-frequency radiation [124, 196].

When electrons exchange energy by Compton scattering with an isotropic field of radiation, they will be heated at the rate

$$W^+ = \frac{\sigma_{\text{T}} c^2}{8\pi m_e} \int_0^\infty \frac{\Sigma_\nu^2}{\nu^2} d\nu. \tag{112}$$

This is the same expression as (95), but without the term responsible for spontaneous scattering. Accordingly, the stationary electron distribution will be Maxwellian with the temperature

$$kT = \frac{c^3}{32\pi\Sigma} \int \frac{\Sigma_\nu^2}{\nu^2} d\nu, \tag{113}$$

where it has been assumed that the electrons cool by inverse Compton scattering. If the effective width of the spectrum $\Delta\nu \sim \nu$, then (113) may be written as

$$T \sim \frac{T_b}{4}. \quad (114)$$

The expressions above are valid in the nonrelativistic limit $kT \sim kT_b \ll m_e c^2$. According to (114), electrons could be heated up to relativistic temperatures $kT \sim m_e c^2$ at a relatively low $T_b \sim 10^{10}$ K, but this is in fact a gross overestimation. An accurate relativistic treatment of the problem (see [73] and references therein) demonstrates that for $kT_b \gg m_e c^2$ the resulting electron momentum distribution will be nonthermal, unless relaxation processes can rapidly Maxwellize the plasma. As we have shown in [142], plasma can be heated only up to mildly relativistic temperatures $kT \lesssim 0.1 m_e c^2 \sim 10$ – 100 keV in the presence of low-frequency, isotropic radiation of temperature $T_b \sim 10^{11}$ – 10^{12} K typical of powerful extragalactic radio sources.

In the situation where plasma is irradiated by an external source, the radiation field will be strongly anisotropic and the steady-state electron momentum distribution will be characterized by two temperatures [193]:

$$\begin{aligned} kT_{\parallel} &= \frac{3c^3}{512\pi\Sigma} \int \frac{\Sigma_{\nu}^2}{\nu^2} d\nu \left(\frac{R}{r}\right)^4, \\ kT_{\perp} &= \frac{3c^3}{128\pi\Sigma} \int \frac{\Sigma_{\nu}^2}{\nu^2} d\nu \left(\frac{R}{r}\right)^2, \\ \text{since} \\ W^+ &= \frac{3\sigma_T c^2}{64\pi m_e} \int \frac{\Sigma_{\nu}^2}{\nu^2} d\nu \left(\frac{R}{r}\right)^2, \\ W^- &= 4 \frac{kT}{m_e c} \sigma_T \Sigma. \end{aligned} \quad (115)$$

Here R represents the characteristic size of the radiation source, $r \gg R$ is the distance between the source and the site of the plasma heating, and $\Sigma_{\nu} \sim (R/r)^2$ represents the local radiation spectral density. In terms of the radiation brightness temperature at the source surface $T_b(R)$,

$$\begin{aligned} T_{\parallel} &\sim \frac{3}{64} \left(\frac{R}{r}\right)^6 T_b(R), \\ T_{\perp} &\sim \frac{3}{16} \left(\frac{R}{r}\right)^4 T_b(R). \end{aligned} \quad (116)$$

These steep dependences on distance result from the greatly reduced efficiency of the induced Compton heating in an anisotropic field as compared to the isotropic situation. Not only the energy density of the radiation drops with moving away from the source, but also only narrow-angle induced scatterings are possible since the radiation is collimated in a narrow beam ($\alpha \lesssim r/R$).

Coulomb collisions will tend to isotropize the system of electrons, imparting a unique temperature to it. This thermalization can actually take place if the characteristic heating time of electrons, $t_{\text{heat}} = kT/W^+$, is shorter than the characteristic time for Coulomb collisions [156]

$$t_{e-e} = 5 \times 10^{12} (\ln A/20)^{-1} \left(\frac{kT}{m_e c^2} \right)^{3/2} N_e^{-1} \text{ s} . \quad (117)$$

If, as we have assumed so far, the heating is driven by the induced Compton process and the cooling is due to inverse Compton scattering, the heating time will be equal to the Compton cooling time given by (47). However, in the presence of a more efficient cooling mechanism the stationary electron temperature and consequently t_{heat} will be reduced. For example, bremsstrahlung losses, with $W_{\text{ff}}^- \approx 10^{-27} N_e T^{1/2} \text{ erg s}^{-1}$, will dominate Compton cooling when $T^{1/2} \Sigma / N_e < 10^{-4} \text{ K}^{1/2} \text{ erg}$.

Induced Radiation Force

In continuation of the discussion started in §1.3, let us consider the force exerted by a radiation field on an electron at rest. By definition, this force is equal to the rate of change of the electron momentum,

$$\mathbf{f} = \left\langle \frac{\Delta \mathbf{p}}{\Delta t} \right\rangle = \int \Delta \mathbf{p} \frac{I_\nu(\nu, \boldsymbol{\Omega})}{h\nu} \frac{d\sigma}{d\Omega'} [1 + n(\nu', \boldsymbol{\Omega}')] d\nu d\boldsymbol{\Omega} d\boldsymbol{\Omega}' , \quad (118)$$

where $\Delta \mathbf{p} = h(\nu \boldsymbol{\Omega} - \nu' \boldsymbol{\Omega}')/c$, ν' is given by (5), and $d\sigma/d\Omega'$ is the Thomson scattering cross section (11). When only spontaneous scattering is taken into account [the term $n(\nu', \boldsymbol{\Omega}')$ is omitted] and the recoil frequency shift is neglected ($\nu' = \nu$), we come to the familiar expression

$$\mathbf{f}_{\text{sp}} = \frac{\sigma_{\text{T}} \mathbf{q}}{c} , \quad (119)$$

where $\mathbf{q} = \int I_\nu(\nu, \boldsymbol{\Omega}) \boldsymbol{\Omega} d\nu d\boldsymbol{\Omega}$ is the radiation flux.

There would be no additional contribution from induced Compton scattering to the force (119) if the photon frequency remained unchanged after scattering. Indeed, taking into account the term $n(\nu', \boldsymbol{\Omega}')$ in (118) but assuming $\nu' = \nu$ as before, we find that the contribution of the induced effect is proportional to $\int n(\boldsymbol{\Omega}) n(\boldsymbol{\Omega}') (\boldsymbol{\Omega} - \boldsymbol{\Omega}') [1 + (\boldsymbol{\Omega} \boldsymbol{\Omega}')^2] d\boldsymbol{\Omega} d\boldsymbol{\Omega}' = 0$.

However, in reality the photon frequency diminishes by a tiny amount, $\Delta\nu \sim -h\nu^2/m_e c^2$, during a scattering event, which gives rise to induced radiation pressure [95]:

$$\begin{aligned} \mathbf{f}_{\text{ind}} = & \frac{3\sigma_{\text{T}}}{16\pi m_e c} \int [1 + (\boldsymbol{\Omega} \boldsymbol{\Omega}')^2] \left[\frac{I_\nu(\nu, \boldsymbol{\Omega}) I_\nu(\nu, \boldsymbol{\Omega}')}{\nu^2} \right. \\ & \left. + I_\nu(\nu, \boldsymbol{\Omega}) \frac{\partial}{\partial \nu} \left(\frac{I_\nu(\nu, \boldsymbol{\Omega}')}{\nu^3} \right) \nu^2 \right] d\nu d\boldsymbol{\Omega} d\boldsymbol{\Omega}' . \end{aligned} \quad (120)$$

One can derive the above expression with the help of the approximation

$$n(\nu', \boldsymbol{\Omega}') = n(\nu, \boldsymbol{\Omega}') + \frac{h\nu^2}{m_e c^2} (1 - \boldsymbol{\Omega}\boldsymbol{\Omega}') \frac{\partial n(\nu, \boldsymbol{\Omega})}{\partial \nu} . \quad (121)$$

The full force acting on the electron is of course

$$\mathbf{f} = \mathbf{f}_{\text{sp}} + \mathbf{f}_{\text{ind}} . \quad (122)$$

If an anisotropic radiation field is produced by a distant source, then the force will be [95]

$$\mathbf{f} = \mathbf{f}_{\text{sp}} + \mathbf{f}_{\text{ind}} = \frac{\sigma_{\text{T}} \mathbf{q}}{c} + \frac{\sigma_{\text{T}} c^2}{16\pi m_e} \int \frac{\Sigma_{\nu}^2}{\nu^2} d\nu \left(\frac{R}{r} \right)^2 \frac{\mathbf{r}}{r} . \quad (123)$$

Induced light pressure rapidly decreases with the distance from the source: $\mathbf{f}_{\text{ind}} \sim r^{-6}$, as compared to $\mathbf{f}_{\text{sp}} \sim r^{-2}$. In terms of the brightness temperature (123) may be written as

$$\mathbf{f} \approx \mathbf{f}_{\text{sp}} \left[1 + \frac{kT_{\text{b}}(R)}{m_e c^2} \left(\frac{R}{r} \right)^4 \right] , \quad (124)$$

where $T_{\text{b}}(R)$ is the radiation brightness temperature at the source surface.

We note that (124) correctly describes the force acting on an electron moving with velocity $v \ll c$ provided that the radiation spectrum is not too narrow, which means that its effective width must be larger than the characteristic Doppler frequency shift (taking into account that only small-angle induced scatterings with $\theta \lesssim R/r$ are possible far from the source),

$$\frac{\Delta\nu}{\nu} \gg \frac{v}{c} \frac{R}{r} . \quad (125)$$

2.6 Photon Production Mechanisms

Compton scattering conserves the number of photons. In actual situations there will always be processes operating that produce new photons and absorb photons. Among such mechanisms are free-free processes and double Compton scattering, considered below.

Bremsstrahlung

Bremsstrahlung (free-free emission) is the radiation associated with the acceleration of electrons in the electrostatic fields of ions and the nuclei of atoms. We shall restrict our consideration below to the case of hot ionized gas with a Maxwellian distribution of electron velocities.

The spectral emissivity of thermal plasma at frequency ν is given by

$$\begin{aligned}\epsilon_{\nu}^{\text{ff}} &= \sqrt{\frac{8}{3\pi}} \alpha \sigma_{\text{T}} h c \left(\frac{m_e c^2}{kT} \right)^{1/2} \exp\left(-\frac{h\nu}{kT}\right) g(\nu, T) N_e \sum N_i Z_i^2 \\ &= 6.8 \times 10^{-38} T^{-1/2} \exp(-x) g(T, x) N_e \sum N_i Z_i^2 \\ &\quad \text{erg cm}^{-3} \text{ s}^{-1} \text{ Hz}^{-1},\end{aligned}\tag{126}$$

where $x = h\nu/kT$, $\alpha = 2\pi e^2/hc \approx 1/137$ is the fine-structure constant, $\sigma_{\text{T}} = 6.65 \times 10^{-25} \text{ cm}^2$ the Thomson cross section, T the plasma temperature (in K), N_e the electron number density (in cm^{-3}) and N_i the number density of ions of charge Z_i (in cm^{-3}). Finally, $g(T, x)$ is the Gaunt factor, for which accurate approximations in the broad parameter range $1 \leq Z_i \leq 28$, $6.0 \leq \log T \leq 8.5$, $-4 \leq \log x \leq 1$ have been presented by Itoh et al. [77].

There is also the related process of bremsstrahlung absorption. The corresponding absorption coefficient α_{ν}^{ff} and photon mean free path λ_{ff} are related to the volume emissivity given by (126) through Kirchhoff's law:

$$\alpha_{\nu}^{\text{ff}} = \frac{1}{\lambda_{\text{ff}}} = \frac{\epsilon_{\nu}^{\text{ff}}}{4\pi} \frac{c^2}{2h\nu^3} \left[\exp\left(\frac{h\nu}{kT}\right) - 1 \right].\tag{127}$$

It follows that λ_{ff} is smaller than $\lambda_{\text{T}} = (\sigma_{\text{T}} N_e)^{-1}$, the photon mean free path for Thomson scattering, if

$$T^{-7/2} \sum N_i Z_i^2 < 1.7 \times 10^{-2} \frac{x^3}{g(x)} (1 - e^{-x})^{-1}.\tag{128}$$

The frequency-integrated bremsstrahlung emissivity is given by

$$\epsilon_{\text{ff}} = \int \epsilon_{\nu}^{\text{ff}} d\nu = 1.43 \times 10^{-27} T^{1/2} g(T) N_e \sum N_i Z_i^2 \text{ erg cm}^{-3} \text{ s}^{-1},\tag{129}$$

where $g(T) \approx 1.3$ (see [78] for a more accurate description). We may compare the plasma energy losses due to bremsstrahlung with those due to inverse Compton cooling:

$$\epsilon_{\text{Comp}} = 1.34 \times 10^{-23} \Sigma N_e T \text{ erg cm}^{-3} \text{ s}^{-1}.\tag{130}$$

The latter expression follows directly from (59), Σ is the radiation energy density, and it is assumed that $\langle h\nu \rangle \ll kT$, so that Compton heating is unimportant. Thus, the Compton cooling will dominate over the free-free cooling when

$$\Sigma^{-1} T^{-1/2} \sum N_i Z_i^2 < 1.0 \times 10^4,\tag{131}$$

i.e. in rarefied, high-temperature plasma.

Kompaneets [87] wrote down the kinetic equation describing the joint action of Compton scattering and free-free emission and absorption, including the corresponding induced processes:

$$\frac{\partial n}{\partial t} = \frac{a}{x^2} \frac{\partial}{\partial x} x^4 \left[\frac{\partial n}{\partial x} + n(1+n) \right] + \frac{K_{\text{ff}}(x)e^{-x}}{x^3} [1 - (e^x - 1)n], \quad (132)$$

where the rate of the Compton processes is specified by the parameter

$$a = \frac{kT}{m_e c^2} \sigma_T N_e c = 3.4 \times 10^{-24} N_e T \quad (133)$$

and of the free-free processes by the parameter

$$K_{\text{ff}}(x) = 1.22 \times 10^{-12} N_e^2 T^{-7/2} g(T, x), \quad (134)$$

where we have assumed for simplicity a hydrogen plasma ($\sum N_i Z^2 = N_e$).

The quantity $K_{\text{ff}}(x)$ is proportional to the square of the electron density. In most of the problems involving a rarefied plasma $K(x)$ can be completely neglected, or neglected everywhere except in a small region $x < x_0$, with x_0 given by

$$x_0^{\text{ff}} = \sqrt{\frac{K_{\text{ff}}(x_0^{\text{ff}})}{4a}} \approx 3 \times 10^5 N_e^{1/2} T^{-9/4} \sqrt{g(T, x_0)}. \quad (135)$$

For $x \leq x_0 < 1$, free-free processes dominate (the bremsstrahlung contribution to the Kompaneets equation grows like x^{-3} as $x \rightarrow 0$) and the Rayleigh-Jeans spectrum $n(x) = 1/x$ is maintained, but for $x > x_0$, Compton scattering causes photons to move upward along the frequency axis.

Modified Blackbody Spectrum

Compton scattering on free electrons plays a major role in the formation of emission spectra of accretion disks. The standard thin disk [148] is composed of three parts differing in physical properties. In the outer zone, the opacity is determined by free-free absorption and other mechanisms. In the intermediate and inner regions (the latter may be absent at low accretion rates), the reverse situation takes place: electron scattering gives the main contribution to opacity for typical photons. Electron scattering dominates absorption also in the hot atmospheres of bursters. The radiation emergent from such regions has a nonthermal spectrum.

Consider the formation of radiation spectra in an accretion disk. In its outer zone, a Planck spectrum is formed (since the optical depth $\tau_{\text{ff}} \gg 1$), with the flux emergent from the surface given by

$$F_\nu(x) = \pi B_\nu(x) = \frac{2\pi h}{c^2} \left(\frac{kT}{h} \right)^3 \frac{x^3}{e^x - 1}, \quad \text{where } x = \frac{h\nu}{kT}, \quad (136)$$

and we have assumed that the disk (at a given radius) may be considered an isothermal atmosphere. In reality, the spectrum at a given frequency ν forms at an optical depth $\tau_{\text{ff}}(\nu) \sim 1$ below the surface, which is characterized by its own temperature, so the actual spectrum will somewhat deviate from (136).

In the intermediate region, photons at sufficiently high frequencies, such that $\lambda_{\text{ff}} > \lambda_{\text{T}}$, where $\lambda_{\text{ff}}(\nu)$ is given by (127), may undergo many scatterings before escaping from the surface. Let N be the total number of scatterings experienced by a photon. Then, the total zigzag path of the photon will be $\Delta s(\nu) = N\lambda_{\text{T}}$. At the same time, the distance traversed by the photon in the vertical direction will be smaller, $\Delta z(\nu) = N^{1/2}\lambda_{\text{T}}$. Since typically $\Delta s \sim \lambda_{\text{ff}}(\nu)$, we find that $N(\nu) \sim \lambda_{\text{ff}}(\nu)/\lambda_{\text{T}}$ and

$$\Delta z(\nu) \sim [\lambda_{\text{ff}}(\nu)\lambda_{\text{T}}]^{1/2}. \quad (137)$$

The surface brightness of the disk at a specified frequency ν represents summed bremsstrahlung emission from the layer $0 \leq z \leq \Delta z(\nu)$. In the case of a homogeneous and isothermal atmosphere with temperature T , the emergent flux at high frequencies is given by [46, 146]

$$F_{\nu}(x) \approx \pi B_{\nu}(x) \left[\frac{\lambda_{\text{T}}}{\lambda_{\nu}^{\text{ff}}(x)} \right]^{1/2} = \text{const} \sqrt{N_{\text{e}}} T^{5/4} \frac{x^{3/2} e^{-x}}{(1 - e^{-x})^{1/2}}, \quad \lambda_{\nu}^{\text{ff}} \gg \lambda_{\text{T}}. \quad (138)$$

The dependence (138) is called a *modified blackbody spectrum*.

The overall emergent spectrum, including the low-frequency region where $\lambda_{\nu}^{\text{ff}} < \lambda_{\text{T}}$, is approximately given by

$$F_{\nu} \approx \pi B_{\nu} \left(\frac{\tau_{\nu}^{\text{ff}}}{\tau_{\nu}^{\text{ff}} + \tau_{\text{T}}} \right)^{1/2} \left[1 - \exp(-\sqrt{\tau_{\nu}^{\text{ff}}(\tau_{\nu}^{\text{ff}} + \tau_{\text{T}})}) \right]. \quad (139)$$

Here $\tau_{\text{T}} \gg 1$ and τ_{ν}^{ff} are the vertical optical depths of the disk for Thomson scattering and free-free absorption, respectively. One can distinguish three spectral zones:

- In the region $\nu < \nu_1$ where $\tau_{\text{ff}} \gg \tau_{\text{T}}$, the spectrum is blackbody, $F_{\nu} = \pi B_{\nu}$. For this region one usually has $h\nu \ll kT$, so a Rayleigh–Jeans spectral distribution, $F_{\nu} \sim \nu^2$, results.
- In the region $\nu_1 < \nu < \nu_2$ where $\tau_{\text{ff}} \ll \tau_{\text{T}}$ and the *effective optical depth* $\tau^* \equiv \sqrt{\tau_{\text{ff}}\tau_{\text{T}}} \gg 1$, F_{ν} is given by (138). If additionally $h\nu \ll kT$ in this transition region, then $F_{\nu} \sim \nu$, and the width of the region is $\nu_2/\nu_1 \sim \tau_{\text{T}}$.
- For $\nu > \nu_2$, when the two inequalities $\tau_{\text{ff}} \ll \tau_{\text{T}}$ and $\tau^* \ll 1$ are simultaneously satisfied, the atmosphere becomes translucent (photons are never absorbed) and F_{ν} assumes the $\exp(-h\nu/kT)$ shape of the thermal bremsstrahlung emissivity curve.

In the hot, radiation-dominated inner zone of the disk, the energy of a typical photon changes appreciably due to the Doppler and recoil effects during multiple electron scatterings: $\Delta\nu/\nu \sim N(4kT/mc^2) \sim \tau_{\text{T}}^2(kT/mc^2) > 1$. As a result, the Comptonization spectrum

$$F_{\nu}(x) \sim x^3 e^{-x} \quad (140)$$

is formed [70].

Double Compton Effect

When a photon is scattered by an electron,

$$\gamma_1 + e \rightarrow \gamma'_1 + e',$$

there is a small but finite probability that an additional, soft photon γ_2 will be emitted:

$$\gamma_1 + e \rightarrow \gamma'_1 + \gamma_2 + e',$$

just as in the elastic scattering of an electron by a proton,

$$e + p \rightarrow e' + p',$$

there is a small but finite probability of photon emission:

$$e + p \rightarrow e' + p' + \gamma,$$

which is the bremsstrahlung process.

In bremsstrahlung, the photon production probability is proportional to the square of the plasma density, but in the case of double Compton emission it is proportional to the product of the electron density N_e and the photon density N_γ . Hence if $N_\gamma \gg N_e$, the double Compton effect could become an important source of photons.

In the nonrelativistic case ($h\nu_1 \ll m_e c^2$, $v \ll c$), the cross section for emission of a photon of frequency $\nu_2 \ll \nu_1$ is given by [15]

$$d\sigma_{\text{DC}} = \frac{4\alpha}{3\pi} \left(\frac{h\nu_1}{m_e c^2} \right)^2 (1 - \cos \theta_1) \frac{d\nu_2}{\nu_2} d\sigma_C, \quad (141)$$

where θ_1 is the scattering angle for the first photon,

$$d\sigma_C = \frac{3}{8} \sigma_T (1 + \cos^2 \theta_1) d\cos \theta_1 \quad (142)$$

represents the Thomson differential scattering cross section, and $\alpha = 1/137$ is the fine-structure constant. Integrating (141) over all scattering angles gives

$$d\sigma_{\text{DC}} = \frac{4\alpha}{3\pi} \sigma_T \left(\frac{h\nu_1}{m_e c^2} \right)^2 \frac{d\nu_2}{\nu_2}. \quad (143)$$

Note that the cross-section (143) is of the same order in α as the bremsstrahlung cross section [15].

When the constraint $\nu_2 \ll \nu_1$ is relaxed, the more general formula [60]

$$d\sigma_{\text{DC}} = \frac{2\alpha}{3\pi} \sigma_T \left(\frac{h\nu_1}{m_e c^2} \right)^2 F(w) \frac{d\nu_2}{\nu_1} \quad (144)$$

applies, where

$$F(w) = w(1-w) \left[\frac{1+(1-w)^2}{w^2} + \frac{1+w^2}{(1-w)^2} + w^2 + (1-w)^2 \right], \quad w = \frac{\nu_2}{\nu_1}. \quad (145)$$

The function $F(w)$ is symmetric around $w = 1/2$, i.e. $F(w) = F(1-w)$, which is expected because a specification of the energy of one outgoing photon determines that of the other, their total being fixed. The normalization of $F(w)$ chosen in (144) requires that this formula be used for $0 \leq \nu_2 \leq \nu_1/2$. In the limit $w \rightarrow 0$, $F(\nu_2/\nu_1)/\nu_1 \rightarrow 2/\nu_2$, and (144) reduces to (143).

The volume emissivity of ionized plasma due to double Compton scattering (without the induced process) can thus be expressed by

$$\begin{aligned} \epsilon_\nu^{\text{DC}}(\nu \equiv \nu_2) &= N_e c \int_{\nu_2}^{\infty} h\nu \frac{d\sigma_{\text{DC}}}{d\nu} \frac{\Sigma_\nu(\nu_1)}{h\nu_1} d\nu_1 \\ &\approx \frac{4}{3} \frac{\alpha}{\pi} \sigma_T \frac{h^2}{m_e^2 c^3} N_e \int_{\nu_2}^{\infty} \Sigma_\nu(\nu_1) \nu_1 d\nu_1. \end{aligned} \quad (146)$$

If we consider blackbody radiation ($\Sigma_\nu = 8\pi h\nu^3 c^{-3} [\exp(h\nu/kT + \mu) - 1]^{-1}$), then

$$\begin{aligned} \epsilon_\nu^{\text{DC}}(\text{BB}) &= 2.66 \times 10^2 \alpha h c \sigma_T \left(\frac{kT}{hc} \right)^3 \left(\frac{kT}{m_e c^2} \right)^2 N_e \\ &= 4.4 \alpha h c \sigma_T \left(\frac{kT}{m_e c^2} \right)^2 N_\gamma(\text{BB}) N_e \end{aligned} \quad (147)$$

in the range $h\nu_2 \ll kT$, for which the lower limit of integration in (146) may be set equal to zero. Here $N_\gamma(\text{BB}) = 60.4(kT/hc)^3$ is the photon number density. Since $\Sigma_\nu(\nu_1)$ falls off exponentially when $h\nu_1 > kT$, the form of the expression (146) indicates that $\epsilon_\nu^{\text{DC}}(\nu_2)$ similarly should decline exponentially for $h\nu_2 > kT$.

For a Wien spectrum with $T_r = T$ ($\Sigma_\nu = 8\pi h\nu^3 c^{-3} [\exp(h\nu/kT + \mu)]^{-1}$, $\mu \gg 1$), we find that

$$\epsilon_\nu^{\text{DC}}(\text{Wien}) = 5.1 \alpha h c \sigma_T \left(\frac{kT}{m_e c^2} \right)^2 N_\gamma(\text{Wien}) N_e f(x), \quad (148)$$

where $N_\gamma(\text{Wien}) = 50.1(kT/hc)^3 e^{-\mu}$, and

$$f(x) = e^{-x} \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} \right) \approx 1 - \frac{x^5}{120} + \dots \quad (149)$$

is the frequency correction, which becomes important when $h\nu_2 \gtrsim kT$.

One can therefore estimate for Bose-Einstein spectral distributions ($\Sigma_\nu = 8\pi h\nu^3 c^{-3} [\exp(h\nu/kT + \mu) - 1]^{-1}$) the ratio of the emissivities due to double Compton scattering (in the limit $h\nu \ll kT$) and due to bremsstrahlung as

$$\frac{\epsilon_\nu^{\text{DC}}(\text{BE})}{\epsilon_\nu^{\text{ff}}} \approx \frac{5}{g(T, \nu)} \left(\frac{kT}{m_e c^2} \right)^{5/2} \frac{N_\gamma(\text{BE})}{N_e}. \quad (150)$$

It is possible to add to the Kompaneets equation a term representing double Compton emission and absorption, similarly as we did before for the bremsstrahlung processes [98, 173]:

$$\frac{\partial n}{\partial t} = \frac{a}{x^2} \frac{\partial}{\partial x} x^4 \left[\frac{\partial n}{\partial x} + n(1+n) \right] + \frac{K_{\text{DC}}(x)}{x^3} [1 - (e^x - 1)n] , \quad (151)$$

where $x = h\nu/kT$ and

$$K_{\text{DC}}(x) = \frac{4\alpha\sigma_{\text{T}}N_{\text{e}}c}{3\pi} \left(\frac{kT}{m_{\text{e}}c^2} \right)^2 \int_0^\infty [1 + n(x_1)]n(x_1)x_1^4 dx_1 . \quad (152)$$

The (151) is strictly valid in the soft-photon limit, i.e. at frequencies $\nu \ll \nu_1 \ll m_{\text{e}}c^2/h$, where ν_1 represents typical photon frequencies contributing to the integral in (152); [24, 31] describe frequency and mildly-relativistic temperature corrections to this expression.

In the case of blackbody radiation ($T_{\text{r}} = T$),

$$K_{\text{DC}} \approx 11.0\alpha\sigma_{\text{T}}N_{\text{e}}c \left(\frac{kT}{m_{\text{e}}c^2} \right)^2 = 4.6 \times 10^{-35} N_{\text{e}}T^2 , \quad (153)$$

Therefore, the double Compton effect will be an important process in comparison with Compton scattering at frequencies below

$$x_0^{\text{DC}} \approx \sqrt{\frac{K_{\text{DC}}}{4a}} = 1.8 \times 10^{-6} T^{1/2} , \quad (154)$$

which may be compared with the corresponding frequency for bremsstrahlung (135).

The astrophysical role of the double Compton effect has been considered [59], with specific applications to the universe [24, 37], stellar interiors [173], and high-temperature astrophysical plasma [98, 171].

3 Comptonization in Bounded Plasma Clouds

In early attempts to calculate the spectra of X-ray sources, the results of the cosmologically important problem about Comptonization in an unbounded homogeneous medium (see §2.2) were naively carried over to the situation prevailing in a spatially bounded plasma cloud, where the distribution of photons with respect to the time when they escape from the source plays a key role. Different photons will undergo a differing number of collisions there, decisively affecting the radiation spectrum formed through Comptonization and emerging from the plasma cloud.

3.1 Spatial Problem

The importance of solving the spatially limited problem was recognized simultaneously and independently by Katz [82], Shapiro et al. [149] and Pozdnyakov et al. [130]. In the first two papers the analysis relied on a solution of the stationary Kompaneets equation ([82] adopted a numerical approach while [149] solved it analytically for a single set of parameter values), whereas the calculations in [130] were performed by the Monte-Carlo method. Naturally, very similar results were obtained: in the case of a low-frequency ($h\nu \ll kT_e$) photon source the radiation emerging from the cloud was found to have a power-law spectrum at low frequencies ($h\nu < kT_e$) but an exponential cutoff in the range $h\nu > 3kT_e$.

The next step was taken by Sunyaev and Titarchuk [163], who solved analytically the problem of the Comptonization of low-frequency ($h\nu \ll kT_e$) radiation in an isothermal, nonrelativistic ($kT_e \ll m_e c^2$) plasma cloud having a substantial optical depth with respect to Thomson scattering ($\tau_0 \gg 1$). In this case the diffusion approximation will correctly describe how the photons are distributed over their escape time, or equivalently, over the number of scatterings u they experience within the source. The average value of u is of order τ_0^2 , and the probability of a photon being scattered many more times than average falls off exponentially with increasing u :

$$P(u \gg \bar{u}) = A \exp(-u/\bar{u}) . \quad (155)$$

On the other hand, as follows from (100), the frequency of a photon will increase from ν_0 to ν after a typical number

$$u = \frac{1}{3} \frac{m_e c^2}{kT} \ln \frac{\nu}{\nu_0} \quad (156)$$

of inverse Compton scatterings, provided that $h\nu \ll kT$. The probability distribution (155) together with the law (156) lead to the emergence of a power-law spectrum. A more accurate proof will be presented below.

The behavior here is similar to the familiar Fermi statistical acceleration mechanism, which gives rise to a power-law spectrum for the same reason. Note that as the optical depth of the cloud increases, multiple scatterings become more probable and the radiation spectrum flattens out.

3.2 Distribution of Photons over the Escape Time

Homogeneous Sphere

Consider a spherical cloud of radius R filled with ionized gas of density N_e and temperature T . The plasma and radiation interact only via Compton scattering. The optical depth of the cloud with respect to Thomson scattering $\tau_0 = \sigma_T N_e R \gg 1$. There is a source of photons somewhere in the cloud. At

the moment $t = 0$ an instantaneous flare of the source occurs. By solving the problem of photon diffusion in the cloud, one can determine the distribution $P(t)$ of photons over the time of escape from the cloud. This solution was found in [163] and is described below.

We assume that the photon source is situated at the center of the cloud. It is convenient to introduce dimensionless time $u = \sigma_T N_e c t$, characterizing the number of collisions experienced by a photon in the cloud. In the diffusion problem $u \gg 1$ and it may be regarded as a continuous variable rather than a discrete parameter. The average photon escape time $\bar{t} = \int_0^\infty t P(t) dt = R\tau_0/2c$, and the average number of scatterings $\bar{u} = \tau_0^2/2$. The peak of $P(t)$ lies near $t = 0.3R\tau_0/c$ or $u_0 = 0.3\tau_0^2$. When $u \gg u_0$, we have the asymptotic expression

$$P_1(u) = \frac{\pi^2}{3(\tau_0 + 2/3)^2} \exp \left[-\frac{u\pi^2}{3(\tau_0 + 2/3)^2} \right], \quad (157)$$

and when $1 \ll u \ll u_0$ the asymptote⁴

$$P_2(u) = \frac{3\sqrt{3}}{2\sqrt{\pi}} \frac{\tau_0^3}{u^{5/2}} \exp \left(-\frac{3\tau_0^2}{4u} \right). \quad (158)$$

Interesting is the case where the sources of photons are distributed according to the law

$$\phi(\tau) = \frac{\tau_0}{\pi\tau} \sin \frac{\pi\tau}{\tau_0}. \quad (159)$$

This is an intermediate case between those of uniform distribution of sources and the central source. In this case $P(u)$ is very simple, because it is an eigenfunction of the diffusion equation:

$$P(u) = \frac{\pi^2}{3(\tau_0 + 2/3)^2} \exp \left[-\frac{\pi^2 u}{3(\tau_0 + 2/3)^2} \right] = \beta e^{-\beta u}, \quad (160)$$

where

$$\beta = \frac{\pi^2}{3(\tau_0 + 2/3)^2}. \quad (161)$$

The average number of scatterings experienced by photons in the source $\bar{u} = \beta^{-1}$.

Disk

If in a homogeneous disk the sources of photons are distributed in the plane of symmetry or homogeneously over its volume, then $P(u)$ differs slightly from the formulae for a spherical cloud. If sources are distributed according to the eigenfunction of the diffusion equation, then [163]

$$P(u) = \beta e^{-\beta u} \quad \text{and} \quad \beta = \frac{\pi^2}{12(\tau_0 + 2/3)^2}. \quad (162)$$

Here τ_0 corresponds to the half-thickness of the disk.

⁴ Lightman et al. [99] pointed out a misprint in the formula published in [163]

3.3 Solution of the Stationary Equation of Comptonization

When the probability of photon escape from the plasma cloud is $P(u) = \beta \exp(-\beta u)$, the Comptonization problem can be reduced to the solution of the stationary Kompaneets equation

$$\frac{1}{x^2} \frac{d}{dx} x^4 \left(\frac{dn}{dx} + n \right) = \gamma n - \frac{\gamma f(x)}{x^3}. \quad (163)$$

Sunyaev and Titarchuk [163] have solved this equation by reducing it to Whittaker's equation. On the left-hand side of (163) stands the differential Kompaneets operator (see §2.2), which describes the Doppler diffusion of photons in frequency and their downward motion along the frequency axis due to the recoil effect. The induced process is neglected. On the right-hand side, the first term describes the diffusion of photons in space and the second allows for the presence of photon sources with a spectrum $f(x_e)$ in the cloud. As before $x = h\nu/kT$. The parameter $\gamma = \beta m_e c^2 / kT_e$, in particular

$$\gamma = \frac{\pi^2}{3} \frac{m_e c^2}{(\tau_0 + 2/3)^2 kT} \quad (164)$$

if the geometry is spherical, while

$$\gamma = \frac{\pi^2}{12} \frac{m_e c^2}{(\tau_0 + 2/3)^2 kT} \quad (165)$$

in the case of a disk.

Comptonization of Low-Frequency Radiation in Hot Plasma

If the characteristic frequency of the radiation from the source $\nu_0 \equiv x_0 kT/h \ll kT/h$, then (163) has the solution

$$F_\nu(x) = Ax^{\alpha+3}, \quad (166)$$

for the flux density at $x \ll x_0$, and the solution

$$F_\nu(x) = Bx^3 e^{-x} \int_0^\infty t^{\alpha-1} e^{-t} \left(1 + \frac{t}{x} \right)^{\alpha+3} dt \quad (167)$$

at $x \gg x_0$, with [149]

$$\alpha = -\frac{3}{2} + \sqrt{\frac{9}{4} + \gamma}. \quad (168)$$

The integrals in (167) reduce to gamma functions in two limits. For $x_0 \ll x \ll 1$ (when recoil plays a negligible role compared to the Doppler effect), the emergent spectrum is a power law:

$$F_\nu(x) = Cx^{-\alpha}, \quad (169)$$

The spectral index α depends only on the electron temperature and optical depth of the plasma cloud, not on its internal distribution of photon sources. That is quite natural, because after having been scattered $\sim \tau_0^2$ times the photons completely forget where they were born. When $\gamma \rightarrow 0$, the spectrum becomes flat in the region $x_0 \ll x \ll 1$, with $\alpha \rightarrow 0$.

At high frequencies ($x \gg 1$), when recoil dominates, a Wien spectrum forms:

$$F_\nu(x) = Dx^3 e^{-x}. \quad (170)$$

If $\gamma \ll 1$, the Wien spectrum extends over most of the spectrum, also into the region $x < 1$.

The significance of the solution (167) was recognized once this comparatively simple expression proved to fit perfectly the X-ray spectrum of the famous black hole candidate Cygnus X-1 [162].

Cloud Luminosity

In an infinite homogeneous medium, the radiation energy density increases with time according as $\Sigma(y) = \varepsilon_0 \exp(4y)$. This law is correct only when the Doppler effect dominates. On multiplying (163) by x^3 and integrating over x , we find that when the luminosity of the low-frequency sources of photons is L_0 , the total luminosity of the plasma cloud will be

$$L = L_0 \frac{\gamma}{\gamma - 4} = L_0 \frac{\alpha(\alpha + 3)}{(\alpha - 1)(\alpha + 4)}. \quad (171)$$

This solution is only true when $\gamma \gg 4$ ($\alpha \gg 1$), or more accurately:

$$\frac{\gamma}{\gamma - 4} \ll \frac{4}{5} \left(\ln \frac{1}{x_0} + 1.5 \right). \quad (172)$$

For example, if $x_0 = 10^{-3}$, we have $\gamma \gg 4.7$. The rate of the energy loss by all the electrons in the cloud is equal to $L - L_0$.

When $\gamma \rightarrow 0$, the emergent radiation will have a nearly Wien spectrum, with $\overline{h\nu} = 3kT_e$ and $L/L_0 \rightarrow 3x_0^{-1+\gamma/3}$. Indeed, since the number of photons is conserved and $L_0 = N_\gamma h\nu_0$, then $L_{\max} = 3N_\gamma kT_e$ and $L_{\max}/L_0 = 3/x_0$. If the sources of low-frequency emission had a Planck spectrum, we would obtain $L_0 = 2.7N_\gamma kT_r$ and $L_{\max}/L_0 = T/0.9T_r$.

Comptonization of High-Energy Photons in Cold Plasma

Another problem of astrophysical interest is the Comptonization of high-energy photons in a cloud of cold plasma ($T = 0$). In the limit $h\nu_0 \gg kT$, (163) reduces to

$$\frac{1}{z^2} \frac{d}{dz} z^4 n - \beta n = -\frac{\beta f(z)}{z^3}, \quad (173)$$

where $z = h\nu/m_e c^2$.

Equation (173) has the following solution [163]:

$$F_\nu(z) = \frac{\beta}{z} \exp(-\beta/z) \int_z^\infty f(\xi) \exp(\beta/\xi) \frac{d\xi}{\xi}. \quad (174)$$

For a monochromatic radiation source, with $f(z) = z_0 \delta(z - z_0)$, we find that

$$F_\nu(z) = \begin{cases} \beta z^{-1} \exp[-\beta(1/z - 1/z_0)], & \nu < \nu_0 \\ 0, & \text{otherwise.} \end{cases} \quad (175)$$

It is important to note that this solution is only valid in the case of photon sources distributed according to the law (159). However, it correctly describes the exponential shape of the spectrum at frequencies $\nu_0 - \nu \gg \tau_0^2 h\nu_0^2/m_e c^2$ for any distribution of sources.

For a power-law spectrum of seed photons, $f(z) = Az^{-\alpha}$ with $\alpha > 0$, the emergent spectrum in the region $z \gg \beta$, i.e. at $(h\nu/m_e c^2)\tau_0^2 \gg 1$ is

$$F_\nu(z) = \frac{A\beta}{\alpha} z^{-\alpha-1}, \quad (176)$$

i.e. the spectral index increases by unity. On the other hand, the scattering does not affect the power-law spectrum in the region $z \ll \beta$.

3.4 Solution by the Convolution Method

The solutions presented in §3.3 were obtained by solving the stationary Kompaneets equation (163), which had been written down for the specific case of photon sources distributed according to the law (159). Nevertheless, as we already mentioned before, the shape of the emergent spectrum should not depend on the distribution of seed photons within the cloud if the photons composing the spectrum have experienced $u \gtrsim \tau_0^2$ scatterings. However, this may be untrue for some parts of the emergent spectrum. In particular, if the initial spectrum is a narrow line, then the contribution to the emergent spectrum of photons that have undergone only a few scatterings in the cloud will be significant or even dominant near the position of the input line.

A more accurate solution can be obtained [29, 108] by direct convolution of $P(t)$, the distribution of outgoing photons over the escape time, with the solution $I_\nu(\nu, t)$ of the Kompaneets equation for the infinite medium (84):

$$F_\nu(\nu) = \int_0^\infty I_\nu(\nu, t) P(t) dt. \quad (177)$$

We shall restrict ourselves here to an application of this method to the problem of Comptonization of high-energy photons in a cold plasma cloud,

because in this case ($kT \ll h\nu_0 \ll m_e c^2$) analytical treatment is possible. In the opposite limit ($h\nu_0 \ll kT \ll c^2$), one needs to resort to a numerical integration. Both cases are discussed in detail in [163].

In the Thomson limit ($h\nu \ll m_e c^2$), scattering is characterized by the Rayleigh angular diagram, and according to the Compton formula (6) the average increase in photon wavelength after u scatterings in a cold plasma will be

$$\lambda = \lambda_0 + \lambda_C u, \quad (178)$$

where λ_0 is the initial wavelength.

Suppose that the sources in the cloud emit the monochromatic line $f(\nu) = A\nu\delta(\nu - \nu_0)$. To a first approximation we may consider the line to remain monochromatic as it shifts downwards in frequency with each successive scattering. This is exactly what is predicted by the Kompaneets equation (see §2.4). In reality the line broadens to $[\lambda_0, \lambda_0 + 2\lambda_C]$ already after the first scattering, but we shall see below that the main cause of the profile broadening is the dispersion in the number of scatterings undergone by photons emerging from the cloud. Therefore, we can relate the emergent photon frequency with the number of scatterings:

$$u = \frac{\lambda - \lambda_0}{\lambda_C} = \frac{m_e c^2}{h} \left(\frac{1}{\nu} - \frac{1}{\nu_0} \right). \quad (179)$$

Accordingly, the emergent spectrum will be

$$F_\nu = A\nu \frac{dN_\gamma}{d\nu} = A\nu \frac{dN_\gamma}{du} \left| \frac{du}{d\nu} \right| = \frac{Am_e c^2}{h\nu} P \left(\frac{m_e c^2}{h\nu} - \frac{m_e c^2}{h\nu_0} \right). \quad (180)$$

Using the formulae for $P(u)$ (see §3.2 and [163]) it is easy to determine the line profile for any distribution of sources over the cloud. For example, in the case of a spherical cloud with a central source, the emergent line profile will peak at $\lambda_{\max} = \lambda_0 + 0.3\lambda_C\tau_0^2$, and will have exponentially declining wings at $\lambda - \lambda_0 \gg 0.3\lambda_C\tau_0^2$ and $\lambda - \lambda_0 \ll 0.3\lambda_C\tau_0^2$. The line width is $\sim \lambda_C\tau_0^2$. For comparison, the solution (175) of the stationary Kompaneets equation correctly describes the long-wavelength exponential wing ($\lambda - \lambda_0 \gg \lambda_C\tau_0^2$), but not the short-wavelength one. This is due to the fact that emergent photons with $\lambda \approx \lambda_0$ have experienced only a few scatterings in the cloud ($u \ll \tau_0^2$). In fact the output spectrum is extremely sensitive in the region $\lambda - \lambda_0 \ll \lambda_C\tau_0^2$ to the distribution of sources – see [163].

In a more accurate treatment, Illarionov et al. [74] and Lightman et al. [99] have taken into account the dispersion due to the scattering angle in the wavelength shift for a fixed number of scatterings. The solutions obtained by these authors become noticeably different from the one described above within $\sim \tau_0$ Compton wavelengths from the position of the input line, where the dispersion due to the variable angle of scattering is important in comparison with the dispersion due to the variable number of scatterings. In particular, in this

region of the emergent spectrum there are signatures of unscattered and once and twice scattered photons. However, at $\lambda - \lambda_0 \gtrsim \tau_0 \lambda_C$, the solution (180) is always a good approximation.

3.5 Double Compton Effect as Source of Low Frequency Photons

Consider a homogeneous, isothermal plasma cloud optically thin to free-free absorption but with a large Thomson depth $\tau = \sigma_T N_e R \gg 1$. We may consider two limiting cases [70]: a) if the parameter $y = (kT_e/m_e c^2)\tau^2 \ll 1$, Comptonization will have little influence on the radiation spectrum; b) if $y \gg 1$, the spectrum inside the cloud will be practically independent of the photon source spectrum and approximate a Wien law $\Sigma_\nu = A\nu^3 \exp(-h\nu/kT_e)$, where the constant A depends solely on the number of photons emitted by the cloud during the mean photon escape time.

Case $y \ll 1$

Bremsstrahlung radiation will be emitted uniformly over the cloud, and on solving the diffusion equation

$$\frac{1}{r^2} \frac{d}{dr} r^2 \frac{N_\gamma}{dr} + \frac{3\sigma_T N_e}{c} \epsilon_\nu = 0, \quad (181)$$

we find that the density of free-free photons at the center of the cloud is

$$N_\gamma = \frac{\epsilon_\nu R}{2c} \left(\tau + \frac{4}{3} \right), \quad (182)$$

and photons will escape from it on a time scale $R\tau/2c$. For $h\nu/kT \gtrsim 1$, it follows from (126), (146) that at the center of the cloud

$$\frac{\epsilon_\nu^{\text{DC}}}{\epsilon_\nu^{\text{ff}}} = \frac{2}{3} \frac{\alpha}{\pi} \frac{kT}{m_e c^2} y \left(1 + \frac{h\nu}{kT} \right). \quad (183)$$

Clearly $\epsilon_\nu^{\text{DC}} \ll \epsilon_\nu^{\text{ff}}$ if $y < 1$. Exactly the same result is obtained in the time-dependent problem of an infinite, homogeneous medium whose photon population grows with time.

Case $y \gg 1$

At the center of the cloud, $\Sigma_\nu = Ax^3 e^{-x}$. The radiation energy $\Sigma = \int \Sigma_\nu d\nu = 6AkT/h$, while the photon density $N_\gamma = \Sigma_\nu/3kT = 2A/h$.

According to (148), the rate of double Compton production of soft photons at the center of the cloud is

$$\frac{dN_\gamma}{dt} = \int_{x_0}^{\infty} \frac{\epsilon_\nu^{\text{DC}}}{h\nu} d\nu \approx \frac{2}{9} \frac{\alpha}{\pi} \sigma_T N_e c \frac{\Sigma}{m_e c^2} \frac{kT_e}{m_e c^2} \left(24 \ln \frac{1}{x_0} + 50 \right); \quad (184)$$

$x_0 \ll 1$ corresponds to the frequency at which the photon absorption rate through the double Compton effect or by free-free absorption is comparable with the rate at which photons upscatter along the frequency axis.

Photon production by the double Compton effect will play a significant role if Compton scatterings can yield a single photon during $\langle t \rangle = R\tau/2c$, the characteristic time scale for a photon to emerge from the cloud. Therefore, it is necessary that

$$\frac{8\pi}{\pi} \left(\frac{kT_e}{m_e c^2} \right)^2 \tau^2 \ln \left(\frac{8}{x_0} \right) \gtrsim 1. \quad (185)$$

If $x_0 \sim 10^{-4}$ – 10^{-2} , the quantity $(kT_e/m_e c^2)^2 \tau^2 = 5$ – 10 , while in order for a Wien spectrum to develop we must have $y = (kT_e/m_e c^2) \tau^2 > 1$. Thus, the double Compton mechanism of photon production can sustain the Comptonization process in very hot, optically thick clouds.

On the other hand, double Compton photon production will surpass the contribution of bremsstrahlung processes only if the source is particularly luminous and compact. Indeed, the cloud will have a luminosity $L \approx (4\pi/3)R^3\Sigma/\langle t \rangle = (8\pi/3)R^2 c\Sigma/\tau$. Comparing ϵ_ν^{DC} with ϵ_ν^{ff} , and replacing the Σ in ϵ_ν^{DC} by L , we find that the double Compton effect will predominate if

$$\frac{L}{L_{\text{Edd}}} \gg 0.7 \frac{m_e}{m_p} \frac{R}{R_S} \left(\frac{m_e c^2}{kT_e} \right)^{3/2} g(x_0), \quad (186)$$

where L_{Edd} is the Eddington luminosity, given by (38), $R_S = 2GM/c^2$ is the Schwarzschild radius and $g(x_0) \sim 10$ is the bremsstrahlung Gaunt factor.

The estimates above demonstrate that in a cloud with $kT_e \sim 25$ keV, the double Compton effect will be important only if the cloud luminosity is near-Eddington and the plasma has great optical depth, $\tau \gtrsim 10$.

3.6 Monte Carlo Calculations of Comptonization Spectra

The analytic solution described above is only applicable when the Thomson depth of the cloud $\tau_0 \gg 1$. In this case, the spatial propagation of photons within the cloud can be considered a diffusion process. This approximation breaks down when the cloud becomes more transparent with respect to Thomson scattering, when $\tau_0 \lesssim 3$. Spectra formed via Comptonization of low-frequency radiation in an optically thin or moderately thick cloud of hot plasma can be computed very efficiently by Monte-Carlo methods. Pozdnyakov et al [132] were the first to develop and successfully apply a Monte-Carlo code to solving Comptonization problems. Another advantage of the Monte Carlo approach is that it can be applied with equal success to situations in which the plasma is relativistic. For comparison, the analytic solution of Sunyaev and Titarchuk is valid only in the nonrelativistic limit ($kT_e \ll m_e c^2$).

3.7 Bulk Comptonization

During the process of thermal Comptonization, low-frequency photons receive energy from electrons rapidly traveling in random directions. In many astrophysical situations, the scattering medium may be undergoing substantial bulk motions. Blandford and Payne [19] have shown that in a nonuniform fluid flow, e.g. converging or diverging, the photons will receive more energy from the the bulk motion of the scattering electrons than from their random thermal motions if the bulk speed u is larger than the typical thermal velocity: $u \gtrsim (3kT/m_e c^2)^{1/2}$.

The nonuniformity of the flow plays a crucial role in this problem, since electrons must have different velocities relative to each other in order for photons to be capable of attaining energy as they undergo successive scatterings. To illustrate this point, let us consider the extreme situation in which a cloud of cold ($T = 0$) ionized gas is moving as a whole with a constant velocity. It is obvious that in this case no Comptonization will result, because from the point of view of an observer moving with the flow, all the electrons are at rest.

In the case where the scattering medium is optically thick to Thomson scattering and the motions involved are nonrelativistic ($u \ll c$, $kT \ll m_e c^2$), the propagation of photons through the plasma can be considered in the diffusion approximation, and a Fokker-Planck equation similar to that of Kompaneets (84) results [19]:

$$\begin{aligned} \frac{\partial n}{\partial t} + \mathbf{u} \cdot \nabla n - \nabla \cdot \left(\frac{c}{3\sigma_T N_e} \nabla n \right) &= \frac{1}{3} (\nabla \mathbf{u})_\nu \frac{\partial n}{\partial \nu} \\ + \frac{\sigma_T N_e h}{m_e c} \frac{1}{\nu^2} \frac{\partial}{\partial \nu} \nu^4 \left(n + \frac{kT}{h} \frac{\partial n}{\partial \nu} \right) &+ f_\nu. \end{aligned} \quad (187)$$

Here $n = (1/4\pi) \int n(\nu, \mathbf{\Omega}) d\mathbf{\Omega}$ is the photon occupation number averaged over all directions $\mathbf{\Omega}$ in the nearly isotropic radiation field, f_ν is the source term, and we ignored induced effects. The first two terms on the left-hand side of (187) govern the spatial advection of the radiation field induced by dynamics, while the third term describes the diffusion of photons throughout the medium. The terms on the RHS determine the evolution of n in the energy space; they account for the heating of radiation by compression (or cooling by expansion) and the heating and cooling by thermal Comptonization. Whether advection or diffusion establishes depends essentially on the competition among the left-hand side terms. Advection dominates diffusion if $\tau_0 u/c \gg 1$; the opposite case $\tau_0 u/c \ll 1$ defines the static diffusion regime. Here τ_0 is the characteristic optical depth.

For the case $u = 0$ and a stationary situation, (187) reduces to the Kompaneets equation with an additional diffusion term that accounts for the effect of photon escape (see §3.3).

Energy Exchange

Multiplying (187) by ν^3 and integrating over ν , we obtain the equation governing the radiation energy density [19],

$$\frac{\partial \Sigma}{\partial t} + \mathbf{u} \nabla \Sigma - \nabla \left(\frac{1}{3\sigma_{\text{T}} N_{\text{e}}} \nabla \Sigma \right) = -\frac{4}{3} (\nabla \mathbf{u}) \Sigma + \frac{4kT}{m_{\text{e}} c} \sigma_{\text{T}} N_{\text{e}} \Sigma - \frac{\sigma_{\text{T}} N_{\text{e}} h}{m_{\text{e}} c} \int \nu \Sigma_{\nu} d\nu + F, \quad (188)$$

where F is the frequency-integrated emissivity.

From this equation we find the characteristic time scales for Compton heating, Compton cooling, and compressional heating (or expansion cooling):

$$t_{+} = \frac{1}{4\sigma_{\text{T}} N_{\text{e}}} \frac{m_{\text{e}} c}{kT}, \quad (189)$$

$$t_{-} = \frac{1}{\sigma_{\text{T}} N_{\text{e}}} \frac{m_{\text{e}} c}{h\nu}, \quad (190)$$

$$t_{\text{b}} = \frac{3}{4\nabla \mathbf{u}}. \quad (191)$$

Bulk Comptonization dominates thermal Comptonization when $t_{+} \gg t_{\text{b}}$. In typical situations (see examples below), the velocity scale-length is $\sim c/\sigma_{\text{T}} N_{\text{e}} u$, which leads to the condition $u \gtrsim (3kT/m_{\text{e}})^{1/2}$ for the relative importance of bulk acceleration.

Comptonization in a Radiation Dominated Shock

In many astrophysical contexts one encounters the braking of plasma in a radiation field. Among these problems are the dissipation of perturbations in the early universe, critical and supercritical accretion onto neutron stars [10, 39, 147], and supercritical, spherically symmetric accretion by black holes [13].

The process in question has a number of distinctive features. If the plasma is dominated by radiation, it will decelerate as photons are scattered by the electrons (at the densities and temperatures typical of the radiation-dominated case, scattering will generally prevail over absorption processes). The photons will, on the average, accumulate energy through the Doppler effect. When the energy of a photon has become high enough, part of it will be transmitted to the electrons by the Compton recoil effect. As a consequence the electrons will undergo Compton heating. Through this process the protons will play a passive role. Acting as the main reservoir of kinetic energy, and aided by the magnetic or electrostatic field, the protons will drag the electrons through the photon gas, heating it as well as the electrons; but the protons themselves will become heated only in the last instance, through their collisions with the electrons.

Blandford and Payne [20] investigated the problem of the interaction of radiation with plasma in a radiation-dominated, plane-parallel shock, assuming a negligible electron temperature. In this case, most of the momentum flux will be converted into radiation pressure over a length-scale $\sim (c/u)$ Thomson optical depths [as results from balancing convection of the radiation by the background medium with diffusion, i.e. equating the second and the third terms on the left-hand side of (187)]. The relative velocity across one optical depth $du/d\tau \sim u^2/c$, and since a typical photon undergoes $\sim (c/u)^2$ scatterings in crossing the shock, there will result a net gain in energy of order unity from the bulk acceleration. This is similar to a cosmic-ray mediated shock [18, 43].

By solving (187) with the thermal-Comptonization terms on the right-hand side neglected and applying boundary conditions that result from the appropriate shock solution (see, e.g. [202]), [20] found an analytic, steady-state solution for the spectrum of radiation transmitted through the shock. For incident monochromatic radiation of frequency ν_0 , the resulting spectrum $F_\nu(\nu)$ at frequencies $\nu \gg \nu_0$ is power-law with an index

$$\alpha = \frac{(M^2 - 1/2)(M^2 + 6)}{(M^2 - 1)^2}, \quad (192)$$

where M is the Mach number of the shock. In the strong-shock limit ($M \gg 1$), $\alpha \rightarrow 1$.

Inclusion of Temperature

Lyubarsky and Sunyaev [102] extended the analysis of Blandford and Payne by relaxing the assumption $T = 0$ and considering the thermal Comptonization within the shock as well. They applied the general (187) to the problem under consideration, taking into account the thermal-Comptonization terms on the right-hand side. Transforming to the variables $x = h\nu/kT$ and $\tau = \sigma_T \int N_e dr$, this equation becomes for the one-dimensional steady-state problem

$$\frac{m_e c^2}{kT_e} \left(-\frac{1}{3} \Delta_\tau n + \frac{1}{c} (u \nabla_\tau) n \right) = -\delta \frac{\partial n}{\partial x} + \frac{1}{x^2} \frac{\partial}{\partial x} x^4 \left(\frac{\partial n}{\partial x} + n \right). \quad (193)$$

Here $\nabla_\tau \equiv \partial/\partial\tau$, and the parameter $\delta = -(m_e c/3kT)(du/d\tau) \sim -(m_e c^2/3kT)(u/c)^2$ is assumed to be known from solution of the problem of plasma braking in a radiation dominated shock. Since we are dealing with a compressible medium, the quantity δ is positive. Equation (193) admits of a separation of variables if $du/d\tau = \text{const}$. We proceed to consider this case.

In standard fashion, by setting $n(\tau, x) = A(\tau)N(x)$ we arrive at the pair of equations

$$\frac{m_e c^2}{kT} \left(\frac{1}{3} \Delta_\tau A - \frac{1}{c} (u \nabla_\tau) A \right) = -\gamma A, \quad (194)$$

$$\frac{1}{x^2} \frac{d}{dx} x^4 \left(\frac{dN}{dx} + N \right) - \delta x \frac{dN}{dx} = \gamma N. \quad (195)$$

We are interested in the function $N(x)$. Since in the problem under consideration the radiation energy density greatly exceeds the thermal energy density of the plasma, the electron temperature tends to adjust itself to the stationary value $kT = (h/4\Sigma) \int \nu \Sigma_\nu d\nu$, determined by the balance between Compton heating and cooling. We can then find a relation between the separation constant γ and the parameter δ by multiplying (195) by x^3 and integrating from 0 to ∞ . In this way we obtain $\gamma = 4\delta$.

The solution of (195) can be expressed in terms of the Whittaker function. At frequencies x above the characteristic frequency x_0 of a soft-photon source, the emergent spectrum will have the form

$$N(x) = x^{(\delta-1)/2} e^{-x/2} W_{2+\delta/2, \sqrt{9+10\delta+\delta^2}/2}(x). \quad (196)$$

The Whittaker function has the convenient integral representation

$$W_{\lambda, \mu}(x) = \frac{x^{1/2-\mu} e^{-x/2}}{\Gamma(\mu - \lambda + 1/2)} \int_0^\infty e^{-t} t^{\mu-\lambda-1/2} (x+t)^{\mu+\lambda-1/2} dt, \quad (197)$$

where $\Gamma(z)$ is the gamma function. In our case $\mu = \sqrt{9+10\delta+\delta^2}/2$, $\lambda = 2 + \delta/2$ (remember that $\delta > 0$).

At low frequencies ($x \ll 1$), the spectra conform to a power law with a spectral index

$$\alpha = \frac{1}{2} \left(\sqrt{9+10\delta+\delta^2} - 3 - \delta \right). \quad (198)$$

At low temperatures (as $\delta \rightarrow \infty$), the index $\alpha \rightarrow 1$, in agreement with Blandford and Payne's solution. In the high-temperature limit ($\delta \rightarrow 0$), we arrive at the problem of thermal Comptonization in a finite medium, with the effective Comptonization parameter $(kT/m_e c^2) \tau_0^2 \sim (kT/m_e c^2)(c/u)^2 \sim 1/\delta \gg 1$. Accordingly, $\alpha \rightarrow 0$.

At high frequencies (as $x \rightarrow \infty$), the solution asymptotically approaches

$$F_\nu(x) \propto x^3 N(x) \propto x^{3+\delta} e^{-x}. \quad (199)$$

We see that the exponential cutoff caused by the recoil effect is effectively shifted to a higher frequency, $h\nu_{\text{cut}} \sim (3 + \delta)kT_e$, as compared to the case of thermal Comptonization, when a Wien spectrum $F_\nu \propto x^3 \exp(-x)$ with $h\nu_{\text{cut}} \sim 3kT$ is formed. This is the result of the combined operation of bulk and thermal Comptonization in the shock.

Lyubarsky and Sunyaev's solution given above is, as is typical of Comptonization problems, independent of the coordinates of the source of soft photons, because the spectrum of interest to us (at $\nu \gg \nu_0$) is formed by those photons which have been scattered far more times than the average number of scatterings in the plasma cloud. Moreover, despite the fact that the solution above was obtained assuming $du/d\tau = \text{const}$, the spectrum formed by

photons that have survived a long time in the shock will obviously depend little on the particular velocity distribution; it will instead be determined by some average value of $du/d\tau \sim u^2/c$. This value can usually be found by solving the dynamical problem of plasma deceleration.

The spectral shape described by the solution (196) – a power law with a small spectral index ($0 < \alpha < 1$) and an exponential cutoff at high energies resembles the spectra actually observed from accretion-powered X-ray pulsars in binary systems.

Spherical Accretion Flow

Bulk Comptonization can also be important during spherical, supercritical accretion of gas onto a black hole. In this case, photons can be accelerated by the converging flow of the accreting gas. This problem was first studied by Blandford and Payne [21]. If the gas, accreting at a rate $\dot{M}$, is in free-fall, then the radial Thomson scattering optical depth to infinity from a radius r is

$$\tau(r) = \frac{1}{2} \left(\frac{\dot{M}}{\dot{M}_{\text{cr}}} \right) \left(\frac{R_{\text{S}}}{r} \right)^{1/2}, \quad (200)$$

where $\dot{M}_{\text{Edd}} = 4\pi GMm_{\text{p}}/\sigma_{\text{T}}c$ is the Eddington critical accretion rate, $R_{\text{S}} = 2GM/c^2$ is the Schwarzschild radius and M is the mass of the black hole. In the case of supercritical accretion ($\dot{M} > \dot{M}_{\text{Edd}}$), there exists a well-defined region of the flow for which $\tau(r) > 1$ and from which photons must escape diffusely. This outward diffusion of the radiation is inhibited by its inward convection by the scattering electrons. The velocity of the inflowing electrons eventually becomes so large that photons are convected inward more rapidly than they can diffuse outward. The radius at which this occurs is the trapping radius r_{tr} [133], defined by

$$\frac{u(r_{\text{tr}})}{c} \tau(r_{\text{tr}}) = \frac{1}{3} \quad (201)$$

Most of the energy radiated to infinity is produced in the vicinity of the trapping radius. In escaping diffusely from r_{tr} , the photons undergo $\sim (c/u_{\text{tr}})^2$ scatterings [here $u_{\text{tr}} = u(r_{\text{tr}})$], each one giving on the average a fractional energy increase $\sim (u_{\text{tr}}/c)^2$ and a total average increase of order unity. The emitted radiation spectrum will have a power-law shape at high frequencies.

Assuming that the accreting plasma is cold ($T = 0$), the Fokker–Planck equation (187) applied to the case of a steady radial flow reduces to

$$\tau^* \frac{\partial^2 n}{\partial \tau^{*2}} + \left[\frac{d(\ln u r^2)}{d(\ln \tau^*)} - \tau^* \right] \frac{\partial n}{\partial \tau^*} + \frac{1}{3} \left[\frac{d(\ln u r^2)}{d(\ln \tau^*)} \right] \nu \frac{\partial n}{\partial \nu} = 0, \quad (202)$$

where

$$\tau^* = 3 \frac{u(r)}{c} \tau(r) = \frac{3\dot{M}\sigma_T}{4\pi m_p c r} . \quad (203)$$

(note that the trapping radius corresponds to $\tau^* = 1$). The equation has an analytic solution if the velocity changes with radius according to the law $u \propto r^{-\beta}$; $\beta = 1/2$ corresponds to the case of free-fall. For a monochromatic source of photons of frequency ν_0 located at a given depth τ_0^* , the emergent spectral flux is given by [21]

$$F_\nu \propto \frac{\tilde{x}}{(1 - \tilde{x})^{4-\beta}} \exp \left[-\frac{\tau_0^*}{1 - \tilde{x}} \right] ,$$

$$\tilde{x} = \left(\frac{\nu}{\nu_0} \right)^{-3/(2-\beta)} . \quad (204)$$

The spectrum has a power-law shape at high frequency ($\nu \gg \nu_0$), with an index

$$\alpha = \frac{3}{2 - \beta} . \quad (205)$$

In the free-fall case, $\alpha = 2$.

The total emergent luminosity of the source for the case of free-fall is

$$L = L_0 \left[1 + \frac{1}{3} \tau_0^* (1 + \tau_0^*) \right] e^{-\tau_0^*} , \quad (206)$$

where L_0 is the intrinsic luminosity of the source of low-energy photons. One can see that the source luminosity declines exponentially when the injection radius becomes less than the trapping radius, i.e. when $\tau_0^* \gg 1$. The maximum energy amplification, $L = 1.36L_0$, occurs when $\tau_0^* = 1.21$.

Inclusion of Temperature

One can allow for the plasma temperature in the present problem in the same way as we did when treating the case of a plane-parallel shock. Spatial photon diffusion and energy transfer can be decoupled if the following conditions are satisfied: (1) the temperature T is constant throughout the medium, and (2) the radial velocity is proportional to the free-fall velocity: $u = lc(R_S/r)^{1/2}$ (here l is the dimensionless parameter). The solution for the emergent spectrum was found by Colpi [35]; in the region $\nu \gg \nu_0$ it depends on two parameters: the location of the source of soft photons τ_0^* and

$$\eta = \frac{\dot{M}_{\text{Edd}}}{\dot{M}} l^2 \frac{m_e c^2}{kT_e} = \frac{t_+}{t_b} , \quad (207)$$

with the time scales t_+ and t_b defined by (189), (191).

The resulting spectrum has an approximately power-law shape over the range $h\nu_0 \ll h\nu \ll kT$, with an index

$$\alpha = \frac{1}{2} \left\{ [(\eta - 3)^2 + 20\eta]^{1/2} - 3 - \eta \right\}. \quad (208)$$

The index increases from 0 to 2 as η grows from 0 to ∞ . Large values of η can be achieved, for a fixed dynamics ($0 < l < 1$), at low electron temperatures or small accretion rates. In the limit $\eta \rightarrow \infty$, (208) gives the same value of the spectral index ($\alpha = 2$) as found for a cold electron plasma accreting in free fall. A further consequence of (208) is the softening of the spectrum as the radial velocity increases at fixed accretion rate. This effect is mainly determined by the decrease of the electron density due to mass conservation.

As in the case of a radiation-dominated shock, the power-law spectrum extends up to $h\nu_{\text{cut}} \sim 3kT$ if $\eta \rightarrow 0$ (high temperatures) but $h\nu_{\text{cut}} \gg kT$ if $\eta \rightarrow \infty$ (low temperatures). At $h\nu \gtrsim h\nu_{\text{cut}}$, the spectrum falls off exponentially.

Inclusion of Inner Boundary and Relativistic Effects

The early efforts to calculate the radiation spectrum emergent from a spherically symmetric converging flow ignored the presence of the inner boundary in the problem, i.e. it was assumed that photons could random-walk into the region about $r = 0$ where the electron density grows without limits. In reality, the flow is truncated at a finite radius, which is the radius of the event horizon in the case of a central black hole. The inner boundary can have a large influence on the outgoing spectrum, particularly in the case of small optical depths. Another major shortcoming of these studies is that they are based on nonrelativistic formalism, although it is obvious that the general and special relativistic effects must play an important role in the vicinity of a black hole.

4 Interaction of X-Rays with Partially Ionized Media

In the preceeding section we have considered the interaction of high-energy photons via Compton scattering with free electrons in a fully ionized plasma as a formation mechanism of spectra of X-ray sources. The only other mentioned radiative processes were bremsstrahlung and double Compton scattering, which were pointed out as possible sources of low-frequency photons for the Comptonization. However, in many astrophysical environments, X-rays interact with a gas that is neutral or only partially ionized, which causes other radiative mechanisms to come into play and may have a significant impact on the emergent radiation spectra.

In particular, photoabsorption may become a more important source of opacity than Compton scattering for photons with energies $h\nu \lesssim 10 \text{ keV}$. This point is central to the problem of Compton reflection in Galactic black hole candidates and Active Galactic Nuclei (AGN), which we consider in

§4.1 below. Also, the scattering on electrons bound in atoms is substantially different from the scattering on free electrons in the photon energy range $h\nu \lesssim$ a few keV. This motivates our discussion in §4.2 of the scattering of X-ray lines in molecular clouds.

4.1 X-Ray Reflection

An X-ray binary consists of a compact X-ray source – a neutron star or a black hole, and a normal optical star. An appreciable amount of X-rays emitted by the compact secondary may be reflected and reprocessed by the extended atmosphere of the primary. Also, in a large fraction of X-ray binaries as well as in luminous AGN, there is a geometrically thin, optically thick accretion disk extending inwards almost to the compact object, which is a supermassive black hole in the AGN case. The disk will intercept and reprocess a large fraction of X-rays produced in its innermost, hottest region or/and on the surface of a neutron star. Therefore, spectroscopy, timing analysis and polarimetry of the reflected X-ray component can give us unique information on the geometrical and physical properties of accreting X-ray sources.

In most cases, the reflection of X-rays from a stellar photosphere or an accretion disk can be treated using the approximation of a plane-parallel atmosphere, because the characteristic height of the media is much smaller than the characteristic curvature. This considerably simplifies the calculations.

Reflection by the Atmosphere of a Normal Star

Basko and Sunyaev [9] and Basko, Sunyaev and Titarchuk [11] have demonstrated that in a close binary system up to 30% of the X-ray source radiation reaching the surface of the normal star is reflected. The remaining 70% is absorbed and subsequently reradiated as optical and ultraviolet radiation.

The X-rays are absorbed through the photoionization of hydrogen, helium and the K-electrons of heavy elements. This process is effective for low energy photons, but its cross section rapidly decreases with increasing frequency: $\sigma_{\text{ph}} \propto \nu^{-3}$ (except near the absorption edges, where the cross section changes abruptly). In a weakly ionized plasma of normal cosmic abundance, the Thomson scattering cross section $\sigma_{\text{T}} = 6.65 \times 10^{-25} \text{ cm}^2$ exceeds the photoionization cross section per hydrogen atom at $h\nu \gtrsim 10 \text{ keV}$ [109]. Thus, the total absorption (true absorption plus scattering) cross section of X-rays of frequency ν is to a first approximation given by

$$\sigma(\nu) = \sigma_{\text{T}} \left[1 + \left(\frac{10 \text{ keV}}{h\nu} \right)^3 \right], \quad (209)$$

and the albedo of a single scattering is approximately

$$\lambda(\nu) = \frac{\sigma_T}{\sigma(\nu)} = \left[1 + \left(\frac{10 \text{ keV}}{h\nu} \right)^3 \right]^{-1}. \quad (210)$$

Note that the high degree of ionization of helium and heavy elements such as C, N, O, Ne in the X-ray irradiated atmosphere somewhat increases the photoionization cross section and moves the point at which $\sigma_{\text{ph}} \approx \sigma_T$ to energies below 10 keV.

The ionization of hydrogen, which supplies the most of the free electrons, has in practise little effect on both the photoabsorption cross section (since it is mainly the heavy elements which are active in the photoabsorption of photons with $h\nu > 1 \text{ keV}$) and the scattering cross section. At energies $h\nu > \alpha m_e c^2 \approx 3.7 \text{ keV}$ (here $\alpha \approx 1/137$ is the fine structure constant) the photon wavelength is less than the Bohr radius, and the scattering of hard X-rays from hydrogen and helium atoms leads to a tearing off a bound electron, since the recoil energy $\sim h\nu(h\nu/m_e c^2)$ is then greater than the ionization potential of hydrogen (13.6 eV). Consequently the differential cross section for scattering from electrons bound in hydrogen atoms is the same as for free electrons (see §4.2) for a further discussion of this subject).

Thus, the fate of X-ray photons striking the photosphere of the normal star depends on their initial energy: at $h\nu \lesssim 10 \text{ keV}$ they are absorbed and transformed into soft (in particular optical radiation); at $h\nu \gg 10 \text{ keV}$ a considerable fraction of the incident photons is reflected. The energy of hard X-rays can be absorbed not only through photoionization, but also as a consequence of recoil by Compton scattering: $\Delta\nu \sim -h\nu^2/m_e c^2$. The recoil effect acts in two ways, both leading to a decrease in the resulting energy albedo: at every scattering, part of the photon energy is transferred to the electron, and so after $\sim m_e c^2/h\nu$ scatterings the photon loses a considerable part of its initial energy; also, the probability of photoabsorption, which increases with decreased photon energy, increases after each scattering. Note that the X-ray heated stellar atmosphere has a temperature of $T \lesssim 2 \times 10^4 \text{ K}$ [9], and the Doppler frequency shift by scattering can thus be neglected, since $kT \ll h\nu(h\nu/m_e c^2)$.

Energy Albedo as a Function of the Incident Photon Energy

Basko et al. [11] have numerically solved a nonrelativistic (Thomson-limit) equation of X-ray transfer in a plane-parallel atmosphere, taking into account Compton recoil and photoabsorption and making the simplifying assumption that the scattering is isotropic. In particular, they have calculated the energy albedo $A(\nu_0, \mu_0)$ of the atmosphere as a function of the incident photon energy $h\nu_0$ and incident angle $\theta \equiv \arccos(\mu_0)$ (with respect to the normal to the atmosphere). Consider a monochromatic beam of photons: $I_0(\nu, \mu) = f_0 \delta(\nu - \nu_0) \delta(\mu - \mu_0)$, with $-1 \leq \mu_0 \leq 0$; the energy albedo is defined as the ratio of the output to the input total flux:

$$A = \frac{\int_0^\infty d\nu \int_0^1 d\mu \mu I_\nu(\nu, \mu)}{f_0}. \quad (211)$$

It turns out that for normally falling X-rays ($\mu_0 = -1$), the albedo reaches a maximum of $\approx 45\%$ at $h\nu_0 \sim 50$ keV and declines rapidly at $h\nu_0 \lesssim 20$ keV and at $h\nu_0 \gtrsim 300$ keV, due to photoabsorption and Compton recoil, respectively. The albedo increases somewhat for larger angles of incidence.

White, Lightman and Zdziarski [187] performed Monte Carlo simulations of the purely Compton reflection (neglecting photoabsorption) of hard X-rays and gamma-rays (with energies up to ~ 15 MeV) by a cold electron-scattering atmosphere. Interestingly, their relativistic result for $A(\nu_0)$, described by an approximate analytic expression, is not very different from the nonrelativistic result of [11] in the spectral region $50 \text{ keV} \lesssim h\nu \lesssim 500 \text{ keV}$, where the effect of photoabsorption is negligible. On the other hand, in the low-frequency range $h\nu \lesssim 10 \text{ keV}$, where the scattering can be considered coherent, a good approximation for the albedo is provided by the classical result for an atmosphere of normal chemical composition (see e.g. [154]).

X-Ray Scattering in the Accretion Disk in Neutron Star Low-Mass X-Ray Binaries

In a low-mass X-ray binary (LMXB) with a weakly magnetized ($H \lesssim 10^8$ G) neutron star, about half [161] of the total X-ray luminosity released via accretion originates in a narrow boundary layer of the disk [128] or in a flow spreading on the surface of the neutron star [75] (the remaining fraction is emitted by the disk). Furthermore, when the accreted matter at the neutron star surface reaches a critical density of $\sim 10^9 \text{ g cm}^{-2}$, a thermonuclear flash occurs, accompanied by a powerful X-ray burst. Since the accretion disk reaches to the neutron star, it must intercept and re-emit a significant fraction of the central X-ray radiation both during X-ray bursts and between them.

Following Lapidus and Sunyaev [94], we can estimate the fraction of the neutron star radiation intercepted by the accretion disk. Let R and H be the neutron star radius and the half-width of the emitting zone, respectively. We know that between bursts $H \ll R$ (however, H becomes comparable to R when the luminosity approaches the Eddington critical value, see [75]), and $H = R$ during a burst. Now, if the flux of radiation from the unit of neutron star surface area per unit of solid angle is

$$\frac{dF}{dS d\Omega} = \mu I(\mu) = I_0 \mu (a_1 + a_2 \mu + a_3 \mu^2 + \dots), \quad (212)$$

then the total flux in all directions from the upper hemisphere is

$$F_{\text{tot}} = I_0 R^2 \frac{H}{R} (2\pi)^2 \left(\frac{a_1}{2} + \frac{a_2}{3} + \frac{a_4}{4} + \dots \right). \quad (213)$$

A simple trigonometric calculation gives the fraction of radiation flux reaching the disk:

$$\begin{aligned}
 & \frac{F_{\text{down}}}{F_{\text{tot}}} \\
 & \approx \frac{a_1(\pi/2) \cos \theta (1 - \cos \theta/2) + a_2(2/3)(\theta \cos \theta + 2/3 - \sin \theta + \sin^3 \theta/3)}{\cos \theta [a_1 \pi + a_2(2\pi/3)]} \\
 & \approx \frac{1}{2} - \frac{a_1 + 8a_2/3\pi}{4a_1 + 8a_2/3} \frac{H}{R} \text{ for } H \ll R,
 \end{aligned} \tag{214}$$

where $\theta = \arccos(H/R)$.

We obtain that during a burst ($H = R$), $F_{\text{down}}/F_{\text{tot}} = 1/4$ and ≈ 0.23 if the emissivity of the neutron star surface obeys the Lambert law [$I(\mu) = \text{const}$] or the Chandrasekhar–Sobolev law for a pure electron scattering atmosphere [$I(\mu) \approx 1 + 2.06\mu$], respectively. In reality the fraction of radiation falling on the disk should be somewhat higher because of the curvature of photon trajectories in the strong gravitational field of the neutron star [94]. In the case of a narrow boundary layer on the surface of the neutron star ($H \ll R$), $F_{\text{down}}/F_{\text{tot}} \rightarrow 0.5$, as expected.

The scattering and reprocessing of the illuminating X-rays occurs mainly in the central region of the disk of several neutron star radii. According to the standard accretion disk theory, plasma in this region has a temperature of $kT \sim 1 \text{ keV}$ [148]. Furthermore, if the illuminating X-ray flux is higher than the disk intrinsic flux, as is the case during bursts, the plasma can be heated up to the characteristic Compton temperature of the external radiation, $kT \sim \text{a few keV}$. In either case, the gas is expected to be almost completely ionized, and photoelectric absorption of X-rays can be neglected compared with Compton scattering.

The Rossi X-ray Timing Explorer (RXTE) detections of millisecond periodic and quasi-periodic X-ray flux oscillations from dozens of LMXBs have demonstrated that the neutron stars in these systems are rapidly rotating, with spin frequencies between 300 and 600 Hz (see [180] for a review). These brightness oscillations are likely produced by spin modulation of emission from a few localized regions on the neutron star surface. We should learn much more than we know now about the geometry and physical processes taking place on rapidly rotating neutron stars from future huge X-ray observatories such as XEUS or dedicated timing missions, which will be capable of resolving the waveforms of individual X-ray oscillations.

We [143] investigated a possible role of X-ray scattering in the accretion disk in forming oscillation profiles. Since the innermost part of the disk is rotating with a huge speed $\sim 0.5c$, photons emitted by the neutron star and reflected by the disk will be Doppler-boosted in the direction of the disk rotation. As a result, a relatively weak pulse of scattered emission should reach an observer a quarter of a full cycle ahead of the main pulse coming directly from the stellar surface. A detection/non-detection of this signature

would be a proof/disproof that a standard disk extends all the way down to the neutron star. Furthermore, it should be possible to uncover LMXBs in which the disk rotates in the opposite sense with respect to the neutron star (see [151] on possibilities of formation of such systems): because the scattered emission is then expected to lag behind the primary signal.

Compton Reflection in AGN and Black Hole Candidates

The X-ray spectra of luminous AGN such as Seyfert galaxies and quasars consist of several components. A hard power-law component extends to high energies above 100 keV and a soft X-ray excess is often observed below 1 keV. A hardening of the power-law continuum above 10 keV and an emission line of iron near 6.4 keV are attributed to a further component, the reflection spectrum. Similar spectra are characteristic of Galactic black hole candidates in their low state. The standard interpretation invokes a hard X-ray source illuminating an optically thick, geometrically thin accretion disk; the observer sees both direct (power-law) and reflected hard X-ray emission together with soft X-rays from the disk. The reflected spectrum is mainly produced by Compton scattering and fluorescence in the disk.

The reflection spectrum, characterized by a broad bump between ~ 10 keV and a few 10^2 keV, can to a first approximation be described by the product of the input power spectrum with the monochromatic energy albedo $A(\nu)$ calculated by Basko et al. [11]. Those computations were carried out in the Thomson limit and pertained to the case of a cold atmosphere, when the total opacity is practically parameter-independent and approximately given by (209). However, in the case currently under consideration, the illuminating spectrum is a power law extending above 100 keV and possibly to gamma-ray energies, which makes it necessary to work with the Klein–Nishina scattering cross section in order to get accurate results. Furthermore, the zone of the accretion disk responsible for the reflection may be strongly ionized, partly as a result of external irradiation by hard X-rays. Therefore, the reflection spectrum below ~ 10 keV will generally depend on the temperature and ionization parameter of the reflecting medium.

White et al. [187] and Lightman and White [100] performed Monte Carlo simulations of the Compton reflection of X-rays and gamma-rays by a cold ($T = 0$) plane-parallel atmosphere, taking into account both electron scattering in the relativistic regime and photoabsorption, and complemented these computations with nonrelativistic analytic estimates. The results of this work were formulated in terms of a Green's function $G(\nu, \nu_0)$, which is defined as the probability that a photon injected with frequency ν_0 will emerge from the medium with a frequency in the interval $[\nu, \nu + d\nu]$. Thus, for an incident photon spectrum N_{in} , the reflected spectrum N_{out} is given by

$$N_{\text{out}}(x) = \int_x^\infty G(x, x_0) N_{\text{in}}(x_0) dx_0, \quad (215)$$

where $x = h\nu/m_e c^2$. The lower integration limit in (215) arises from the fact that scattering from cold electrons always produces an increase in photon wavelength.

Note that we already dealt with Green's functions in this review. In particular, the Compton scattering kernel considered in §2.1 is the Green's function for a single-scattering problem. Another example of a Green's function can be found in §3.4, where we considered the Comptonization of high-energy photons in an optically thick cloud of cold gas, which has a close relation to the problem currently under consideration.

We summarize the results of [100, 187] below. For photon energies $h\nu < 15$ keV, i.e. $x < 0.03$, Compton scattering can be considered elastic and the Green's function is well fit by

$$G(x, x_0) = \frac{1 - \epsilon^{1/2}}{1 + \epsilon^{1/2}} \delta(x - x_0) , \quad (216)$$

where $\epsilon = \sigma(\nu)/[\sigma(\nu) + \sigma_T]$, and $\sigma(\nu)$ is the photoionization cross section. At higher energies, $h\nu > 15$ keV, the scattering cannot be treated as elastic but another approximation is possible:

$$G(x, x_0) = W(x, x_0) G_C(x, x_0) . \quad (217)$$

Here

$$W(x, x_0) = \exp \left[10^{-5} \left(\frac{1}{4x_0^4} - \frac{1}{4x^4} \right) \right] \quad (218)$$

gives the probability that a photon of initial energy x_0 has reached the energy x (after several scatterings) without being absorbed. As can be seen, photon absorption is negligible for $h\nu > 50$ keV. $G_C(x, x_0)$ is the Green's function for pure electron scattering with no absorption:

$$\begin{aligned} G_C(x, x_0) &= x^{-2} G_0(\Delta y, y_0), \quad y = \frac{1}{x}, \quad \Delta y = y - y_0, \\ G_0(\Delta y, y_0) &= \begin{cases} B[(y_0 + 2)/(y_0 + \Delta y)]^\beta, & \Delta y < 2 \\ A(\Delta y)^{-3/2} (\Delta y_c / \Delta y)^\alpha, & 2 < \Delta y < \Delta y_c \\ A(\Delta y)^{-3/2}, & \Delta y_c < \Delta y, \end{cases} \\ \Delta y_c &= 10^3 - y_0, \\ \alpha &= -0.30 y_0^{-0.51} + 0.06 y_0^{-0.824}, \\ \beta &= 0.37 - 1.0 y_0^{0.85}, \\ A &= 0.56 + 1.12 y_0^{-0.785} - 0.34 y_0^{-1.04}, \\ B &= \frac{1 - A\{2 + [(\Delta y_c/2)^{1/2+\alpha} - 1](1/2 + \alpha)\}/(\Delta y_c)^{1/2}}{y_0^{1-\beta} (y_0 + 2)^\beta [(1 + 2/y_0)^{1-\beta} - 1]/(1 - \beta)} \quad \alpha \neq -1/2 \\ &= \frac{1 - A[2 + \ln(\Delta y_c/2)]/(\Delta y_c)^{1/2}}{y_0^{1-\beta} (y_0 + 2)^\beta [(1 + 2/y_0)^{1-\beta} - 1]/(1 - \beta)}, \quad \alpha = -1/2. \end{aligned} \quad (219)$$

The normalization

$$1 = \int_0^\infty G_C(x, x_0) dx. \quad (220)$$

reflects the fact that the number of photons is conserved by scattering. In the nonrelativistic regime ($x_0 \ll 1$, or $y_0 \gg 1$) $G_0(\Delta y, y_0)$ is independent of energy and can be conveniently approximated by the simple expression [99, 187]

$$G_{\text{nr}}(\Delta y) \approx \begin{cases} 0.10, & \Delta y < 2 \\ 0.56(\Delta y)^{-3/2}, & \Delta y > 2. \end{cases} \quad (221)$$

The ionization parameter determines the shape of the spectrum below ~ 15 keV.

The Green's function given by (216)–(219) was obtained on the assumption that the incident photons are supplied by an optically thin source covering the plane-parallel atmosphere [with the intensity distribution $I_0(\mu) = \text{const}$ for $-1 \leq \mu \leq 0$] and upon averaging the emergent radiation over all viewing angles. Magdziarz and Zdziarski [104] have improved on these results by computing and tabulating Green's functions for Compton reflection as a function of the viewing angle. There are significant differences (of the order of 20%) between the angle-dependent reflection spectra and the averaged one. In particular, the face-on reflected spectrum in the case of the $\alpha = 1$ incident power law is both significantly harder in the 10–30 keV range and softer above 30 keV than the angle-averaged spectrum.

4.2 Scattering of X-Ray Lines on Neutral Hydrogen and Helium

The scattering of X-ray photons by hydrogen atoms is discussed in detail in a number of monographs and reference books. The laws of conservation of momentum and energy for the scattering of a photon by a free electron moving with a given velocity uniquely relate the final frequency of the photon to the geometry of the scattering – see §1.1. In the case of scattering by a bound electron in a hydrogen atom, additional factors complicate the process: finite binding energy of the electron and motion of the electron in the atom. Since the energy levels of the electron are discrete, the change in the photon frequency cannot take arbitrary values; also because of the random nature of electron motion in the atom, the amount of energy transferred to the photon is no longer a unique function of the scattering angle.

As we know from §2.1, even a low temperature ($kT \sim 1$ eV) of free electrons has a noticeable effect on the spectrum of the scattered emission: the single-scattering line profile is smeared by the Doppler effect. Note that in this case, the electron velocity is $v \sim 400$ km/s. The characteristic velocity of the electron in a hydrogen atom is $v \sim \alpha c \sim 2000$ km/s ($\alpha = 1/137$ is the fine-structure constant), so this velocity should significantly affect the amount of energy transferred to the electron by a scattering photon. The resulting ambiguity in the energy transfer does not violate the conservation

laws, because the heavy nucleus with negligible kinetic energy can carry away the necessary momentum.

Depending on the final state of the electron, the scattering of a photon by a hydrogen atom can be divided into three channels:

- Rayleigh (coherent) scattering: $\gamma_1 + H = \gamma_2 + H$. The frequency of the photon remains essentially unaltered, and only the direction of its motion changes. The recoil effect is smaller than for the scattering by a free electron by a factor of $\sim m_p/m_e$.
- Raman scattering: $\gamma_1 + H = \gamma_2 + H(n, l)$, where $H(n, l)$ denotes one of the excited states of the hydrogen atom. The photon energy decreases by the excitation energy of the corresponding level: $h\nu_2 = h\nu - E_{n,l}$ and the Raman satellites of the line appear.
- Compton scattering: $\gamma_1 + H = \gamma_2 + e^- + p$, which is accompanied by ionization of the atom. The photon energy decreases by the ionization potential of the atom, and the kinetic energy of the electron after scattering: $h\nu' = h\nu - 13.6 \text{ eV} - E_e$. The kinetic energy of the proton can be disregarded.

Note that in the nonrelativistic limit ($h\nu \ll m_e c^2$), the sum of the differential cross sections for the three channels is exactly equal to the Thomson differential cross section: $(d\sigma/d\Omega)_{\text{Th}} = 0.5r_e^2(1 + \cos^2 \theta)$.

Below we briefly discuss each of these three channels. A more detailed discussion on the scattering by the hydrogen atom and references to the original papers can be found in [44]. The following notation is used below:

$$\nu, \nu', \mathbf{k} = \Omega \frac{h\nu}{c}, \mathbf{k}' = \Omega' \frac{h\nu'}{c} \quad (222)$$

are the initial and final frequencies and momenta of the photon, $\Delta\nu = \nu - \nu'$, $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ are the changes of the photon frequency and momentum, $\chi = \mathbf{q}/\hbar$, $a = r_B/\hbar$, r_B is the Bohr radius, θ is the scattering angle.

Rayleigh Scattering

Hydrogen Atom

For Rayleigh scattering, the final state of the electron coincides with its initial (ground) state. Thus, Rayleigh scattering occurs without a change in the frequency of the photon, but with a change in the direction of its motion. The motion of the atom as a whole compensates for the change of the photon momentum. For the scattering of photons with energy $h\nu$ much greater than the characteristic binding energy of the electron in the atom ($E_b \approx 13.6 \text{ eV}$) but with a wavelength much longer than the characteristic atomic size ($c/\nu \gg r_b$), the differential scattering cross section in the Thomson limit is given by the expression

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega} \right)_{\text{Th}}. \quad (223)$$

At energies of the order of 1–10 keV the wavelength of the photon is comparable to the atomic size, and the expression for the cross section takes the form (see, e.g. [44])

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega} \right)_{\text{Th}} \left[1 + \left(\frac{1}{2} qa \right)^2 \right]^{-4}. \quad (224)$$

It can be seen from (224) that Rayleigh scattering plays an important role for $qa \lesssim 1$, i.e. for $(2\pi r_b \nu / c) \sqrt{2(1 - \cos \theta)} \lesssim 1$. For X-ray photons, the initial momentum of the photon is large, and the condition $qa \lesssim 1$ means scattering at small angles $\theta \lesssim 1/qa$. For $qa \gg 1$, the cross section for Rayleigh scattering falls off as $(qa)^{-8}$.

Hydrogen Molecule and Helium Atom

An important property of Rayleigh scattering is the possibility of coherent scattering of photons by electrons which are concentrated in a small volume (e.g. in an atom) of characteristic size l . In classical electrodynamics, the parameter $x = l\chi$, the characteristic phase shift between the waves scattered by different electrons, plays a major role. The scattering cross section for $x \ll 1$ is proportional to Z^2 , where Z is the number of electrons. For $x \gg 1$, the scattering by individual electrons occurs independently, and the cross section is simply proportional to Z . The same relationship holds in quantum mechanics.

Under astrophysical conditions, coherent scattering can appreciably increase the importance of elements with $Z > 1$ compared to atomic hydrogen (due to the factor Z per electron for small-angle scattering). For normal cosmic abundances, the contribution of neutral atoms and weakly ionized ions of heavy elements is not too large: summation over all elements increases the cross section for forward scattering by a factor of ~ 1.5 per hydrogen atom. The largest correction ($\sim 40\%$) is introduced by helium. Obviously, the increase in the cross section for Rayleigh scattering by molecular hydrogen and helium may be significant in huge molecular clouds which scatter emission from X-ray sources.

Raman Scattering

For Raman scattering, the final state of the electron corresponds to one of the excited discrete levels. In this case, the photon energy changes by the excitation energy of the appropriate level. For the hydrogen atom, the photon energy decrement is $13.6(1 - 1/n^2)$ eV, where n is the principal quantum

number of the excited level. For X-ray photons, the scattering cross section (with excitation of level n) is given by [145]

$$\left(\frac{d\sigma}{d\Omega}\right)_n = \left(\frac{d\sigma}{d\Omega}\right)_{\text{Th}} \frac{2^8}{3} \frac{(qa)^2}{n^3} \left[3(qa)^2 + \frac{(n^2 - 1)}{n^2}\right] \times \frac{[(n-1)^2/n^2 + (qa)^2]^{n-3}}{[(n+1)^2/n^2 + (qa)^2]^{n+3}}. \quad (225)$$

For X-ray photons, the contribution of Raman scattering to the total cross section is not large. At very small scattering angles, $qa \ll 1$, the cross section $(d\sigma/d\Omega)_n \propto (qa)^2$, and Rayleigh scattering dominates, while at large angles, $qa \gg 1$, and the cross section for Raman scattering falls off as $(qa)^{-8}$. Raman scattering gives the largest contribution when $qa \approx 1$; for 6.4 keV photons, this corresponds to a scattering angle of $\sim 30^\circ$.

Note again that for the scattering of a monochromatic line with energy $h\nu$, a set of monochromatic lines with energies $h\nu' = h\nu - \Delta E_n$, $n = 1, 2, \dots$ arises. This makes it possible to observe the 10.2-eV energy gap (the energy corresponds to the $1s-2p$ transition in hydrogen) below the energy of the initial line. The scattered photons cannot appear in this gap because of the law of conservation of energy.

Compton Scattering

In the case of Compton scattering, the final state of the electron corresponds to one of the continuum states. For the scattering by a free electron at rest, the energy of the scattered photon is uniquely related to the scattering angle by formula (5). For the scattering by a bound electron, this relation breaks down even if the atom or molecule at the initial time was at rest. This is because the photon is essentially scattered by an electron with a certain momentum, rather by an electron at rest. In this case, the law of conservation of momentum is not violated, because the nucleus carries the momentum away. The possibility of this treatment of the scattering process (the so-called impulse approximation) for a change of the photon energy $\Delta h\nu \gg E_b$ was discussed in detail by Eisenberger and Platzman [44]. An analog of the Compton scattering by a bound electron in this approximation is the Compton scattering by a moving electron. It is easy to show that a simple expression for the change of the photon energy follows from the laws of conservation of energy and momentum,

$$\Delta h\nu = \frac{q^2}{2m_e} + \frac{\mathbf{q}\mathbf{p}_0}{m_e}, \quad (226)$$

where $\mathbf{p}_0$ is the initial momentum of the electron in the atom. Note that the first and second terms in (226) correspond to ordinary recoil and the Doppler effect, respectively. The broader the distribution of electrons in momentum, the greater the deviations in the change of the photon energy compared to (5).

For bound electrons, the momentum distribution plays the same role as the temperature does for free electrons. The left wing of the line scattered by free electrons in plasma with temperature ~ 13.6 eV resembles the result of Compton scattering by a neutral atom.

It is possible to derive exact analytical expressions for atomic hydrogen. For X-ray photons, the Compton scattering cross section is given by the expression [44, 62]

$$\begin{aligned} \frac{d\sigma}{dh\nu d\Omega} &= \left(\frac{d\sigma}{d\Omega} \right)_{\text{Th}} \left(\frac{\nu}{\nu'} \right) \int \langle |M_{fi}|^2 \rangle \delta(E_f - E_i - \Delta h\nu) \frac{p^2}{2\pi^2} dp \\ \langle |M_{fi}|^2 \rangle &= \frac{\pi^2 8^3 a^2}{p} \left(1 - e^{-2\pi/pa} \right)^{-1} \exp \left[\frac{-2}{pa} \tan^{-1} \left(\frac{2pa}{1 + q^2 a^2 - p^2 a^2} \right) \right] \\ &\quad \times \left[q^4 a^4 + \frac{1}{3} q^2 a^2 (1 + p^2 a^2) \right] [(q^2 a^2 + 1 - p^2 a^2)^2 + 4p^2 a^2]^{-3}, \\ p^2/2m &= -|E_b| + \Delta h\nu. \end{aligned} \quad (227)$$

For multielectron atoms, the impulse approximation can be used to calculate the spectrum of the scattered emission (for an energy change $\gg E_b$),

$$\begin{aligned} \frac{d\sigma}{dh\nu d\Omega} &= \left(\frac{d\sigma}{d\Omega} \right)_{\text{Th}} \frac{1}{(2\pi)^3} \int \delta \left(\Delta E - \frac{q^2}{2m} - \frac{\mathbf{q}\mathbf{p}_0}{m_e} \right) P(\mathbf{p}_0) d^3 p_0 \\ &= \left(\frac{d\sigma}{d\Omega} \right)_{\text{Th}} J(\mathbf{q}\mathbf{p}_0), \end{aligned} \quad (228)$$

where $P(\mathbf{p}_0)$ is the probability of finding the electron with momentum $\mathbf{p}_0$ in the initial state. The quantity $J(q) = J(\mathbf{q}\mathbf{p}_0)$ is called the Compton profile. There are extensive tables that give Compton profiles calculated for multi-electron atoms (see, e.g. [23]).

At lower energies the Rayleigh and Raman scatterings increase considerably in importance, as do the distortions for the Compton scattering.

Scattering by Molecular Hydrogen and Atomic Helium

Molecular Hydrogen

For the scattering by molecular hydrogen, the principal differences from the case of atomic hydrogen arise for small-angle scattering. First, coherent (Rayleigh) scattering by small angles will be enhanced due to the factor Z^2 . Second, the structure of electron terms differs somewhat from the structure of the levels in the hydrogen atom. In particular, the gap between the unshifted line (Rayleigh scattering) and the line arising from the Raman scattering with excitation of the first electron term is close to 11 eV as compared to 10.2 eV for the hydrogen atom.

Compton scattering by large angles is very similar to the scattering by atomic hydrogen. In particular, the recoil profile is smeared due to the distribution in initial electron momentum.

Helium

For the scattering by a helium atom, the Rayleigh scattering increases in importance and the structure of the lines corresponding to the Raman scattering changes significantly. In particular, the gap between the ground level and the first excited level is ~ 20 eV. Note, that at energies ~ 6 keV, the wavelength of X-ray photons $\lambda \sim 2$ Å is comparable to the atomic size, and the parity selection rule is not strict. Since the electron is more strongly bound in the helium atom, the distribution in electron momentum is appreciably broader than the distribution for atomic and molecular hydrogen. Hence, the left wing of the scattered line will be smeared more strongly.

Vainstein et al. [179] have performed numerical calculations of the differential cross section for the scattering by atomic helium using the ATOM code [178].

The presence of an energy gap that is twice as wide as that for the hydrogen atom and the noticeably different scattered-line profile gives us hope that we will be able to determine the helium abundance in the scattering medium by analyzing the scattered emission. Note that even for multiple scattering, the photons scattered by helium cannot fall in this energy gap.

Allowance for the Structure of Fluorescent Lines and for the Energy Resolution of X-Ray Detectors

In the preceding examples, we considered the 6.4 keV monochromatic line. In order to calculate the actually observed spectrum of the scattered $K\alpha$ emission, it is necessary to examine more closely the structure of iron fluorescent lines and the finite resolution of X-ray detectors. Two lines ($K\alpha_1$ and $K\alpha_2$) with energies of 6.404 and 6.391 keV and relative intensities 2:1 make the largest contribution to the fluorescent emission of neutral iron atoms (see, e.g. [8]). Interpolation of experimental data indicates that the intrinsic width of these lines is ~ 2.65 and 3.2 eV, respectively, although theoretical calculations predict slightly lower values of ~ 1.5 eV [135]. Fairly accurate measurement of the intrinsic width of each of these components will be accessible to the HTXS observatory.

Models

Let us consider several simple models using the scattering of the fluorescent $K\alpha$ line of iron (6.4 keV) as an example.

Monochromatic Source

All major changes in the spectrum of the scattered emission are clearly seen in the case of scattering in an optically thin medium.

Note again that the distortions of the left wing of the line scattered by neutral hydrogen and free electrons with temperature of ~ 10 eV are similar. Thus, under typical astrophysical conditions, the line profile is smeared nearly always: at low temperatures, electrons are bound in atoms, and the low-frequency wing is smeared due to the momentum distribution of bound electrons, while at high temperatures, electrons are free, and the smearing results from the Maxwellian distribution of electron momenta. Note that under typical astrophysical conditions (the interstellar medium, stellar atmospheres, accretion disks), hydrogen is completely ionized even at temperatures of ~ 1 eV. Consequently, there is an interval of temperatures ~ 1 – 5 eV at which the smearing is not so significant as in the case of higher and lower temperatures.

If the cloud is inhomogeneous or the source is not isotropic, then certain scattering angles will dominate, and the profile of the scattered emission will thus change. In particular, for a cloud illuminated by a distant monochromatic source, the recoil profile will be determined by the relative positions of the cloud, source, and observer.

The discovery of a giant molecular cloud [7] in the direction of the strong hard X-ray source 1E1740.7–2942 suggests that this source is surrounded by dense molecular gas. Millimeter observations indicate that the Thomson depth of the cloud may reach $\tau_T \sim 0.2$. Sunyaev et al. [170] have pointed out that in this case the cloud must scatter up to 20% of the emission from the source if it lies at the center of the cloud. The source 1E1740.7–2942 is highly variable; the characteristic time scale of the variability is close to half a year, according to GRANAT observations. The X-ray flux from this source at minimum light decreases at least by a factor of 5–10 [32], which significantly facilitates observations of the X-ray emission scattered by molecular hydrogen. It is obvious that along with scattering, the interstellar gas must photoabsorb X-rays and strongly emit in fluorescent lines of iron and other heavy elements.

Since the optical depth of the molecular cloud for Thomson scattering is fairly large (~ 0.2), it is hoped that new-generation X-ray spectrometers will be capable of detecting the second-order effect—recoil due to the scattering of the iron fluorescent line formed within the cloud by molecular hydrogen. This effect is proportional to the square of the cloud optical depth, i.e. up to 20% of the photons in the fluorescent line will show an appreciable decrease in their energy compared to unscattered photons.

Observations of the recoil effect make it possible to pinpoint, in principle, the position of the source in the cloud. The recoil profile strongly suggests that we are dealing with the scattering by molecular or atomic hydrogen. The abundance of the latter is low, because no intensity peak in the 21-cm line has been detected in this direction. A detailed analysis of the recoil profile also allows us to derive the helium abundance in the cloud.

Galactic Center Region

Another obvious example is the Galactic Center region as a whole. GINGA observations have revealed a bright diffuse X-ray source in the central region of the Galaxy that intensely emits in the resonance line of the helium-like ion of iron with energy of ~ 6.7 keV. The ART-P telescope aboard the GRANAT satellite has localized five compact X-ray sources within 100 pc of the Galactic center, including a weak variable source with a hard X-ray spectrum within 1 arcmin of the well-known radio source Sgr A* [120]. The ART-P X-ray map of the Galactic Center region shows that the angular distribution of the hard diffuse emission is in good agreement with the CO brightness distribution which reflects the distribution of molecular clouds [106].

Sunyaev et al. [160] noted that such an angular distribution of the diffuse emission may result from the scattering of emission from compact sources, which were bright in the past, by the gas of the molecular clouds surrounding the Galactic Center. It is obvious that if Sgr A* or any compact binary source in this region had a luminosity of 10^{39} ergs/s 100–400 years ago, then we would observe now a bright diffuse component of the scattered emission. Sunyaev et al. [160] predicted that if the diffuse component arises from the scattering by molecular hydrogen, then molecular clouds must be bright in the 6.4 keV fluorescent line.

This prediction has been confirmed by ASCA observations [90] that have revealed a bright fluorescent line of iron in the direction of the largest molecular complexes Sgr B, Sgr A, and Sgr C. In addition, the ASCA observations have lent support to the presence of diffuse emission in the resonance lines of helium-like iron with an energy of ~ 6.7 keV. There is thus the problem of scattering of the observed $K\alpha$ line by the gas of the same cloud in which the fluorescent photons are produced. Furthermore, molecular complexes must scatter the emission in the lines of highly ionized iron that illuminates the cloud from outside.

With the advent of a new generation of X-ray telescopes with high sensitivity and energy resolution of 1–10 eV, observations of the recoil profile may become a major source of information on the amount and distribution of neutral and molecular hydrogen in the Galactic Center region.

Active Galactic Nuclei

A major goal of the new generation of X-ray telescopes is the spectroscopy of AGNs. The spectra of a significant fraction of these objects are known to exhibit strong absorption at low energies which is interpreted as due to the passage of their emission through the gas and dust torus that surrounds the central source. The Thomson depth of these sources may be of the order of unity or larger. Since the matter in the gas and dust torus is neutral, the observed line profile will be distorted by the effects considered above.

Another important subject of research is the line profile formed in accretion disks around galactic nuclei. The Doppler shift causes the line to broaden, allowing the line profile to be used for diagnosing the motion of matter in accretion disks. The scattering by neutral matter in a disk can also contribute to the distortions of the line profile.

Huge concentrations of molecular gas of mass $M \sim 10^{11} M_{\odot}$ were detected in quasars located at redshifts ~ 2.3 and 4.7 [117, 118, 155]. It is of interest to measure the He/H ratio at such large redshifts.

10^4K Plasma in the Vicinity of QSOs and AGNs

Gas clouds with an appreciable optical depth for Thomson scattering, in which hydrogen is completely ionized while helium is single ionized, are observed in the vicinity of QSOs and AGNs. This makes it possible to observe the scattering by hydrogen-like ions of helium with a characteristic energy gap of 40.8 eV .

In conclusion, we note that the Raman lines must also arise from the scattering by other (heavier) elements. The major factors than determine the intensity of the Raman lines (in the case of an appreciable optical depth for Thomson scattering) are the abundance of a given element and the presence of levels whose excitation energy is comparable to the characteristic recoil energy for the scattering by a free electron at rest. From this point of view, of particular interest may be young supernova remnants with an overabundance of heavy elements.

5 6.4-keV Fluorescent Emission from Molecular Clouds in the Galactic Center

The central $\sim$ square degree of our Galaxy is known to host a powerful diffuse X-ray source with a luminosity of $\sim 10^{37}\text{ erg/s}$ [182]. The spectral shape of the X-ray continuum is consistent with thermal emission from an optically thin hot plasma at a temperature of about 10 keV . The GINGA satellite has discovered intense emission in the 6.7-keV resonance line of helium-like iron [89, 190]. ASCA observations [91] have revealed a number of X-ray lines in the $1\text{--}7\text{ keV}$ energy range which are attributed to helium- and hydrogen-like ions of Si, S, Ar, Ca, and Fe. The simultaneous existence of the emission lines of iron and lighter elements indicates that the hot plasma in the Galactic Center is not in collisional ionization equilibrium, i.e. it cannot be characterized by a single temperature.

The ART-P telescope aboard the GRANAT satellite has localized five compact X-ray sources within 100 pc of the Galactic Center, including a weak variable source with a hard X-ray spectrum within 1 arcmin of the well-known radio source Sgr A* [120]

The X-ray surface brightness distribution is elongated along the Galactic plane and, particularly at higher energies, $\gtrsim 12$ keV, roughly follows the angular distribution of CO emission in the 2.6-mm line [106]. It has been suggested [106, 160] that this higher-energy component may result from the Thomson scattering of X-ray emission from nearby compact sources, which were bright in the past, by the dense gas of the molecular clouds. Based on such a scenario, [106, 160] predicted that the molecular clouds must be bright in the 6.4-keV fluorescent line. This prediction has been confirmed by ASCA observations [90] that have revealed a bright fluorescent line of iron in the direction of the largest molecular complexes Sgr B, Sgr A, and Sgr C. The molecular complex Sgr B2 located ~ 40 arcmin east of the Galactic Center turns out to be particularly bright in the 6.4-keV line.

5.1 Surface Brightness Distribution of the Neutral and Ionized Iron Line Emission

One of the most prominent spectral features is the complex of iron lines in the 6.4–7.0 keV energy range. The ASCA observations have shown [91] that this complex of spectral lines can be resolved into two distinct components:

- 6.4-keV $K\alpha$ line of neutral iron resulting from reprocessing of X-ray emission by neutral or weakly ionized gas.
- Blends of lines from highly ionized iron (mostly He-like and H-like) in the 6.6–7.0 keV range.

The presence of these two components indicates that both neutral and highly ionized gas contribute to the observed emission. The surface brightness distribution and equivalent width of the two components are essentially different.

The line emission from both neutral and ionized iron concentrates toward the Galactic plane and roughly follows the brightness distribution of CO emission. However, there is no global correlation between line and integrated CO emission on angular scales of $\sim$ a few arcmin.

The brightness distribution of emission from highly ionized iron is approximately symmetric with respect to the Galactic Center. No strong variation of the equivalent width of the 6.7- and 6.9-keV lines has been found with typical values of ~ 400 and ~ 200 eV, respectively.

On the contrary, the surface brightness distribution of the 6.4-keV line is strongly asymmetric, with the most of the emission originating at positive Galactic longitudes. The flux and equivalent width of the 6.4-keV line peak towards the Sgr B2 complex (the equivalent width ~ 1 keV) and the Sgr A/Radio Arc region (~ 0.5 keV). These two bright spots are connected by a “bridge” of 6.4-keV emission with an averaged value of the equivalent width of ~ 0.3 keV. The average value of the equivalent width at negative Galactic longitudes is about twice smaller, ~ 0.15 keV.

5.2 Sgr B2 Giant Molecular Cloud

The brightest spot on the 6.4-keV line map is associated with the Sgr B2 giant complex of molecular clouds. It is also bright in continuum X-ray emission as well as in the lines of heavily ionized iron ions (H- and He-like). The continuum emission spectrum differs from the measured spectra of emission from other regions and has a shape typical of spectra of reflected emission from an optically thick medium. Infrared and millimeter observations have provided an estimate of the mass of molecular gas in the Sgr B2 complex of $\sim 4 \cdot 10^6 M_\odot$ and indicated ongoing star formation (see, e.g. [57]).

A comparison of the surface brightness distribution for the 6.4-keV line with that of ^{13}CO emission integrated over the $+40 - +80$ km/s velocity range shows that these two distributions correlate fairly well. It is therefore plausible to assume that the 6.4-keV emission is indeed related to molecular gas of the Sgr B2 complex. However, the peak of the 6.4-keV emission does not coincide with either of the Sgr B2 cores and is offset by $\sim 1-2$ arcmin approximately in the direction to the nucleus of the Galaxy. On the other hand, the maximum of the 6.4-keV emission nearly coincides with the maximum of the $60\mu\text{m}$ IRAS map. Not all molecular cloud complexes that are visible well in molecular lines and on dust emission maps are bright in the 6.4-keV line emission. The Sgr B1 complex clearly visible on the IRAS $60\mu\text{m}$ map does not manifest itself in the fluorescent emission.

A remarkable feature of the X-ray emission in the direction of the Sgr B2 complex is large equivalent width and luminosity of the 6.4-keV line. The equivalent width of the line, ≈ 1 keV, is consistent with the expected value for a situation where only scattered emission and no direct emission is observed (assuming the solar abundance of iron and a moderate optical depth $\tau_T \lesssim 1$) [158, 177]. It therefore suggests that the direct emission from a source illuminating the molecular gas of the Sgr B2 complex does not contribute significantly to the observed continuum.

Neither the ambient diffuse emission nor any of the compact sources observed in the region are luminous enough to account for the observed luminosity of the Sgr B2 complex in the 6.4-keV line, $L_{6.4} \sim 4 \cdot 10^{34}$ erg/s. Therefore, there are two major possibilities:

- A strongly variable X-ray source located either inside or outside the Sgr B2 molecular cloud or
- A heavily obscured source(s) located inside the cloud, for example, associated with star forming regions found in the the cloud cores (see, e.g. [49]).

Luminosity of a Source of the Primary Radiation

The flux in the 6.4-keV line from a cloud exposed to a continuum radiation is given by the expression

$$F_{6.4} = \frac{\Omega}{4\pi D^2} n_{\text{Fe}} r Y \int_{7.1}^{\infty} I(E) \sigma_{\text{ph}}(E) dE \text{ phot s}^{-1} \text{ cm}^2, \quad (229)$$

where Ω is the solid angle subtended by the cloud at the location of the primary source, D is the distance to the observer, $n_{\text{Fe}} r$ is the column density of the cloud expressed in terms of the number of iron atoms, $I(E)$ is the spectrum of the primary source (in units of phot/s/keV). Since the photoabsorption cross section $\sigma_{\text{ph}}(E)$ is a steep function of energy, the 6.4-keV flux depends mainly on the source flux at $\sim 7\text{--}9$ keV. It is convenient to express the 6.4-keV flux via the source luminosity at 8 keV in a 8 keV-wide energy range,

$$F_{6.4} = \phi \cdot 10^7 \frac{\Omega}{4\pi D^2} \frac{\delta_{\text{Fe}}}{3.3 \cdot 10^{-5}} \tau_{\text{T}} L_8 \text{ phot s}^{-1} \text{ cm}^2, \quad (230)$$

where ϕ is a factor of the order of unity, depending (weakly) on the shape of the source spectrum. For bremsstrahlung emission, this factor changes from 1 to 1.3 when the temperature increasing from 5 to 150 keV. The parameter L_8 characterizes the luminosity of the source in the standard X-ray band. For example, for bremsstrahlung spectra with temperatures between 5 and 150 keV, L_8 corresponds to 40–45 of the source luminosity in the 1–20 keV band. Thus the source luminosity required to produce the observed 6.4-keV flux is

$$L_8 \approx 6 \cdot 10^{38} \left(\frac{F_{6.4}}{10^{-4}} \right) \left(\frac{0.1}{\tau_{\text{T}}} \right) \left(\frac{\delta_{\text{Fe}}}{3.3 \cdot 10^{-5}} \right)^{-1} \left(\frac{R}{100 \text{ pc}} \right)^2 \text{ erg s}^{-1}, \quad (231)$$

where R is the distance from the source to the cloud. The above crude estimate assumes that the source is well outside the cloud and $\tau_{\text{T}} \ll 1$. Although high enough, this value is still much below the Eddington limit 10^{44} erg/s for a $\sim 10^6 M_{\odot}$ black hole that is thought to be residing in the Galactic Center [50], and even a rather short (lasting, say, several days) flare at the Eddington level could provide the required flux. Note that if the duration of the flare, Δt , is shorter than the light crossing time of the cloud, r/c , the above estimate should be multiplied by a factor $\sim r/c\Delta t$. In other words, for a very short flare, it is the product $L\Delta t$ (luminosity $\times$ duration) which determines the 6.4-keV flux [160].

A less luminous object is required if one assumes that the primary source of continuum emission was located close to or inside the Sgr B2 complex and faded away some time (~ 10 years) ago. For a source embedded into a uniform cloud, the required luminosity is

$$L_8 \approx 6 \cdot 10^{35} \left(\frac{F_{6.4}}{10^{-4}} \right) \left(\frac{0.1}{\tau_{\text{T}}} \right) \left(\frac{\delta_{\text{Fe}}}{3.3 \cdot 10^{-5}} \right)^{-1} \text{ erg s}^{-1}. \quad (232)$$

For a hard spectrum (e.g. bremsstrahlung with $kT \sim 100$ keV) the 1–150 keV luminosity is a factor of ~ 7 larger than L_8 , but it is still consistent with the observed luminosities of X-ray Novae with hard spectra. This estimate should also be increased if the source was bright during a period of time shorter than the light crossing time of the cloud.

5.3 X-Ray Archaeology: Activity of Sgr A* in the Recent Past

As suggested in [91, 106, 160], a primary candidate for an illuminating source external to the cloud is the supermassive black hole located at the Galactic Center. A conservative upper limit on the present luminosity of this object is $\sim 10^{36}$ erg/s, which corresponds to $\sim 10^{-8}$ of the Eddington luminosity for a $\sim 2 \cdot 10^6 M_\odot$ black hole. In order to account for the observed 6.4-keV line flux from the Sgr B2 complex, the nucleus of the Galaxy must have had luminosity of $\sim 10^{39}$ erg/s ~ 200 – 300 years ago (assuming a duration of the outburst $\Delta t \sim 10$ – 50 years).

In the case of such a short outburst, a parabola with focus at Sgr A* denotes positions with similar propagation times from the source (Sgr A*) to the cloud and then to the (distant) observer. The size of the parabola is determined by the time elapsed since the outburst. Therefore, the fluorescent photons which are observed at a given moment of time were produced in neutral matter located at the surface of the parabola. Molecular clouds located either inside or outside the parabola cannot contribute to the observed reprocessed emission. This may provide an explanation for the above-mentioned lack of a correlation between the $K\alpha$ line and CO emission and, in particular, for the fact that some of the giant molecular clouds of mass of $\sim 10^5$ – $10^6 M_\odot$ are dim in the reprocessed emission.

Bright Spots

If the flare is short compared to the light-crossing time of the cloud, then the observed surface brightness at a given moment will be determined not by the total optical depth of the cloud, but rather by the density of the cloud at the surface of the parabola. The surface brightness is defined by the integral $\int (I/4\pi r^2) n dl$ over the line of sight. The integration limits are defined by two parabolas corresponding to the beginning and the end of the flare. One can write a simple expression for the surface brightness (flux from the solid angle $d\Omega$) of the 6.4-keV line emission,

$$S = 7 \cdot 10^{-6} \left(\frac{n}{10^{15} \text{ cm}^{-3}} \right) \left(\frac{\Delta t}{1 \text{ year}} \right) \left(\frac{L_8}{10^{39}} \right) \left(\frac{100 \text{ pc}}{x} \right)^2 \times \left(\frac{d\Omega}{(15'')^2} \right) \frac{\eta^2}{1 + \eta^2} \text{ phot s}^{-1} \text{ cm}^{-2}, \quad (233)$$

where Δt is the duration of the flare, $\eta = x/ct$, x is the projected distance from the source to the bright spot, t is the time elapsed since the flare.

The above formula (scaled to the angular resolutions of the XMM and JET-X on Spectrum-X-Gamma) shows that with an integration time of 10^5 s and with the effective area of ~ 300 – 3000 cm^2 at 6.4 keV, these instruments will be capable of tracing the density variations in the cloud. The estimated size of dense condensations in the Sgr B2 cloud of ~ 0.5 – 0.3 pc (see, e.g. [181])

is well matched with the angular resolution of these telescopes. Note that the energy resolution of a typical X-ray CCD is sufficient for searching for bright spots. Thus, if the Sgr B2 cloud was indeed illuminated by a short flare, then one can expect very strong variations (up to three orders of magnitude according to the data on molecular line tracers of high density) in the surface brightness of the 6.4-keV flux across the cloud image on the angular scales corresponding to the size of nonuniformities in the cloud, 10–20". If on the contrary, the flare lasted a sufficiently long time, then the surface brightness distribution would reflect the total optical depth of the cloud (in a given line of sight). In this case, the distribution will be substantially smoothed because of the large contribution to the total scattering mass of the extended cloud envelopes.

6 X-Ray Emission from Supernova 1987A

The outburst of the supernova 1987A in the LMC has once again drawn attention to the problem of Comptonization of high-frequency radiation in a cold plasma cloud which is optically thick for Thomson scattering. There are several possibilities for the source of hard photons in the central part of the cloud. We mention three of them here. a) The detection of radioactive ^{56}Co is accompanied by the emission of gamma-rays with energies ranging from 511 keV to 3.2 MeV. b) A young pulsar may be radiating similar to the pulsar in the Crab nebula, but possibly with a shorter period and a harder spectrum. c) Hard radiation may be emitted by cosmic rays which are accelerated by the young pulsar in the inner cavity of the envelope. The fate of all of the hard photons is more or less identical. The photons lose their energy rapidly after several Compton scatterings, and the energy falls to 100 keV. Subsequently, they diffuse spatially through the plasma cloud as they undergo Compton scattering off electrons (both free and those which are bound in atoms). In each scattering off an electron at rest, the photon energy is reduced because of the recoil effect: photons begin to flow down along the frequency axis. In this problem, the photons undergo a large variety of number of scatterings in the cloud. Hence, the spectrum of emission which emerges from the cloud must be a broad continuum.

At sufficiently low frequencies, photoabsorptions on the K-shells of heavy elements come into play. In the first instance, this is due to the iron group. This effect leads to a sharp cut-off in the spectrum. This problem was posed in the context of a supernova envelope and solved by a Monte Carlo method. A similar problem was considered independently elsewhere. The principal result of the present article is an analytic solution of this problem. This solution will be obtained using the Fokker–Planck approximation, and it yields quite good agreement with the numerical results for the photons which emerge from the envelope with energies $h\nu \leq 200$ keV. At these low energies, the initial energy of the photons plays practically no role.

6.1 Analytic Solution of the Problem

Transport Scattering Cross Section

The cross section which enters in the spatial diffusion coefficient, $D = c/3\sigma_{\text{tr}}N_e$ takes into consideration the fact that small-angle scatterings cause almost no change in the photon frequency. For scattering off electrons at rest

$$\sigma_{\text{tr}}(\nu) = \sigma_{\text{T}}(\nu)\phi(\nu) = \int (1 - \cos \theta) d\sigma_{\text{C}}(\nu \rightarrow \nu') , \quad (234)$$

where

$$\begin{aligned} d\sigma_{\text{C}} = & \frac{3}{4}\sigma_{\text{T}} \frac{m_e c^2}{h\nu} \\ & \times \left[\frac{\nu'}{\nu} + \frac{\nu}{\nu'} + \left(\frac{m_e c^2}{h\nu'} - \frac{m_e c^2}{h\nu} \right)^2 - 2 \left(\frac{m_e c^2}{h\nu'} - \frac{m_e c^2}{h\nu} \right) \right] \frac{d\nu'}{\nu} \end{aligned} \quad (235)$$

is the differential cross section for Compton scattering; ν is the photon frequency prior to scattering; ν' is the frequency after scattering; θ is the scattering angle.

Integrating (234) over ν' from $\nu/(1 + 2h\nu/m_e c^2)$ to ν we obtain

$$\frac{8}{3}x^4\phi(x) = (3 + 4x - x^2)\ln(1 + 2x) + 2x^4/(1 + 2x)^2 + 2x(x^2 - x - 3) , \quad (236)$$

where $x = h\nu/m_e c^2$. For $x \ll 1$, we have

$$\phi(x) \approx 1 - \frac{14}{5}x + \frac{81}{10}x^2 + \dots \quad (237)$$

We can compare this with the well-known expansion of the Klein–Nishina cross section: $\sigma_{\text{KN}} \approx \sigma_{\text{T}}(1 - 2x + \dots)$. From this we find that, even in the first order term in the x -expansion, a difference is showing up between what we are using and the Rayleigh scattering coefficient. For the function $\phi(x)$, the following approximation is valid with an uncertainty of no more than 2% for energies below 1 MeV:

$$\phi(x) = (1 + 2.8x - 0.44x^2)^{-1} . \quad (238)$$

The Evolution of Photon Energy with Time

is determined by Compton recoil:

$$\frac{dx}{dt} = N_e c \int (x' - x) d\sigma_{\text{C}}(x \rightarrow x') , \quad (239)$$

where $d\sigma_{\text{C}}$ is given by (235). Integrating, we find

$$\alpha(x) = \frac{1}{\sigma_T N_e c} \frac{dX}{dt} = \frac{3}{8x^2} \left[(x^2 - 2x - 3) \ln(1 + 2x) - \frac{2x}{1 + 2x} \left(1 + x - \left(1 - \frac{4}{3} \frac{x^4}{1 + 2x} \right) / (1 + 2x) \right) + 6x \right]. \quad (240)$$

In the limit of small x this reduces to

$$\alpha(x) \approx x^2 \left(1 - \frac{21}{5}x + \frac{147}{10}x^2 + \dots \right). \quad (241)$$

Expression (241) can be approximated well by the formula

$$\alpha(x) = x^2 / (1 + 4.6x + 1.1x^2). \quad (242)$$

The number of scatterings which a photon undergoes during the time required to alter its energy from x_0 to x ($< x_0$) is equal to

$$u = \int_x^{x_0} \sigma_C(x) N_e c \frac{dt}{dx} dx = \int_x^{x_0} \frac{\sigma_C}{\sigma_T \alpha(x)} dx. \quad (243)$$

We can use an approximate expression for the Compton scattering cross section: $\sigma_C(x) = \sigma_T(1 + x)/(1 + 3x + 0.64x^2)$, this is valid for $x < 2$. Combining this with (242), we obtain

$$u \approx \frac{1}{x} - \frac{1}{x_0} + 2.6 \ln \frac{x_0 + 0.36}{x + 0.36} + 0.12 \ln \frac{x_0 + 4.33}{x + 4.33}. \quad (244)$$

The Photon Distribution as a Function of Time of Escape from the Spherically Symmetric Cloud

The photon distribution as a function of time of escape from the spherically symmetric cloud has been derived in the limit of Thomson scattering. In the diffusion approximation, the probability $P(u)du$ that a photon escapes from the cloud after undergoing a number of scatterings between $u = \sigma_T N_e c t$ and $u + du$ (where t is the time which has elapsed prior to escape) is given by the following series:

$$P(u) = \sum \lambda_k \sin \lambda_k \tau_0 \exp(-\lambda_k^2 u / 3), \quad (245)$$

where the eigenvalues λ_k are determined by the equation

$$\tan \lambda_k \tau_0 = -\lambda_k \tau_0 / (1 - 3\tau_0/2). \quad (246)$$

The probability $P(u)$ for photon escape from the cloud as a function of its optical depth for Thomson scattering, $\tau_0 = \sigma_T N_e R$, has been calculated by a Monte Carlo method, assuming a central point source of photons. When the optical depth of the cloud is large ($\tau_0 \gg 1$), the escape probability is

determined by a single parameter, namely, the characteristic photon diffusion time:

$$t_0 \approx \frac{R^2}{4D} \approx \frac{\sigma_{\text{tr}}}{\sigma_{\text{T}}^2} \tau_0 / N_e c \approx \tau_0^2 / \sigma_{\text{T}} N_e c . \quad (247)$$

In the real problem, the transport cross section is frequency-dependent. This results in a substantial alteration in the distribution of photons according to the number of scatterings that they have experienced. The initial photon energies (847 keV and 2.6 MeV) correspond to energies of gamma-ray lines from ^{56}Co . Now, in the non-relativistic case, the photon distribution according to escape time t is determined totally by the quantity $t/t_0 = (4/3)u/\tau_0^2$, in the real problem, this distribution is determined by the quantity

$$\begin{aligned} \frac{4}{3} u_{\text{eff}} / \tau_0^2 &= \int_0^t dt / t_0 \\ &= \frac{4}{3} \tau_0^{-2} \int_{x_0}^x \frac{\sigma_{\text{T}}}{\sigma_{\text{tr}}(x)} \frac{dx}{\alpha(x)\phi(x)} \sigma_{\text{T}} N_e c dt = \frac{4}{3} \tau_0^{-2} \int_{x_0}^x \frac{dx}{\alpha(x)\phi(x)} . \end{aligned} \quad (248)$$

Here, x characterizes the photon energy at the time of escape; x_0 is the initial energy; and their relation to the escape time t is given by the equation

$$t = \frac{x}{x_0} \frac{dx}{\sigma_{\text{T}} N_e c \alpha(x)} . \quad (249)$$

Using (3) and (5), we obtain the following result for energies $h\nu_0 < 1$ MeV:

$$u_{\text{eff}} \approx \frac{1}{x} - \frac{1}{x_0} + 7.4 \ln \frac{x_0}{x} + 13.54(x_0 - x) . \quad (250)$$

Notice that u_{eff} characterizes the escape time of photons from the cloud, and is different from the number of scatterings u that the photons actually experience. It is tempting to assume that $P(u_{\text{eff}})$ has the same form as $P(u)$ in the non-relativistic diffusion problem, assuming the Thomson cross section. Then the escape probability after u scatterings is determined by the formula

$$dP = P(u_{\text{eff}}) \frac{du_{\text{eff}}}{dx} \frac{dx}{du} du = P(u_{\text{eff}}) \frac{\sigma_{\text{T}}}{\sigma_{\text{C}}} \frac{du}{\phi} . \quad (251)$$

7 Accretion onto Black Holes and Neutron Stars

7.1 Introduction

One of the most important properties of accreting black holes in our Galaxy was discovered by Riccardo Giacconi and the *Uhuru* Team in 1971, when they discovered the spectral transition of Cyg X-1 from the soft to the hard state (Tananbaum et al. 1972). Simultaneously, a radio source appeared in the

vicinity of Cyg X-1. Radio observations permitted its localization with high accuracy and the identification of the X-ray source with a bright star of the 9th magnitude. Immediately thereafter, measurements of its optical spectrum showed that this star is member of a 5.6-day non-eclipsing binary with an optically invisible companion (Bolton 1972). Lyuty et al. (1973) interpreted the observed ellipsoidal variations in the brightness of the optical star as a result of the gravitational influence of a nearby black hole invisible in optical light. Today Cyg X-1 is the best-known steadily accreting black hole in our Galaxy. Now we have a list with more than 12 excellent black-hole candidates and many of them show similar soft- to hard state transitions (Tanaka & Shibazaki 1996). Recently, Cyg X-1 experienced the third transition from a hard to a soft state in 18 years.

Such transitions became a signature of black holes. Today we know that all galactic black-hole candidates show a very soft X-ray spectrum. As predicted by standard accretion theory, this is a multicolor disk spectrum (cf. Shakura & Sunyaev 1973) or a power-law hard X-ray spectrum with a Wien-type decay at high energies formed due to comptonization (Sunyaev & Trümper 1979, Sunyaev & Titarchuk 1980). Sometimes we do not even see the high frequency decay yet. Therefore, usually when a newly discovered X-ray transient shows an extremely hot tail in its X-ray spectrum, we immediately refer to it as a black-hole candidate.

Neutron stars without magnetic fields and black holes have practically the same gravitational potential and must show many similarities. Nevertheless, we know now that they have very different X-ray spectra and variability characteristics. One of the great surprises of the last 15 years of observations is the discovery that neutron stars also exhibit soft- to hard-state transitions (Fig. 2).

Neutron stars with small magnetic fields usually have spectra which are significantly harder than the spectra of multicolor accretion disks around black-hole candidates in a high/soft state. But their spectra are usually much softer than the spectra of black-hole candidates in the hard/low state. Sometimes we observe hot tails in the persistent flux of X-ray bursters. However, spectra of these hot tails from neutron stars are much steeper than in the case of black holes and contain a smaller fraction of the source luminosity. It seems that now we know the reason. In the case of black-hole accretion we only see the radiation of accretion disk – plus, maybe, the corona above it (Galeev et al. 1979) or the advection flow with even smaller accretion efficiency (Narayan & Yi 1995). In the case of neutron stars we have an object with a solid surface. Therefore, part of the gravitational energy of the accreting matter must be released in an extended accretion disk, and another part in the narrow boundary layer in the vicinity of the neutron star where accreting matter is decelerating from the Keplerian velocity (of the order of half the velocity of light) to the velocity of rotation at the equator of the neutron star. The surface of the star is able to produce enough soft protons

for comptonization to cool down the hot parts of the disk and boundary layer to temperatures below 20 keV (Sunyaev & Titarchuk 1989). The physics of the boundary layer permits us to explain the strong differences between the radiation spectra of accreting black holes and neutron stars. It also predicts a strong difference in the characteristic variability timescales of the X-ray flux from black holes and neutron stars (see below).

7.2 Efficiency of Accretion onto a Rapidly Rotating Neutron Star

The recent discovery of quasi-periodic oscillations (QPO) with frequencies of the order of 500–600 Hz during the nuclear bursts on the surface of a neutron star appears to be very strong evidence of neutron-star rotation with the same frequency, or with periods of the order of 1.6–2 ms (Strohmayer et al. 1998). This interpretation is natural for a nuclear burning front propagating on the surface of a rapidly rotating neutron star. A bright front region manifests itself as a hot spot giving rise to the QPO. It is important that for a given neutron star the QPO frequency remains the same from burst to burst.

The efficiency of accretion onto neutron stars is higher (usually) than the efficiency of accretion onto black holes. The reason is obvious: in the case of a black hole we have an event horizon and an effective energy release and the release of the observed radiation flux might occur only in the accretion flow well beyond the event horizon. In the case of a neutron star without a strong magnetic field part of the energy is released in the extended accretion disk and another part is liberated in the narrow boundary layer near the surface of the neutron star. In Newtonian mechanics energy release in the boundary layer is equal to

$$L_s = \frac{1}{2} \frac{GM\dot{M}}{R_*} \left(1 - \frac{f}{f_k}\right)^2,$$

or is equal to the energy liberated in the disk

$$L_d = \frac{1}{2} \frac{GM\dot{M}}{R_*}$$

in the case of a slowly rotating compact star. Here and below M is the gravitational mass of the star, R_* is its radius, $f_* = \frac{1}{2\pi} \sqrt{\frac{GM}{R_*^3}}$ the cyclic keplerian frequency near the its surface, f is the frequency of stellar rotation and $\dot{M}$ is the accretion rate.

The problem becomes much more complicated in the case of General Relativity. Kerr metrics is not applicable to the case of rapidly rotating neutron star because the mass distribution within the star is no longer spherically symmetric. There is a strong quadrupole component in the mass distribution. Fortunately, there is an exact solution of the GR equations for the case when the mass distribution has a quadrupole component. Using this solution, Sibgatullin & Sunyaev (2000) plotted the dependence of the energy release

due to the accretion onto a neutron star as a function of the rotation frequency of that star (Fig. 3). The existing GR solution permits us to find the efficiency of the energy release only in the case when the spin directions of the neutron star and accretion disk are parallel or anti-parallel. Unfortunately, the problem with an arbitrary angle between the axes of rotation of the neutron star and the accretion disk is much more complicated.

The energy release efficiency drops rapidly with increasing frequency in the case of corotation and increases rapidly towards high frequencies of counter rotation. The ratio of the disk luminosity to the luminosity in the boundary layer or in the spreading layer near the surface of the star also strongly depends on the frequency of rotation. It is close to 1 for the case of corotation with $f = 600$ Hz and decreases up to 0.2 in the case of counter rotation with the same frequency.

For frequencies of corotation higher than 550 Hz a gap between the marginally stable orbit in the accretion disk and the radius of the star does not exist; then the disk is in contact with the surface of the neutron star. For lower frequencies of corotation and in the case of counter rotation for the EOS FPS and $M = 1.4 M_\odot$ there is a gap $R_m - R_* \approx [1.44 - 3.06(f/\text{kHz}) + 0.843(f/\text{kHz})^2 + 0.6(f/\text{kHz})^3 - 0.22(f/\text{kHz})^4]$ km. In the most interesting case of corotation the gap is very narrow and the thickness of the boundary layer or the height of the spreading layer usually exceeds the dimension of the gap. However, in the case of counter rotation (negative values of f) the gap could be sufficiently large that it has to be taken into account.

The energy release efficiency due to accretion onto a counter-rotating neutron star may reach very large values up to 0.67 Me^2 for the case of a neutron star with baryonic mass $m = 2.1 M_\odot$ for $f = 1.5 \text{ kHz}$ and the EOS FPS. Obviously, such a high energy release efficiency is connected with the spin down of the rapidly (counter) rotating star. This efficiency is much higher than that of disk accretion onto a Kerr black hole. In the case of corotation the energy release efficiency, due to accretion onto a Kerr black hole, is higher than in the case of counterrotation. This is reversed in the case of accretion onto a neutron star.

7.3 Structure of the Boundary Layer

The problem of disk accretion onto a neutron star without a magnetic field is two-dimensional. The height of an accretion disk at low accretion rates and luminosities ($0.01 < L/L_{\text{Edd}} < 0.3$) is small in comparison with the radius of the neutron star. Here and below $L_{\text{Edd}} = \frac{4\pi GMm_p}{\sigma_T}$ is the critical Eddington luminosity. The angular rotation frequency Ω in the disk is close to keplerian and increases when matter approaches the neutron star. In the boundary layer the matter velocity must decrease to the velocity of rotation at the neutron-star surface and then matter must be redistributed over its

equipotential surface. This surface is defined by the common influence of gravity and centrifugal forces. It is obvious that there must be a ring where Ω reaches its maximum, $d\Omega/dR = 0$. There are two possible approaches to consider the matter flow beyond this point. We could assume that the boundary layer is described by the same equations as those valid for the accretion disk or we could consider the motion of matter in the spreading layer as belonging to the surface of the neutron star. We tried to investigate both of these approaches in one-dimensional approximations. In the paper by Popham & Sunyaev (2000) we computed the structure and properties of the boundary layer considering it as a part of the disk. In the case of a low accretion rate or $L \sim 0.01 L_{\text{Edd}}$, the height of the disk in the “neck” between the accretion disk and the boundary layer is close to only 40 meters and the extension of the boundary layer about 1.5 km. The situation drastically changes when we go to the case of high accretion rates with a luminosity close to the critical Eddington luminosity. The height of the neck between the boundary layer and the accretion disk in this case exceeds 2 km and the boundary layer extends up to 2 neutron-star radii.

A more natural approach was considered by Inogamov & Sunyaev (1999). This approach uses the shallow water or hydraulic approximation. It assumes that the thickness of the spreading layer on the surface of the neutron star is less than the circumference of the neutron-star equator $H \ll 2\pi R_*$. This approach assumes that matter entering the equatorial ring with a very high rotational velocity of the order of $0.5c$, where c is the velocity of light. Then the matter begins to spiral slowly towards the poles losing its kinetic rotation energy due to turbulent friction with the dense underlying layer.

The thickness of the spreading layer is highest in the vicinity of the equator and decreases towards the poles. This means that matter is moving down the hill under the influence of gravity, the centrifugal force and the light pressure force. The problem is extremely interesting. We are dealing with radiation dominated plasma when the radiation pressure strongly exceeds the matter pressure. The sound speed is close to $0.1 - 0.15c$. Radiative viscosity is also much stronger than the viscosity of plasma. The solution of the set of hydrodynamic equations results in the following picture (see Inogamov & Sunyaev 1999 for details). Two bright belts equidistant from the equator appear on the surface of the neutron star due to disk accretion.

The energy release in the vicinity of the equator is very low because there centrifugal forces compensate gravity with high precision. Therefore, any substantial radiation flux could destroy the structure of the thin spreading layer. Fortunately, advection takes the radiation energy density and transports it to the bright belts above and below the equator.

In these bright belts the rotational velocity of the spreading matter becomes low enough to permit the existence of a large radiation flux comparable to the critical Eddington flux $q_0 = \frac{m_p c^3}{2\sigma_T R_g} \left(\frac{R_g}{R_*}\right)^2 = 10^{22} \frac{\text{W}}{\text{m}^2}$, where R_g is the

gravitational radius. This flux value is comparable to radiation fluxes achieved in the most intense petawatt laser facilities (Perry 1996, Budil et al. 2000).

We are dealing here with a critical Eddington flux even in the case of a low luminosity of the neutron star ($0.01 < L/L_{\text{Edd}} < 1$). The surface of the bright belts is small and the high radiation flux from the narrow belts is consistent with the low luminosity of the star.

The matter in the spreading layer is practically levitating. The difference between the gravitational force and the centrifugal- and radiation pressure force is close to $(1 - 3) \times 10^{-3}$ of gravity. At higher longitudes the rotational velocity of matter and the velocity of the flow along the meridian decreases and the flow becomes subsonic, cool, dense and very slow.

One of the most interesting predictions of the theory of the spreading layer is the strong dependence of the matter column density in the spreading layer on the accretion rate or the luminosity of the neutron star. In the case of a low luminosity the levitating layer in the bright belts is optically thin against Thompson scattering $\tau_T \sim 2$. Under these circumstances it is impossible to radiate the energy released due to accretion at low temperatures. Comptonization forms hard tails. In the case of a high luminosity the bright belt has a large column density (up to 10 kg/cm^2). Then free-free processes and comptonization form Bose-Einstein type spectra inside the spreading layer and the resulting spectrum is much softer than in the case of low luminosity.

7.4 Time Variability in the Accretion Disk and in the Boundary Layer

All instabilities existing in the accretion disk modulate the flow of matter onto the neutron-star surface. Therefore, we could expect that the majority of the types of variability we observe in accreting black holes must manifest themselves in accreting neutron stars with characteristic timescales proportional to the mass of the accreting object (see e.g. Shakura & Sunyaev 1976, Wijnands & van der Klis 1999). The spreading layer on the surface of the neutron star is the source of additional high-frequency instabilities (see the discussion in Sunyaev & Revnivtsev 2000). Their origin is obvious – the matter in the bright belts is radiation dominated, levitating, the height is smaller than in the region of the main energy release in the accretion disk, the sound velocity is huge and corresponding sound frequencies are very high.

Sunyaev & Revnivtsev (2000) compared the power density spectra of 9 black holes and 9 neutron stars observed by RXTE in their low/hard state. There is a very strong difference. In the power density spectra of accreting neutron stars with a weak magnetic field significant power is contained at frequencies close to one kHz. At the same time, most Galactic accreting black holes demonstrate a strong decline in the power spectra at the frequencies higher than 10–50 Hz. In principle this might open an additional way to distinguish the accreting neutron stars from black holes in X-ray transients

(we do not mention in this paper the well-known differences: X-ray bursts or X-ray pulsations).

The simplest assumption is that the characteristic frequencies in the power spectra of the sources scale as M^{-1} (Shakura & Sunyaev 1976). This scaling law is valid for e.g. the keplerian frequency in the vicinity of the marginally stable orbit, the thermal and secular instabilities of the accretion disk in the region of main energy release, and the Balbus-Hawley instability. However, this assumption does not account for the observed difference in the high frequency variability between neutron stars and black holes.

References

1. F.A. Aharonyan, A.M. Atoyan: *Astrophys. & Space Sci.* 79, 321 (1981)
2. J. Arons: *Astrophys. J.* 164, 437 (1971)
3. J.P. Babuel-Peyrissac, G. Rouvillois: *J. Quant. Spectr. Rad. Transf.*, 10, 1277 (1970)
4. N.A. Bahcall, S.P. Oh: *Astrophys. J.* 462, L49 (1996)
5. N.A. Bahcall, J.P. Ostriker, S. Perlmutter, P.J. Steinhardt: *Science* 284, 1481 (1999)
6. T. Bai: *Solar. Phys.* 62, 113 (1979)
7. J. Bally, M. Leventhal: *Nature* 353, 234 (1991)
8. W. Bambinek et al.: *Rev. Mod. Phys.* 44, 716 (1972)
9. M.M. Basko, R.A. Sunyaev: *Astrophys. Space Sci.* 23, 117 (1973)
10. M.M. Basko, R.A. Sunyaev: *MNRAS* 175, 395 (1975)
11. M. Basko, R.A. Sunyaev, L.G. Titarchuk: *Astron. Astrophys.* 31, 249 (1974)
12. M. Basko: *Astrophys. J.* 223, 268 (1978)
13. M.C. Begelman: *MNRAS* 187, 237 (1979)
14. C.L. Bennett: *Astrophys. J.* 464, L1 (1996)
15. V.B. Beresteskii, E.M. Lifshitz, L.P. Pitaevskii: *Quantum Electrodynamics, Landau and Lifshitz Course of Theoretical Physics* (2nd ed., Pergamon, Oxford 1982)
16. M. Birkinshaw: *Phys. Rep.* 310, 97 (1999)
17. G.S. Bisnovatyi-Kogan, Ya. B. Zel'dovich, R.A. Sunyaev: *Sov. Astron.* 15, 17 (1971)
18. R.D. Blandford: *Astrophys. J.* 238, 410 (1980)
19. R.D. Blandford, D.G. Payne: *MNRAS* 194, 1033 (1981)
20. R.D. Blandford, D.G. Payne: *MNRAS* 194, 1041 (1981)
21. R.D. Blandford, D.G. Payne: *MNRAS* 196, 781 (1981)
22. G.R. Blumenthal, R.J. Gould: *Rev. Mod. Phys.* 42, 237 (1970)
23. F. Briggs, L. Mendelson, J. Mann: *Atomic Data and Nuclear Data Tables* 16, 202 (1975)
24. C. Burigana, L. Danese, G. De Zotti: *Astron. Astroph.* 246, 49 (1991)
25. S.M. Carroll, W.H. Press, E.D. Turner: *Ann. Rev. Astron. & Astrophys.* 30, 499 (1992)
26. A. Cavaliere, R. Fusco-Femiano: *Astron. & Astrophys.* 49, 137 (1976)
27. A. Challinor, A. Lasenby: *Astrophys. J.* 499, 1 (1998)
28. S. Chandrasekhar S.: *Radiative Transfer* (Dover, New York, 1950).

29. G. Chapline, J. Stevens: *Astrophys. J.* 184, 1041 (1973)
30. G.V. Chibisov: *Sov. Astron.* 16, 235 (1972)
31. J. Chluba et al.: in preparation
32. E.M. Churazov et al.: *Astrophys. J.* 407, 752 (1993)
33. E.M. Churazov, R.A. Sunyaev, L. Vainshtein: in preparation
34. S.A. Colgate: *Astrophys. J.* 195, 493 (1975)
35. M. Colpi: *Astrophys. J.* 326, 223 (1988)
36. L. Danese, G. De Zotti: *Nuovo Cimento* 7, 277 (1977)
37. L. Danese, G. De Zotti: *Astron. & Astrophys.* 107, 39 (1982)
38. L.P. David, C. Jones, W. Forman: *Astrophys. J.* 445, 578 (1995)
39. K. Davidson: *Nature Phys. Sci.* 246, 1 (1973)
40. P.A.M. Dirac: *MNRAS* 85, 825 (1925)
41. A.G. Doroshkevich, Ya.B. Zel'dovich, I.D. Novikov: *Sov. Phys.-JETP* 26, 408 (1968)
42. D.M. Eardley, A.P. Lightman, N.I. Shakura, S.L. Shapiro, R.A. Sunyaev: *Comments Astrophys.* 7, 151 (1978)
43. D. Eichler: *Astrophys. J.* 229, 419 (1979)
44. P. Eisenberger, P.M. Platzman: *Phys. Review A* 2, 415 (1970)
45. R. Fabbri: *Astrophys. Space Sci.* 77, 529 (1981)
46. J.E. Felten, M.J. Rees: *Astron. Astrophys.* 17, 226 (1972)
47. D.J. Fixsen, E.S. Cheng, J.M. Gales, J.C. Mather, R.A. Shafer, E. Wright: *Astrophys. J.* 473, 576 (1996)
48. W. Forman, C. Jones: *Ann. Rev. Astron. Astrophys.* 20, 547 (1982)
49. R. Gaume et al.: *Astrophys. J.* 449, 663 (1995)
50. R. Genzel, D. Hollenbach, C. Townes: *Rep. Prog. Phys.* 57, 417 (1994)
51. I.M. George, A.C. Fabian: *MNRAS* 249, 352 (1991)
52. G. Ghisellini, I.M. George, A.C. Fabian, C. Done: *MNRAS* 248, 14 (1991)
53. M. Ghibilisco: *Astrophys. & Space Sci.* 249, 189 (1997)
54. M. Gierlinski, A.A. Zdziarski, C. Done, W. Johnson, K. Ebisawa, Y. Ueda, F. Haardt: *MNRAS* 288, 958 (1997)
55. M. Gierlinski, A.A. Zdziarski, J. Poutanen, P.S. Coppi, K. Ebisawa, W.N. Johnson: *MNRAS* 309, 496 (1999)
56. V.L. Ginzburg, L.M. Ozernoy: *Sov. Astron.* 42, 943 (1965)
57. M.A. Gordon, U. Berkemann, P.G. Mezger, R. Zylka, C.G.T. Haslam, E. Kreysa, A. Sievers, R. Lemke: *Astron. Astrophys.* 280, 208 (1993)
58. R.J. Gould: *Am. J. Phys.* 39, 911 (1971)
59. R.J. Gould: *Ann. Phys.* 69, 321 (1972)
60. R.J. Gould: *Astrophys. J.* 285, 275 (1984)
61. P.W. Guilbert, M.J. Rees: *MNRAS* 233, 475 (1988)
62. A. Gummel, M. Lax: *Ann. Phys.* 2, 28 (1957)
63. J.M. Jauch, F. Rohrlich, *The theory of photons and electrons* (2nd ed., Springer, New York 1976)
64. F. Haardt: *Astrophys. J.* 413, 680 (1993)
65. E.R. Harrison: *Phys. Rev. Lett.* 18, 1011 (1967)
66. E.R. Harrison: *Ann. Rev. Astron. & Astrophys.* 11, 155 (1973)
67. S. Hatchett, R. Weaver: *Astrophys. J.* 215, 285 (1977)
68. M. Heitler: *The Quantum Theory of Radiation* (Clarendon Oxford 1960)
69. D.G. Hummer, D. Mihalas: *Astrophys. J.* 150, L57 (1967)
70. A.F. Illarionov, R.A. Sunyaev: *Sov. Astron.* 16, 45 (1972)

71. A.F. Illarionov, R.A. Sunyaev: *Sov. Astron.* 18, 413 (1975)
72. A.F. Illarionov, R.A. Sunyaev: *Sov. Astron.* 18, 691 (1975)
73. A.F. Illarionov, D.A. Kompaneets: *Sov. Phys. JETP* 44, 930 (1977)
74. A.F. Illarionov, T. Kallman, R. McCray, R. Ross: *Astrophys. J.* 228, 279 (1979)
75. N.A. Inogamov, R.A. Sunyaev: *Sov. Astron. Lett.* 25, 269 (1999)
76. N. Itoh, Y. Kohyama, S. Nozawa: *Astrophys. J.* 502, 7 (1998)
77. N. Itoh, T. Sakamoto, S. Kusano, S. Nozawa, Y. Kohyama: *Astrophys. J. Suppl.* 128, 125 (2000)
78. N. Itoh, T. Sakamoto, S. Kusano, Y. Kawana, S. Nozawa: *Astron. Astrophys.* 382, 722 (2002)
79. I.I. Ivanov: *Radiative Transfer and Celestial Body Spectra*, Nauka, Moscow (1969)
80. M. Jaroszynski, M.A. Abramowicz, B. Paczynski: *Acta Astron.* 30, 1 (1980)
81. C. Jones, W. Forman: *Astrophys. J.* 276, 38 (1984)
82. J.I. Katz: *Astrophys. J.* 206, 910 (1976)
83. D.S. Kershaw, M.K. Prasad, J.D. Beason: *J. Quant. Spectr. Rad. Transf.* 36, 273 (1986)
84. I.R. King: *Astron. J.* 67, 471 (1962)
85. A. Kogut, A.J. Banday, C.L. Bennett, K.M. Gorski, G. Hinshaw, G.F. Smoot, E.L. Wright: *Astrophys. J.* 464, L5 (1996)
86. E.W. Kolb, M.S. Turner: *The Early Universe* (Addison–Wesley, Reading, MA 1990)
87. A.S. Kompaneets: *Soviet Phys.–JETP* 4, 730 (1957)
88. I. Kovner: *Astron & Astrophys.* 141, 341 (1984)
89. K. Koyama, H. Awaki, H. Kunieda, S. Takano, Y. Tawara, S. Yamauchi, I. Hatsukade, F. Nagase: *Nature* 339, 603 (1989)
90. K. Koyama: *New Horizon of X-ray Astronomy*, FSS-12, 181 (Univ. Acad., Tokyo 1994)
91. K. Koyama et al.: *Publ. Astron. Soc. Japan* 48, 249 (1996)
92. L.D. Landau, E.M. Lifshitz: *Quantum Mechanics*, Landau and Lifshitz Course of Theoretical Physics (Pergamon, Oxford 1958)
93. L.D. Landau, E.M. Lifshitz: *The Classical Theory of Fields*, Landau and Lifshitz Course of Theoretical Physics (4th ed., Pergamon, Oxford 1975)
94. I.I. Lapidus, R.A. Sunyaev: *MNRAS* 217, 291 (1985)
95. E.V. Levich, R.A. Sunyaev, Y.B. Zeldovich: *Astron. Astrophys.* 19, 135 (1972)
96. E.V. Levich, R.A. Sunyaev: *Soviet Astron.* 15, 363 (1971)
97. E.M. Lifshitz, L.P. Pitaevskiy: *Physical Kinetics*, Landau and Lifshitz Course of Theoretical Physics (Pergamon, Oxford 1981)
98. A.P. Lightman: *Astrophys. J.* 244, 392 (1981)
99. A.P. Lightman, D.Q. Lamb, G.R. Rybicki: *Astrophys. J.* 248, 738 (1981)
100. A.P. Lightman, T.R. White: *Astrophys. J.* 335, 57 (1988)
101. A. Loeb, F. McKee, O. Lahav: *Astrophys. J.* 374, 44 (1991)
102. Yu.E. Lyubarsky, R.A. Sunyaev: *Sov. Astron. Lett.* 8, 330 (1982)
103. P. Madau, C. Thompson: *Astrophys. J.* 534, 239 (2000)
104. P. Magdziarz, A.A. Zdziarski: *MNRAS* 273, 837 (1995)
105. P. Maltby, E. Avrett, M. Carlsson et al.: *Astrophys. J.* 306, 284 (1986)
106. M. Markevitch, R.A. Sunyaev, M.N. Pavlinsky: *Nature* 364, 40 (1993)
107. T. Matsuda, H. Sato, H. Takeda: *Prog. Theor. Phys. (Japan)* 46, 416 (1971)
108. S. Miyamoto: *Astron Astrophys.* 63, 69 (1978)

109. R. Morrison, D. McCammon: *Astrophys. J.* 270, 119 (1983)
110. J.C. Mather: *Astrophys. J.* 420, 439 (1994)
111. S.M. Molnar, M. Birkinshaw: *Astrophys. J.* 523, 728 (1999)
112. D.I. Nagirner, J.Poutanen: *Astrophys. & Space Phys. Rev.* ed. by R.A. Sunyaev (Harwood Academic Publishers, Chur 1994) 9, 1
113. R. Narayan, R. Mahadevan, E. Quataert: in *The Theory of Black Hole Accretion Discs* ed. M.A. Abramowicz et al. (Cambridge Univ., Cambridge 1998), 148
114. C.B. Netterfield et al.: *Astrophys. J.*, submitted (2001); astro-ph/0104460
115. P.D. Noerdlinger: *Astrophys. J.* 192, 529 (1974)
116. S.L. O'Dell: *Astrophys. J.* 243, L147 (1981)
117. K. Ohta et al: *Nature* 382, 426 (1996)
118. A. Omont et al: *Nature* 382, 428 (1996)
119. B. Paczynski, P.J. Wiita: *Astron. Astrophys* 88, 23 (1980)
120. M.N. Pavlinsky, S.A. Grebenev, R.A. Sunyaev: *Astrophys. J.* 425, 110 (1994)
121. D.G. Payne: *Astrophys. J.* 237, 951 (1980)
122. P.J.E. Peebles: *Phys. Rev. D.* 1, 397 (1970)
123. P.J.E. Peebles: *Principles of physical cosmology* (Princeton Univ. Press, Princeton 1993)
124. J. Peyraud: *J. de Phys.* 29, 88 (1968)
125. E.S. Phinney: in *Superluminal Radio Sources* ed. by J.A. Zensus & T.J. Pearson (Cambridge Univ., Cambridge 1987), 301
126. T. Piran: *Astrophys. J.* 257, L23 (1982)
127. G.C. Pomraning: *The Equations of Radiation Hydrodynamics* (Pergamon, Oxford 1973)
128. R. Popham, R.A. Sunyaev: *Astrophys. J.* 547, 355 (2001)
129. J. Poutanen, R. Svensson: *Astrophys. J.* 470, 249 (1996)
130. L.A. Pozdnyakov, I.M. Sobol, R.A. Sunyaev: *Sov. Astron. Lett.* 2, 55 (1976)
131. L.A. Pozdnyakov, I.M. Sobol, R.A. Sunyaev: *Astron. Astrophys.* 75, 214 (1979)
132. L.A. Pozdnyakov, I.M. Sobol, R.A. Sunyaev: *Astrophys. & Space Phys. Rev.* ed. by R.A. Sunyaev (Harwood Academic Publishers, Chur 1983) 2, 189
133. M.J. Rees: *Phys. Scripta* 17, 193 (1978)
134. R.R. Ross, R. Weaver, R. McCray: *Astrophys. J.* 219, 292 (1978)
135. S.I. Salem, P.L. Lee: *Atomic Data and Nuclear Data Tables* 18, 234 (1976)
136. C.L. Sarazin: *X-ray Emissions From Clusters of Galaxies* (Cambridge Univ. Press, Cambridge 1988)
137. S.Y. Sazonov, R.A. Sunyaev: *Astron. Lett.* 24, 553 (1998)
138. S.Y. Sazonov, R.A. Sunyaev: *MNRAS* 310, 765 (1999)
139. S.Y. Sazonov, R.A. Sunyaev: *Astron. Astrophys.* 354, L53 (2000)
140. S.Y. Sazonov, R.A. Sunyaev: *Astron. Lett.* 26, 494 (2000)
141. S.Y. Sazonov, R.A. Sunyaev: *Astroph. J.* 543, 28 (2000)
142. S.Y. Sazonov, R.A. Sunyaev: *Astron. Lett.* 27, 481 (2001)
143. S.Y. Sazonov, R.A. Sunyaev: *Astron. Astrophys.* 373, 241 (2001)
144. L. Schiff: *Quantum mechanics* (McGraw-Hill, New-York 1955)
145. P. Schnait: *Ann. Physik* 21, 9 (1934)
146. N.I. Shakura: *Sov. Astron.* 16, 532 (1972)
147. N.I. Shakura: *Sov. Astron.* 18, 259 (1974)
148. N. I. Shakura, R.A. Sunyaev: *Astron. & Astrophys.* 24, 337 (1973)
149. S.L. Shapiro, A.P. Lightman, D.M. Eardley: *Astrophys. J.* 204, 187 (1976)

150. A.I. Shestakov, D.S. Kershaw, M.K. Prasad: *J. Quant. Spectr. Rad. Transf.* 40, 755 (1988)
151. N.R. Sibgatullin, R.A. Sunyaev: *Astron. Lett.* 26, 699 (2000)
152. M. Sikora: *MNRAS* 197, 529 (1981)
153. J. Silk: *Astrophys. J.* 151, 459 (1968)
154. V.V. Sobolev: *Light Scattering in the Planet Atmospheres*, Nauka, Moscow (1972)
155. P.M. Solomon, D. Downes, S.J.E. Radford: *Astrophys. J.* 398, L29 (1992)
156. L. Spitzer: *Physical Processes in the Interstellar Medium*, (Wiley, Chichester 1978)
157. B.E. Stern, J. Poutanen, R. Svensson, M. Sikora, M.C. Begelman: *Astrophys. J.* 449, L13 (1995)
158. R.A. Sunyaev, E.M. Churazov: *MNRAS* 297, 1279 (1998)
159. R.A. Sunyaev, M.R. Gilafnov, E.M. Churazov: *Truemper Symposium*
160. R.A. Sunyaev, M. Markevich, M. Pavlinsky: *Astrophys. J.* 407, 606 (1993)
161. R.A. Sunyaev, N.I. Shakura: *Sov. Astron. Lett.* 12, 117 (1986)
162. R.A. Sunyaev, J. Truemper: *Nat.* 279, 506 (1979)
163. R.A. Sunyaev, L.G. Titarchuk: *Astron. & Astrophys.* 86,121 (1980)
164. R.A. Sunyaev, Ya.B. Zel'dovich: *Astrophys. Space Sci.* 7, 20 (1970)
165. R.A. Sunyaev, Ya.B. Zel'dovich: *Comments Astrophys. Space Phys.*, 4, 173 (1972)
166. R.A. Sunyaev, Ya.B. Zel'dovich: *Astron. Astrophys.* 20, 189 (1972)
167. R.A. Sunyaev: *Soviet Astron. Lett.*, 6, 213 (1980)
168. R.A. Sunyaev, Ya.B. Zel'dovich: *Ann. Rev. Astron. & Astrophys.* 18, 537 (1980)
169. R.A. Sunyaev, Ya.B. Zel'dovich: *MNRAS* 190, 413 (1980)
170. R.A. Sunyaev et al.: *Astrophys. J.* 383, L49 (1981)
171. R. Svensson: *MNRAS* 209, 175 (1984)
172. Y. Tanaka, N. Shibazaki: *Ann. Rev. Astron. & Astrophys.* 34, 607 (1996)
173. K.S. Thorne: *MNRAS* 194, 439 (1981)
174. L.G. Titarchuk: *Astrophys. J.* 434, 570 (1994)
175. A.I. Tsygan: *Astroph. & Space Sci.* 77, 187 (1981)
176. M.S. Turner: *Physics Rep.* 333, 619 (2000)
177. L. Vainshtein, R.A. Sunyaev: *Sov. Astron. Lett.* 6, 673 (1980)
178. L. Vainshtein, V. Shevelko: *Program ATOM for calculation of atomic characteristics*, Preprint of the Lebedev Physical Institute No 19 (Moscow 1983)
179. L.A. Vainshtein, R.A. Sunyaev, E.M. Churazov: *Sov. Astron. Lett.* 5–6, 323(1998)
180. M. Van der Klis: *Ann. Rev. Astron. Astrophys.* 38, 717 (2000)
181. P. de Vicente, J. Martin-Pintado, T.L. Wilson: *Astron. Astrophys.* 320, 957 (1997)
182. M. Watson et al.: *Astrophys. J.* 250, 142 (1981)
183. S. Weinberg: *Gravitation and Cosmology* (Freeman, New York 1972)
184. S. Weinberg: *Rev. Mod. Physics* 61, 1 (1989)
185. R. Weymann: *Astrophys. J.* 145, 560 (1966)
186. R. Weymann: *Astrophys. J.* 147, 887 (1967)
187. T.R. White, A.P. Lightman, A.A. Zdziarski: *Astrophys. J.* 331, 939 (1988)
188. D.A. White, A.C. Fabian: *MNRAS* 273, 72 (1995)
189. P.J. Wiita: *Comm. Astrophys.* 9, 251 (1982)

190. S. Yamauchi et al.: *Astrophys. J.* 365, 532 (1990)
191. A.A. Zdziarski, J. Poutanen, J. Mikolajewska, M. Gierlinski, K. Ebisawa, W.N. Johnson: *MNRAS* 301, 435 (1998)
192. Ya.B. Zel'dovich: *Sov. Physics Usp.* 9, 602 (1967)
193. Ya.B. Zel'dovich, E.V. Levich: *Sov. Phys.-JETP* 28, 1287 (1969)
194. Ya.B. Zel'dovich, N.I. Shakura: *Sov. Astron.* 13, 175 (1969)
195. Ya.B. Zel'dovich, R.A. Sunyaev, *Astrophys. & Space Sci.* 4, 301 (1969)
196. Ya.B. Zel'dovich, E.V. Levich: *Soviet Phys.-JETP* 11, 35 (1970)
197. Ya.B. Zel'dovich, R.A. Sunyaev: *Astrophys. Space Sci.* 9, 368 (1970)
198. Ya.B. Zel'dovich, E.V. Levich, R.A. Sunyaev: *Soviet Phys.-JETP* 35, 733 (1972)
199. Ya.B. Zel'dovich: *Sov. Phys. Usp.* 18, 79 (1975)
200. Ya. B. Zel'dovich, A.F. Illarionov: *Sov. Phys.-JETP* 38, 643 (1975)
201. Ya.B. Zel'dovich, I.D. Novikov: *The Structure and Evolution of the Universe*
202. Ya.B. Zel'dovich, Yu.P. Raizer: *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Academic Press, New York 1967)
203. M. Zombeck: *Handbook of Space Astronomy & Astrophysics* (Cambridge University, Cambridge 1990)

Index

- 1E 1740-29, “great annihilator” 87,
88, 191
1E1740.7-2942 262
3C273 92
- Absorption 20, 151
 bound-free *see also* Photoioniza-
 tion, 20–21
 cross section *see* Cross section,
 photoelectric absorption
 photoabsorption rate *see* Photoab-
 sorption rate
 photoelectric 98–100, 102, 118
Absorption edge 102
Absorption index 112
Absorption line 54
Abundance of element 61, 62, 70, 72
Accretion 207, 274
Accretion disk 7, 199, 208, 230, 236,
253, 254, 273–275
Accretion flow 247
Active galactic nucleus 8, 62, 158, 190,
191, 199, 214, 249, 254, 263
Active shield 119
Advanced Compton Telescope
176–180
Analog-to-digital converter 123
Anderson, C. 86, 89
Anger camera 135, 137–138, 165, 175
Angular momentum
 orbital 37
 spin 37
 total 30
Angular momentum operator 26
Angular momentum state 32
Angular resolution 132, 174, 176, 177
Annihilation 95, 97, 114–116, 118, 119
Annihilation line 86–87, 89, 152–153,
190
Annihilation radiation 117
Anti-coincidence veto system (INTE-
GRAL) 140
Anti-Compton shield 119
Anticoincidence electronics 119
Anticoincidence shield 135, 153, 166,
175, 178
Antimatter 86
ARIEL V 160
Aristotle 159
ASCA 4, 93
ASE-MIT 89
ATHENA 178–179
Atomic units 26, 48, 50
Atoms
 multi-electron 24–41
Attenuation
 coefficient 98
 mass 100–101
 coefficient, linear 98
 gamma rays 122
 mass 100
 probability 100
Attenuation, interstellar 62
Auger decay 43
Auger electron 102
Autoionization 43, 53–57, 79
 rate 54–55
- Background
 cosmic diffuse gamma-ray 85
Backprojection 171–172
Backscatter peak 118
Ballistic deficit (detectors) 130, 149
Balmer α transition 29
Band gap 136–137, 142–147

- BATSE 91–92, 151, 154
- Bayes' theorem 172
- Bayesian inference 172
- Becquerel, H. 89
- Beer's law 113
- BeppoSAX 85, 89, 90, 93
- Bethe integral 47
- Beutler-Fano absorption profile 54
- BGO *see* Detectors, scintillators, inorganic
- Bias resistor 122
- BL Lac object *see* Blazar
- Black hole 7, 8, 199, 209, 247, 249, 254, 267, 272–278
- Blackbody 57, 59
- Blackbody radiation 233, 234
- Blackbody spectrum 231
- Blazar 8, 209
- Bohr radius 26, 48, 257
- Boltzmann factor 49, 73
- Born approximation 46, 47
- Bose-Einstein distribution 219, 233
- Bose-Einstein spectrum 277
- Bound-bound transition *see* Transition
- Bound-free absorption *see* Absorption
- Boundary layer 199, 274–277
- Bragg angle 110–112, 182–185
- Bragg condition 109, 182, 184
- Bragg diffraction 182
- Bragg lens 181
- Bragg relation 185
- Bremsstrahlung 51, 95, 199, 228, 232, 241, 267
- Bremsstrahlung absorption *see* Absorption, bremsstrahlung
- Bremsstrahlung emissivity 229
- Bremsstrahlung escape 117
- Brightness temperature 228
- Brookhaven Gas Detector Group 130
- Calorimeter 176–178
- Camera obscura 159
- Capella 75, 78
- Capillary concentrator 181
- Cascade
 - radiative *see* Radiative cascade
- Cataclysmic variable (CV) 7
- Central field approximation 31–33, 35, 37
- Chandra 4, 66, 75–78, 80
- Chandrasekhar limit 7
- Chandrasekhar-Sobolev law 253
- CLAIRE 189
- Classical harmonic oscillator 15–17
- Clebsch-Gordan coefficients 37
- Close coupling calculation 46
- Coded field of view
 - fully 163, 166
 - partially 163, 166
- Coded mask 151, 152, 158–168
- Coded mask telescope 151, 158
- Collimator 151
 - anticollimator 151, 154
 - bigrid 154–155, 157
 - modulation 151
 - multi-layer 157
 - multi-pitch 157
 - Oda 154
 - rastering 151
 - rotating modulation 151, 152, 155–158
 - scanning 152–153
- Collision strength 48
- Collisional deexcitation *see* Deexcitation, collisional
- Collisional excitation *see* Excitation, collisional
- Collisional ionization *see* Ionization, collisional
- Collisions
 - electron-ion 41–57, 60
- Column density, absorption 62
- Commutation relation 22
- Compact object 7, 96, 158, 164–167, 190
- COMPTEL 86–90, 92, 96, 135, 138, 170, 173–176
- Compton acceleration 207–209
- Compton continuum 104, 117
- Compton cooling 244, 246
- Compton cooling (inverse) 221
- Compton drag 207–208
- Compton edge 104, 118
- Compton effect 100, 103, 135, 169
 - double 232–234, 241–242

- Compton equation 170
- Compton Gamma Ray Observatory (CGRO) 89–91, 173, 175
- Compton heating 221, 244, 246
- Compton heating (induced) 221, 226
- Compton parameter 218
- Compton reflection 249, 256
- Compton scatter angle 133, 170–171
- Compton scattering *see* Scattering, Compton
- Compton telescope 97, 122, 131, 135, 148, 149, 168–180
- Compton temperature 253
- Compton wavelength 201
- Comptonization 199, 210, 217, 237, 241–249, 273
- Conduction band 136–137, 143
- Conduction, thermal 68
- Configuration 32, 39
 - ground-state 38–40
- Configuration interaction 40, 54
- Constellation-X 80, 182
- Continuity equation 10
- Continuum
 - radiative recombination *see* Radiative recombination continuum
- Continuum state 42, 49, 54
- Coronae 6, 60, 61
 - solar 6
- Corrections, relativistic 30
- COS-B 89, 91
- Cosmic abundances *see* Abundance of element
- Cosmic microwave background 214, 222
- Cosmic rays 84, 89, 93, 95, 118, 145
- Coulomb collisions 227
- Coulomb law 10
- Coulomb potential 27, 29, 30, 33, 39, 46
 - screened 31, 39
- Coulomb-Born method 46
- Coupling
 - intermediate 37–40
 - jj 37–40
 - LS 37–40
 - Russell-Saunders 37
- Crab Nebula 95, 154, 189
- Crab pulsar 88
- Critical angle (reflection) 114
- Cross section 21, 100
 - atomic 113
 - Beer's law *see* Beer's law
 - bound-bound transition 21
 - collisional excitation 47, 48, 56
 - collisional ionization 50
 - Compton 104, 201, 270, 271
 - differential
 - bound-free absorption 21
 - photoionization 21
 - scattering 44
 - Klein-Nishina 104–107, 178, 202, 270
 - Pair production
 - Bethe-Heitler 116
 - photoelectric 63
 - photoelectric absorption 100, 102
 - photoionization 52, 250
 - radiative recombination 51–53
 - scattering 17, 206
 - scattering by atomic electrons 106
 - coherent 108
 - incoherent 107
 - Thomson 102, 105, 214, 229, 232, 250
 - Thomson exchange *see* Cross section, collisional ionization
- Cross-correlation function 156
- Crystal diffraction lens 182–191
- Crystals
 - mosaic 110, 185
 - perfect 110
- Curie, P. 83
- CYCLONE 158
- Cyclotron frequency 90
- Cyclotron lines 90, 158, 166
- Cyg X-1 272
- Cygnus region 88, 97
- Cygnus X-1 238
- Dark matter 9
- Darwin model 110, 187
- Darwin term 30
- Darwin width 110
- Debye-Waller factor 110
- Decay rate
 - spontaneous emission 24

- Decoding 156–157, 161–163
- Deconvolution 171–172
- Deexcitation
 - collisional 42, 49
 - rate 59
 - rate coefficient 49
- Degeneracy 33, 48, 49, 53, 56, 58
- Degenerate levels 38
- Degenerate states 27
- Density 5, 60, 72
- Detailed balance 41, 49, 52, 55
- Detector coincidence 172
- Detector response 117, 119–121
- Detectors 121–149
 - CCD 4
 - efficiency 119
 - gas-filled counters 122–133
 - Geiger-Müller counters 84, 123–124
 - ion chambers 128–131
 - ionization chambers 123
 - microstrip gas counters 127
 - multi-wire proportional counters 126
 - position-sensitive proportional counters 126–128
 - proportional counters 123–126
 - time projection chambers 131–133
 - Germanium 86, 153, 158, 166, 178, 189, 190
 - phonon 122
 - scintillation 121, 154, 175
 - scintillators 121, 122, 133–143, 145, 153
 - inorganic 101, 136–140, 147
 - organic 133–136
 - semiconductor 122, 143–149
 - junction 144
 - CdTe 147–149
 - Germanium 144–147
 - semiconductor diode 121
 - solid state 101
 - strip 178–180
- Diagnostics
 - density 72, 74, 78
 - line 70–80
 - temperature 72, 74, 75, 79
- Dielectric constant 112
- Dielectronic capture 43
 - rate coefficient 55–56
- Diffraction
 - Bragg 182–183
 - crystal 151
 - efficiency, crystals 186
 - Laue 111–112, 182–185
- Diffraction coefficient 111–112
- Dipole moment
 - electric 13
 - magnetic 14, 23, 41
- Dirac equation 30
- Distorted wave approach 47
- Dithering 163, 167
- Doppler broadening 17
- Doppler effect 201, 238, 244, 256, 259
- Doppler frequency shift 228
- Drift region (detectors) 127

- Eddington critical accretion rate 247
- Eddington flux 276
- Eddington luminosity 206–209, 252, 267, 268, 275
- Efficiency
 - full energy peak 119
- EGRET 91
- EINSTEIN 95
- Electric dipole 22
- Electric dipole forbidden *see* Transition, forbidden
- Electric dipole term (E1) 13
- Electric dipole transition *see* Transition, electric dipole
- Electric field 10–11
- Electric quadrupole term (E2) 14, 23, 29
- Electric quadrupole transition *see* Transition, electric quadrupole
- Electron escape 117
- Electron exchange 33–35, 74
- Electron leakage 118
- Electron radius, classical 102, 113, 116
- Electron shielding 35
- Electron-electron interaction 46
- Electron-electron repulsion 31, 33, 36, 38
- Electron-electron repulsion energy 36
- Electron-hole pair 136, 142–148

- Electron-ion collisions *see* Collisions, electron-ion
- Emission 20
 - induced 21
 - spontaneous 16, 23–24
 - stimulated 52
- Emissivity 59
 - line 61
- Encoding 156, 161
- Energy eigenstates 18
- Energy resolution 93, 120–121, 126, 128–132, 135, 138–140, 145–146, 148, 166, 174, 177
- Energy transfer 48
- Equilibrium 57–70
 - coronal 60–62
 - photoionization 62
 - thermodynamic 41, 49, 52, 55
 - local 58–60
- Event circle 170
 - reduced 177
- Exchange energy 36
- Excitation
 - collisional 42
 - coefficient 61
 - effective rate coefficient 61
 - rate coefficient 48
- EXPLORER 1 89
- EXPLORER 11 89, 91
- extinction depth, primary 111
- extinction, primary 111
- Fano factor 121, 126, 128, 146
- Fermion 32, 35
- Field length 68
- Fine-structure constant 29, 102, 116
- Flares 6, 61
- Fluorescence 78, 117, 133–134, 254, 261, 264–269
- Fluorescence, prompt 133
- Forbidden transition *see* Transition, forbidden
- Form factor
 - atomic *see* Scattering factor, atomic
- Formation temperature *see* Temperature of formation
- Frank–Condon principle 134
- FRESNEL 192
- Fresnel lens 181, 191–192
 - phase- 191–192
- Fresnel zone plate 151, 160–161, 191–192
- Frisch grid 148
- Full energy peak (spectral) 117
- Full energy peak efficiency *see* Efficiency, full energy peak
- G ratio 74
- Galactic bulge 97
- Galactic center 86–87, 89, 190, 263–269
- Galactic disk 97
- Galactic plane 152
- Galaxy clusters 5, 9, 60, 214
- Galaxy, elliptical 5
- Gamma rays
 - cosmic diffuse 85, 89
 - polarization 168, 172
- Gamma-ray bursts 85, 88, 89, 209
 - afterglows 85, 89
- Gamma-ray lines 86
- Gamma-ray sources
 - angular scales 96
 - line fluxes 96
 - number of 93, 94
- Gas gain (detectors) 125
- Gas multiplication (detectors) 124–128
- Gauge 10
 - Lorentz 10, 19
- Gaunt factor 48, 51, 53
- Geiger–Müller tubes *see* Detectors, gas-filled counters, Geiger–Müller counters
- Geometric optics 149–168
- Ghost mirror image 157
- Ginga 90
- Graded shield 119
- GRANAT 87, 89, 91, 137, 151, 165
- Grand unified model (AGN) 8
- Grazing angle, critical 180–182
- Grazing incidence 180–182
- Great annihilator *see* 1E 1740–29
- Grids 154
- H-like ions 71–74
- Hamiltonian 17–18, 25

- central potential 26, 30
- He-like atoms 33
- multi-electron atom 37
- perturbing 25, 31, 55
- radiation- 19–20
- spherically symmetric 26
- zeroth-order 25, 31
- Hartree-Fock equations 36
- Hartree-Fock method 36
- He-like transition *see* Transition, He-like
- HEAO-3 87, 89, 95, 152–153
- Helium-like atoms 33–35
- Her X-1 90
- HERO 181
- Hess, V. 89, 90
- HESSI 151, 158
- Hole (detectors) 136
- Holography 160
- Hydrogen-like ions 25–29
- Hydrogenic approximation 51
- IBIS (INTEGRAL) 138, 142, 148–149, 151, 167–169
- Identical particles 32
- Image prior 172
- Impact parameter 48, 50
- Index of refraction *see* Refractive index
- INTEGRAL 97, 151
- Intercombination transition *see* Transition, intercombination
- Interference 182
- Interference, destructive 110
- Interstellar medium 54
- Intrachuster gas 9, 60
- Ion pairs (detectors) 122, 128
- Ion saturation (detectors) 123
- Ionization
 - collisional 42, 49–51, 60
 - rate coefficient 50, 61
- Ionization balance 58
- Ionization energy 128
- Ionization parameter 64, 67
- Ionization potential 50, 51, 56, 64, 75
- ISGRI (INTEGRAL) 138, 148–150, 167–169
- JEM-X (INTEGRAL) 128, 151
- Jet
 - relativistic 8
- K-shell ions 38
- K-shell transition *see* Transition, K-shell, 70
- Kirchhoff's law 229
- Kirkpatrick/Baez optics 181
- Klein-Nishina cross section *see* Cross section, Klein-Nishina
- Klein-Nishina limit 206
- Kompaneets equation 217–223, 234, 237, 239–240
- L-shell ions 38
- L-shell transition *see* Transition, L-shell
- Lagrange multipliers 36
- Lagrange's equation 19
- Lagrangian 19
- Laguerre polynomial 27
- Lambert law 253
- Larmor formula 14, 15
- Laue lens 181, 183–191
- Leakage current 143–144
- Legendre polynomial 27
- LHC (Large Hadron Collider, CERN) 140
- Line
 - absorption 21
 - emission 21
- Line power 61
- Line profile
 - Lorentzian 15, 17
- Line width
 - natural 16
- Lobster eye telescope 181
- Local thermodynamic equilibrium *see* Equilibrium, thermodynamic, local
- Lorentz force 10, 19
- LXeGRIT 132, 179–180
- Mössbauer effect 99
- Magnetic dipole term (M1) 14, 23, 29
- Magnetic dipole transition *see* Transition, magnetic dipole
- Magnetic field 6, 10–11

- Mass attenuation *see* attenuation, mass
- Mass thickness 101
- Matrix element 17, 22, 35, 37, 40, 49, 54, 55
- MAX 189–191
- Maximum entropy method 172
- Maxwell's equations 10, 11
- Maxwell-Boltzmann distribution 58
- Maxwell-Boltzmann factor 58
- Maxwellian velocity distribution 49, 50, 52, 55, 58
- Mean free path 95, 101, 203
 - gamma rays 121
- MEGA 176–178
- Microquasar 87, 191
- Miller indices 109, 184
- Milne relation 53
- MIR station 130
- MISO telescope 169
- Modular Detection Units 149
- Modulating Aperture Systems 149–168
- Modulation
 - spatial 151
 - temporal 151
- Momentum transfer 48
- Momentum, canonical 19
- Mosaic width 111, 185
- Mosaicity 185
- Multi-configuration calculations 40
- Multilayer mirror 181–182
 - extremely broad band 182
 - uniform period multilayer 182
- Multipole expansion 12–15, 40, 47
 - quantum- 22–23
- Narrow lines 166
- NCT 178
- Neutron capture 86
- Neutron star 7, 90, 158, 272–278
- NGC 1068 66, 68, 75
- Noise, statistical (detectors) 120
- Nova Muscae 88
- Novae 97
- Nucleosynthesis 84, 97, 177
- Occultation transform imaging 152–154
- Opacity 59
- Operator
 - momentum 22
 - spin 41
- Optical depth 5, 247
 - effective 231
- Optically thick 59
- Optically thin 60, 64
- Orthopositronium 86
- Oscillator strength 22–24, 29, 35, 75
 - dipole absorption 48
- OSO-3 89, 91
- OSO-7 86
- OSSE 87, 88, 90, 92, 151, 153
- Pair plasma 87
- Pair production 95, 98–100, 114–116, 118
- Parametric potential method 36
- Parity 40
- Parseval's theorem 12
- Partial wave expansion 47
- Partition function 58
- Passive material 118
- Pauli exclusion principle 32, 35
- Penning effect 126
- Penning mixtures (detectors) 125
- Photoabsorption 60, 249, 251
- Photoabsorption rate 52
- Photoabsorption spectrum 54
- Photocathode 141
- Photodiode 122, 133, 141–143
 - avalanche 143
 - PIN 138, 142–143
- Photoelectric absorption *see* Absorption, photoelectric
- Photoelectric effect 95, 99, 102–103, 141, 150
 - nuclear 99
- Photoelectric heating 63, 66
- Photoexcitation 16
- Photoionization *see also* Absorption,
 - bound-free, 43, 250
 - continuum 54
 - inner shell 78
 - rate 69
- Photomultiplier 141–142
- Photomultiplier tube 122, 132, 133, 135, 137, 140, 141, 153, 165, 166

- Photon 17, 18
- Photon flux 20
- Photopeak (spectral) 117
- PICsIT (INTEGRAL) 138–139, 142, 149, 167–169
- Pinhole camera 150, 159–161
- Planck spectrum 199, 218, 230, 238
- Planck's constant 17
- Plasma frequency 113
- Point spread function 171
- Polarization
 - dielectric 112
- Polarization vector 20
- Positron 86
- Potential
 - binding of atom 19
- Power-law spectrum 237, 247, 248, 254, 273
- Poynting vector 12, 13, 20
- Procyon 75
- Pulsars 88, 247
- Pulse height analysis 123
- Pulse height spectrum
 - differential 119
- Pulse shape discrimination 147, 173, 176

- Quantum efficiency 142–143
- Quantum indistinguishability 31–32
- Quantum number 32, 38–40
 - principal 27, 32, 57
- Quantum optics 151, 168–180
- Quantum radiation theory 17–24
- Quantum states 17
- Quasar 8, 199, 254
 - type 1 8
 - type 2 8
- Quasi-periodic oscillations 274
- Quench gas (detectors) 125

- R ratio 75
- Racah algebra 37
- Radiation belts 89
- Radiation field 59, 60, 62
- Radiation force 204–206, 227–228
- Radiation reaction 15
- Radiative cascade 64, 74, 75
- Radiative decay
 - spontaneous 59, 60
- Radiative decay rate
 - classical 16
- Radiative recombination *see* Recombination, radiative
- Radiative recombination continuum 65–67
- Radiative transfer 58–59
- Radioactivity
 - cosmic 87
- Raman lines 264
- Ranger 85
- Ranger 3 89
- Ranger 5 89
- Rayleigh scattering *see* Scattering, Rayleigh
- RE J0317-853 86
- Recoil effect 200, 223, 244, 246, 251, 259, 262
- Recoil electron 103, 175, 177, 178
- Recoil energy 251
- Recombination 64
 - 3-body 42, 51
 - coefficient 63
 - dielectronic 53–57, 60, 79
 - rate 56
 - radiative 43, 51–53, 60, 69
 - rate coefficient 52
 - rate coefficient 61
- Reflection 112–114, 151
 - external 113
 - X-ray 250
- Refraction 112–114
- Refractive index 112, 133, 180–182
- Refractive index decrement 112–114
- Refractive optics 114
- Relativistic corrections *see* Corrections, relativistic
- Repulsion energy, electrostatic 35
- Response matrix 119
- Retarded time 11, 13
- Richardson–Lucy reconstruction 172
- Ritz variational method 25
- Ritz variational principle 36
- Roche lobe 7
 - overflow 7
- Rocket flights 6, 84
- Roentgen, G. 89
- ROSAT 90, 93, 95

- Rutherford, E. 83, 89
- Rydberg 50
- Rydberg constant 26
- Rydberg level 57, 58

- Saha equation 53, 56, 58
- SAS-2 89, 91
- Satellite lines 79
- Scalar potential 10, 19
- Scattering
 - by atomic electrons 98, 99, 105–112
 - coherent 106–108
 - incoherent 106–107
 - by free electrons 103
 - coherent 99
 - incoherent 99
 - coherent 105, 151
 - Compton 95, 98, 99, 103–105, 118, 169, 199, 230, 235, 254, 257, 259–260
 - Compton, multiple 118
 - Delbrück 99
 - elastic 105, 200
 - in crystal 108
 - incoherent 151
 - induced Compton 213, 223
 - inelastic 44
 - inverse Compton 95, 201, 221, 226, 235
 - mosaic crystal 111
 - mosaic crystals 110–112
 - nuclear 99
 - coherent 99
 - incoherent 99
 - Raman 257, 259, 260
 - Rayleigh 98, 99, 202, 257–258, 260, 261
 - Thomson 99, 105, 199–201, 235, 243, 247
- Scattering amplitude 44
- Scattering factor
 - atomic 108–113
- Scattering function
 - incoherent 106
- Schroedinger equation 17, 25, 30–32, 45
- Schwarzschild radius 208, 242, 247
- Scintillation 166
- Scintillation efficiency 133

- Scintillator (aperture) 151
- Scintillators *see* Detectors, scintillators
- Sco X-1 4, 89
- Secondary radiation (detectors) 117
- Selection rules 23, 40–41, 47
- Self-consistent field method 36
- Semiconductor (aperture) 151
- Semiconductor diode *see* Detectors, semiconductor diode
- Seyfert 1 galaxy 8
- Seyfert 2 galaxy 8, 66
- Seyfert galaxies 254
- Sgr A* 263, 264, 268
- Sgr B2 266–268
- Shadowgram 159–161
- Shell (atomic) 38
 - closed 39
- Shield leakage 169
- Shocks
 - stellar winds 6
 - supernovae 7
- SIGMA 87–89, 91–92, 137, 151, 165
- Single configuration approximation 32
- Singlet state 33–35, 44
- Slater determinant 32, 35–37, 40
- Small Magellanic Cloud 72
- SMM 86–89
- Snell's law 114
- Solar flares 85–86, 88, 89, 93, 137
- Source function 59
- South Atlantic Anomaly 130
- Spatial Modulation 158–168
- Spectroscopic notation 38–40
- Spectrum
 - photon 119
- Spherical harmonics 26
- SPI (INTEGRAL) 87, 97, 140, 146, 151, 165–167
- Spin 33, 36
 - electron 27
 - intrinsic 23
 - orbital 23
 - photon 17
- Spin state 32, 44
- Spin-orbit interaction 41
- Spin-orbit term 30, 38

- Spontaneous emission *see* Emission, spontaneous
- Spring constant 15
- Stars
 - early-type 6
 - late-type 6
 - winds 6, 7
- Stellar coronae *see* Coronae
- Stirling machine 145, 147
- Structure factor
 - atomic 106
 - crystal 109
 - geometric 110
 - electronic 106
- Sum rule *see* Thomas-Reiche-Kuhn sum rule
- Sun 6, 86, 152, 158
- Supernova remnants 5, 6, 60, 158
 - Cas A 72, 88–90
 - RX J0852-46, “Vela Junior” 88, 90
 - SNR 1E0102-72.3 72
- Supernovae 97, 190, 199
 - SN 1987A 87–89, 97, 269
 - SN 1991T 88, 90, 96, 97
- SWIFT 151

- Temperature 57, 61, 72
- Temperature of formation 62
- Temporal Modulation 152–158
- TGRS 87
- Thermal instability in photoionized plasmas 67–70
- Thermodynamic equilibrium *see* Equilibrium, thermodynamic
- Thomas-Fermi potential 35
- Thomas-Reiche-Kuhn sum rule 22
- Thomson depth 241, 242
- Thomson limit 204
- Thomson scattering *see* Scattering, Thomson
- Three-body recombination *see* Recombination, 3-body
- TIGRE 178
- Time of flight 173, 176
- Time reversal 41
- Townsend avalanche (detectors) 124
- Townsend coefficient 125
- Transition
 - allowed 40
 - bound-bound 21–22
 - electric dipole 29, 40, 64, 74
 - electric quadrupole 41
 - forbidden 23, 40, 47, 74
 - strictly 23
 - He-like 74–75
 - intercombination 74
 - K-shell 5, 71
 - iron 78–80
 - K-shell transition 29
 - L-shell 5, 29, 70–71
 - iron 75–78
 - Lyman α 28
 - Lyman series 26, 71–74
 - magnetic dipole 41
 - radiative 40–41
 - resonance 74
- Transition rate 18, 20
 - bound-bound 21
 - bound-free absorption 20
- Transmission function 161
- Transmission function (grid) 155, 156
- Triplet state 33–35, 44
- Two-photon decay 23

- UHURU 89
- Uhuru 272
- Ultraviolet radiation field 75
- Uncertainty principle 20, 54
- Uniformly redundant array (URA)
 - 163, 165, 168
 - hexagonal 166

- V2 rocket 84–85
- Valence band 136–137, 143
- Valence shell 35
- Vector potential 10, 19
- Vela region 88, 97
- VELA satellites 85, 89
- Very Large Array (VLA) 87
- Veto shield 168
- Vignetting 164
- Villard, P. 83–85, 89
- Virial temperature 5
- Virial theorem 29
- Volume emission measure 62

- W-value (detectors) 128
- Wave 11

- monochromatic 12, 19, 21
- plane 12, 44
- spherical 44
- Wave function
 - antisymmetric 32, 33, 40, 44
 - continuum 44, 46, 55
 - determinantal 32
 - electrons 25
 - single electron 37, 40
 - single particle 31
 - zeroth-order 31
- Wave optics 151, 180–192
- White dwarfs 7, 54, 86
- Whittaker function 246
- Whittaker's equation 237
- Wien spectrum 233, 238, 246
- Wolter-I telescope 181
- Workfunction (detectors) 141
- X-ray binaries 7, 62, 190, 214
 - high-mass (HMXB) 7
 - low-mass (LMXB) 7, 252
- X-ray binary 5
- X-ray escape peak (detectors) 118
- XEUS 80
- XMM-Newton 4, 66, 68, 72, 80
- Z-value (detectors) 133, 136

Printing: Strauss GmbH, Mörlenbach
Binding: Schäffer, Grünstadt