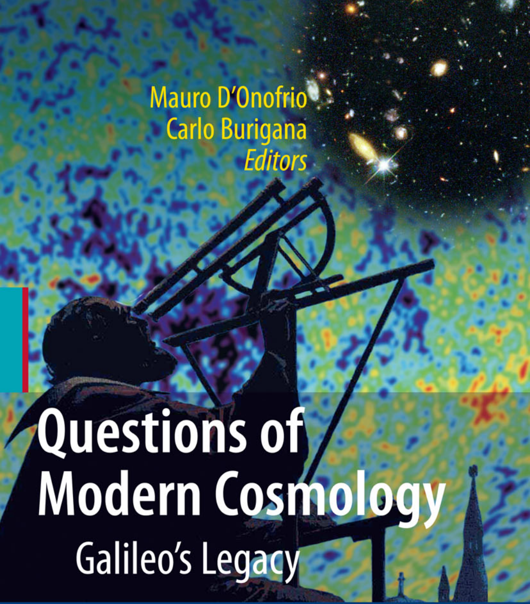

Mauro D'Onofrio
Carlo Burigana
Editors

Questions of Modern Cosmology

Galileo's Legacy

Springer

Questions of Modern Cosmology

Mauro D'Onofrio • Carlo Burigana
Editors

Questions of Modern Cosmology

Galileo's Legacy

Springer

Editors
Mauro D'Onofrio
Università di Padova
Dip. to. di Astronomia
Vicolo Osservatorio, 3
35122 Padova
Italy
mauro.donofrio@unipd.it

Carlo Burigana
INAF Bologna
Ist. di Astrofisica Spaziale e
Fisica Cosmica (IASF)
Via Gobetti, 101
40129 Bologna
Italy
burigana@iasfbo.inaf.it

ISBN 978-3-642-00791-0 e-ISBN 978-3-642-00792-7
DOI 10.1007/978-3-642-00792-7
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009928103

Cover design: WMXDesign GmbH

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

In memory of the men and women who spent their lives for a science free from prejudices, conditioning, and interferences.

Foreword

This book represents a modern tribute to Galileo. The way in which the authors, Mauro D'Onofrio and Carlo Burigana, illustrate the present understanding of the Universe through interviews in which physicists, astrophysicists, and cosmologists highlight their pros and cons on the current models takes our mind immediately to Galileo's opera "Dialogue over the two world maxima systems". In the "Dialogue", three characters are philosophizing on the world "Tolemaic-Aristotelic" system against the innovative "Copernican" system, Salviati, and Simplicio – both scientists – and the noble Venetian Sagredo. The latter represents the moderate reader, while Salviati is the spokesman of the Galileian Copernican ideas and Simplicio is the rigid defender of the traditional and dogmatic "scholastic" doctrine. It is not the case that Simplicio means "simple" as well as "silly". But Simplicio is not at all silly, while is trivial the incapacity of the "scholastic" model to open itself to novelties. "Questo modo di filosofare tende alla sovversione di tutta la filosofia naturale, ed al disordinare e mettere in conquasso il cielo e la Terra e tutto l'universo" (This way of philosophizing subverts the whole of natural philosophy and puts out of order the sky, the earth, and the entire universe) says Simplicio in the first day of the Dialogue.

After four centuries, "Questions of Modern Cosmology – Galileo's Legacy" has many more characters, all well known and respected scientists, representing their views on the cosmological and astrophysical observations, their interpretations, the cosmological paradigms, and possible alternatives. Is there among them any modern Simplicio or are we all representing Simplicio today? It is up to you reader, as a contemporary Venetian patrician Sagredo, to discover it and to be ready to accept any new idea; even the one that the Big Bang is not the beginning of everything but just one among many events.

But looking more deeply at the Copernican model of Galileo's time, it was a brilliant, but only a phenomenological description of the Solar System. It is after the discoveries of Kepler's laws and the interpretation of Newton that the description of the Solar System had a full comprehensive physical explanation. In terms of analogies between Galileo and today, as is evident from the pages of D'Onofrio's and Burigana's book, the current ΛCDM cosmological model is sufficiently satisfactory from the phenomenological point of view, but the full physical description is still missing.

It may be argued that the most important scientific discoveries in cosmology are referable to observations and interpretations by the methods of astrophysics and fundamental physics. In this book, the balance between observational cosmology and the interpretation of the data is well made in Chaps. 2 and 3. Furthermore, the many outstanding contributors interviewed present a complete and exhaustive panorama between the present scenario and the future challenges (Chaps. 2, 3, and 5).

The authors have in mind that the scientific method and the way to do with science is not unique and even the "alternative" ideas are presented without any obscurantistic bias (Chap. 4 and Concluding remarks).

This book is aimed at students and at colleagues in astrophysics, cosmology, and physics, and also at any reader interested in this intriguing field of science.

The way to unfold the most recent discoveries and theories in cosmology, via a number of different voices, makes the book a wonderful story and switches the curiosity of the reader, pushing him to read page after page to discover the wonderful picture of the whole Universe.

Bologna,
May 2009

Nazzareno Mandolesi

Preface

This book is dedicated to *Galileo Galilei*, the Italian man who 400 years ago first pointed a telescope toward the luminous objects visible in the night sky, realizing the famous astronomical discoveries (published in the "Sidereus Nuncius"), which within few years prompted the first scientific revolution, today known as the "Copernican Revolution". A new vision of the Universe was emerging at that time in the garden of Galileo in Padova. His life and scientific discoveries are so famous and important that during 2009, declared by United Nations the year of Astronomy, astronomers from all countries will celebrate his life through meetings, congresses, workshops, lectures, seminars, and a lot of other activities.

Here, in Italy, the astronomical community is working hard for the 2009 celebrations of the International Year of Astronomy. In this context, at the beginning of 2007, we proposed this book to the Springer publisher to give our personal tribute in honor of Galileo. To us, the idea of reviving the Galileo "spirit", by engaging in a discussion with many contemporary scientists about the state-of-the-art of the present day cosmology and the role and character of this science today in our society, was very attractive. For this reason, we realized a series of interviews with many astronomers and physicists from all over the world, with the aim of summarizing the most important and significant advances made by cosmology over the past century and at the beginning of the new millennium. We have tried to interpret, as well as possible, the fighting spirit of Galileo, opening the debate to alternative ideas and strongly favoring the empirical approach to the scientific problems.

We are living in a very strange scientific epoch: theorists are pushing themselves to the limits of pure speculation in the search of a "Theory of Everything", able to reconcile quantum mechanics, general relativity, cosmology, and particle physics. At the same time, experimental and observational projects are getting bigger and better at demonstrating the existence of the theoretically predicted dark matter and dark energy, which would constitute up to $\sim$95% of the energy density, understanding the very early stages of the Universe and reconstructing its evolution. In this context, the parallelism/contrast with the situation at Galileo's time is intriguing. At his epoch, astronomical observations prompted the Copernican Revolution through the discoveries of Jupiter's satellites, of the phases of Venus, and so on. Today, it is the lack of a firm observational or experimental identification of nonordinary types of matter that may trigger a profound shift of scientific paradigm. It should be recognized, in fact, that we still have to catch the essence of the unknown forms of energy and matter that fill the Universe but escape our theories and experiments, despite the great efforts of recent space missions, extensive observational programs in all frequency bands, and dedicated experiments of fundamental and astroparticle physics, and, in parallel, the great theoretical work done to explain them in a unified vision.

In general, in the simplest currently accepted standard model(s) of cosmology, there are relatively few free parameters required to fit the whole set of available data. They are linked to the density of the different forms of matter and energy, to the expansion rate of the Universe, to the kinds and statistics of primordial perturbations, and to the physical processes that occurred in the early stages of structure formation. The model works very well, but no one knows why some of these parameters have the values they do.

In addition to this, it is our aim here to discuss another important parallelism between Galileo's epoch and the present, that is enclosed in the following questions. Galileo was condemned for his ideas and several people fought against the progress of the new scientific vision of the Universe. Are we living with the same things today? How much space is there for alternative ideas that may prompt a new scientific revolution? Are we close to a deeper understanding of the current paradigm or to a new scientific revolution, or rather, are physics and astrophysics going through a profound crisis?

This book is written, as a collection of interviews, not only for all those people with a solid scientific background in cosmology and particle physics, but also for astronomers and physicists not necessarily expert in these fields of research. Here can be found the tentative answers of the scientific community to the challenges posed by cosmology to the standard physical paradigms of the last century. We hope also that it will be of some interest to readers attracted in particular by the way in which our science evolves and reflects on its principles, methods, and self-organization.

Certainly, the selection of the topics addressed by the interviews and the formulation of the posed questions reflect in same way our personal views on the most important aspects of modern cosmology. Our choice for this project, however, was that of avoiding as much as possible any comment to the single interviews, merits and responsibilities of which entirely belong to the colleagues who kindly accepted them, leaving to ourselves some general remarks for the final chapter.

We hope the reader will be satisfied by the discussion started here.

We thank all the colleagues who have kindly accepted to be involved in this project. We greatly appreciate not only their specific answers to our questions, which created a panoramic view of current cosmology, but also their efforts to enter deeply into the spirit of this book, critically discussing the various scientific aspects they addressed.

We also thank Matteo Genghini and Enrico Franceschi for their informatic assistance that helped us in the exchange of the book material during the various stages of interaction with our colleagues. We particularly thank Simone Zaggia for his help in solving some Latex problems, Marcel Clemens for improving the quality of English, Paola Marziani for her editorial suggestions, and Gabriele Umbriaco for his graphic design of the front cover of the book.

We also thank the women in our lives, Simonetta and Marilena, who patiently sustained us throughout the development of this work, and the little Margherita who wants back her daddy Mauro.

Finally, we acknowledge our Springer Editor, Ramon Khanna, for having believed in this project from its beginning.

Venezia,
May 2009

Mauro D'Onofrio
Carlo Burigana

Contents

List of Authors

Amedeo Balbi Dipartimento di Fisica, Università di Roma Tor Vergata,
Via della ricerca scientifica, 00133 Roma, Italy,
e-mail: amedeo.balbi@roma2.infn.it
Author's contribution in Sect. 3.7.

Matthias Bartelmann Institut für Theoretische Astrophysik,
Universität Heidelberg, Albert-Überle-Straße 2, 69120 Heidelberg, Germany,
e-mail: mbartelmann@ita.uni-heidelberg.de
Author's contribution in Sects. 2.14 and 5.9.

Charles L. Bennett Department of Physics and Astronomy,
The Johns Hopkins University, 3400 N. Charles Street, Baltimore,
MD 21218-2686, USA, e-mail: cbennett@jhu.edu
Author's contribution in Sects. 2.8.2, 3.7.1, and 5.4.

Martin Bucher Laboratoire de Physique Théorique, Université Paris-Sud,
91405 Orsay, France, e-mail: bucher@th.u-psud.fr
Author's contribution in Sects. 3.4 and 5.3.5.

Massimo Capaccioli INAF - Osservatorio Astronomico di Capodimonte,
VST Center, Via Moiariello 16, 80131 Napoli, Italy, e-mail: mc@na.infn.it
Author's contribution in Sect. 5.4.1.

Salvatore Capozziello Department of Physics, Universitá di Napoli Federico II,
Via Cintia, 80126 Napoli, Italy, e-mail: capozzie@na.infn.it
Author's contribution in Sect. 4.7.2.

Cesare Chiosi Dipartimento di Astronomia, Universià of Padova,
Vicolo Osservatorio 3, 35122 Padova, Italy, e-mail: cesare.chiosi@unipd.it
Author's contribution in Sects. 2.15 and Sects. 4.12.1

Peter Coles School of Physics and Astronomy, Cardiff University,
Queen's Buildings, 5 The Parade, Cardiff CF24 3AA, UK,
e-mail: Peter.Coles@astro.cf.ac.uk
Author's contribution in Sects. 3.5.2, 3.5.4, and 3.7.2.

Alan Dressler Observatories of the Carnegie Institution of Washington, Pasadena, CA 91101, USA, e-mail: dressler@ociw.edu
Author's contribution in, Sect. 2.11.

Ruth Durrer Université de Genève, Departement de Physique Théorique, 24 quai E. Ansermet, 1211 Genève 4, Switzerland, e-mail: Ruth.Durrer@physics.unige.ch
Author's contribution in, Sect. 3.3 and 5.10.

Isabella Gioia INAF Istituto di Radioastronomia, Via P. Gobetti 101, 40129 Bologna, Italy, e-mail: gioia@ira.inaf.it
Author's contribution in Sects. 2.12 and 5.8.

Günther Hasinger Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße, Postfach 1312, 85741 Garching, Germany, e-mail: ghasinger@mpe.mpg.de
Author's contribution in Sects. 2.9 and 5.8.

Charles R. Lawrence Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, USA, e-mail: Charles.R.Lawrence@jpl.nasa.gov
Author's contribution in Sect. 5.3.3.

Malcolm Longair Cavendish Laboratory, Room 918 Rutherford Building, JJ Thomson Avenue, Cambridge CB3 0HE, UK, e-mail: m.s.longair@mrao.cam.ac.uk
Author's contribution in Sects. 4.3, 4.4, and 4.5.1.

Giuseppe Longo Department of Physical Sciences, University Federico II, via Cintia 9, 80126 Napoli, Italy, e-mail: longo@na.infn.it
Author's contribution in Sect. 5.4.1.

Juan Francisco Macias-Perez Laboratoire de Physique Subatomique et de Cosmologie, 53, rue des Martyrs, 38026 Grenoble Cedex, France, e-mail: macias@lpsc.in2p3.fr
Author's contribution in Sects. 2.8.3–2.8.5 and 5.3.4.

Piero Madau University of California, Santa Cruz, CA 95064, USA, e-mail: pmadau@ucolick.org
Author's contribution in Sects. 2.10 and 5.5

Philip D. Mannheim Department of Physics, University of Connecticut, Storrs, CT 06269, USA, e-mail: mannheim@uconnvm.uconn.edu
Author's contribution in Sect. 4.7.5.

Paola Marziani INAF, Osservatorio Astronomico di Padova, Vicolo Osservatorio 5, 35122 Padova, Italy, e-mail: paola.marziani@oapd.inaf.it
Author's contribution in Sect. 2.5 and 4.12.1.

John C. Mather Astrophysics Science Division, NASA/GSFC, Code 665, Observational Cosmology, Greenbelt, MD 20771, USA, e-mail: john.c.mather@nasa.gov
Author's contribution in Sects. 2.8.1, 3.5.1, and 5.3.1.

Francesca Matteucci Università degli Studi di Trieste,
via G.B. Tiepolo 11, 34131 Trieste, Italy, e-mail: matteucci@oats.inaf.it
Author's contribution in Sect. 2.4.

Moti Milgrom Department of Condensed Matter Physics, Weizmann Institute,
Rehovot 76100, Israel, e-mail: moti.milgrom@weizmann.ac.il
Author's contribution in Sects. 4.7 and 4.7.1.

Pavel Naselsky Niels Bohr Institute,
Blegdamsvej 17, 2100 Copenhagen, Denmark, e-mail: naselsky@nbi.dk
Author's contribution in Sect. 3.6.

Keith Olive William I, Fine Theoretical Physics Institute,
School of Physics and Astronomy, University of Minnesota, 116 Church Street SE,
Minneapolis, MN 55455, USA, e-mail: olive@physics.umn.edu
Author's contribution in Sects. 2.7.1 and 2.7.2 and 4.9.

Thanu Padmanabhan IUCAA, Post Bag 4, Pune University Campus,
Ganeshkhind, Pune 411 007, India, e-mail: paddy@iucaa.ernet.in
Author's contribution in Sects. 4.5.2, 4.7 and 4.12.1.

John Peacock Institute for Astronomy, Royal Observatory,
Blackford Hill, Edinburgh EH9 3HJ, UK, e-mail: jap@roe.ac.uk
Author's contribution in Sects. 2.2, 4.11, 4.12.1, and Sect. 5.4.

Francesca Perrotta INAF, Osservatorio Astronomico di Trieste,
Via G.B. Tiepolo 11, 34131 Trieste, Italy, e-mail: perrotta@oats.inaf.it
Author's contribution in Sects. 3.2 and 4.6.

Luisa Pigatto INAF, Osservatorio Astronomico di Padova,
Vicolo Osservatorio 5, 35122 Padova, Italy,
e-mail: luisa.pigatto@oapd.inaf.it
Author's contribution in Sect. 4.2.

Lucia A. Popa ISS, Institute for Space Sciences,
Bucharest-Magurele 76900, Romania, e-mail: lpopa@venus.nipne.ro
Author's contribution in Sect. 3.8 and 5.11.

Rafael Rebolo Instituto de Astrofisica de Canarias,
38200 La Laguna, Tenerife, Spain, e-mail: rrl@iac.es
Author's contribution in Sects. 4.13 and 5.3.2.

Wolfang Reich Max-Planck-Institut für Radioastronomie,
Auf dem Hügel 69, 53121 Bonn, Germany, e-mail: wreich@mpifr-bonn.mpg.de
Author's contribution in Sects. 2.8.6 and 5.2.

Pierre Marie Robitaille Department of Radiology, The Ohio State University,
130 Means Hall, 1654 Upham Drive, Columbus, OH 43210, USA,
e-mail: robitaille.1@osu.edu
Author's contribution in Sect. 2.8.7.

Michael Rowan-Robinson Imperial College, Blackett Laboratory, Prince Consort Road, London SW7 2BZ, UK, e-mail: m.rrobinson@ic.ac.uk
Author's contribution in Sect. 2.16.

Luigi Secco Dipartimento di Astronomia, Università di Padova, Vicolo Osservatorio 3, 35122 Padova, Italy, e-mail: luigi.secco@unipd.it
Author's contribution in Sect. 4.10.

David N. Spergel Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544 USA, e-mail: Dns@astro.princeton.edu
Author's contribution in Sect. 3.5.3.

Gary Steigman Department of Physics, The Ohio State University, Columbus, OH 43210, USA, e-mail: steigman@mps.ohio-state.edu
Author's contribution in Sect. 2.7.

Matthias Steinmetz Astrophysical Institute of Potzdam, An der Sternwarte 16, 14482 Potsdam, Germany, e-mail: msteinmetz@aip.de
Author's contribution in Sect. 2.13.1 and 5.6.

Kandaswamy Subramanian Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, India, e-mail: kandu@iucaa.ernet.in
Author's contribution in Sect. 3.9.

Jack W. Sulentic Department of Physics and Astronomy, University of Alabama, Box 870324, Tuscaloosa AL 35487-0324, USA, e-mail: jsulenti@bama.ua.edu
Author's contribution in Sects. 2.6 and 4.12.2.

Gerard t'Hooft Spinoza Instituut, Leuvenlaan 4, Postbus 80.195, 3508 TD Utrecht, The Netherland, e-mail: G.tHooft@phys.uu.nl
Author's contribution in Sects. 4.8 and 4.12.2.

Massimo Turatto INAF, Osservatorio Astronomico di Catania, Via S. Sofia 78, 95123 Catania, Italy, e-mail: Massimo.Turatto@oact.inaf.it
Author's contribution in Sects. 2.3 and 5.7.

Nicola Vittorio Dipartimento di Fisica, Università di Roma Tor Vergata, Via della ricerca scientifica, 00133 Roma, Italy, e-mail: nicola.vittorio@roma2.infn.it
Author's contribution in Sect. 3.5.2.

Simon D.M. White Max Planck Institut für Astrophysik, Karl Schwarzschild Straße 1, 85740 Garching, Germany, e-mail: swhite@mpa-garching.mpg.de
Author's contribution in Sects. 2.13, 4.12.2, and 5.6.

Author's Acknowledgements

C. Bennett: I acknowledge the WMAP science team for their collective efforts from the WMAP concept, through the design, build, launch, mission operations, data analysis, and the public dissemination of the data.

S. Capozziello: I acknowledge Vincenzo Cardone, Sante Carloni, Arturo Stabile and Antonio Troisi for suggestions, comments and results achieved on the presented topics.

P. Coles: Most of the work I have described here was done in collaboration with others. I am particularly grateful to two former students of mine, Patrick Dineen and Lung-Yih Chiang, and to a current student Rockhee Sung, for permission to include work we did jointly together.

R. Durrer: I thank Martin Kunz for clarifying discussions and for helping me with Fig. 3.5 (in Chap. 3) and Mark Hindmarsh for a constructive conversation on the problem of string loop size. This work is supported by the Swiss National Science Foundation.

I. Gioia: I thank the hospitality of the Institute for Astronomy of the University of Hawaii where most of this work was completed. I enjoyed fruitful discussions with M. Branchesi, L. Feretti, F. Govoni, P. Henry, M. Murgia, B. Tully. I acknowledge partial financial support from contracts ASI-INAF I/023/05/0, 088/06/0 and DA-030.

C. Lawrence: The Planck mission has been developed by a large number of people. I have taken advantage of the work of too many to name individually, but would like to thank J.-L. Puget and N. Mandolesi, the HFI and LFI PIs, respectively, and J.-M. Lamarre and M. Bersanelli, the HFI and LFI Instrument Scientists, respectively, for their essential work on Planck, the many authors of the Planck Bluebook for a great compilation of the Planck science programme, from which I have borrowed freely, and Bill Jones, Dave Pearson, Ben Wandelt, and Martin White for their help with figures and many other things. Any and all mistakes are my responsibility alone.

G. Longo: Most of the work has been done in collaboration with the VO-Neural team and within the frameworks of the VOTech European Consortium and of a 2008 bi-lateral Italy–USA agreement sponsored by the Italian Bureau for Foreign Affairs.

J. Mather: The COBE team included the Science Working Group (SWG), over 100 supporting scientists and software engineers, and the total team list included over 1,600 people, mostly at Goddard Space Flight Center and at Ball Aerospace where the helium cryostat was made. The SWG membership and their roles (and current affiliations) are as follows: J.C. Mather, GSFC, Project Scientist and FIRAS Principal Investigator; M.G. Hauser, HST-ScI, DIRBE Principal Investigator; G.F. Smoot, UC Berkeley, DMR Principal Investigator; C.L. Bennett, Johns Hopkins University, DMR Deputy PI; N.W. Boggess, GSFC-ret, Deputy Proj. Scientist; E.S. Cheng, Conceptual Analytics, Deputy Proj. Scientist; E. Dwek, GSFC; S. Gulkis, JPL; M.A. Janssen, JPL; T. Kelsall, GSFC-ret, DIRBE Deputy PI; P.M. Lubin, UCSB; S.S. Meyer, University of Chicago; S.H. Moseley, GSFC; T.L. Murdock, Frontier Technology Inc.; R.A. Shafer, GSFC, FIRAS Deputy PI; R.F. Silverberg, GSFC; R. Weiss, MIT, Chairman of SWG; D.T. Wilkinson, Princeton (deceased); E.L. Wright, UCLA, Data Team Leader. The project manager was Roger Mattson (deceased) and his deputy was Dennis McCarthy.

K. Olive: This work was supported in part by DOE Grant No. DE-FG02-94ER-40823.

L. Secco: I dedicate this contribution to the memory of Prof. Niccolò Dallaporta who has been for me Master in Science and in Life. Prof. N. Dallaporta introduced to me these topics giving the feeling of how much useful is to muse on them in order to direct the scientific research itself and to realize of what lies beyond it.

G. Steigman: The research of the author is supported at The Ohio State University by a grant (DE-FG02-91ER40690) from the U.S. Department of Energy.

M. Turatto: I thank Stefano Benetti ed Enrico Cappellaro for suggestions, comments and continuous discussion on the topics discussed in this book.

Acronyms

ACBAR	Arcminute Cosmology Bolometer Array Receiver
ACT	Atacama Cosmology Telescope
ADEPT	Advanced Dark Energy Physics Telescope
AKARI	Previously known as ASTRO-F or IRIS - InfraRed Imaging Surveyor
ALHAMBRA	Advanced Large Homogeneous Area Medium Band Redshift Astronomical
ALMA	Atacama Large Millimeter/submillimeter Array
AMR	Age Metallicity Relationship
AP	Anthropic Principle
APM	Automatic Plate Measuring
APS	Angular Power Spectrum
ARGO	Astrophysical Radiation with Ground-Based Observatory
ARCADE	Absolute Radiometer for Cosmology, Astrophysics, and Diffuse Emission
ASCA	Advanced Satellite for Cosmology and Astrophysics
ASTRON	ASTRonomisch Onderzoek in Nederland (The Netherlands Foundation for Research in Astronomy)
AstroWISE	Astronomical Wide-field Imaging System for Europe
BAO	Baryonic Acoustic Oscillations
BB	Black Body
BBN	Big Bang Nucleosynthesis
BH	Black Hole
BICEP	Background Imaging of Cosmic Extragalactic Polarization
BLR	Broad Line Region
BOOMERanG	Balloon Observations Of Millimetric Extragalactic Radiation and Geophysics
B-Pol	B-Polarization Satellite Mission
BRAIN	Background RAdiation INterferometer
CAIN	CAmera INfrarroja
CBI	Cosmic Background Imager
CCD	Charge Coupled Device
CCSNe	Core–Collapse SNe

CDM	Cold Dark Matter
CERN	Conseil Européen pour la Recherche Nucléaire – European Organization for Nuclear Research
CFHT	Canada–France–Hawaii Telescope
CFHTLS	Canada–France–Hawaii Telescope Legacy Survey
CHANDRA	Chandra X-ray Observatory in honor of the late Indian-American Nobel laureate, Subrahmanyan Chandrasekhar
CℓOVER	Cℓ ObserVER
CMB	Cosmic Microwave Background
CMD	Color-Magnitude Diagram
CMBPol	CMB Polarization (Mission)
CMR	Color-Magnitude Relation
COBE	Cosmic Background Explorer
COBRAS–SAMBA	COsmic Background Radiation Anisotropy Satellite – SAtellite for Measurement of Background Anisotropies
COMBO17	Classifying Objects by Medium-Band Observations – a spectrophotometric 17-filter survey
COSMOGRAIL	COSmological MOnitoring of GRAvItational Lenses
COSMOSOMAS	COSMOlogical Structures On Medium Angular Scales
COSTAR	Corrective Optics Space Telescope Axial Replacement
CPU	Central Processing Unit
CTA	Cherenkov Telescope Array
CUORE	Cryogenic Underground Observatory for Rare Events
CXC	Chandra X-ray Center
DASI	Degree Angular Scale Interferometer
DE	Dark Energy
DES	Dark Energy Survey
DIRBE	Diffuse Infrared Background Experiment
DM	Dark Matter
DMR	Differential Microwave Radiometers
DPOSS	Digital Palomar Observatory Sky Survey
DUNE	Dark UNiverse Explorer
(E-)ELT	(European) Extremely Large Telescope
EMIR	Espectrógrafo Multiobjeto InfrRrojo
ENO	European Northern Observatory
eROSITA	extended Röntgen Survey with an Imaging Telescope Array
EOR	Epoch of Reonisation
ESA	European Space Agency
ESO	European Southern Observatory
ESSENCE	Equation of State SupErNovae trace Cosmic Expansion
FastICA	Fast Independent Component Analysis
FIRAS	Far Infrared Absolute Spectrophotometer
FIRS	Far Infra-Red Survey
FLRW	Friedmann–Lamaître–Robertson–Walker
FOSA	First Order Smoothing Approximation

FOV	Field of View
FP	Fundamental Plane
FWHM	Full Width Half Maximum
GAIA	Global Astrometric Interferometer for Astrophysics
GC	Globular Cluster
GGC	Galactic Globular Cluster
GPS	Global Positioning System
GR	General Relativity
GRB	Gamma Ray Burst
GROND	Gamma-Ray burst Optical/Near-infrared Detector
GUT	Grand Unification Theory
HB	Horizontal Branch
HDF	Hubble Deep Field
HDM	Hot Dark Matter
HEAO	High Energy Astrophysical Observatory
HEGRA	High Energy Gamma Rays Astronomy
HFI	High Frequency Instrument
HRD	Hertzsprung–Russell Diagram
HST	Hubble Space Telescope
HST-ACS	Hubble Space Telescope Advanced Camera for Survey
HST-ScI	Hubble Space Telescope Science Institute
IAC	Instituto de Astrofísica de Canarias
ICM	Intra Cluster Medium
IF	Isochrone Fitting
IGEX	International Germanium EXperiment
IGM	Inter Galactic Medium
ILC	Internal Linear Combination
IMF	Initial Mass Function
INTEGRAL	International Gamma Ray Astrophysics Laboratory
IRAS	Infrared Astronomical Satellite
ISM	Inter Stellar Medium
ISW	Integrated Sachs–Wolfe
IVOA	International Virtual Observatory Alliance
JAXA/ISAS	Japan Aerospace Exploration Agency/Institute of Space and Astronautical Science
JDEM	JointDarkEnergyMission
JPL	Jet Propulsion Laboratory
JWST	James Webb Space Telescope
KATRIN	KArlsruhe TRItium Neutrino
KM3NET	km^3 NEutrino Telescope
LAMBDA	Legacy Archive for Microwave Background Data Analysis
LCOGT	Las Cumbres Observatory Global Telescope
LHC	Large Hadron Collider
LFI	Low Frequency Instrument
lhs	left hand side

LIGO	Laser Interferometer Gravitational-Wave Observatory
LISA	Laser Interferometer Space Antenna
LOFAR	Low Frequency Array
LOSS	Lick Observatory SN Search
LSS	Large Scale Structure
LSST	Large Synoptic Survey Telescope
LWA	Long Wavelength Array
MACHO	MAssive Compact Halo Objects
MAGIC	Major Atmospheric Gamma-ray Imaging Cherenkov
MAX	Millimeter wavelength Anisotropy eXperiment
MAXIMA	Millimeter Anisotropy eXperiment IMaging Array
MCMC	Monte Carlo Markov Chain
MEM	Maximum Entropy Method
MFD	Mean-Field Dynamo
MHD	Magneto Hydro Dynamics
MOND	MOdified Newtonian Dynamics
MSAM	Medium-Scale Anisotropy Measurements
NASA	National Aeronautics and Space Administration
NeXT	New X-ray Telescope (Mission)
NFPS	NOAO Fundamental Plane Survey
NIST	National Bureau of Standards
NLS1	Narrow Line Seyfert 1
NMR	Nuclear Magnetic Resonance
NuStar	Nuclear Spectroscopic Telescope Array
NVO	National Virtual Observatory
OGC	Open Galactic Cluster
OGLE	Optical Gravitational Lensing Experiment
OSIRIS	Optical System for Imaging and Low/intermediate-Resolution Integrated Spectroscopy
PanSTARRS	Panoramic Survey Telescope And Rapid Response System
PRONAOS	PROgramme National d'AstrOnomie Submillimétrique
PQ	Palomar Quest
PSB	Polarization Sensitive Bolometer
PSPC	Position Sensitive Proportional Counters
QAS	Quasar Absorption System
QCD	Quantum Cromo Dynamics
QED	Quantum Electro Dynamics
QSO	Quasi Stellar Object
QUaD	QUEST at DASI
QUEST	Q and U Extra-galactic Sub-mm Telescope
QUIJOTE-CMB	Q, U, I JOint TEnerife CMB
RAVE	Radial Velocity Experiment
RGB	Red Giant Branch
rhs	right hand side
ROSAT	ROentgen SATellite

rpm	rounds per minute
SAO	Smithsonian Astrophysical Observatory
SAX	Satellite per Astronomia X
SCUBA	Submillimetre Common-User Bolometer Array
SDSS	Sloan Digital Sky Survey
SED	Spectral Energy Distribution
SEGUE	Sloan Extension for Galactic Understanding and Exploration
SF	Star Formation
SFR	Star Formation Rate
SHOES	Supernovae and H_0 for the Dark Energy Equation of State
SIM PQ	Space Interferometry Mission Planet Quiqui
SIRTF	Space Infrared Telescope Facility
SKA	Square Kilometer Array
SMICA	Spectral Matching Independent Component Analysis
SNAP	SuperNova Acceleration Probe
SN	Supernova
SNe	Supernovae
SNFactory	Nearby Supernova Factory
SNLS	Supernova Legacy Survey
SPACE	SPectroscopic All-sky Cosmic Explorer
SSP	Single Stellar Populations
SPT	South Pole Telescope
STEP	Space Test of the Equivalence Principle
SVM	Service Vehicle Module
SZA	Sunyaev-Zel'dovich Array
SZE	Sunyaev-Zel'dovich Effect
TeVeS	Tensor–Vector–Scalar (theory)
THEMIS	Télescope Héliographique pour l'Étude du Magnétisme et des Instabilités Solaires
TopHat-BAM	TopHat-Balloon-borne Anisotropy Measurement
UKIDSS	UKIRT Infrared Deep Sky Survey
UKIRT	UK Infrared Telescope
ULIRG	Ultraluminous Infrared Galaxy
VISTA	Visible and Infrared Survey Telescope for Astronomy
VLA	Very Large Array
VLBI	Very Long Baseline Interferometry
VLT	Very Large Telescope
VSA	Very Small Array
VST	VLT Survey Telescope
WFCAM	Wide Field Camera
WIMP	Weak Interacting Massive Particles
WINGS	WIde-field Nearby Galaxy-clusters Survey
WMAP	Wilkinson Microwave Anisotropy Probe
XCS	XMM Cluster Survey
XLF	X-ray Luminosity Function

XEUS	X-ray Evolving Universe Spectroscopy (Mission)
XMM-Newton	X-ray Multi Mirror Satellite
XTF	X-ray Temperature Function
2dF	2-Degree Field
2dFGRS	2-Degree Field Galaxy Redshift Survey

Chapter 1
Introduction

Mauro D'Onofrio and Carlo Burigana

> Est enim Galaxia nihil aliud, quam innumerarum Stellarum coacervatim consitarum congeries: in quamcumque enim regionem illius Perspicillum dirigas, statim Stellarum ingens frequentia sese in conspectum profert, quarum complures satis magn, ac valde conspicu, videntur; sed exiguarum multitudo prorsus inexplorabilis est[1].

Four-hundred years ago, Galileo Galilei, an Italian, first described with these words the Milky Way galaxy in the Sidereus Nuncius. The discoveries that he announced are so famous that, we can say, he opened the doors of observational cosmology.

This book is aimed at celebrating this event during 2009, proclaimed by United Nations the International Year of Astronomy. We have tried to achieve this essentially in two ways: first by reviewing the most important results in cosmology in the past years, and second by opening a joint discussion among cosmologists and astrophysicists on the present situation of our science, both from the point of view of the self-organization of the scientific community and from exploring the external influences that may affect the development of science.

The book is written by adopting a new style with respect to the classical astrophysical books. Here physicists, astrophysicists, and cosmologists have been interviewed with several questions aimed at highlighting the pros and cons of the current cosmological models, at describing the standard cosmological scenario and the most appealing alternative ideas, and at looking forward in the future, discussing the next challenges addressed by space missions, new telescopes, new surveys, etc.

Since Galileo, telescopic exploration of the sky has completely changed our knowledge of the visible Universe. The last century, in particular, has seen enormous progress in cosmology. Today, a long time seems already to have passed since the theory of General Relativity (GR) by Einstein and the measured expansion law of galaxies by Hubble, providing, for the first time, the evidence of an evolving Universe. What was simply absurd one century ago is now a robust empirical fact.

---

[1] The Galaxy is nothing else but a mass of innumerable stars planted together in clusters. Upon whichever part of it you direct the telescope straightway, a vast crowd of stars presents itself to view; many of them are tolerably large and extremely bright, but the number of small ones is quite beyond determination. English translation of the Sidereus Nuncius by Edward Stafford Carlos in the electronic edition of the book published by Cultural Heritage Langauge Technologies. See Sidereus in web page list.

M. D'Onofrio and C. Burigana (eds.), *Questions of Modern Cosmology: Galileo's Legacy*, DOI 10.1007/978-3-642-00792-7_1,

The exponential growth of astrophysics and cosmology is so significant that the last $30 \div 50$ *years will possibly be remembered as their "golden age". After Friedmann and Lemâitre's discovery of the solutions for Einstein's equations describing the expansion of the Universe, we saw a number of theoretical and observational breakthroughs, many of them driven by the exploration of the Cosmos at frequencies not in the visible band. We recall some of them below, without assigning a classification of importance.*

First, the discovery of the Cosmic Microwave Background (CMB) radiation by Penzias and Wilson in 1965, which supported the Big Bang hypothesis and confirmed the prediction by Gamov of an hot early Universe. Second, the Big Bang Nucleosynthesis (BBN) theory, which describes the production of the light atomic elements during the first minutes after the Big Bang, largely due to the work of Gamov, Fowler, Hoyle, Wagoner, and Peebles. Third, inflation, the model proposed by Guth during the 1980s to solve some critical aspects of the standard Big Bang paradigm (the flatness and horizon problems). Fourth, the recent measurements of the CMB properties providing a direct view of the Universe at the epoch of matter–radiation decoupling: the accurate determination of the Planckian shape of its spectrum and the detection of its anisotropies by two instruments onboard the Cosmic Background Explorer (COBE) led by Mather and Smoot, the discovery of an almost flat Universe by the Balloon Observations of Millimetric Extragalactic Radiation and Geophysics (BOOMERanG) team, and the accurate all-sky mapping of CMB anisotropies, thanks to the Wilkinson Microwave Anisotropy Probe (WMAP) led by Bennett. Fifth, the compelling evidences coming from different fields of astrophysics, indicating a Universe almost completely filled by dark matter (DM) and dark energy (DE), in which the matter we see is only $\sim$*4% of the global energy density. Sixth, let us recall the enormous effort done to establish the Large Scale Structure (LSS) of the Universe, which, since the seminal works of Abell and Zwicky, has now given a clear picture of the distribution of galaxies and clusters. Seventh, stellar evolution studies, which led to an understanding of the chemical enrichment of galaxies and established a reliable astrophysical clock of the evolving Universe. Eighth, the incredible progress in the definition of the cosmological distance ladder, with particular reference to the exceptional results coming from type Ia Supernovae (SNe), which prompted the transition from the cold dark matter (CDM) to the* Λ*CDM paradigm. Last, but not least, the investigation of the properties of the Universe at the time of its structure formation, which shed light on the dark and dawn ages of its evolution.*

In almost every field of astrophysics and cosmology, we have seen incredible progresses, thank also to the parallel growth of technology. Many of them will be the subject of this book, in particular, those realized over the last 10–20 years.

The currently accepted cosmological model, also known as the "Concordance Model" for the converging values of some of its parameters derived in different astrophysical contexts (SNe, LSS, CMB, etc.), represents in the same way the synthesis of these recent achievements. In this model, the Universe is isotropic and homogeneous on large scale and its dynamical evolution is described by GR. Its energy content today is mainly in the form of baryons for $\sim$*4%, DM for* $\sim$*20%, and*

DE for $\sim$76%. The backward extrapolation in time of its evolution implies that the Universe was extremely hot and dense in its early phases and that quantum fluctuations immediately after the Big Bang are at the origin of observed inhomogeneities. Also, a salient feature of the model is that the LSS of the Universe and the formation of the cosmic structures from primordial inhomogeneities were driven essentially by gravity.

We have tried to follow in this book the main events that occurred during the history of the Universe, highlighting the standard and alternative explanations of the empirical evidence.

The road that brought us to the current standard Big Bang picture and Friedmann–Lamaître–Robertson–Walker (FLRW) ΛCDM scenario has seen the cooperative efforts of astrophysicists, cosmologists, and physicists. The link among these fields of research is now so tight that the Universe itself can be regarded as a physics laboratory, in which all physical phenomena at all possible energy scales can be observed. For this reason, in the context of this book, we have used the word "cosmology" in its widest sense, considering it like a puzzle in which each piece of knowledge contributes to the whole picture. It will emerge from the various contributions to this book that every field of astrophysics brings its cosmological information via its own "transfer function". That is, the signals originating from each cosmic source, discrete of diffuse, arrive at the observer only after they have been modified by a series of physical processes.

Some questions are in the background of all interviews, even if not always explicitly formulated. The reader can form his own opinion on them from the various points of view included in the book. One of these key questions is, "Are we living in the 'golden age' of cosmology?" If we judge from the exceptional success of the Concordance Model, the answer is certainly yes. However, if we look deeply at the details of the theories and observations, we discover that a lot of things are still unexplained; first of all, what the DM and DE are and why they seem to be so abundant in the Universe. For many people this is a very uncomfortable situation for a scientific theory. While experiments and observations continuously improve the precision of the determination of cosmological parameters – the building blocks of the present cosmological model – we are still ignorant as to what the Universe is made of.

What would Galileo think of such a situation? Could we trust our theories to the point of believing in something we are currently unable to see experimentally? Have we forgotten Galileo's lesson of a robust empirical approach in doing science?

Many people believe that we are approaching a new revolution in physics. The symptoms of such a crisis seem to come from different sides of fundamental physics. The connection between gravity and quantum mechanics is far from being completed, and the String theory, on which several scientists have invested their life's research, does not seem to have found the final solution. The cosmological constant and the DE problems are two more ingredients of such a crisis. The same can be said for DM, being the weak interacting particles, responsible of the observed surplus of gravitational interaction, not detected yet. The so-called "Anthropic Solution", and the connected "Why Now?" problem, also addressed in this book, are two *further examples of such a crisis. Are we really close to understand the nature of the unknown ingredients of the Concordance Model and the physics of the early Universe or rather are we approaching a shift of paradigm? Or, more dramatically, are we traveling in the middle of a desert? These are further underlying questions in the background of all our interviews.*

Some space is also given in the book to the roles of neutrinos and magnetic fields for cosmology as examples of the above mentioned links between cosmology and fundamental physics.

Being a tribute to Galileo, to his life and scientific experience, and to its role in the change of cosmic vision, part of this book is dedicated to those theories and ideas that are not in line with the standard ΛCDM paradigm. We have included some interviews with scientists who propose alternative explanations of gravity, of the microwave background, and of redshift.

Having Galileo in mind, we also decided to open our pages to a joint discussion among the interviewed scientists about the role and influences of society in the development of science. In which ways sociological, economical, and political conditions as well as ethical and religious ideas may impinge on the progress of science? How have scientific communities organized themselves? Do we have today forms of ostracism vs. "heretical" ideas? All these questions have been posed. Although we are conscious that these are difficult questions, not so close to the interest and work of scientists, and object of specialistic analyses, we believe it is important to hear the points of view of researchers on such issues. The reader will certainly deduce from the various answers that an ample reflection on these themes is mandatory even today for a scientific community, because even if the idea of freedom is now typically accepted by democratic countries, different forms of pressure can still operate on scientists and on the new generations of researchers.

The parallelism between today and the epoch of Galileo is another 'lite motif' in our interviews. According to the idea at the basis of the title of the book, we propose to make evident the self-examination made by scientists, explicitly invited through our interviews to look at the foundation of their scientific activity, their role in the present society, and the perspectives for the future. We hope that the "Galilean spirit" will emerge from the whole set of interviews, given the scientists involved in this enterprise.

The various arguments discussed in this book are organized as follows.

Chapter 2 collects interviews describing the fundamental observations and data interpretations that support the current cosmological scenario or that raise doubts about it, with particular emphasis to the discoveries since $\sim$1990. The main goal of the questions posed in this chapter is the delineation of a path in which the results from experiments and observations most relevant for the comprehension of the Universe and its structures can be critically reviewed and compared with the main predictions coming from the theoretical framework of the standard model.

Chapter 3 reviews the most relevant fields of theoretical cosmology through a series of interviews concerning the global properties of the Universe (geometry, evolution, fate, structure), the stages of its evolution (from inflation and symmetry breaking phase transitions originating defects to the structure formation epoch, *passing through the plasma era, recombination, dark age, reionization, dawn age), its constituents (including ordinary matter, neutrinos, DM, DE), its large scale structure as developed from primordial perturbations, and the main forces driving structure evolution (gravity, electromagnetism).*

Our tribute to Galileo is detailed in Chap. 4. Here, two types of interviews are collected: those dedicated to a critical discussion of some fundamental physical aspects crucial for cosmology, in both the standard framework and alternative models, and those addressing the link between science and society. From one side, the standard cosmological framework still needs to be firmly confirmed through the identification of its fundamental constituents and of the signatures left over from the very early phases of the Universe. On the other hand, alternative ideas are continuously proposed to overcome the problem of the missing detection of DM and DE. In this context, we opened a discussion about the boundary conditions influencing astrophysical research and about the possibility that today we are close to a turning-point of current physics and cosmology.

What should we expect in the future from the point of view of the further verification or falsification of the current paradigm? Chapter 5 illustrates the forthcoming and future projects planned in various fields of astrophysics and cosmology and their scientific promises. The interviews emphasize the technological aspects of these projects, including hardware development intrinsic to space missions, astronomical facilities, extensive observational campaigns, and strategies of data analysis, computer science, etc., and their tight link with the corresponding scientific goals. At the same time, they provide an exciting view of the great efforts done by the scientific community in times like these.

Finally, we present our concluding remarks in Chap. 6.

Chapter 2
Fundamental Cosmological Observations and Data Interpretation

Contributions by Matthias Bartelmann, Charles L. Bennett, Carlo Burigana, Cesare Chiosi, Mauro D'Onofrio, Alan Dressler, Isabella Gioia, Günther Hasinger, Juan Francisco Macias-Perez, Piero Madau, Paola Marziani, John Mather, Francesca Matteucci, Keith Olive, John Peacock, Wolfgang Reich, Pierre-Marie Robitaille, Michael Rowan-Robinson, Gary Steigman, Matthias Steinmetz, Jack W. Sulentic, Massimo Turatto, and Simon D.M. White

2.1 Outline of the Chapter

The last decades have seen an exponential increase of experiments and observations aimed at establishing the structure, the evolution, and the constituents of the Universe, covering essentially the whole electromagnetic spectrum.

The interviews collected in this chapter describe the main evidence at the base of the currently accepted cosmological scenario and discuss the interpretation of experimental data that support it. Some alternative (heretical?) ideas on the interpretation of redshifts and Cosmic Microwave Background (CMB) radiation are also included.

We start presenting the most important change of cosmological paradigm that has occurred in the last two decades: the transition from the Cold Dark Matter (CDM) to the ΛCDM models, described in Sect. 2.2 by John Peacock and in Sect. 2.3 by Massimo Turatto from two different points of view. Far from being a simple re-parameterization of the various kinds of energy contents of the Universe, the new scenario has dramatically impacted on our vision of the fate of the Universe and of its "recent" dynamical evolution and structure formation. It raises the problem of the cosmological constant with the related question of the dark energy (DE) content, and the alternatives to these possibilities that are linked to fundamental physics, from early Universe processes to gravitational theories.

The final step towards the new cosmological paradigm largely relies on the study of type Ia Supernovae (SNe), the pros and cons of which, in cosmological and astrophysical context, are described by Massimo Turatto in Sect. 2.3 and by Francesca Matteucci in Sect. 2.4. The reliability of these objects as distance indicators is clearly the center of our interest in their interviews. How robust is the present indication of an accelerating Universe coming from SNe? How far can we trust in such indicators at redshifts so far from that of nearby type Ia SNe?

Such discussions inevitably bring us to seek new and more powerful distance indicators. Do we have any? In Sect. 2.5, we ask Paola Marziani why the luminous quasars cannot be useful distance indicators for tracing the structure of the Universe, despite the fact that they can be observed up to redshifts ($z \sim 6$) larger than that of

M. D'Onofrio and C. Burigana (eds.), *Questions of Modern Cosmology: Galileo's Legacy*, DOI 10.1007/978-3-642-00792-7_2,

type Ia SNe. Of course, speaking of quasars, it was natural for us to explore the well known question of the anomalous redshifts observed in the past by Halton Arp for some of these objects. We discuss such problems with Jack Sulentic in Sect. 2.6.

Discussion on the empirical cornerstones of the Concordance Model develops along the following lines.

We begin with the cosmological nucleosynthesis, presented in Sect. 2.7, in which we ask the points of view of Keith Olive and Gary Steigman on the standard Big Bang Nucleosynthesis (BBN) theory and on the empirical tests proving this scenario. The interview with Keith is more focused on nuclear reactions and empirical tests of the BBN, while Gary will more closely follow the link between nucleosynthesis process and the expansion of the Universe, addressing the complementary information coming from BBN, CMB, and Large Scale Structure (LSS) and the possible alternatives to the standard theory.

The second big cornerstone, the CMB, has been reviewed by the Nobel Laureate John Mather, by Charles Bennett, and by Juan Francisco Macias-Perez. John and Charles delineated the most important aspects of the Cosmic Background Explorer (COBE) and the Wilkinson Microwave Anisotropy Probe (WMAP) all-sky space missions. These projects have strongly impacted on our ideas on the evolution of the Universe. Their interviews can be found in Sects. 2.8 and 2.8.2. John will review the cosmological context before and after COBE, the great success of the mission, and the experiments on-board COBE. Charles, after a brief discussion of the cosmological questions left open by COBE, will introduce us to the aims and strategy of WMAP, presenting its main scientific achievements. Juan reviews for us the fundamental results from balloon-borne experiments realized in the years between COBE and WMAP (see Sect. 2.8.3).

One key aspect for the interpretation of microwave data is to understand in which way the foreground signals coming from the Milky Way and extragalactic sources are separated from the truly cosmological CMB signal. We start by describing the properties of the far infrared foreground and of dust emission, mainly observed at high resolution through balloon experiments. Juan will address these questions in Sect. 2.8.4, where he gives a concise description of the algorithms of component separation. He also presents an overview of the properties of the Inter Stellar Medium (ISM) in Sect. 2.8.5. Later on, Wolfgang Reich will remind us the characteristics of the radio foreground in Sect. 2.8.6, presenting the most recent all-sky radio maps painstakingly obtained both in total intensity and polarization and discusses the role of magnetic fields in our Galaxy.

Despite the big success of CMB experiments, the interpretation of the CMB data is not uniformly accepted yet. For this reason, in the spirit of this book, we decided to give space even to the more radical opposition. Pierre-Marie Robitaille will give his point of view on CMB in Sect. 2.8.7. He will discuss, in particular, the interpretation of the Planck and Kirchhoff data and the origin of the CMB monopole signal.

From the truly diffuse CMB background, we then move ourselves towards the complex problem of X-ray background, the astrophysical sources responsible for it, and the contribution of X-ray astronomy to the development of the current *cosmological scenario. These themes are addressed here by Günther Hasinger in Sect. 2.9, where he points also out the effect of AGN evolution and massive Black Hole (BH) feedback.*

Having brought the discussion on primordial BHs, we continue our interviews with the difficult problem of how such primeval objects and structures emerged from the dark era, when the Universe entered in its nonlinear phase. Piero Madau will review for us the story of the dark ages and the appearance of first stars in Sect. 2.10. He also addresses the question of the feedback from BHs in galaxy formation in Sect. 2.10.3.

The sections that follow are mainly dedicated to the cosmological information coming from the "nearby" Universe, dominated by the presence of self-gravitating structures of dark and visible matter. The largest of such structures are galaxy clusters. In Sect. 2.11, Alan Dressler will describe the general properties of clusters and their importance for the current cosmology. In particular, he revisits the information that can be extracted from the Morphology–Density relation, which he discovered some years ago, the problem of the bias in the measurements of mass, and the role of the scaling relations, such as the Fundamental Plane and the Tully–Fisher relations, in the cosmological context.

Galaxy clusters are indeed important tracers of the mass distribution in the Universe. In particular, the relationships between X-ray luminosity and temperature, and between temperature and mass, due to the hot gas in the Intra Cluster Medium (ICM), can give us a measure of the clustering of matter around such structures. Isabella Gioia, in Sect. 2.12, will provide a multifrequency view of the properties of galaxy clusters and their use for extracting information on the cosmological parameters.

Of course, speaking of clusters, it was inevitable to address one of the most crucial problems of present cosmology: the dark matter (DM). Does it exist and what is it? How is it distributed around galaxies and clusters? Simon White and Matthias Steinmetz have reviewed the problem in Sect. 2.13 and in related subsections, discussing also the contribution to these studies coming from cosmological simulations. While Simon presents the empirical evidence that call for the existence of DM, Matthias focuses his interview on the pros and cons of the standard CDM scenario.

The lensing phenomenon is closely linked to the problem of DM. Although predicted by Einstein and observed for the first time by Eddington during the solar eclipse of 1919, the cosmological exploitation of lensing started only during the last 10 years. What are its properties and what is its role in determining the distribution of dark mass around structures? Matthias Bartelmann in Sect. 2.14 will review the physical concepts at the basis of our expectations for cosmology coming from weak and strong lensing.

The interviews so far included mainly concern some fundamental categories of physical observables providing cosmological information. These are mass and type of matter, both visible and dark, in the form of astronomical objects or diffuse components, their geometrical (angular and redshift–distance) distributions, and their spectral energy properties. The category considered in the next section is instead directly related to the exploitation of cosmic time evolution of stars, that, *as well known, are the best known astrophysical clocks. Cesare Chiosi will illustrate the constraints coming from stellar evolution in determining the time scale of the Universe. In Sect. 2.15, he revisits the problem of the stellar ages from the perspective of star clusters and that of the integral color of galaxies.*

In the final interview of this chapter, the cosmological question of the value of the Hubble constant H_0 from various observables is addressed. Although the discovery of the expansion of the Universe by Hubble was one of the starting points of modern cosmology, we decided to move this discussion to the end of the chapter, because H_0 is now based on various kinds of observables, from astronomical objects to diffuse cosmic background, and the analysis of discrete objects assumes the formation of astronomical structures described in the previous sections. Michael Rowan-Robinson reviews the present situation of H_0 in Sect. 2.16.

Let us start with the interview of John Peacock, who will now clarify the epoch of transition from the CDM to the LCDM scenarios.

2.2 From CDM to ΛCDM Paradigm

Dear John *(Peacock)***, recent surveys of galaxies and clusters have significantly contributed to our current knowledge of the structure of the Universe and of its evolution. Would you like to summarize here the most important results coming from such studies in the context of observational cosmology?**

Structure in the visible Universe has driven cosmological enquiries from very early on. Once the basic nature of galaxies as stellar systems in motion had been established by the work of Slipher and Hubble, astronomers were inevitably led to ask how galaxies had originated. Even before the detection of the CMB, the theory of gravitational collapse from small initial fluctuations had been worked on extensively, and many of the key elements of modern understanding were in place by the early 1970s, including the idea of characteristic patterns in cosmic structure being imprinted by the transition between an early radiation dominated era and matter domination. These ideas were developed further during the 1970s, at the same time as the first comprehensive attempts at quantifying the inhomogeneities in the galaxy distribution: the correlation-function programme whose results were summarized in the hugely influential book by Peebles [392].

Peebles's book marked a true turning point in the subject, as it was almost immediately followed by two key theoretical advances. The most fundamental was the development of inflation, which proposed the heroic vision of a quantum origin for cosmic structure. Furthermore, the simplest prediction of inflation was seen to be fluctuations that were adiabatic (equal fractional perturbations to the matter density and the photon number density), and nearly scale-invariant in character (metric perturbations that were fractal-like, so that deviations from flatness were of equal magnitude in each logarithmic range of spatial wavelength). The other main step was the idea that the main matter component might be a collisionless relic particle. Initially, the main candidate of interest was the massive neutrino, but it was rapidly appreciated that this idea was a dead end: randomly directed streaming of neutrinos while they are relativistic would erase all galaxy-scale structures [55]. This led to the simpler idea of CDM , in which the particle is massive and any small-scale damping is unobservably small [393]. The consequences of this for the CMB were quickly worked out [54, 577]. Thus, in an astonishing burst of activity, all the elements were put in place of the theoretical picture that still applies today, a quarter of a Century later.

The CDM model made two specific predictions: (1) that there would be a break in the power spectrum of matter fluctuations at a length around $c \times t$ at the time of matter–radiation equality; (2) that the characteristic angular scale of CMB fluctuations would depend on space-time curvature, being around 1° for a flat Universe, but moving to smaller angles for an open Universe. As we know, it has been possible to use these predictions to measure the character of the Universe with astounding precision. The history of this is rather interesting, and shows that galaxy surveys really led the way in the 1990s, although by now the CMB data give us much the most important and accurate constraints. In 1990, for example, only upper limits on CMB fluctuations existed – but these were still important. The small-scale limits on 10-arcmin scales from the Caltech group were rather stringent, and it was already clear that CDM models that were heavily open could be excluded. If flat models were preferred, it came down to a choice between the $\Omega_m = 1$ Einstein–de Sitter Universe, or one that was vacuum dominated, satisfying $\Omega_m + \Omega_v = 1$. Throughout the 1980s, most cosmologists would have plumped for the former alternative – based largely on worries about the fine-tuning involved in a small cosmological constant that becomes important only around the present. But by 1990, strong evidence had accumulated in favor of this alternative. The CDM spectrum contains a break at the horizon scale of matter–radiation equality, and observations of this break scale allow the combination $\Omega_m h$ to be pinned down (where $h \equiv H_0/(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}})$ is the usual dimensionless Hubble parameter at the present time, i.e., the reduced Hubble constant). A lower matter density implies relatively larger fluctuation on large scales, and clear evidence for these was seen in the projected clustering properties of the Automatic Plate Measuring (APM) galaxy survey, which was produced from scans of UK Schmidt Telescope photographic plates. The preferred value of $\Omega_m h \simeq 0.2$ argued for a low matter density, even though the Hubble constant was not so well known then as it is now. The only way of reconciling the small-scale CMB constraint with the requirement of a low matter density from galaxy clustering was to assume that the Universe was dominated by a cosmological constant – so that the sum of this vacuum energy and the matter content yielded a flat Universe. This argument was made with admirable clarity in a Nature paper by Efstathiou et al. [147].

Although the logic was impressive at the time, there was still considerable resistance to the conclusion. I certainly remember being deeply unhappy with the idea of a fine-tuned vacuum density, and spent the early 1990s looking for ways out: basically trying to see if nonlinear evolution and scale-dependent bias could help. It was abundantly clear that $\Omega_m = 1$ was dead, but I was philosophically more attracted to the idea that the Universe might be open (with, say, $\Omega_m \simeq 0.4$ and a low Hubble parameter) than to accept the reality of vacuum domination. These slightly open models looked a progressively less good match to the data as time went by, although their rejection was delayed by the first Perlmutter et al. paper [402] on the supernovae Hubble diagram, which conclusively rejected what we would now regard as the correct vacuum-dominated model. It was only in 1997, just at the time I was finishing my textbook on cosmology, that it became clear that the SN story was changing and also falling in line with the picture that LSS+CMB had been painting since 1990. At that stage, I abandoned any further resistance to the idea of nonzero Λ – but it is interesting to see how the suggestion has arisen that the "discovery of DE" came out of a clear sky with the 1998 SN papers. I certainly feel fortunate in the alignment of the timing of all this with finishing my book: if I'd been more efficient and got it done when the publishers first wanted, it would have been horribly out of date within a couple of years. As it was, I was able to in effect write a first-hand witness account of the birth pangs of the present standard model. Although we know many numbers much more accurately than we did 10 years ago, it is astonishing how little has changed over the past decade in terms of our basic set of ideas. I cannot decide if this is a cause for celebration or depression; certainly, the nature of the subject has altered out of all recognition from the glory days of the 1980s, and the kind of creative cosmological speculation that was common then is less easy to carry off today.

Thank you John. Indeed the last 10 years have seen enormous progress in both observations and theories, but not yet decisive to establish the cosmological scenario. Could we interpret this as a symptom of a crisis? We will address this question later in this book. For the moment we are interested in highlighting the transition from the CDM to the ΛCDM from the point of view of type Ia SNe at high redshifts.

2.3 Type Ia SNe as Probe of the Paradigm Shift

Dear Massimo *(Turatto)*, **since the discovery of the cosmic acceleration of the Universe prompted by observations of high-z SNe-Ia, the ΛCDM scenario has been confirmed by CMB experiments, in particular by WMAP. Could you please briefly review the role of SNe as cosmological tracers?**

Supernovae are celestial objects that, even if for short time, shine as bright as their entire host galaxy ($M_B \sim -19$), making them detectable up to cosmological distances with large telescopes and modern detectors. For this reason, SNe are unmatched probes of the different evolutionary conditions of the Universe. In particular, the subclass of SNe called of type Ia, which can be recognized for the characteristic light curve and spectral features (see [561, 562] for a modern SNe taxonomy), has the specific property of having a relatively small dispersion of luminosity at maximum light, making them unique distance indicators. As explained in detail by Francesca Matteucci in the Sect. 2.4, type Ia SNe are the outcome of thermonuclear explosion of White Dwarfs reaching the Chandrasekhar limit by accretion of material from a companion. The fact that in first approximation they are similar one to each other is therefore not surprising.

In the 1990s, with the improved photometric calibration it became evident that type Ia SNe show a significant diversity at optical wavelengths (hence are not strictly standard candles), but it was also found that simple relations exist between the shapes of the light curves and the absolute magnitudes at maximum [402, 405, 449]. Therefore, in analogy to what happens with Cepheids, type Ia SNe can be recovered as distance indicators (*standardizable* candles). This method has proved to be very effective, but has a major caveat, the lack of a satisfactory theoretical interpretation. The fact that type Ia SNe might indeed be much better standard candles in the near-IR [296], where also the reddening is much less than a problem, is of little help in the current context because of the drift in the luminosity peak with redshift to even longer wavelengths, with the known observational complications and of the lower luminosity of type Ia SNe in the IR. If this finding will be confirmed, the James Webb Space Telescope (JWST) might exploit the near-IR properties of type Ia SNe.

The Hubble diagram built with the shape-corrected luminosities of nearby type Ia SNe has a dispersion less than 0.2 mag, that is, 10% in distance (see [316] and references therein). A direct output of the Hubble diagram is the determination of the Hubble constant once the absolute magnitude of type Ia SNe is known independently. Making use of the calibration of nearby type Ia SNe by Cepheids, a recent determination provides $H_0 = 73 \pm 4(\mathrm{stat}) \pm 5(\mathrm{sys})\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ [451] (see also Sect. 2.16 by Michael Rowan-Robinson). The Hubble diagram of type Ia SNe shows evidence for a *local bubble* with local ($v < 7{,}000\,\mathrm{km\,s^{-1}}$) expansion velocities larger than the cosmic average [258, 600].

But type Ia SNe allow to go further and to explore the cosmic expansion rate up to look-back times of about two-third the age of the Universe ($z \sim 1.5$). In 1998, two competing SN teams [403, 450] announced the independent discovery that type Ia SNe at $z \sim 0.5$ are fainter than predicted in an empty Universe and, therefore, the expansion of the Universe is accelerated, possibly due to the presence of a new (dark) energy component, which opposes the gravitational pull or a modification of gravitation theory. Combining SN measurements with those obtained from the LSS and CMB, a so-called Concordance Model has emerged, in which the Universe is flat and filled with about 4% baryons, 20% DM, and 76% DE (see [183] for a recent review). These claims are now supported by better statistics provided by the Supernova Legacy Survey (SNLS) and the Equation of State Supernovae trace Cosmic Expansion (ESSENCE) collaborations [16, 591], which have measured light curves for several hundred type Ia SNe in the $0.3 < z < 0.9$ range. In addition, a SN search carried out with the Hubble Space Telescope (HST) confirmed that the higher-redshift ($z > 1$) Universe is decelerating (see Fig. 2.1), and was able to sample the transition from a deceleration in the past to the current acceleration [452]. The SN data together with those from Baryonic Acoustic Oscillations (BAO) and CMB have been used to constrain the equation of state of the DE ($w = P/(\rho c^2)$) (see Figs. 14 and 15 of [293]). Presently the available data are consistent with $w = -1$ with no time variation, though other models cannot be excluded.

The cosmological results just mentioned are getting more robust as larger data set are collected but the number of issues remain, both technical and astrophysical, which limit the accuracy of the results (see [316] for an extensive overview).

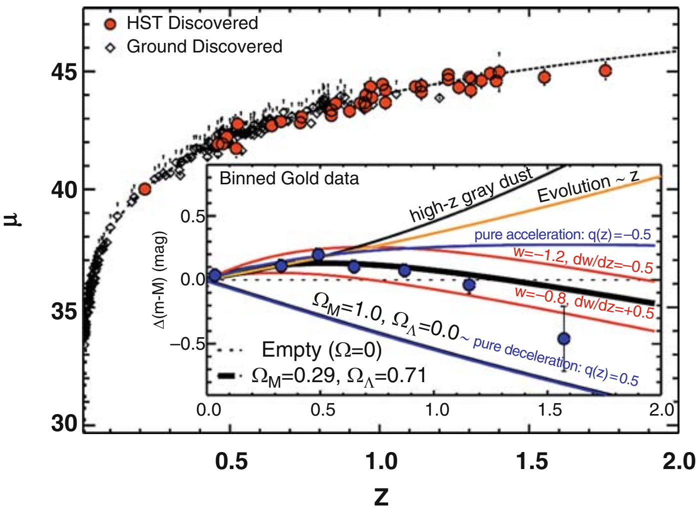

Fig. 2.1 SN Ia Hubble diagram. SNe Ia from ground-based discoveries in the gold sample are shown as *diamonds*, HST-discovered SNe Ia are shown as *filled symbols*. Overplotted is the best fit for a flat cosmology: $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$. *Inset*: Residual Hubble diagram and models after subtracting empty universe model. The gold sample is binned in equal spans of $n\Delta z = 6$, where n is the number of SNe in a bin and Δz is the redshift range of the bin. From [452]

Detailed treatments of the systematics have shown that more accurate photometric calibration especially when dealing with mosaic detectors or with data coming from different telescopes are needed. Moreover, it is important to monitor the stability and the uniformity of photometric systems as well as the atmospheric temporal and spatial variations. Even small uncertainties at the edges of the photometric band passes can affect significantly the K-corrections of the feature-rich SN spectra or the computation of light curve models based on spectral templates. Different light curve fitting methods used so far seem to give consistent results but still can be the source of systematic biases. The problem of removal of host galaxy contamination has not been solved, as well as the contamination of the SN samples by possible type Ia SNe impostors, like bright SNIb/c [32, 549]. Several are the astrophysical issues deserving further investigation, among which are the gravitational lensing and the existence of peculiar velocity fields. The presence of noncanonical ($R_V \sim 2$) reddening laws in the host galaxies [154, 297] might be a serious concern especially in the early Universe. Last but not least, there is the critical assumption that lies at the basis of the observational cosmology with SNe, that is, the observational properties of SNe do not change with redshift. The currently available data show that the spectroscopic and light curve behaviors of type Ia SNe at low and high redshifts are rather similar [16, 50, 452, 591] and the cosmological parameters derived in different host environments [541] or SN subsamples [293] are consistent, but still this is an issue that must be continuously investigated. Not only that, before drawing any firm conclusion on the nature of our Universe, we need a full theoretical comprehension of these fascinating, but largely unknown objects, and to answer basic but still unsolved questions. What is the true nature of the progenitors? What is(are) the explosion mechanism(s)? What drives the observed diversity? What is(are) the parent Population(s)?

Do there exist other types of SNe that can be of interest for cosmology?

Contrary to the thermonuclear type Ia SNe, Core-collapse SNe (CCSNe), descending from stars more massive than 8–10 $M_\odot$, are far from being standard candles and show huge variations in the peak luminosity and in the shape of the light curve, due to different configurations of the exploding stars at the moment of the explosion and to different energetics of the explosion itself. They range from the faint ($M_V \sim -14$) SNe [389] to the bright ($M_R \sim -22$) and hyper-energetic SNe like 2005ap and 2006gy [5,432,519]. Nevertheless, some CCSNe can be used as distance indicators by mean of the Expanding Photosphere Method (EPM [278]), which allow to derive their distance independent of the calibration of lower rungs of the distance ladder. The uncertainty in the determination of dilution factor, which accounts for the difference of the SN spectra from that of a Black Body (BB), seems to overcome by the modern incarnation of the method (Spectral-fitting Expanding Atmosphere Model (SEAM) [25]), which exploits a detailed spectral modeling for each object. An empirical method has been recently developed [226], which makes use of the luminosity of the extended plateau characterizing the light curve of the hydrogen-dominated type IIP.

CCSNe are attractive in cosmology for other reasons. Because of the short evolutionary life-times of their progenitors, (<30Myr) the determination of the rate of their explosion provides a direct measurement of the on-going star formation rate (SFR) for an assumed initial mass function (IMF). The study of the SN rate as a function of the redshift thus provides a trace of the star formation history (SFH). The most recent determinations confirm the steep increase of CCSN rate (SFR) by a factor of 3 for a look-back time of 3 Gyr [65].

Stripped-envelope CCSNe (SNIb/c), that is, those whose progenitors have lost the H (SNIb) and He (SNIc) layers by massive stellar winds or by interaction with a nearby companion star, have been recently associated to the Gamma-ray burst (GRB) of long duration. In particular, the association seems to hold for high-energy, broad-lined events like the type Ic SN 1998bw ($E > 10^{52}$ ergs) [190, 252]. A continuous distribution of properties seems to connect these *hypernovae* to less energetic objects like 1994I, passing through intermediate objects like SNe 2002ap and 2006aj.

Finally, overall SNe of all types are the major contributors to the chemical enrichment of the Universe by returning to the ISM the heavy elements synthesized during the hydrostatic and explosive burning. The impact of various SNe types on the chemical evolution of galaxies are extensively discussed by Francesca Matteucci in the next section.

Thank you Massimo. It seems that the uncertainties on the distance of SNe is still a matter of controversy today: going below a 10% uncertainty on distance is very difficult for several reasons. Furthermore, as you say correctly, can we exclude a *luminosity evolution of type Ia SNe with redshift? As we do not really know the details of the explosion mechanisms and the nature of the progenitors, can we trust in them as good standard candles? Let us ask Francesca Matteucci about that.*

2.4 SNe Physics and the ΛCDM Scenario

Dear Francesca *(Matteucci)*, the current lack of a full theoretical explanation of the physics of the explosion of SNe may be potentially dangerous for the cosmological Concordance Model. Can you explain why?

The most dangerous fact for the cosmological Concordance Model is the possibility that type Ia SNe are not standard candles. This could happen if the mechanism of explosion would be different among the progenitors of such SNe, including the possibility that high redshift type Ia SNe are different from the local ones. This would therefore challenge the assumption that type Ia SNe can be considered as standard candles. What about the explosion mechanism? What do we know? It is commonly believed that type Ia SNe originate from the explosion, by carbon-deflagration, of a carbon oxygen white dwarf, which has reached its limiting mass for stability, namely the Chandrasekhar mass ($M_{\rm Ch} \sim 1.44 {\rm M}_{\odot}$), as illustrated in Fig. 2.2, where a possible scenario for the progenitor of type Ia SNe is presented. So, each SN Ia should be the result of the explosion of a fixed mass. This would ensure that the maximum luminosity of such SNe is always the same as it was believed until a few years ago, when type Ia SNe with different maximum luminosity were discovered.

However, Phillips [405] pointed out that there is a significant intrinsic dispersion in the absolute magnitudes at maximum light of local type Ia SNe. This result was interpreted to arise from a possible range of masses for the progenitors or from variations in the explosion mechanism. Both interpretations could cast doubt on the use of type Ia SNe as very accurate standard candles, particularly at large redshifts where Malmquist bias[1] could be an important effect. It has then been proposed that some SNe type Ia could be the result of the explosion of a sub-Chandrasekhar white dwarf ($0.6 - 1.0\, M_{\odot}$, see, e.g., [592]). Variations in the explosion mechanism could also produce a dispersion in the absolute magnitudes [275]; besides deflagration, other possible explosion mechanisms are detonation, delayed detonation, pulsating delayed detonation, and tamped detonation, although carbon-deflagration is preferred as it produces the right amount of chemical elements observed in SN spectra. However, there has been shown to exist a correlation between the maximum absolute magnitude of SNe Ia and the rate of decline of their luminosity after the

[1] The Malmquist bias is a selection effect in observational astronomy. In particular, if a sample of objects (e.g., galaxies, quasars, etc.) is flux-limited, then the observer will see an increase in average luminosity with distance. This is of course because the less luminous sources at large distances will not be detected. The solution is to use a sample that is not magnitude limited such as a volume limited sample.

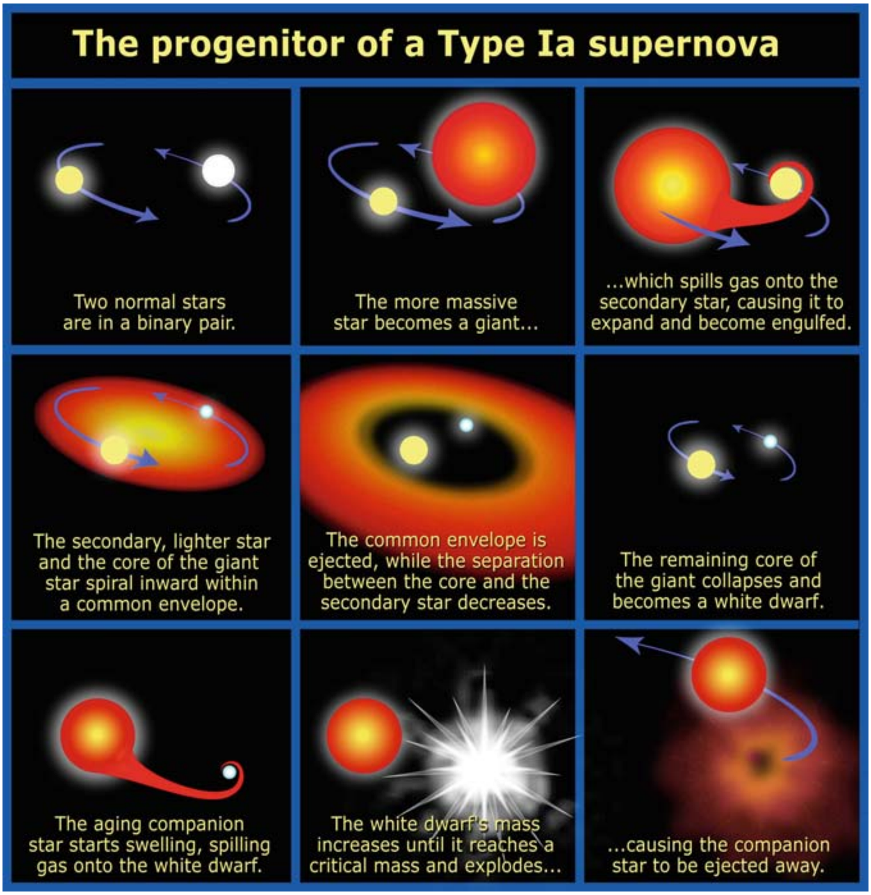

Fig. 2.2 The progenitor of a type Ia SN in the context of the single-degenerate model: here we have a binary system made of a C–O white dwarf plus a normal star. Illustration credit: NASA, ESA, and A. Feild (HST-ScI)

maximum, which therefore allows one to calculate the maximum luminosity in any case, and therefore to retain the SNe Ia as standard candles.

The other challenge to the Concordance Model of cosmology is the possibility that high redshift type Ia SNe are different from local ones, and therefore that the correlation between the maximum luminosity and the rate of decline is no more applicable. Howell et al. [246] pointed out that there are basically two groups of type Ia SNe, as suggested also by the studies of Mannucci et al. [336, 337]: one group is made of prompt SNe exploding on short timescales (less than 0.1 Gyr) and they are intrinsically more luminous and with broader light curves, the other group is made of SNe, which take several Gyrs to explode since the birth of the progenitors, have narrower light curves, and are less luminous. A possible interpretation is that the prompt SNe originate from white dwarfs with more massive progenitors than the delayed SNe. In particular, progenitors with main sequence masses between 5 and 8 $M_\odot$ should be related to the prompt SNe. They are brighter and produce more ^{56}Ni, whose decay is responsible for the observed light curve. As the cosmic star formation rate increases by a factor of 10 from redshift $z = 0$ to $z = 1.5$ [245], at high redshift, these SNe will be an order of magnitude more common, and even more so if it is assumed that ellipticals and spheroids formed preferentially at high redshift, as many observations seem to indicate. Therefore, it may be necessary to apply corrections for the evolution of SN Ia properties with redshift (beyond the correction for light curve shape established for local type Ia SNe), and future studies requiring increasing precision must take into account the effects of an evolving type Ia SN population.

While the elements (O, Ne, Mg, Si, S) are mainly produced in type II SNe in relatively short time-scales $(0.01 \div 0.03$ Gyr), the Fe-peak elements are produced in SNe Ia in a longer time-scale, which has been estimated to be $0.3 \div 1.0$ Gyr and more through chemo-dynamical modeling of galaxies. Recently, high redshift ($z > 6$ or larger) quasars have been discovered with high iron abundance. Do you think that the time scale for metal enrichment may fall in contrast with standard cosmological scenario if more quasars with high metallicity will be discovered at similar or even higher redshifts? Could it be a serious challenge to the Standard Model?

The most common interpretation of the abundance ratios in galaxies is the "time-delay" model, namely the delay with which some stars restore their nuclear products into the ISM relative to other stars. In particular, type II SNe, which originate from core-collapse of massive stars ($M \geq 10\,M_\odot$), restore their main nuclear products (O, Ne, Mg, Si, S, Ca) and a fraction of Fe on timescales of the order of $0.01 \div 0.03$ Gyr. On the other hand, type Ia SNe, which are believed to originate from exploding carbon–oxygen white dwarfs in binary systems, restore Fe, which is their main nuclear product, into the ISM on a large range of timescales going from 0.035 Gyr to a Hubble time. The typical timescale for the Fe enrichment from SNe of type Ia, which we can define as the time of the maximum for the type Ia SN rate, is a function of the assumed progenitor model for type Ia SNe and of the star formation history. Therefore, it varies from galaxy to galaxy, as different SFHs characterize the Hubble Sequence . This timescale can vary from $\sim$0.3 Gyr in ellipticals and bulges to $\sim$1 Gyr in the local disk and to $\sim$4–5 Gyr in irregular dwarf galaxies. These timescales have been evaluated [348] by assuming the "single-degenerate scenario" for the progenitor of type Ia SNe and that the star formation rate is decreasing in intensity and increasing in length going from early to late type galaxies, in agreement with observational evidence [273]. The single degenerate scenario for the progenitor of type Ia SNe suggests that a binary system made of a carbon–oxygen white dwarf plus a younger star can produce a SN Ia. In fact, when the younger companion evolves and becomes a red giant, it starts losing material onto the white dwarf, which can reach the Chandrasekhar mass limit and explode catastrophically leaving no remnant. This thermonuclear explosion produces $\sim 0.6 \div 0.7\,M_\odot$ of Fe plus minor quantities of the elements from C to Si. In this progenitor scenario, the most massive binary system that can contribute to a type Ia SN is made of an 8 $M_\odot$ plus a companion of roughly the same mass. This implies that the minimum time for the explosion of the first type Ia SNe is no longer than 0.035–0.040 Gyr. Another possible scenario for the progenitors of type Ia SNe is the double-degenerate one: in other words, two carbon–oxygen white dwarfs of roughly 0.7 $M_\odot$ merge after loosing angular momentum due to gravitational wave emission. In this case, the clock to the explosion is given by the lifetime of the originally smaller mass, as in the single degenerate scenario, plus the gravitational time delay, namely the time necessary to bring together the two white dwarfs that will give rise to a Chandrasekhar mass and explode as in the previous scenario. The nucleosynthesis products are the same in the two scenarios. The minimum time for the explosion of these systems will therefore be given by the lifetime of an 8 $M_\odot$ plus the minimum gravitational time delay (0.001 Gyr, see [219]). Having said that, we see that the minimum timescale for the explosion of SNe type Ia is practically the same for all galaxies and in any scenario (the difference is only 0.001 Gyr), whereas the timescale for the maximum in the SN rate, which is relevant for chemical enrichment, changes according to the SFH, as already discussed. Observational evidence for such prompt type Ia SNe has been recently provided by [336, 337]. These authors suggested, in fact, that roughly 50% of all type Ia SNe explode soon after stellar birth, in a time of the order of 10^8 years, whereas the remaining 50% of type Ia SNe explode on a much wider timescale distribution going from times larger than 10^8 years to a Hubble time (14 Gyr). They reached this conclusion by studying the dependence of the SN Ia rate on the colors of parent galaxies and the enhancement of the SN Ia rate in radio-loud early type galaxies. The fraction of prompt SNe Ia suggested by Mannucci and collaborators is higher than the fraction generally assumed in modeling the type Ia SN rate. In particular, the type Ia SN rate, in the framework of the single degenerate scenario (e.g., [220, 348]) predicts a fraction of prompt type Ia SNe not larger than 13% of the total.

To answer the question, we can say that there is no problem in explaining the high Fe abundance observed in high redshift ($z > 6$) quasars as long as we assume that galaxy formation, in particular the formation of large ellipticals, started at redshift $z \gg 6$. In the following we explain why Quasars are generally hosted by massive ellipticals and the quasar phenomenon is attributed to matter falling into a central BH (see Fig. 2.3). The chemical abundances measured from the broad emission lines in quasars indicate supersolar abundances of several chemical elements including Fe (e.g., [138, 225, 335]).

It has been shown by [347] that in massive ellipticals, the hosts of quasars, when a strong starburst is assumed together with a standard IMF [484] and the single degenerate scenario for the progenitors of type Ia SNe, the interstellar gas can reach solar Fe abundances on timescales of the order of 0.1 Gyr from the beginning of star formation (note that in such a galaxy the relevant timescale for Fe enrichment from type Ia SNe is ~0.3 Gyr), and that at 1 Gyr the Fe abundance has already reached ten times the solar value, as is shown in Fig. 2.4, where we present the evolution of the abundances of several chemical species in the ISM of a large elliptical. In this model, it was assumed that the initial strong starburst ends when a galactic wind develops.

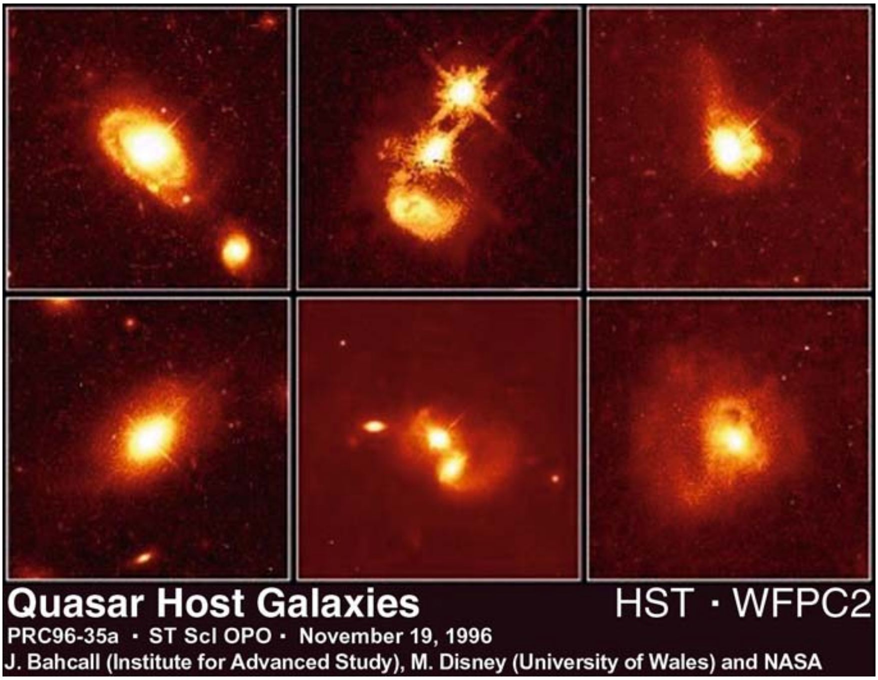

Fig. 2.3 Images of quasar host galaxies from the HST. The host galaxies are generally ellipticals

During the starburst, type II SNe already produce a non-negligible fraction of Fe and most of the oxygen on timescales not longer than 0.03 Gyr. Actually, in a strong starburst like this one, type II SNe can produce by themselves enough Fe to enrich the gas up to the solar value. The wind occurs when the gas thermal energy, due to the energy deposited by SN (II and Ia) explosions, equates the potential well of the gas. The galactic wind carries away all the residual gas and then the elliptical galaxy evolves passively. At this point, the gas restored by the dying stars goes to feed the central BH and the quasar phase starts.

In Fig. 2.4, the occurrence of the wind is identified by the discontinuity in the curves (between 10^8 and 10^9 years). As one can see, after the occurrence of the galactic wind, the Fe abundance reaches values as high as 10 times solar, while the O abundance grows up to an abundance in excess of solar before the winds and remains constant and lower than the Fe abundance after the wind. This is due to the fact that after the wind, which removes most of the gas from the galaxy, the star formation (SF) process stops and therefore the O production, which is related to the short living massive stars, also stops. On the other hand, Fe continues to be produced by type Ia SNe, which continue to explode until the present time. This scenario for the evolution of the gas in ellipticals hosting quasars seems to reproduce not only the high Fe abundance at high redshifts but also the observed constancy of the Fe and other element abundances in quasars as functions of redshift.

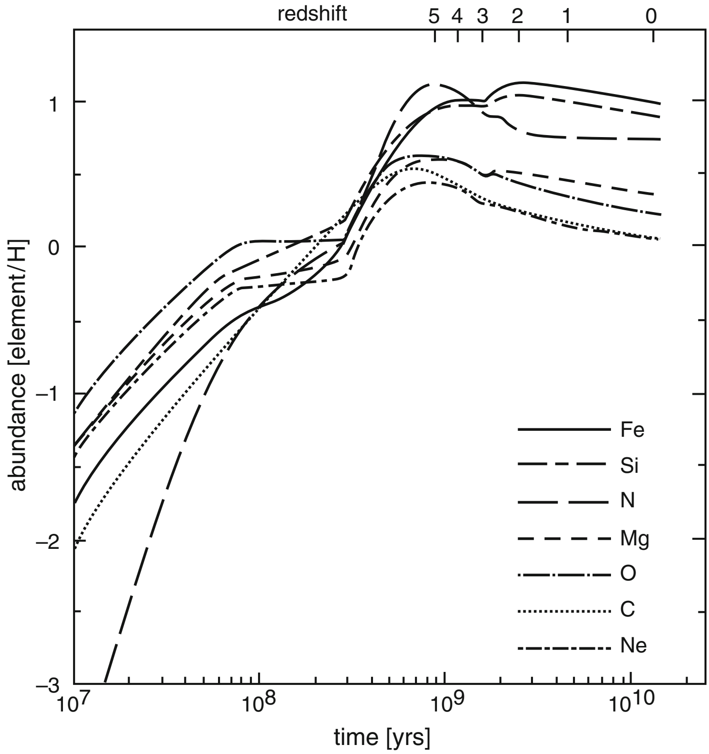

Fig. 2.4 Predicted time-dependence of several chemical elements, as indicated in the figure, relative to the solar abundances (0 means solar, 1 means 10 times solar, etc.) for an elliptical galaxy of $10^{12}\,M_\odot$ luminous matter. The adopted IMF is from [484] and the cosmology is the standard one with an assumed redshift of galaxy formation of $z = 10$. The time at which the abundances show a discontinuity refers to the time of occurrence of a galactic wind. From [347]

The results of [347] have been confirmed in the following years by the models of [213, 414, 472]. Moreover, the rapid increase of the Fe abundance would further be strengthened if the type Ia SN rate suggested by [336] were adopted. In this case, in fact, the increase of the Fe abundance in time would be even faster than in Fig. 2.4, because of the fraction of prompt type Ia SNe higher than in the rate adopted by [347], as explained before.

Before concluding, it is important to note that the cosmic age of 1 Gyr (the time at which the Fe abundance is maximum) corresponds, in the concordance cosmology ($\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 65$), to $z = 5$, if the redshift of galaxy formation is assumed to be $z_f = 10$. Therefore, there is no problem in explaining the high Fe abundances observed in quasars at redshift $z = 6$ and beyond.

In summary, the timescale of metal enrichment is not in conflict with the standard cosmological scenario as long as the redshift of galaxy formation for massive ellipticals is set at $z_f \geq 10$.

Thank you very much Francesca. It is remarkable how the studies of metallicities in the local Universe and in distant galaxies and high redshift Quasi Stellar Objects (QSOs) might represent a piece of the puzzle in the understanding of cosmic structures.

It is natural to ask now why extremely bright and high redshift objects like QSOs cannot be used as standard candles. Paola Marziani will explain us why not.

2.5 Cosmology with Quasars

Dear Paola *(Marziani)*, **quasars are among the most luminous sources in the Universe, and their optical luminosity is, in most cases, stable over periods of several years. Quasars have been discovered up to $z > 6$. Type Ia SNe are, in comparison, much dimmer sources, and even the most recent studies employ SNe only up to redshift $z \approx 1.9$. Why quasars have never been effectively used as standard candles?**

The question you are asking me is both challenging and embarrassing. It is challenging because a good standard candle needs to have a known, well-defined luminosity with a small intrinsic dispersion around an average value. Or, at least, a standard candle should be based on a calibration of a measurable property that tightly correlates with luminosity. Good standard candles, especially in a cosmological context, should then be easily recognizable and highly luminous.

2.5.1 The Challenge

There is no doubt that the last two properties are met by quasars. Quasars emit a fairly univocal spectrum, with prominent broad emission lines in the optical and in the UV range. And no doubt they can be very, very luminous: their absolute magnitude reaches $M_{\rm B} \approx -30$, which corresponds to a luminosity 10^4 times that of Messier 31, the Andromeda galaxy. This is unfortunately only a part of the story. If quasars can be the most luminous sources in the Universe that can be stable over periods of several years (as opposed to GRBs), they can also be comparatively faint. We can immediately think of the other extreme, at low luminosity: the famed nucleus of NGC 4395 hosts the least luminous quasar known: its $M_{\rm B} \approx -10$ is just 10 times the luminosity of a typical blue supergiant star [366]. And we know that quasars can have all luminosities in between the two extrema (which are a mind-wobbling 10^8 times apart!), with a luminosity function that is open-ended at low luminosity. Nor it has been possible to identify a flavor of quasars whose luminosity distribution is peaked or even tightly constrained.

Speaking of quasars, everyone naturally thinks of those star-like objects at high redshift. After all, the term quasar comes from *quasi stellar radio source*. Quasi-stellar because the spectrum did not look like that of a star when the first spectroscopic optical observations were carried out in the early 1960s. On the other hand, quasars looked unresolved on the photographic plates, exactly like a star. And the first quasar discovered was a powerful radio source, identified as 3C 273. Let us now make a jump of nearly 50 years. There has been a sort of *luminosity unification* of quasars. Decades of observations with ever-improving instruments found that the emission-line spectrum and the spectral energy distribution of quasars are very similar over a very wide range of luminosity. We see the same lines and almost the same line widths also in the bright nuclei of the relatively nearby galaxy, the one known as Seyfert galaxies (discovered 20 years earlier than quasars but not understood at the time) as well as in luminous quasars. We observe strong and ubiquitous hard X-ray emission.

Figure 2.5 conveys the meaning of these words in term of three spectra of quasars of widely different redshift and luminosity, even if the comparison is restricted to a narrow range around the HI Balmer line Hβ. The spectra show clearly that a very luminous quasar can look like a bright, nearby Seyfert galaxy, albeit it is important to stress that not all quasars look like the ones shown in Fig. 2.5. The luminous

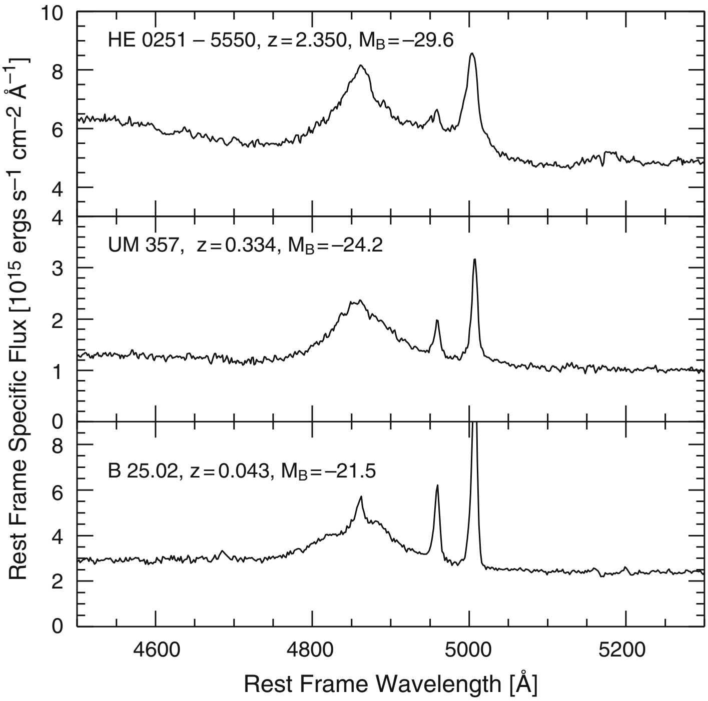

Fig. 2.5 The spectra of three quasars of widely different luminosity and redshift, covering the broad hydrogen Balmer Hβ line and the narrow, forbidden [OIII]$\lambda\lambda$ 4959,5007 lines. Note that the HE 0251-5550 is ≈2,000 times more luminous than *B* 25.02. While *B* 25.02 is a local Seyfert 1 galaxy, HE 0251-5550 is a distant quasars seen at a lookback time of 2/3 the age of the Universe. Yet, their Hβ spectra look very similar

nuclei of galaxies showing a quasar-like spectrum have become to be known as AGN, but there is no discontinuity in the luminosity distribution between Active Galactic Nuclei (AGN) and quasars. The distinction originated from the resolving power of the instruments: if a surrounding host galaxy is appreciable then it is an AGN; if not, then it is a quasar. We are presently able to resolve host galaxies up to redshift $z \approx 1$ [248], while the first quasar at $z \approx 0.15$ looked stellar. For much larger redshift, seeing or not seeing an underlying galaxy is still a matter of faith, but the results at $z < 1$ are still remarkable: $z \approx 1$ means a lookback time of roughly 7 Gyr, half the age of the Universe.

The previous digression is in part reassuring, as the ability to resolve the host galaxy is still a much needed confirmation of the assumption that distant quasars are luminous nuclei of galaxies as found locally. However, it also highlights why quasars are so cumbersome if one thinks of their potential use for measuring fundamental cosmological parameters like the Hubble constant H_0, the energy density associated to matter Ω_m, and to the cosmological constant Λ, Ω_Λ. We have a class of sources whose luminosity is spread over an enormous range in luminosity, and whose spectral properties is fairly similar over a large range of redshift, basically from local $z \approx 0$ Seyfert 1 nuclei to the most distant quasars at $z > 6$! And there are more sources of concern.

Quasars are anisotropic sources in most regions of the electromagnetic spectrum. Two main effects contribute to anisotropy: relativistic beaming in radio-loud sources, and obscuring material co-axial with the accretion disk in both radio-loud and radio-quiet AGN. This is the essence of the so-called *Unification Scheme* of quasars and AGN [248]. If a quasar is seen through a thick structure of absorbing gas and dust, its innermost emitting line regions will be obscured and the UV soft-X spectral energy distribution will be strongly affected. We will ignore these obscured AGN (conventionally called type-2) in the following. However, beaming and orientation effects are not yet fully understood as far as their influence on optical/UV spectroscopic properties of *unobscured* AGN (conventionally termed type-1) are concerned. Their occurrence in the main flavor of quasars, the ones that are radio-quiet (about 90% of all quasars), is even more enigmatic: for example, who can tell which radio-quiet AGN are seen pole-on in analogy to radio-loud BL Lacs and optically violently variable radio quasars?

But we know that orientation matters. Let me make just one example. Core-dominated and lobe-dominated quasars, which are thought to be sources whose radio jet is, respectively, almost aligned or grossly misaligned to our line-of-sights, show different Balmer line widths, by a factor $\approx$2. This has remarkable consequences on physical parameters estimation, as I will try to explain later.

For the moment it is important to keep in mind that quasars are such pranksters that they look different if they are seen along different line-of-sights. And that we do not understand well how. Is this enough not to plunge anyone into deep depression? And we are still not done. There is also the embarrassment.

The embarrassing side of your question is that quasars are plentiful: data for $\sim 10^5$ are presently available from the SLOAN Digital Sky Survey (SDSS) Data Release (DR) 6, more than 13,000 with $z \geq 2.3$ (for comparison, you can consider that only 200 quasars were catalogued in 1971!). Quasars are not only luminous sources, much more luminous than type Ia SNe ($M_B > -30$ vs. $M_B \approx -21.7$), they are also variable. The last resort is to look for parameters that we can easily measure and that tightly correlate with luminosity. Is it possible that, in 50 years of intensive quasar research, none has yet found even one suitable parameter?

2.5.2 *Exploiting Quasar Variability*

Let me recall that optical variability has been established as an identifying property of type-1 AGNs for more than three decades. AGN typically show continuum variations by 1–2 magnitudes with timescales ranging from days to years. Broad emission lines have also been found to vary. True, every attempt to find a periodicity in quasar variability patterns failed. A period–luminosity relationship is not even to be mentioned! But can we somehow exploit the variability of quasars?

A key idea is to consider that emission line variations lag the continuum variations with delays ranging from a few days to months in luminous Seyfert 1 nuclei. The cross-correlation function between the continuum and the emission line light curve then measures a time lag Δt_{obs} due to the travel time needed by continuum photons to reach Broad Line Region (BLR). This means that the distance of the BLR from a supposedly point-like, central continuum source can be simply written as

$$r_{BLR} = \frac{c\Delta t_{obs}}{1+z},$$

where the factor $(1 + z)$ reduces the observed time lapse to the time lag in the rest frame of the quasar. The evaluation of Δt_{obs} follows from several assumptions, mostly untested. Some of them may be even physically unreasonable. The coupled effects of a broad radial emissivity distribution, an unknown angular radiation pattern of line emission, and suboptimal temporal sampling of the light curve can cause errors that are difficult to quantify.

At any rate, the basic idea beyond exploiting time delays is to measure the linear size of a chosen structure from light travel times (i.e., in a way independent from H_0), and an angular size from a resolved, direct image of the same structure [156]. If we were able to measure the angular distance of the BLR from the central continuum source, then we could recover the cosmological angular distance d_A between us and the source, which could be written as

$$d_A = \frac{r_{BLR}}{\theta_{BLR}} = f(z \mid H_0, \Omega_\Lambda, \Omega_m, \Omega_k) \approx f(z \mid H_0), \text{if } z \ll 1.$$

But if the determination of r_{BLR} can be fraught by large uncertainty, the measurement of an angular size is even impossible. Trouble is that angular size measurements of the BLR are prohibitive with present-day technology: for the Seyfert 1 nucleus of NGC 5548, one finds $\Delta t \approx 21$ days; at $z = 0.017$, which means an $d_{BLR} \approx 50\,\mu$arcsec. Even if the BLR linear size increases with luminosity as

$\propto L^{0.7}$ [268], resolution better than 10 μarcsec is still required to resolve nearby quasars: for example, 3C 273 has $\Delta t \approx 387$ days at $z = 0.158$, and this delay corresponds right to ≈10 μarcsec. Clearly, the method is not yet applicable to any AGN near and far, although some measurements should become feasible with optical interferometers that are presently under development, like the Space Interferometry Mission Planet (SIM PQ) proposed to NASA [564].

In the meantime, methods based on time delays are probably bound to remain model-dependent until the required angular resolution will be achieved. A tickling attempt of the model-dependent kind has been based on the observed time-scale difference between continuum flux variations at different frequencies. No angular size measurement is required at the expense of a major assumption: the existence of an accretion disk around a massive BH. It is known that variation timescales are shorter at shorter wavelengths, so that the observed time delay $t(\lambda)$ in the optical-UV continuum is wavelength-dependent. The assumption that the optical-UV continuum is emitted by an illuminated, geometrically thin accretion disk allows to scale the disk radial temperature $T(r)$ with $r(\lambda)$, that is, with the delay multiplied by the speed of light [110]. The observed specific flux is then

$$f_\nu \propto t^2 d^{-2} \lambda^{-3}.$$

This relationship can give an H_0-independent distance d, suggesting $H_0 \approx 42 \pm 9$ from data for the Seyfert 1 nucleus of NGC 7469.

The story is different if there is an intervening galaxy between us and the quasar, and especially if the galaxy is not perfectly aligned with the quasar along our line of sight. In this case, a galaxy (or any other massive object like a cluster of galaxies) acting as a gravitational lens yields multiple, asymmetrically displaced images of the quasar. Following the intrinsic light variations of the quasars, one measures different time delays for the displaced images due to the path-length difference between the quasar and the earth, and also due to the gravitational effect on light rays traveling in slightly different potential wells. As a consequence, the computation of H_0 requires model-dependent assumptions on the gravitational potential of the intervening galaxy. The resulting H_0 value is usually below or in agreement [496] with the value obtained from the Cepheids, $H_0 = 72 \pm 8$. As multiple images often show an accessible angular separation, the method is promising and several campaigns (e.g., Supernovae and H_0 for the Dark Energy Equation of State (SHOES), COSmological MOnitoring of GRAvItational Lenses (COSMOGRAIL)) are underway to fully exploit its long-term potential. But I will not dwell on that, as I suppose that your question refers more to the intrinsic properties of quasars.

2.5.3 Quasar Diversity and Quasar Evolution

At high z, we are observing quasars that can be very similar to the AGN we are observing at low z, in terms of line width, prominence of singly ionized iron emission, and equivalent widths of other emission lines [538]. Luminosity effects remain weak and prone to sample biases. As we will see better later, there are samples where the main luminosity correlations (the Baldwin effect) are not significant while several properties correlate with the luminosity-to-mass ratio L/M. This can be actually *measured* as the ratio between total luminosity (i.e., bolometric) and mass of the central compact object of a quasar. It is also important to stress that L/M is proportional to the Eddington ratio, that is, the ratio between bolometric and Eddington luminosity, which is considered, under some conditions, a limiting luminosity for the accretion process. We will exchange L/M and Eddington ratio as synonyms in the following.

The lack of a strong luminosity dependence may reflect a self-similarity in the accretion process, which is as yet not fully understood and exploited for quasar modeling, even if phenomenological analogies between accreting systems with stellar-mass BHs (the so-called "mini-quasars") and the supermassive BHs found in AGN are now recognized [360]. Two daring people [599] even suggested an analogy between quasars and two accreting white dwarves (not even BHs!), showing an optical emission line spectrum excitingly similar to the one of I Zw 1, the prototypical Narrow Line Seyfert 1 (NLS1) nucleus – save for different line widths, and save a factor of 10^8 in the mass of the accreting compact object.

True, it is also known that there is a strong *luminosity evolution* of quasars: we live in an Universe that, locally, does not possess luminous quasars. The most luminous quasars we detect now shone a long time ago, and very far from us, at $z \approx 2$ [448]. But now we see, among galaxies in the local Universe, the signature of that brilliant past: very massive compact objects (even a few billion solar masses) that were once accreting material at a pace large enough to sustain an enormous luminosity, and that are nowadays literally extinct, or accreting at very low Eddington ratio. On the other hand, not very far from us, we see sources accreting close to the Eddington limit, but whose masses are by far less than the ones of the quasar population that was once so luminous. As we shall see later, here is where the quasar spectral diversity comes in. For the moment, let me still consider in some more detail how luminosity seems to affect quasars.

2.5.4 The Baldwin Effect

Baldwin and co-workers noticed almost 30 years ago an inverse correlation between the equivalent width of the CIVλ1549 emission line and the apparent luminosity of bright quasars [21, 22]. Quoting the original 1978 paper [21],

> The data indicate that the luminosity of QSO emission lines increases as the 1/3 power of the continuum luminosity.

In other words, the lines, even if their luminosity increases, become less prominent over the underlying continuum with increasing luminosity. In the origin, the effect was believed to be fairly strong, with the equivalent width of CIVλ1549 decreasing proportionally to $L^{-\frac{2}{3}}$. Jumping to present times, if we focus the attention on lines emitted by ions of ionization potential $\gtrsim$50 eV (which we call high-ionization lines for brevity), then we observe a significant luminosity dependence. However, it is important to stress that the *Baldwin effect* has survived as a much weaker and very loose anticorrelation between specific luminosity and high-ionization lines equivalent width. Claims and counter-claims of a Baldwin effect on the basis of small samples (few tens of objects) are unreliable; the statistical weakness of the Baldwin correlation implies that the effect becomes significant only if a very large range in luminosity is considered, $4 \div 6$ decades, as also confirmed by Monte-Carlo simulations [536]. Results until mid-1999 have led to a standard scenario in which the slope of the Baldwin relationship between logarithm of equivalent width of CIVλ1549 (the most widely studied high-ionization line) and luminosity is ≈ -0.15, and not $-\frac{2}{3}$ as originally thought. The Baldwin effect seems to occur in all measurable high-ionization lines except NVλ1240, and the slope of the anticorrelation increases with the energy needed to create the ionic species from which a given line originates. These results have been basically confirmed by more recent studies based on large quasar samples [115, 218]. The anticorrelation of CIVλ1549 remains very weak, however, and cannot be exploited, as it is, for any cosmological purpose. The left panel of Fig. 2.6 shows the disarming spread of data points in a low-redshift sample.

Yet, neither do I share the pessimistic opinion that sees the Baldwin effect as a stalwart of a dogmatic view in quasar and cosmology research, nor do I share

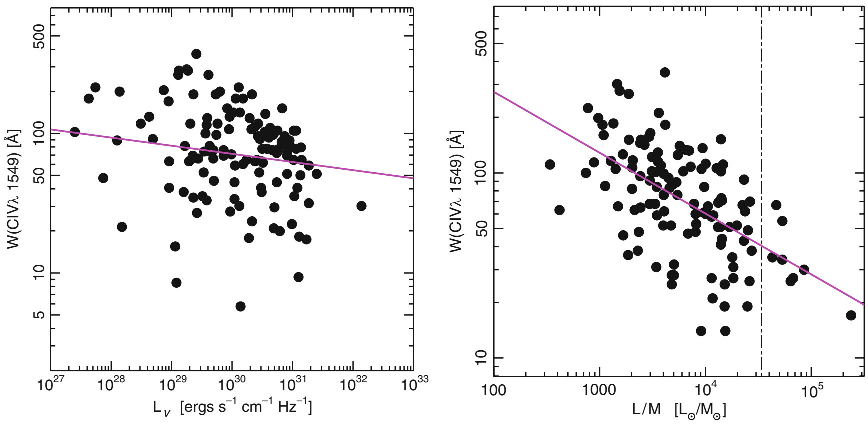

Fig. 2.6 *Left panel*: The weakness of the Baldwin effect in a sample of low-z quasars for which measurements of the CIVλ1549 high ionization line come from archived HST observations [540]. Abscissa is the specific luminosity; ordinate is the rest-frame equivalent width of CIVλ1549. *Right panel*: the Eddington ratio dependent "Baldwin effect", for quasars of the previous sample with a BH mass estimate from the Hβ line width. The abscissa is the luminosity to mass ratio expressed in solar units. log $L/M \approx 4.53$ corresponds to unity Eddington ratio (*dot-dashed vertical line*). We use the term "Baldwin effect" in an improper way here, as it customarily means an inverse correlation with luminosity. The inverse correlation with Eddington ratio is, however, much stronger than the one with luminosity

dogmatic views that have been widely accepted in the past and that considered the Baldwin effect as a "must be". What the Baldwin effect basically tells us is that the quasar spectrum is not fully redshift/luminosity independent. What it says is that quasar spectra become systematically of lower-ionization with increasing redshift/luminosity, as low-ionization lines seem to be even less affected by luminosity, or not affected at all [538]. And that there is a large spread in equivalent width of high-ionization lines for a given specific continuum luminosity. Both aspects cannot be ignored. A few years ago, our group suggested that the Baldwin correlation may mainly reflect a combination of quasar luminosity evolution and of selection effects in the L/M ratio [536].

It is still debated whether the Baldwin effect is primarily evolutionary in its nature, but at this point a parenthesis must be opened to consider that quasar spectra are not all self-similar. Even if luminosity is not what affects spectra at most, there is still a considerable diversity in quasar spectral properties. From the optical/UV spectra, we go from sources of low overall ionization, prominent singly-ionized iron emission, and relatively narrow HI lines, low-equivalent width of CIVλ1549 to objects with weaker or almost absent FeII emission, broader lines, prominent high-ionization CIVλ1549. If we consider a CIVλ1549 equivalent width versus luminosity plot for low-z quasars, then we see that a considerable source of scatter is added by some low CIVλ1549 equivalent width sources at $\sim 10 \div 30$Å that tend to blur the Baldwin relationship. These sources are of the "low-ionization" kind, and include NLSy1s. Just a few years ago, two groups working on the CIVλ1549 emission feature from archived HST observations realized that CIVλ1549 equivalent width correlates much more strongly with L/M rather than with luminosity [19,27]. Figure 2.6 shows the luminosity and the Eddington ratio correlation side-by-side for the same sample [540].

Actually, Eddington ratio seems to be relevant not only for physical conditions but also for the dynamics of the BLR gas. Low CIVλ1549 equivalent width sources can show prominent blueshifted CIVλ1549 with respect to the quasar rest frame, while sources with more prominent CIVλ1549 show no large shifts, at least at low-z [540]. Almost all low-redshift quasars belong to a sequence in the plane defined by the Full Width Half Maximum (FWHM) of Hβ and by the prominence of FeII emission in the optical spectrum. The main variable that seems to govern this sequence is again Eddington ratio [536]. We can safely conclude that sources with lower ionization spectra, including NLSy1s, are the ones radiating at higher Eddington ratio, although quantifying each quasar's Eddington ratio from spectral parameters is still an open issue.

2.5.5 Exploiting the Luminosity-to-Mass Ratio

It is now possible to glimpse a way out from the impasse. No strong dependence on luminosity, but several, easily measurable spectral parameter correlate pretty well with L/M. If we can find a very tight correlation with maybe a linear combination of observed spectral parameters and L/M, and we are so clever to measure the central BH mass, then it is obvious that we could retrieve a redshift-independent value of luminosity [340]. It is not yet possible to do that in a meaningful way. What we miss here is the equivalent of an Hertzsprung–Russell Diagram (HRD) for quasars. We have a view that is sketchy at best. Mass estimates have become an easy exercise, applied even to samples of tens of thousands of quasars, but they rely on two major assumptions: (1) that the gas motion giving rise to the line Doppler broadening is predominantly virial, and (2) that the size of the BLR correlates with optical or UV luminosity following a power law of index $0.5 \div 0.7$ [268]. This means that the M can be simply written as

$$M \propto r_{\rm BLR}{\rm FWHM}^2 \propto L_{\rm opt}^{\alpha}{\rm FWHM}^2,$$

where the FWHM of a suitable line is considered, preferentially a low-ionization line like the Balmer Hβ or MgIIλ2800 [352]. These mass estimates have a statistical value as the inferred uncertainty for individual sources (a factor of 3 at best) is still large. One can then apply a bolometric correction to retrieve the Eddington ratio, an approach seemingly rough but relatively stable as a matter of fact [341].

AGN have proved harder to understand than main sequence stars, whose physical characteristics are determined mainly by a single parameter, that is, mass, to a lesser extent by metallicity and age, and to virtually no extent by orientation. A 2D parameter space (the HRD) is sufficient to characterize both main sequence and nonmain sequence stars. It has become obvious that this is not possible for AGN, even for those with broad emission lines. The aspect-dependent phenomenology due to accretion of matter constrained in an accretion disk demands that at least an aspect parameter θ be taken into account. We think that, for typical quasars, the orientation parameter θ varies from a few degrees to $45° \div 60°$, beyond which the object appears as an obscured, type-2 source, and its BLR view is hidden from us [565]. Orientation matters in the width of the emission lines – this is known since almost 20 years and has been confirmed by several later studies – but we still do not know how to estimate the θ angle in individual sources, with the exception of a few very special objects [537]. We can be confident that orientation effects are a factor $\approx$2 in radio-loud sources, but we are afraid that the effects could be much larger if line emission is constrained in a strongly flattened system. This could be the case of at least some radio-quiet quasars but, at present, none knows for sure.

2.5.6 Guessing Further...

Where do we go from here? Is this all quasars can tell us? We do not have a satisfactory theory that connects accretion parameters of quasars to the structure of the line emitting regions and to measurable spectral properties. However, not even Galileo could count on theory of optics when he mounted his telescope, so please allow me a little explorative calculation. The question is too important to be dropped in this way. After all, we learned a few things since early attempts to exploit the most luminous quasars' Hubble diagram to confirm the expanding Universe scenario [20]. As already pointed out, we have been able to estimate masses and Eddington ratios for large samples of quasars. I know not everyone will agree with me but, as far as the present-day, "best" data are concerned, one can make two cautious statements. The first is that there is no convincing evidence of sources radiating above the Eddington limit. The second is that the extremely large BH masses are not observed. Rather, data are tantalizingly consistent with a maximum BH mass $\sim 5 \times 10^9\,M_\odot$. This mass limit is consistent with the largest spheroids observed in the present day Universe, for which there is a known spheroid–BH mass relationship [539]. Also, the mass function of BHs in quasars at $1.6 \lesssim z \lesssim 2.6$ seems to drop sharply above that value [573].

It is also important to stress that the BH mass follows from a measurement of time delay. Even if we then employ a correlation with luminosity [268], the mass value is independent from H_0, as changing H_0 will change the luminosity but will not affect the mass, that is, changing H_0 will just produce an horizontal shift in the plane r_{BLR} vs. L [550]. Therefore, one is tempted to assume that the most luminous quasars are the ones radiating at their Eddington limit and at their largest mass. In other words, we can estimate the maximum intrinsic luminosity of quasars. So we can derive the apparent magnitude of the brightest sources at each redshift. And this is a function of the cosmological parameters.

Stated this way, the argument is too simplistic. It is known that there is a strong luminosity evolution with cosmic age. But the argument could still be applicable to the most luminous quasars. We know that the luminous quasar population peaked at $z \approx 2$ [448]. Since then, quasars faded and some of them became even almost extinguished by our detection standards. So, sources at $z \ll 2$ should be considered with care: they may radiate well at Eddington ratio (we have very good examples of local AGN radiating close or at Eddington ratio), but the accreting masses could be lower than the assumed maximum mass due to the quasar strong luminosity evolution. If we consider the B photometric band, then one should take into account that $z \approx 1.6$ and $z \approx 3.0$ imply contamination by the strong UV lines of C IV λ1549 and hydrogen Lyα. In addition, shortward of Lyα, quasar spectra often show very complex patterns of narrow absorptions, due to the Lyα absorption by neutral gas clouds between us and the quasars. But why should not we give a look at the brightest quasars at least in the range $1.6 \lesssim z \lesssim 3.0$ once we keep in mind these problems?

It is really intriguing that observational data seem to constrain the cosmography. In Fig. 2.7, the three dashed lines describe the expected magnitudes for sources radiating at Eddington limit as a function of redshift for three different H_0 values (50, 75, 100). A simple k-correction, appropriate for sources at $z \approx 2$, has been assumed. Large H_0 values are not favored; rather, the observed brightest B magnitudes (corrected by Galactic absorption) apparently favor a small value of H_0, $\approx$50, not unlike methods based on gravitational lenses. Data collected from the Hamburg-ESO survey are the most homogeneous and are therefore preferred, but the previous "cosmological" conclusion is not strongly affected if we consider a query for quasars in the 12th edition of the catalogue by Véron-Cetty & Véron [572].

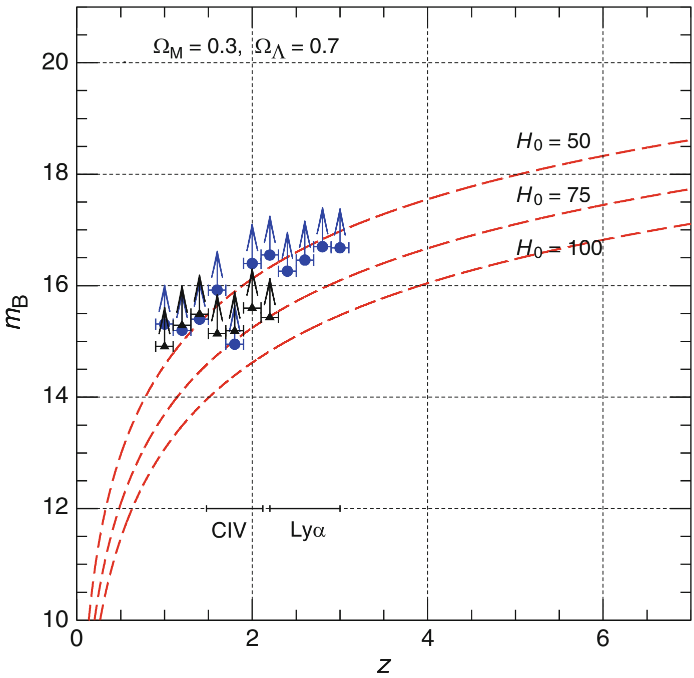

Fig. 2.7 The *dashed lines* show the expected magnitudes of the brightest quasars as a function of redshift for three different values of H_0, assuming a flat Universe with $\Omega_\Lambda = 0.7$ and $\Omega_m = 0.3$, and a "maximum" BH mass $M_{\rm BH} = 3 \times 10^9\,{\rm M}_\odot$. The data points represent the blue magnitudes (corrected by Galactic absorption) of the brightest quasars, in bins of $\delta z = 0.2$. *Blue filled circles*: Hamburg-ESO survey; *grey triangles*: data from the 2006 edition of the Véron-Cetty & Véron catalogue of quasars and AGN [572]. In this case, no data points were plotted for $z > 2.2$, as all brightest sources belonged to the Hamburg-ESO survey. The segments indicate the redshifts at which the strong emission lines of CIV1549 and Lyα are in the B passband

Of course, a cosmological inference can be entirely washed away if quasars are subject to a physical limit in their mass accretion rate, so that very massive BHs never radiate close to Eddington ratio, or, even simpler, if surveys missed the brightest quasars in the sky.

Resuming a cosmological interpretation, if we now compare models with $\Omega_\Lambda = 0$ and $H_0 \approx 50$ fixed with the same observational data of Fig. 2.7, we find that $\Omega_m \rightarrow 1$ is not favored: otherwise we would observe too many bright quasars. In the redshift range $1.6 \lesssim z \lesssim 2.5$, where our comparison seems safer, there is a discrepancy of $\sim$1 magnitude between the observed magnitudes and the prediction for $\Omega_m \approx 1$. But I am not going to show you this diagram.

I am anyway tempted to say that we live in a Universe where quasars become systematically fainter with redshift; the way they do and the way they shine when most luminous suggest a pretty "large" Universe by current standards, consistent with the currently accepted value of Ω_Λ or, if $\Omega_\Lambda = 0$ and $\Omega_{\rm k} = 0$, with an open Universe described by a small value of Ω_m.

Of course, the previous computations are only for illustrative purposes, even if they may not be very far from correct results. Quasar properties pose an enormous breath of challenges, and so it may not now sound surprising if quasars have not been yet used as widely successful cosmological probes. Nonetheless, I hope to have given you a glimpse of how a better understanding of quasar physics and evolution could make a difference for cosmology.

Thanks Paola for this overview of quasar properties. The nature of these objects has been a mystery since their discovery in the 1960s. Anyway, speaking of quasars, it is impossible to avoid the question of the anomalous redshifts observed for some of them. This is strictly connected to the cosmological interpretation of redshift on which we have founded our cosmology. We are now going to interview Jack Sulentic, who has always manifested his doubts about this interpretation.

2.6 The Heretical View on Cosmological Redshifts

Dear Jack *(Sulentic)***, are there viable alternatives to the cosmological explanation of redshift of distant galaxies and quasars?**

This is one of the questions posed in this book dealing with controversial ideas or observations. They were controversial 30 years ago and they are controversial now. Of course, many will say that the answers are well known. That the observational evidence has been refuted. That the subject is closed. I could respond to your question by restating all of the empirical evidence and reviewing all of the alternative ideas. But to what point? They (and technical references) can be found in many places (e.g., books such as "New Ideas in Astronomy" [43] and "The Red Limit" [171] or in the film "Universe: The Cosmology Quest" [357]). I will restate and update some of these results/ideas, but repeating them in detail would be a waste of time. The responses to the posed question will vary greatly from person to person. Some will say that these issues have been settled and others will say that they have never been seriously considered. I like to believe that I am somewhere in the middle and hope my responses reinforce that impression.

Moving on to the question.

There are actually two different questions that should be inferred from such a query:

1. Is there a need for an alternative explanation? Is there empirical evidence that places the standard explanation in doubt?

and

2. Whether needed or not, are there mechanisms capable of producing a pseudo-Doppler redshift?

Perhaps many careerists[2] would refuse to discuss the question beyond calling it nonsense. Others would refuse to decouple the two questions – stating that no

[2] *See the author definition of careerists and Baconians in Chap. 4.*

empirical evidence would be convincing to them unless a suitable physical mechanism were already known. At the risk of being labeled a Baconian, I would argue that empirical evidence should be judged independently of existing beliefs or ideologies. If a need for new mechanisms were felt, then I am sure that we would see many alternatives published in short order and by careerists who would see alternative models as a new avenue to advancement. Instead, should we believe that we live at a special time in the history of science where most of the basic laws have been discovered? Even if it were true one should resist the inclination to approach science this way because it would essentially preclude new discovery.

There have been vocal advocates for the view that empirical evidence does exist – even claims of overwhelming evidence for non-Doppler redshifts (e.g., [13,14,80–82,554]). Are iconoclasts always Baconian or can they also be careerists? Having spent a number of years exploring much of the evidence, I would not describe it as overwhelming but I would argue that it deserves to be taken seriously and tested. Some of those evidence will be discussed later. Much of it involves apparent associations between objects (e.g., galaxies and quasars) with significantly different redshifts.

Perhaps, the fairest thing to say is that the evidence for Doppler redshifts is somewhat more compelling than the evidence against. It is very difficult to look evenhandedly at both sets of evidence. The most compelling evidence for the paradigm is not, in my opinion, evidence for evolution in the Universe. Recall please the earlier mentioned low S/N observations of high z sources using large telescopes. The lack of evidence for evolution in the Universe is what surprises me. We find "old" galaxies at high redshift and we find quasars that are spectroscopically similar over the entire redshift range ($0.1 \lesssim z \lesssim 6$) that they are observed. This includes super-solar heavy element abundances in the highest redshift quasars. It also includes inferred BH masses as large as $\log M_{BH} \sim 9.5\,\mathrm{M}_{\odot}$ at all redshifts. One sees what one wants to see, and if one is a careerist, it is very difficult to see anything wrong with the paradigm that has advanced ones career. If a problem arises (e.g., super solar abundance in $z \sim 6$ objects), agility becomes a necessary careerist skill – add a small additional complexity (i.e., everything interesting happened before $z = 6$, where we could not see it happen because the quasars were enshrouded by dust or something—a new form of DM!) to the standard model and it fits! I think the most impressive support for Doppler redshifts is something much less subject to mixed interpretations. Something like gravitational lenses (and arcs) where we see multiple images of the same (?) high redshift quasar lensed by a lower redshift galaxy that is almost certainly in front of it.

Summarizing my response to the question: It is not clear that there are viable alternatives but few are looking for them. Feynman was not convinced by the evidence for non-Doppler redshifts, but said that, if convinced, he would look for a mechanism involving the correlation of light.

Given the success of theories in explaining so many high-energy phenomena in quasars and related sources, does it make still sense to reason like Faraday when we already have Maxwell's theory available?

What do you mean by success? Most astronomers know about gravity. Some (X-ray astronomers and jet theorists) understand aspects of plasma physics and Magneto Hydro Dynamics (MHD). Few observations allow us to constrain well effects driven by electromagnetic forces. It is still nontrivial to measure magnetic fields in galaxies and quasars. Theorists can build models making sophisticated use of Maxwell's Equations. Again observations of jets are one area where the cross-talk between theory and observation has become quite sophisticated. Maybe there are other areas that I am unfamiliar with.

As BHs are mentioned in the questions and as they are regarded as the central engine driving quasars, perhaps they provide a better way to answer this question, or to illustrate why the answer to the question is "yes". We hypothesize a central supermassive object in quasars and that quasar activity manifests accretion onto this BH. We cannot see it and we cannot distinguish between different BH structures driven by spin, but we think that an accretion disk surrounds most BH and gives rise to much of the line emission. For a brief period, we thought to be able to distinguish Kerr and Schwarzschild configurations via the earlier mentioned 6.4 keV Fe Kα line arising at the inner edge of the accretion disk. But that was with low S/N (ASCA) data. Now we have higher S/N (XMM-*Newton* and *Chandra*) X-ray data and they rarely show the signature that we thought we saw before. Current attempts to prolong this game are sad and desperate. Recall the previously mentioned difficultly in getting time on large telescopes for high S/N observations in extragalactic astrophysics and cosmology? No careers have advanced with negative results or refutations involving better data.

The current desperate game involves estimating the masses of the central BHs, especially, estimating masses for high redshift quasars and comparing them to the local ones. The paradigm says that they should be smaller at earlier times, because they are thought to grow by accretion and mergers. But earlier predictions have a habit of being forgotten if the evidence points in another direction. We use emission line widths (or velocity dispersions) and assume that the lines arise in a virialized distribution of emitting clouds. This seems to work best for the Balmer lines of hydrogen. But the good lines are lost (redshifted out of the visible) at quite modest redshift $z \sim 0.7 \div 0.9$, then what do we do? Use other lines that from their measured properties do not likely arise from the clouds producing the Balmer lines and that show characteristics that throw the virial assumption into doubt? Virial assumption ($2T + P = 0$; here T and P are kinetic and potential energies, respectively)! We are talking here of potential energy and kinetic energy not of Maxwell's equations: Undergraduate physics. This embarrasses careerists who respect, and like to show off, complexity and technical prowess (Albert Einstein also mentioned this specific type of careerist in his tribute to Max Planck, see, e.g., [171]). It does not embarrass Baconians – we are where we are and we try to always move forward without jumping over the problem. We follow the Balmer lines out to $z \sim 3$ [539] and find the same large BH masses that are found locally. In fact, within about $z \sim 1$, the largest BH masses may be smaller than the ones observed at higher redshifts. There is some evidence that this BH obesity problem may extend to $z \sim 6.5$, as far as quasars are currently observed. There are a few areas, for example, modeling the details of jets, where Maxwell's equations can be applied and comparisons can be made between models and observations. But these detailed models do not tell us how jets are produced or why only a few percent of quasars manifest them.

2.6.1 On the Wolf Effect

Dear Jack *(Sulentic)*, **among plasma physicists, the Wolf effect has been considered a possible cause of noncosmological redshifts. Can you please explain why this effect should be relevant under the physical conditions expected for line emission from quasar? Can this effect account for the internal shifts observed between lines emitted by ions of widely different ionization potential?**

Emil Wolf is an impressive scientist who easily satisfies the definition of a truth seeker (Baconian). He was therefore rather naive when he suggested that a mechanism that came to be called the Wolf effect might have an application in astrophysics. This was a mechanism capable of producing non-Doppler shifts in spectral lines. Any application offering this possibility will be dismissed by careerists because any demonstration of a non-Doppler component, however small, could be said to open Pandora's Box. I accept some of the blame for encouraging him to explore such possibilities. I think he was genuinely surprised by the rancor and hostility that greeted his suggestion. The Wolf effect can be included in a general category of scattering mechanisms that can in principal (i.e., given the proper set of physical conditions) shift line emitting photons to longer (or shorter) wavelength. Others include Compton and Raman scattering mechanisms. Compton downshifting of photons is well known especially among X-ray astronomers. All of these mechanisms might play a roll in complex sources like quasars. Compton scattering was invoked [516] to explain a significant (but not cosmological) redshift observed in the 6.4 keV X-ray line discovered in many low redshift quasars in the 1980s. It was not warmly welcomed but now higher S/N spectra reveal that most of the redshifted lines were not real.

Scattering mechanisms do not seem promising as a way to produce all or most of the cosmological redshift. I am not qualified to discuss such models in detail, but the empiricism can provide first-order constraints for such models. Producing small shifts or asymmetries in emission lines is one thing, but shifting the bulk of the photons outside the envelope of the intrinsic (rest frame) line is quite another. A scattering process generally broadens a line and alters its shape. The broad and complex emission lines in quasars offer a tempting target for scattering applications. Electron scattering, for example, almost certainly has some small effect on emission line structure in quasars. Unfortunately, a large fraction of quasars also show one or more narrow emission lines with the same or very similar redshift as the broad ones. Scattering mechanisms also often produce wavelength-dependant shifts. In recent years, we have been able to compare the emission lines at UV, Optical, and IR wavelengths, and we find that all lines in a quasar show the same redshift within a scatter of at most $4 \div 5 \times 10^3\,\mathrm{km\,s^{-1}}$ (UV emission lines do show a systematic blueshift relative to optical lines in perhaps 60% of quasars). In summary, we do not know enough about the physical conditions within the central regions of quasars to rule out scattering mechanisms, but to produce a pseudo-cosmological redshift, they would have to scatter all of the lines from Lyα to Paschen α in a way to produce the same redshift and preserve their intrinsic widths. If I were looking for an astrophysical application of the Wolf effect in quasar astronomy, it would be to explain small scale shifts and asymmetric differences within a source.

In summary, the Wolf effect appears most promising for explaining smaller line shift and shape anomalies. It does not appear promising as a mechanism that might produce non-Doppler redshifts that could mimic observed cosmological redshifts.

2.6.2 Anomalies with Quasars?

Dear Jack *(Sulentic)*, **apparent connections like luminous bridges and tails between sources of widely different redshift (typically, a nearby galaxy and a distant quasar) have been known since long. A seminal case, the connection between NGC 4319 and Markarian 205, provoked a vigorous debate until a post-COSTAR HST image led to claims that the two sources are widely separated by time and space. In the last few years, several new odd sources have been discovered or studied with more advanced instrumentation, notably around NGC 7603, and close to the nucleus of NGC 7319. What is the astrophysical significance of Mark 205 and of the other alignments/superpositions? Can they be dismissed as chances or are they extremely unlikely occurrences that straightforwardly point toward a problem with our current understanding of redshifts?**

NGC4319 + Markarian 205 and NGC7603ab are two of the most famous examples of apparently connected objects with very different redshifts. Perhaps the strongest argument against their physical reality involves their rarity rather than their redshift differences. If all or most of the high redshift quasars are ejected from the nuclei of low redshift galaxies, as Arp has hypothesized, then one might expect to see many more with luminous connections. Of course, one can argue that they are ejected at high velocity but their rarity argues in favor of the chance projection hypothesis. Another problem: Markarian 205 is embedded in a host galaxy of like redshift. If quasars represent young matter ejected from galaxy nuclei as quasars with high non-Doppler redshifts, which subsequently develop a host galaxy, then why do we still see a connection in this case – where host galaxy of the quasar is already developed?

Sources like NGC4319+Markarian 205 are feared by careerists but are viewed with amusement and curiosity by Baconian types. These puzzles make science more fun – it must be a lot less fun when one is constantly worried about keeping an unblemished reputation and career advancement. Fear of these puzzles has been manifested in the way that the counter evidence has been advanced and accepted. Any logically fallacious and/or statistically weak argument against their reality is almost immediately accepted and circulated; the recent culmination of this trend – surprising since there has been no technical discussion of NGC4319-Markarian 205 for almost 20 years – in an above mentioned HST press release declaring the case closed. The data was never published. Five minutes of manipulation (not modification) of the HST data reveals that the luminous "connection" is still there (as presented in [535]). The evidence supporting the reality of the feature has not been refuted or otherwise explained – it is simply disregarded. If the skeptics are so certain of the correctness of their view, then why are they afraid of the data and why would they warn young people away from studying such data? Even I realize that these strange configurations *must* be accidents – and that is all the more reason why I should study them until I understand them past this superficial reason for rejecting them. I am not upset about any of this because I do not work in this area anymore. However, I could not allow the press release to pass – but the cost? No HST time for any project for 13+ years – with 30+ proposals submitted during that time period.

Lets consider these two famous cases in more detail.

NGC4319+Markarian205 involves a low redshift ($z \sim 0.071$) quasar projected inside the arms of a low redshift ($z \sim 0.004$) spiral galaxy [12, 586]. A low surface brightness luminous filament appears to connect the two objects. Figure 2.8 shows an old image of the configuration that was obtained by Chip Arp many decades ago. It does not show the luminous connection but hundreds of pictures that do can be found on the web. It does shows the unusual brightness of Markarian 205 and the spiral arms of Markarian 205. This source is so threatening that its discoverer was motivated to apologize for discovering it [587]. A number of papers have claimed that the connection does not exist and/or is not what it seems to be. All could and have been easily refuted – at least the existence part. The earlier claims are still accepted and the refutations ignored. It matters little what was said in those papers.

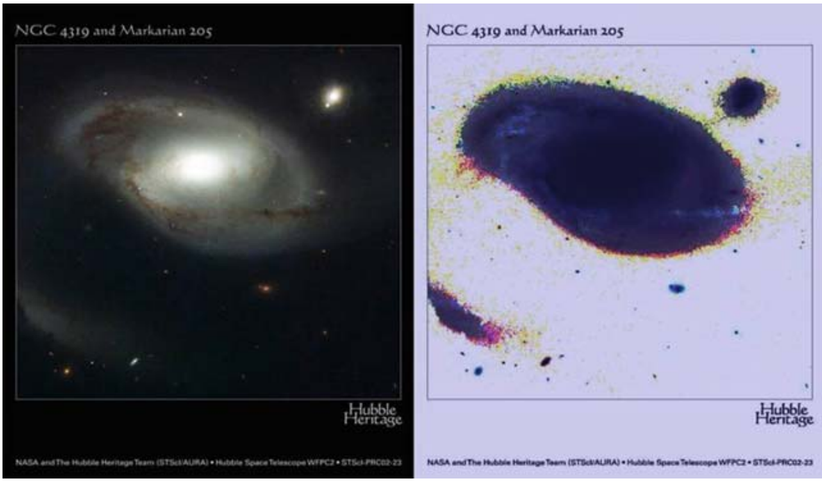

Fig. 2.8 Image of the discordant redshift configuration involving the spiral galaxy NGC4319 and higher redshift quasar Markarian 205. From NASA archive

In other words, the evidence is irrelevant unless it gives the right answer, because we now know so much that evidence alone has no meaning unless it fits into the existing paradigm. Lets assume that in this case this approach is OK. But what about other less blatant cases? What does it say about the way science is conducted? It serves no useful purpose to discuss in detail the different technical papers that have addressed the nature of the filament (see, e.g., [13, 266, 535, 587] or the film "Universe: The Cosmology Quest") as nothing has changed in the past 15 years.

A physical connection between two objects with significantly different redshift would produce a revolution in extragalactic astronomy. Baconians would say that this possibility, however remote, justifies careful study of discordant redshift pairs like the ones mentioned here. Careerists would argue that such configuration must be chance projections meriting no further study. It would be a "waste of telescope time" to study them. Our attempts to study NGC4319+Markarian 205 in more detail with HST were rejected. Perhaps it was thought that people like Arp and Sulentic would be too biased? Some HST time was subsequently given to an amateur astronomer (a high school teacher) who showed us the images he obtained with HST. They confirmed the connection. I warned him that he had got the wrong answer and would find the HST people reluctant to support publication. In fact, the (pre-COSTAR) results were never published. As mentioned earlier, the more recent post-COSTAR images confirm the luminous "connection".

So what is NGC4319+Markarian 205? Assuming the filament is real, and NOT a connection between discordant redshift objects, it is logically either related to "foreground" NGC4319 or "background" Markarian 205. NGC4319 is not alone. A large accordant redshift elliptical galaxy (NGC4291) lies only $\sim$6 arcmin away. This corresponds to a projected physical separation of only a few tens of kiloparsecs. Unless the two galaxies are much more widely separated along the line of sight than their redshifts suggest (they could show very similar redshift and still be separated by a megaparsec), they represent a close pair. NGC4319 does not show a typical Hubble or deVaucouleurs morphology (Fig. 2.8) and the unusual structure could be due to tidal interaction between NGC4291 and NGC4319. This does not, however, explain a high spatial frequency structure like the apparent connection. We observe no other similar features in NGC4319 pointing in random directions. The narrowness favors a tidal feature but at a larger distance and thus associated with Markarian 205. There is a faint object (compact galaxy?) close to Markarian 205 and it apparently shows the same redshift as Markarian 205 [530]. Thus Markarian 205 may not be alone and we know that gravitational interactions can produce tidal bridges and tails. The luminous connection might therefore have nothing to do with NGC 4319 and a lot to do with Markarian 205 and its like redshift neighbor. This interpretation might be testable if suitable telescope time were available. But would it not be a waste of telescope time to confirm what we already know to be the answer? Maybe observing time can be given to the unbiased people responsible for the HST press release, thus protecting astronomy from more biased interpretations.

The other configuration mentioned in the question involves NGC7603ab. This connection between two galaxies is, on many levels, a completely different type of association. The active quasar-like nucleus [289] lies in the assumed "parent" spiral galaxy ($z \sim 0.029$) and the companion involves a smaller early-type (S0 = lenticular) galaxy ($z \sim 0.056$) that shows only a stellar absorption line (i.e., stellar) spectrum. They appear to be connected by a filament that looks like a spiral arm of the parent galaxy. Figure 2.9 shows an image processed version of NGC7603ab using a 5 m Palomar plate obtained by H. Arp in the 1970s. I digitized the photographic image and displayed it in an unconventional way to assess the S/N properties of the lowest light levels, including the "bridge". The plate was scratched but the arm/connection is well seen. Two blobs in that arm/connection can also be seen as small dark spots. Twenty five years ago, we could not observe such faint objects spectroscopically – even with the Palomar 5 m – but I assumed they were HII regions in the spiral arm of NGC7603 and would likely show the same redshift as NGC7603. Arp was not so sure and he was correct! Recent new observations show that they have much higher redshifts ($z \sim 0.245$ and 0.394 [323]). Always a surprise when one looks more deeply with a larger telescope and new technologies.

In Fig. 2.9, the spiral arm/connection appears to terminate at the higher redshift galaxy, but even deeper images show that a much fainter arm extends beyond

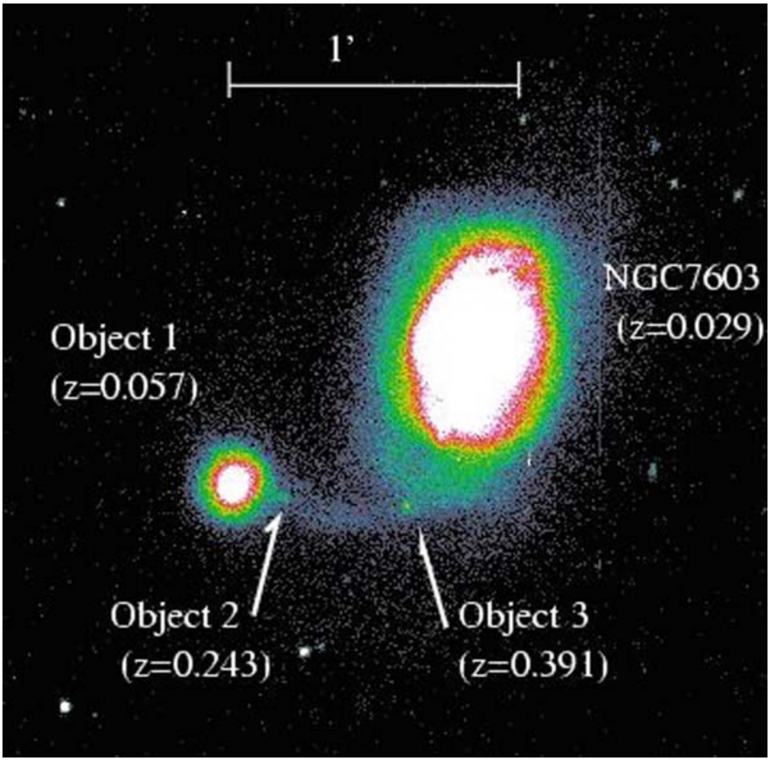

Fig. 2.9 Image of the discordant redshift pair NGC7603ab. The larger galaxy NGC7603a shows a quasar-like nucleus with strong broad emission lines, while the companion shows a higher-redshift stellar absorption line spectrum. Two even higher redshift emission line objects appear as *black dots* on the "bridge" connecting NGC7603ab. Credit [322]

NGC7603b [511], lessening the impression that it is a connection. This result can be used to argue that the two galaxies are unconnected and unrelated. Perhaps a more powerful argument in favor of the connection involves the one-sidedness of the brighter feature interpreted as a spiral arm of NGC7603a. A spiral arm usually has a counterpart on the other side of the galaxy nucleus. Among the thousand brightest spiral galaxies on the sky, this striking asymmetry is almost unique. The next step for a Baconian (we lost the careerists two paragraphs ago) is to try and find another nearby galaxy that can be blamed for this configuration. It should show a redshift similar to NGC7603, allowing us to invoke a tidal interaction as the cause of the asymmetric spiral arm – conventionally NGC7603b and the two higher redshift blobs must be accidental projections along the same direction as that arm. My careerist side fully expected that another galaxy projected near NGC7603 would show a similar redshift. The field was recently included in the SDSS, providing redshift measures for all reasonable candidate galaxies. None of these galaxies show a redshift similar to NGC7603. The single asymmetric arm cannot be explained as a tidal feature because there is no suitable neighbor with whom NGC7603 could interact. A careerist will unblushingly say that the like redshift companion was eaten by NGC7603 just after it produced the asymmetry. Invoking things that we cannot see has become a cottage industry among the careerists (e.g., dark massive objects, dark matter, dark energy, dark galaxies). Naturally this invocation is distasteful to Baconians (after all why do we need observers if most of the Universe is invisible?) – but this explanation may be the correct one.

Moving away from two well-know examples, one can point out that there are now numerous statistical studies that show a clear excess of high redshift quasars in the vicinity of low redshift galaxies (e.g., [470, 534, 603]). The statistical results carry much more weight than any single connected pair of discordant redshift objects. These results are not much discussed unless the cause is attributed to "DM lensing". Always something we cannot see. There is not much room for empiricism in modern astronomy but a wide berth for ideology. No more (or less?) far out would be the hypothesis that discordant redshift quasars are being ejected from the nuclei of lower redshift galaxies [13]. This interpretation involves two new ideas: ejection and discordant redshift compact objects of unknown nature. The ejected objects (quasars?) at the redshift distances of their parent galaxies would not be very luminous, similar to HII regions in the galaxies.

Actually the ejection idea is not so new – a "slingshot" ejection model having been proposed long ago [494]. An amusing recent discovery involves the best example of a (naked) quasar HE0450-2958 projected on the disk of a – fortunately like redshift – galaxy [332]. This discovery recently gave rise to a flurry of papers again, reviving a quasar ejection mechanism and most without crediting the pioneers of this idea (for an exception see [222]). Presumably, because the authors of the original paper showed their disloyalty to the standard paradigm by advocating the mechanism in connection with discordant redshift associations. Such treacherous acts are never forgiven. Have the respectable people who recently revised the quasar ejection hypothesis unwittingly opened Pandora's box halfway? They accept that ejection can occur. Now all that is needed is a model to explain the physical nature of compact quasar ejecta with non-Doppler redshifts. No small order.

In summary, the famous cases involving apparent connected discordant objects deserve detailed study, no matter how certain we are that they are spurious. Our very certainty requires it. At the same time we have many studies showing an excess of higher redshift quasars near lower redshift galaxies. Unless magic (always unseen) effects can explain them, they represent a fundamental challenge to the standard paradigm. If they were not so controversial perhaps, they could have been better utilized to map the DM distribution. Alternatively, an ejection mechanism exists to explain how they got where they are. But no explanation for their discordant nature exists although scattering mechanisms mentioned earlier would be the place to look. Most of these quasars will show both broad and narrow emission lines with the same redshift, so it is clear that any of these potential non-Doppler redshift producing mechanisms must overcome a major challenge.

Thank you Jack.

After the treatment of the main change of paradigm from CDM to ΛCDM and of the use of astronomical candles for mapping the expansion of the Universe, we now move to sections dedicated to the main empirical cornerstones of the standard cosmological model. Having mentioned in previous sections the problem of element abundances as a result of stellar evolution, it is important to understand where the simplest elements come from in the early Universe. We then start with the interviews of Keith Olive and Gary Steigman on BBN, its observational tests, and its relation to the results coming from the analysis of CMB and LSS.

2.7 Cosmological Nucleosynthesis

2.7.1 Theory of Cosmological Nucleosynthesis

Dear Keith *(Olive)*, **cosmological nucleosynthesis is one of the main probe of the Standard Cosmological Model. Could you sketch the fundamental concepts and nuclear reactions involved in BBN theory?**

Element abundances offer several unique probes into physical processes throughout the history of the Universe. Indeed, one of the most fundamental questions in science relates to the chemical origins of the elements and their nuclear isotopes. By far, most of the natural elements are synthesized in stars, but a handful trace their origins back to the first few minutes of the Universe. In fact, BBN offers the deepest reliable probe of the early Universe, being based on well-understood Standard Model physics. Predictions of the abundances of the light elements, D, ^{3}He, ^{4}He, and ^{7}Li, synthesized shortly after the Big Bang are in good overall agreement with the primordial abundances inferred from observational data, thus validating the standard hot Big Bang cosmology (see [173, 382, 582]). This is particularly impressive, given that these abundances span nine orders of magnitude – from ^{4}He/H $\sim$ 0.08 down to ^{7}Li/H $\sim 10^{-10}$ (ratios by number).

All the heavier elements have been synthesized in stars. Abundance patterns and ratios also offer a unique glimpse into the chemical evolution of the Galaxy and Inter

Galactic Medium (IGM). Indeed abundance ratios, observed in systems of varying degrees of metallicity, allow one to trace the star formation history of the Universe, and can in principle determine the very nature of the first stars.

2.7.1.1 BBN Theory in Short

The Universe as described by the Big Bang theory began in an extremely hot, dense, and largely homogeneous state. Today, the Universe is nearly 14 billion years old, but at the time of the formation of the light elements, the Universe had existed only for minutes. Because the Universe cools as it expands, the early Universe was very hot and at the time of BBN, the temperature exceeded 10^{10} K. The density of neutrons and protons was about $10^{17}\,\mathrm{cm}^{-3}$ when nucleosynthesis began. Though small when compared to terrestrial densities, it was far larger than the average density of normal matter in the Universe today, $10^{-7}\,\mathrm{cm}^{-3}$. As will be described later, equilibrium processes governed the production of the light nuclei as the Universe cooled.

The nucleosynthesis chain begins with the formation of deuterium in the process $p + n \rightarrow \mathrm{D} + \gamma$. However, because of the large number of photons relative to nucleons (or baryons), $\eta_{\mathrm{B}}^{-1} = n_\gamma/n_{\mathrm{B}} \sim 10^{10}$, deuterium production is delayed past the point where the temperature has fallen below the deuterium binding energy, $E_{\mathrm{B}} = 2.2\,\mathrm{MeV}$ (the average photon energy in a BB is $\bar{E}_\gamma \simeq 2.7T$). This is because there are many photons in the exponential tail of the photon energy distribution with energies $E > E_{\mathrm{B}}$ despite the fact that the temperature or $\bar{E}_\gamma$ is less than E_{B}. The degree to which deuterium production is delayed can be found by comparing the qualitative expressions for the deuterium production and destruction rates,

$$\begin{aligned} \Gamma_p &\approx n_{\mathrm{B}}\sigma v \\ \Gamma_d &\approx n_\gamma \sigma v\, \mathrm{e}^{-E_{\mathrm{B}}/T} \end{aligned} \tag{2.1}$$

When the quantity $\eta_{\mathrm{B}}^{-1}\exp(-E_{\mathrm{B}}/T) \sim 1$, the rate for deuterium destruction $(\mathrm{D} + \gamma \rightarrow p + n)$ finally falls below the deuterium production rate and the nuclear chain begins at a temperature $T \sim 0.1$ MeV.

In addition to the $p + n \rightarrow \mathrm{D} + \gamma$ reaction, the other major reactions leading to the production of the light elements tritium (T) and ^{3}He are

- $\mathrm{D} + \mathrm{D} \rightarrow p + \mathrm{T}$ $\quad ^3\mathrm{He} + n \rightarrow p + \mathrm{T}$,
- $\mathrm{D} + \mathrm{D} \rightarrow n + {}^3\mathrm{He}$ $\quad \mathrm{D} + p \rightarrow \gamma + {}^3\mathrm{He}$.

Followed by the reactions producing ^{4}He

- $^3\mathrm{He} + \mathrm{D} \rightarrow p + {}^4\mathrm{He}$ $\quad \mathrm{T} + \mathrm{D} \rightarrow n + {}^4\mathrm{He}$.

The gap at $A = 5$ is overcome and the production and destruction of mass $A = 7$ are regulated by

- $^3\mathrm{He} + {}^4\mathrm{He} \rightarrow \gamma + {}^7\mathrm{Be}$, followed by the decay of ^{7}Be

- $^7\mathrm{Be} + e^- \rightarrow {}^7\mathrm{Li} + \nu_e$

as well as the reactions involving ^{7}Li directly

- $\mathrm{T} + {}^4\mathrm{He} \rightarrow \gamma + {}^7\mathrm{Li}$ $\quad {}^7\mathrm{Be} + n \rightarrow p + {}^7\mathrm{Li}$ $\quad {}^7\mathrm{Li} + p \rightarrow {}^4\mathrm{He} + {}^4\mathrm{He}$.

The gap at $A = 8$ prevents the production of other isotopes in any significant quantity.

When nucleosynthesis begins, nearly all the surviving neutrons end up bound in the most stable light element ^{4}He. Heavier nuclei do not form in any significant quantity both because of the absence of stable nuclei with mass number 5 or 8 (which impedes nucleosynthesis via $^4\mathrm{He} + n$, $^4\mathrm{He} + p$, or $^4\mathrm{He} + {}^4\mathrm{He}$ reactions) and the large Coulomb barriers for reactions such as the $\mathrm{T} + {}^4\mathrm{He} \rightarrow \gamma + {}^7\mathrm{Li}$ and $^3\mathrm{He} + {}^4\mathrm{He} \rightarrow \gamma + {}^7\mathrm{Be}$ reactions listed earlier. Hence the primordial mass fraction of ^{4}He, conventionally referred to as Y_p, can be estimated by the simple counting argument

$$Y_\mathrm{p} = \frac{2(n/p)}{1 + n/p} \simeq 0.25\,. \tag{2.2}$$

There is little sensitivity here to the actual nuclear reaction rates, which are important in determining the other "left-over" abundances: D and ^{3}He at the level of a few times 10^{-5} by number relative to H, and ^{7}Li/H at the level of about 10^{-10} (when $\eta_{10} \equiv 10^{10}\eta_B$ is in the range $1 \div 10$).

Historically, BBN as a theory explaining the observed element abundances was nearly abandoned due to its inability to explain *all* element abundances. Subsequently, stellar nucleosynthesis became the leading theory for element production [83]. However, two key questions persisted. (1) The abundance of ^{4}He as a function of metallicity is nearly flat and no abundances are observed to be below about 23%. In particular, even in systems in which an element such as oxygen, which traces stellar activity, is observed at extremely low values (compared with the solar value of O/H), the ^{4}He abundance is nearly constant. This is very different from all other element abundances, with the exception of Li. (2) Stellar sources cannot produce the observed abundance of D/H. Indeed, stars destroy deuterium and no astrophysical site is known for the production of significant amounts of deuterium. Thus we are led back to BBN for the origins of D, ^{3}He, ^{4}He, and ^{7}Li.

The resulting elemental abundances predicted by standard BBN are shown in Fig. 2.10 as a function of η_B [118]. The plot shows the abundance of ^{4}He by mass, Y, and the abundances of the other three isotopes by number. The bands indicate the central predictions from BBN, while their thickness corresponds to the uncertainty in the predicted abundances. The uncertainty range in ^{4}He reflects primarily the 1σ uncertainty in the neutron lifetime.

In the standard model with three neutrino flavors, the only free parameter is the density of baryons that sets the rates of the strong reactions. Thus, any abundance measurement determines η_B, while additional measurements overconstrain the theory and thereby provide a consistency check. BBN has thus historically been the premier means of determining the cosmic baryon density. With the increased

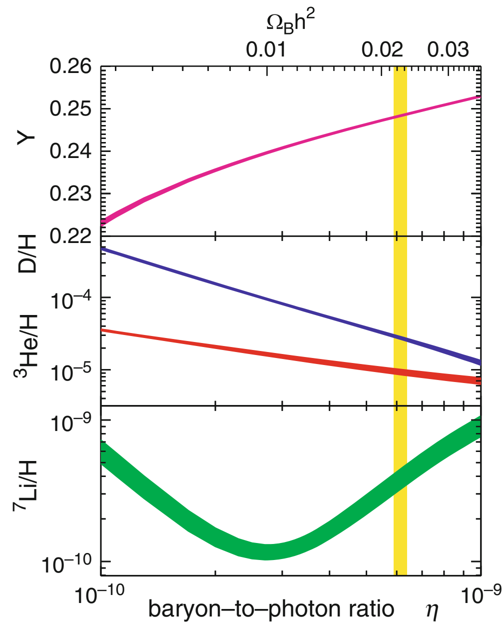

Fig. 2.10 The predictions of standard BBN. Primordial abundances as a function of the baryon-to-photon ratio η_B. Abundances are quantified as ratios to hydrogen, except for ^{4}He, which is given as a mass fraction. The bands give the 1σ uncertainties about the central values of the abundances as a function of η_B, reflecting the uncertainties in the nuclear and weak interaction rates. The vertical band shows the WMAP derived value of η_B. From [118]

precision of microwave background anisotropy measurements, it is now possible to use the CMB to independently determine the baryon density. The third year WMAP data implies [524]

$$\eta_{10} = 6.11 \pm 0.22. \tag{2.3}$$

Equivalently, this can be stated as the allowed range for the baryon mass density expressed today as a fraction of the critical density: $\Omega_B = \rho_B/\rho_{crit} \simeq \eta_{10}h^{-2}/274 = (0.0223 \pm 0.0008)h^{-2}$. This range in η_B is shown as a vertical strip in Fig. 2.10.

The promise of CMB precision measurements of the baryon density suggests a new approach in which the CMB baryon density becomes an input to BBN. Thus, within the context of the Standard Model (i.e., with $N_\nu = 3$), BBN becomes a zero-parameter theory, and the light element predictions are completely determined to be within the uncertainties in η_B and the BBN theoretical errors. Comparison with light element observations then can be used to restate the test of BBN–CMB consistency, or to turn the problem around and test the astrophysics of post-BBN light element evolution [119]. Alternatively, one can consider possible physics beyond the Standard Model (e.g., with $N_\nu \neq 3$) and then use all of the abundances to test such models.

Thank you Keith.

Dear Gary *(Steigman)*, **what are in your opinion the key aspects of the BBN theory and how are the primordial abundances of the light element isotopes**

(D, ^{3}He, ^{4}He, and ^{7}Li) connected with the ratio between the densities of baryons and photons and the expansion rate of the Universe?

The present Universe is observed to be expanding and to be filled with radiation, the CMB radiation whose spectrum is very precisely that of a BB at a temperature of 2.725 K. As the Universe expands, the average density of all its constituents decreases and the temperature of the CMB decreases as well. Conversely, in the past, the density and temperature were higher; the earlier the epoch in the evolution of the Universe, the hotter and denser were its constituents. The very early Universe was a hot, dense primordial soup of all the particles we know (from accelerator experiments and from the "standard model of particle physics"). The physical properties of the Universe, such as its rate of expansion and the temperature and density of its constituents can be tracked quantitatively using the standard Friedmann–Lamaître–Robertson–Walker (FLRW) cosmology, based on Einstein's theory of General Relativity (GR). According to this "standard cosmological model", when the Universe was only a fraction of a second old, the temperature corresponded to a thermal energy in excess of a few mega electron volt and the density was very high, so that collisions among the particles present at that time (neutrons, protons, electron-positron ($e^{\pm}$) pairs, neutrinos and photons – the blue-shifted CMB photons) were very rapid compared to the rate at which the Universe was expanding. It is during this epoch, from a fraction of a second to several minutes, that collisions among the neutrons and protons build the light elements, deuterium, helium-3, helium-4, and lithium-7 in primordial nucleosynthesis.

The calculation of the BBN-predicted primordial abundances involves following the nuclear and weak interactions as they interconvert neutrons and protons and, as they build neutrons and protons into more complex nuclei. For a more detailed description of this physics than is presented here, see my recent review [528]. The predicted abundances depend mainly on the density of nucleons, often called baryons (B). As the Universe is expanding, all densities evolve with time. A useful measure is in terms of the baryon density parameter $\eta_{\rm B}$, formed by the ratio of the number densities of baryon and CMB photons.

Aside from the extra photons produced when the $e^{\pm}$ pairs annihilate in the early Universe, this ratio remains constant as the Universe expands (and cools). The results of the BBN calculation in the standard model are shown as a function of $\eta_{10} = 10^{10}\eta_{\rm B}$ in Fig. 2.10, where the mass fraction of ^{4}He, Y, and the ratios of D, ^{3}He, and ^{7}Li to hydrogen (by number) are shown as a function of η_{10}.

The primordial abundances, especially that of ^{4}He, also depend on the early Universe expansion rate, as measured by the Hubble parameter, H, which, for the standard cosmology, depends on the square root of the energy density, ρ. As, at the time of BBN, the energy density is dominated by the contributions from relativistic particles (photons, $e^{\pm}$ pairs, light neutrinos), any deviation from the standard model, with three flavors of neutrinos (ν_e, ν_μ, ν_τ), may be parametrized by the expansion rate factor, $S \equiv H'/H = (\rho'/\rho)^{1/2}$ or, by the effective number of neutrinos, $\mathrm{N}_\nu \equiv 3 + \Delta\mathrm{N}_\nu$,

$$S \equiv H'/H = (\rho'/\rho)^{1/2} = (1 + 7\Delta\mathrm{N}_\nu/43)^{1/2}. \quad (2.4)$$

As the neutron-to-proton ratio is determined by the weak interactions involving neutrinos (electron type neutrinos and antineutrinos), any asymmetry in the abundance of neutrinos relative to antineutrinos (lepton asymmetry) can affect the BBN abundances. Once again, the ^{4}He abundance is especially sensitive to this deviation from the standard model. The parameter ξ_e provides a measure of the lepton asymmetry (in a similar manner, the baryon density parameter, $\eta_B \equiv 10^{-10}\eta_{10}$, measures the baryon asymmetry). However, if $\xi_e \approx \eta_B \approx 6 \times 10^{-10}$ (see below), there would be no measurable effect on BBN. Only for $|\xi_e| \gtrsim 0.001$ will a lepton asymmetry significantly modify the standard BBN-predicted primordial abundances of the light nuclides.

It is easy to understand the qualitative trends in Fig. 2.10 without having to delve deeply into the details of the standard BBN ($N_\nu = 3$ ($S = 1$), $\xi_e = 0$) calculation. Weak interactions among neutrons, protons, electrons, positrons, and neutrinos (electron neutrinos and antineutrinos) interconvert neutrons and protons. As the neutron is more massive than the proton, the proton is favored over the neutron and the ratio of neutrons to protons, by number, is always ≤ 1. At the time BBN begins in earnest, this ratio is $(n_n/n_p)_{BBN} \approx 1/7$. As the reactions leading to the production of ^{4}He, the most tightly bound of the light nuclei, are very fast compared to the universal expansion rate, virtually all the available neutrons are incorporated into ^{4}He. As a result, the ^{4}He abundance (by mass) is $Y = \frac{2n_n}{n_n+n_p} \approx 1/4$, very nearly independent of the baryon density as may be seen in Fig. 2.10.

As the ^{4}He abundance is so closely tied to the neutron-to-proton ratio at BBN, it is sensitive to the competition between the weak interaction rate and the universal expansion rate [281]. ^{4}He provides an early Universe chronometer.

$$\Delta Y = 0.16(S - 1) \approx 0.013\Delta N_\nu. \tag{2.5}$$

The primordial abundance of ^{4}He also probes a universal lepton asymmetry. For $\xi_e > 0$, there are more ν_e than $\bar{\nu}_e$, so that the abundance of neutrons relative to protons is reduced, reducing the BBN-predicted primordial abundance of ^{4}He [265, 281].

$$\Delta Y \approx -0.23\xi_e. \tag{2.6}$$

As may be seen from Fig. 2.10, Y is a very slowly varying function of the baryon (nucleon) density. For $\eta_{10} \approx 6$, $N_\nu \approx 3$ ($S \approx 1$), and $|\xi_e| \lesssim 0.1$, a very good fit to the BBN-predicted primordial abundance of ^{4}He is [281, 528]

$$Y_P = 0.2485 \pm 0.0006 + 0.0016(\eta_{10} - 6) + 0.013\Delta N_\nu - 0.23\xi_e. \tag{2.7}$$

In contrast to ^{4}He, as the less tightly bound nuclei of D and ^{3}He are being burned to produce ^{4}He, their relic abundances are sensitive to the baryon density at BBN with $(D/H)_P \propto \eta_{10}^{-1.6}$ and $(^3He/H)_P \propto \eta_{10}^{-0.6}$. D and ^{3}He are primordial baryometers. Notice that while ^{4}He, the second most abundant element to emerge from the early Universe, has a BBN-predicted abundance, by number, of order 10% of that of hydrogen, the abundances of D and ^{3}He are predicted to be smaller than that of hydrogen by some 4–5 orders of magnitude.

The only other nuclide produced in an astrophysically interesting abundance is ^{7}Li, whose abundance is smaller than those of D and ^{3}He by another five orders of magnitude. The reason for the very small abundance of ^{7}Li traces to the bottleneck at ^{4}He due to the gap at mass-5: there is no stable nucleus at mass-5. Coulomb-suppressed reactions of ^{4}He with the much less abundant D, ^{3}H, ^{3}He nuclei, guarantees that the primordial abundances of the heavier nuclides are strongly suppressed. Those few reactions that do jump the mass-5 gap lead to mass-7, producing ^{7}Be and ^{7}Li. Later in the evolution of the Universe, ^{7}Be captures an electron and decays to ^{7}Li, the only surviving mass-7 primordial nuclide.

2.7.1.2 Primordial Abundances

Inferring the primordial abundances of the light nuclides from present-day observations of a variety of astronomical objects (stars, H II regions (regions of ionized gas), neutral gas) involves a complex interplay between physics, astrophysics, and astronomy. At each step in the process, statistical errors as well as systematic uncertainties may arise. While the former may be reduced by acquiring large amounts of data, the latter are, by their very nature, difficult to quantify and more data may, or may not, lead to their reduction. These uncertainties must be kept in the forefront of any confrontation between theory (the BBN-predicted abundances) and observation (the observationally inferred primordial abundances). While the bad news, at present, is that limited data sets plague the determination of the primordial abundances of D, ^{3}He, and ^{7}Li, and systematic errors are a cause for concern in determining the primordial abundances of all the light nuclides. For a detailed discussion of current data and the problems and uncertainties associated with inferring the primordial abundances from the observational data, the reader is referred to [528] and references therein. Here, I summarize the results that emerge from that analysis.

2.7.1.3 Deuterium

Because of its simple post-BBN evolution and its notable dependence on the baryon density parameter, Deuterium is the baryometer of choice. As a result of its very weak binding, whenever gas containing D is cycled through stars, deuterium is destroyed. As a result, $(\mathrm{D/H})_{\mathrm{P}} \gtrsim (\mathrm{D/H})_{\mathrm{OBS}}$. The abundances of the "heavy" nuclei (the so-called "metals": C, N, O, ...) provide a measure of the amount of gas that has been processed through stars. In the limit of low metallicity, the observed deuterium abundance should provide an accurate probe of its primordial abundance. Deuterium is best observed by its absorption spectrum as light passes from a background light source (e.g., a QSO) through intervening, neutral gas. The data from seen lines of sight through high-redshift, low-metallicity QSO absorption line systems [277] leads to an estimate [404] of

$$y_{\mathrm{DP}} \equiv 10^{5}(\mathrm{D/H})_{\mathrm{P}} = 2.70^{+0.22}_{-0.20}. \qquad (2.8)$$

For standard BBN, this abundance corresponds to a baryon density parameter $\eta_{10} = 6.0 \pm 0.4$. This BBN-inferred baryon density is in excellent agreement with that inferred from observations of the CMB temperature fluctuation spectrum [524], $\eta_{10} = 6.1 \pm 0.3$, which provide a measure of the universal baryon density some 4×10^5 years after BBN. The standard model is consistent with observations of relics from the Universe at a few minutes and a few hundred thousand years after the beginning of the expansion.

2.7.1.4 Helium-3

^{3}He is observed in Galactic H II regions via the emission from the spin-flip transition at 3.46 cm (the analog of the 21 cm line in hydrogen) from singly ionized ^{3}He. Unfortunately, these Galactic H II regions contain gas that has been processed through several generations of stars and the evolutionary correction required to infer the primordial abundance introduces model-dependent, systematic uncertainties into the inferred primordial value. Following the suggestion [23] that the ^{3}He abundance inferred from observations of the most metal-poor (least processed) Galactic H II regions provide an estimate of (or, an upper bound to) the primordial abundance,

$$y_3 \equiv 10^5 (^3\mathrm{He/H})_\mathrm{P} = 1.1 \pm 0.2. \tag{2.9}$$

For standard BBN, this ^{3}He abundance corresponds to a baryon density $\eta_{10} = 5.6^{+2.2}_{-1.4}$ which, while much more uncertain than that inferred from BBN and D, and from the CMB, is entirely consistent with each of them. Observations of D, ^{3}He, and the CMB provide consistent, independent support for the standard models of cosmology and particle physics.

2.7.1.5 Helium-4

As gas is cycled through generations of stars in post-BBN chemical evolution, the abundance of ^{4}He increases. To minimize the model-dependent evolutionary corrections, the most valuable data are provided by observations of low-metallicity, extra-galactic H II regions where the presence of ^{4}He is revealed via the emission lines produced when ionized helium (and hydrogen) recombines. The good news is that there is a database of some 90 such H II regions, which are useful for minimizing the statistical uncertainties in the inferred value of Y_P. The bad news is that the detailed analyses of the physics and astrophysics of the formation and radiative transfer of the recombination lines in such H II regions have many systematic uncertainties. As a result, current estimates of Y_p vary from $Y_\mathrm{p} = 0.243 \pm 0.001$ [254], inferred from the study of some 80 H II regions, to $Y_\mathrm{p} = 0.249 \pm 0.009$ [380], or $Y_\mathrm{p} = 0.250 \pm 0.004$ [186], or $Y_\mathrm{p} = 0.248 \pm 0.003$ [399] inferred from studies of many fewer H II regions, but with an eye to deal more carefully with systematic corrections and their uncertainties. Following a critical review of these recent results, I have suggested [528] that the current data and analyses are consistent with a primordial abundance

$$Y_P = 0.240 \pm 0.006, \tag{2.10}$$

and a robust upper bound to Y_p of

$$Y_P < 0.251 \pm 0.002. \tag{2.11}$$

Notice that for either the BBN-predicted baryon density found using deuterium [528] or the consistent value inferred from the CMB [524], the standard BBN-predicted abundance of ^{4}He is $Y_p = 0.249 \pm 0.001$, consistent, within the uncertainties, with the primordial abundance inferred from the observational data. However, if these other data are ignored, then the baryon density inferred from standard BBN and ^{4}He alone, $Y_p = 0.240 \pm 0.006$, would be much smaller, $\eta_{10} = 2.8^{+2.0}_{-1.0}$ [528], hinting at a tension between D (and ^{3}He) and ^{4}He (and between the CMB and ^{4}He).

2.7.1.6 Lithium-7

Observations of lithium in the Sun and in the local interstellar gas are of little value in inferring the primordial abundance of ^{7}Li due to the large and highly uncertain evolutionary corrections required to connect the observationally inferred abundances to the BBN-predicted abundance. The only data of value available for inferring the primordial abundance are provided by observations of lithium on the surfaces of the very oldest, most metal-poor stars in the Galaxy. Even these data may need to be corrected for the post-BBN evolution of ^{7}Li [15, 481]. And, it is clear that there are processes at work in these stars which, over their long lifetimes, may have modified their surface abundances via depletion, dilution, or gravitational settling. Ignoring these latter corrections, a primordial abundance [15]

$$[\mathrm{Li}]_P \equiv 12 + \log\,(\mathrm{Li/H})_P = 2.1 \pm 0.1 \tag{2.12}$$

is inferred [15, 481]. In contrast, using the BBN or CMB inferred baryon density, the primordial abundance is predicted to be $[\mathrm{Li}]_P = 2.63^{+0.07}_{-0.08}$ (D + BBN) or, $2.65^{+0.05}_{-0.06}$ (CMB + BBN) which, while consistent with each other, differ from the above estimate by a factor of $\sim$3 or more. On the basis of lithium alone, the standard BBN-predicted baryon density would be $\eta_{10} = 4.0 \pm 0.6$ [528].

However, in an attempt to estimate the correction to the observed lithium abundance due to gravitational settling, observations of stars in a Globular Cluster (of the same age and metallicity) have led to a (model-dependent) higher estimate [292] of

$$[\mathrm{Li}]_P = 2.54 \pm 0.10 \tag{2.13}$$

which, within the errors, is entirely consistent with the standard BBN prediction.

Thank you Gary. Now Keith will enter more deeply into the observational tests of the BBN.

2.7.2 Tests of Cosmological Nucleosynthesis

Dear Keith *(Olive)***, in which way observations of element abundances can probe the BBN theory? Could you explain why, and what are the most significant tests used in present day cosmology?**

Unfortunately, we can not observe element abundances directly at the time of BBN. Abundances are observed at much later epochs, after stellar nucleosynthesis has commenced. The ejected remains of this stellar processing can alter the light element abundances from their primordial values, and also produce heavy elements such as C, N, O, and Fe ("metals"). Thus one seeks astrophysical sites with low metal abundances, to measure light element abundances that are closer to primordial. For all of the light elements, systematic errors are an important and often dominant limitation to the precision of the primordial abundances.

In recent years, high-resolution spectra have revealed the presence of D in high-redshift, low-metallicity quasar absorption systems (QAS), via its isotope-shifted Lyman-α absorption. These are the first measurements of light element abundances at cosmological distances. The six most precise observations of deuterium ([383] and references therein) in QAS give D/H = $(2.83 \pm 0.26) \times 10^{-5}$, where the error is statistical only. These measurements are clearly consistent with the CMB/BBN determined value of the primordial D/H abundance, which is predicted to be

$$(\mathrm{D/H})_p = 2.6 \pm 0.2 \times 10^{-5}. \tag{2.14}$$

^{4}He is observed in clouds of ionized hydrogen (HII regions), the most metal-poor of which are in dwarf galaxies. There is now a large body of data on ^{4}He and carbon, nitrogen, and oxygen (CNO) in these systems ([254] and references therein). The He abundance from this sample of 89 HII regions obtained $Y_\mathrm{p} = 0.2429 \pm 0.0009$ [254]. However, the recommended value is based on the much smaller subset of 7 HII regions, finding $Y_\mathrm{p} = 0.2421 \pm 0.0021$.

^{4}He abundance determinations depend on a number of physical parameters associated with the HII region in addition to the overall intensity of the He emission line. These include the temperature, electron density, optical depth, and degree of underlying absorption. A self-consistent analysis may use multiple ^{4}He emission lines to determine the He abundance, the electron density, and the optical depth. The question of systematic uncertainties was addressed in some detail in [379]. It was shown that there exist severe degeneracies inherent in the self-consistent method, particularly when the effects of underlying absorption are taken into account. The results of a Monte Carlo reanalysis [380] of NCG 346 [397, 398] showed that solutions with no absorption and high density are often indistinguishable (i.e., in a statistical sense they are equally well represented by the data) from solutions with underlying absorption and a lower density. In the latter case, the He abundance is systematically higher. These degeneracies are markedly apparent when the data is analyzed using Monte Carlo methods, which generate statistically viable representations of the observations. When this is done, not only are the He abundances found to be higher, but the uncertainties are also found to be significantly larger than in a direct self-consistent approach. The extrapolated ^{4}He abundance was determined to be $Y_p = 0.2495 \pm 0.0092$. The value of η_B corresponding to this abundance is $\eta_{10} = 6.9^{+11.8}_{-4.0}$ and clearly overlaps with η_{CMB}. Conservatively, it would be difficult at this time to exclude any value of Y_p inside the range 0.232 – 0.258.

The systems best suited for Li observations are metal-poor Pop II stars in the spheroid of our Galaxy. Observations have long shown [525] that Li does not vary significantly in Pop II stars with metallicities $\lesssim$ 1/30 of solar – the "Spite plateau". Recent precision data suggest a small but significant correlation between Li and Fe [480], which can be understood as the result of Li production from Galactic cosmic rays [172, 571]. Extrapolating to zero metallicity, one arrives at a primordial value [481] $\mathrm{Li/H_p} = (1.23 \pm 0.06) \times 10^{-10}$.

The ^{7}Li abundance based on the WMAP baryon density is predicted to be

$$^7\mathrm{Li/H} = 4.3 \pm 0.7 \times 10^{-10}. \tag{2.15}$$

This value is in contradiction with most estimates of the primordial Li abundance. It is a factor of $\sim$3 higher than the value observed in most halo stars, and just about 0.2 dex over the value observed in globular clusters, $^7\mathrm{Li/H} = (2.2 \pm 0.3) \times 10^{-10}$ [57, 58]. There are many possible sources for this discrepancy. Among them lie the possibility that some Li was destroyed or removed from the stellar surface. However, the lack of dispersion in the Li data limits the degree to which depletion can be effective. A very real systematic uncertainty stems from the assumed physical properties of the star that are used to derive the abundance from raw observations. Most important among these is the surface temperature of stars [174, 353]. Nonstandard process may also play a role.

Detailed abundance observations of heavier elements play a key role in building a picture of the chemical history of the Universe. Indeed, chemical evolution at high redshift connects massive star formation in the early Universe to the epoch of reionization, the heavy element abundances of the oldest stars and the high redshift IGM, and the mass outflows associated with galaxy formation. Furthermore, predicted SN rates provide us with an independent probe of the early epoch of star formation. Combining abundance and SN rate predictions allows us to develop an improved understanding of both the cosmic star formation history and of the enrichment of the IGM, as well as to elucidate the nature of Population III ([120], see review of [102]).

Thank you Keith.

Dear Gary *(Steigman)*, how do the abundances of the light elements predicted by BBN compares with the information from CMB and LSS?

In Fig. 2.11 are shown the standard BBN-predicted values of η_{10}, inferred from a comparison with the observationally inferred abundances adopted before, along with their 2σ ranges. From deuterium alone we find $(\eta_{10})_D = 6.0 \pm 0.6$ (at 2σ).

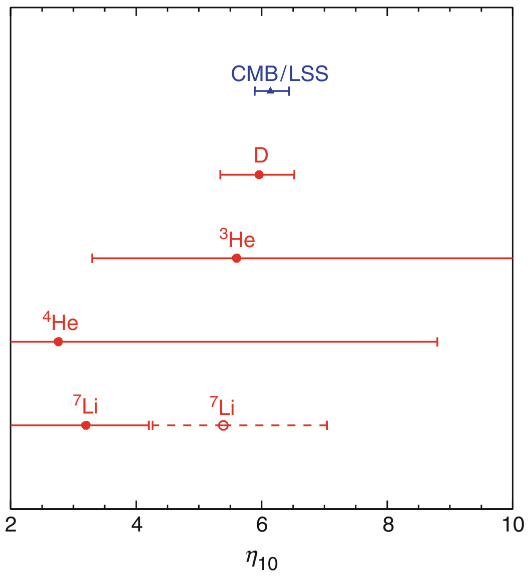

Fig. 2.11 The standard BBN-predicted values of η_{10}, along with their 2σ ranges, corresponding to the adopted primordial abundances (*filled circles*), along with the value inferred from cosmic background radiation and large scale structure data (CMB/LSS: *filled triangle*). The *open circle* and *dashed lines* correspond to the alternate lithium abundance discussed in the text

This value is in excellent agreement with that inferred from the more uncertain abundance of ^{3}He. While the central value of the ^{4}He abundance corresponds to a very different – much smaller – baryon density, as may be seen from Fig. 2.11, it is in agreement with the D, ^{3}He, and the CMB/LSS values at less than 2σ. However, even at 2σ, the lithium abundance adopted in (2.12) [15, 481] is inconsistent with these baryon density determinations, although the higher value corresponding to (2.13) [292] does agree with them. As the range of the density parameter covered in Fig. 2.11 is larger than the range of applicability of the analytic fit described earlier in (2.7), the specific values shown there are derived from a numerical BBN code. For the central value of the deuterium-predicted baryon abundance, the standard BBN-predicted helium abundance is $Y_{\mathrm{p}} = 0.249$, only 1.5σ away from the observationally inferred value $Y_{\mathrm{p}} = 0.240 \pm 0.006$. The lithium abundance poses a greater challenge; for $\eta_{10} = 6.0$, the standard BBN-predicted lithium abundance exceeds $[\mathrm{Li}]_{\mathrm{P}} = 2.6$, which is far from the observationally inferred value of $[\mathrm{Li}]_{\mathrm{P}} = 2.1 \pm 0.1$.

It is noteworthy that the two nuclides that may pose the most serious challenges to standard BBN (^{4}He and ^{7}Li) are those for which systematic corrections, and their corresponding uncertainties, have the potential to change the observationally inferred relic abundances by the largest amounts. The values of η_{10} corresponding to the alternative choices for ^{4}He and ^{7}Li considered above in (2.11) and (2.13) are in much better agreement with the D and ^{3}He determined baryon abundance: $(\eta_{10})_{\mathrm{He}} < 7.8^{+1.9}_{-1.5}$ and $(\eta_{10})_{\mathrm{Li}} = 5.4 \pm 0.6$, respectively.

Observations of the small temperature fluctuations in the cosmic background radiation and of the LSS they seeded currently provide the tightest constraint on the universal abundance of baryons $\eta_{\mathrm{CMB/LSS}} = 6.1 \pm 0.2$ [524], as shown in Fig. 2.11. The observationally inferred relic abundances of D and ^{3}He are in excellent agreement with the standard BBN predictions for this value/range of η_{10}. Depending on the outcome of various systematic corrections for the observationally inferred primordial abundances of ^{4}He and ^{7}Li, they may, or may not, pose challenges to standard BBN. If the tension between D and ^{4}He is taken seriously, it could be a sign of "new physics": $N_\nu \neq 3$ and/or $\xi_e \neq 0$. For example, for $\eta_{10} = 5.7$ and $N_\nu = 2.4$, the BBN-predicted primordial abundances of D and ^{4}He are now in perfect agreement with those inferred from the observational data and, also with that of ^{3}He [528]. However, the BBN-predicted lithium abundance remains very close to $[\mathrm{Li}]_{\mathrm{P}} \approx 2.6$, still a factor of $\sim$3 higher than that inferred from the observations. The same is true for the $\{\eta_{10}, \xi_e\} = \{6.0, 0.034\}$ pair [528].

2.7.2.1 At a Glance

According to the standard model of cosmology, the early Universe was hot and dense and, when it was a few minutes old, nuclear reactions among neutrons and protons synthesized astrophysically interesting abundances of the light elements D, ^{3}He, ^{4}He, and ^{7}Li. In the standard model, these relic abundances depend on only one free parameter, the baryon abundance (the baryon to photon ratio). Self-consistency of the standard model requires that there is a unique baryon abundance, consistent with the observationally inferred primordial abundances of these light elements. For D and ^{3}He, this is the case. As a bonus, this BBN-predicted baryon abundance is in excellent agreement with that inferred from the CMB/LSS. While the BBN-predicted abundance of ^{4}He may be somewhat higher than its observationally inferred value, within the observational errors, there is agreement. Three for the price of one; four, counting the CMB/LSS! However, the BBN-predicted relic abundance of ^{7}Li is a factor of three, or more, higher than its observationally inferred value. While this may provide a challenge to the standard model, it is not unlikely that the resolution of this challenge lies in the uncertain stellar physics associated with the evolution of the surface abundances of the oldest, most metal-poor stars in the Galaxy.

How may the observational verification of primeval nucleosynthesis be affected by post-BBN stellar nucleosynthesis and galactic evolution? Is this contribution known with the accuracy necessary for a robust verification of the cosmological model?

An essential, unavoidable step in comparing the predictions of primordial nucleosynthesis with the observational data is accounting for the chemical evolution of material that has been cycled through stars in the $\sim$14 Gyr since BBN was completed. Account for post-BBN evolution is not separate from, but is a crucial part of the analysis that leads us from the observational data to the inferred, relic abundances.

The post-BBN evolution of deuterium is straightforward, as whenever gas is cycled through stars, deuterium is entirely destroyed. Because of the very small binding energy of the deuteron, whenever deuterium is formed by nuclear reactions in the hot interiors of stars, it is immediately burned to tritium, helium-3, helium-4, and beyond. As a result, as the abundance of deuterium can only have decreased since BBN, *any* deuterium observed *anywhere* in the Universe, at *any* time in its evolution, provides a *lower* bound to the primordial abundance of D. For the same reason, by concentrating on those astrophysical objects (e.g., QSO Absorption Line systems) at high redshift and low metallicity, we can expect to measure an abundance nearly identical with the primordial value.

The post-BBN evolution of ^{3}He is much more complicated than that of D, because when gas containing ^{3}He is cycled through stars, some of the ^{3}He is burned away, some is preserved (not all layers of all stars are hot enough to burn ^{3}He) and, for some stars, new ^{3}He is produced. The result is that the extrapolation of the current data from chemically evolved H II regions in the Galaxy back to the early Universe is uncertain and model-dependent. It is for this reason that ^{3}He is usually given less weight in the comparison between theory and observation. Nonetheless, given the observed abundance of ^{3}He and our best estimate of its chemical evolution, theory and observation are in excellent agreement (see Fig. 2.11).

Stars burn hydrogen to helium (^{4}He). The abundance of ^{4}He observed in the post-BBN Universe has increased from its primordial value. Stars also synthesize the heavier nuclei, "metals" such as C, N, O, ... As the heavy element abundance (metallicity) increases in the course of stellar and galactic evolution, so, too, does the abundance of ^{4}He. Two options are available for accounting for – or avoiding – this inevitable correction required to pass from the observational data to the primordial abundance. One choice is to restrict attention to the very lowest metallicity regions observed and to *assume* that this will minimize the correction for post-BBN production of ^{4}He. The other is to use data from regions of all metallicities and to *extrapolate* to zero metallicity to find the BBN abundance. Each of these approaches has assets and liabilities and, each introduces its own uncertainties into the value of Y_{p} inferred from observations. These corrections, along with their attendant uncertainties, have been used to infer the primordial abundance of ^{4}He listed earlier.

Lithium is observed in the very most metal-poor, oldest stars in the Galaxy, stars with heavy element abundances lower than those in the Sun by factors of a thousand or more. It is expected that for these stars the observed lithium is completely dominated by the relic component from primordial nucleosynthesis. The problem, as discussed earlier, is that these oldest stars in the Galaxy have had the most time to modify their surface material, the material which is observed to infer the stellar lithium abundance. Post-BBN evolution of ^{7}Li is the least of our worries, compared to the uncertainties of stellar structure and evolution, in using the data to infer the ^{7}Li primordial abundance.

The bottom line is that for the two nuclides, D and ^{4}He, which are most valuable in testing the standard model, the post-BBN evolution is likely well-enough understood so that our conclusions are robust.

2.7.3 Alternatives to Standard BBN

Dear Gary *(Steigman)*, **do there exist reliable modifications and alternatives to the standard nucleosynthesis theory? Can the observational data be explained by alternative ideas?**

Alternatives to the standard model are limited only by the creativity and imagination of physicists and cosmologists. BBN is a pillar of modern cosmology, in that the first test any alternative theory must pass is that the correspondingly modified BBN needs to be in agreement with the observationally inferred primordial abundances presented earlier. Many new alternative theories never see the light of day because they fail this test. Nonetheless, there are still large classes of alternative theories that *may* be consistent with the predictions of BBN in the standard model and, therefore, consistent with the observational data.

For example, it could be that the early-Universe expansion rate, H, is modified compared to that in the standard model, as quantified by the expansion rate parameter, S, defined in (2.4) (or, by ΔN_ν). Although the standard value of $S = 1$ ($\Delta N_\nu = 0$, $N_\nu = 3$) is, within the uncertainties, consistent with the abundances, models of new physics or cosmology with nonstandard values are restricted to the range $1.6 \lesssim N_\nu \lesssim 3.3$. Models with N_ν outside this range are excluded.

Or, because of nonstandard physics, it could be that the lepton asymmetry of the Universe exceeds the baryon asymmetry by some nine orders of magnitude ($|\xi_e| \sim 0.1$). Again the effect of such an asymmetry would be to change the neutron-to-proton ratio at BBN and, therefore, to modify the BBN abundance of ^{4}He. This is allowed only for the asymmetry parameter restricted to the narrow range, $-0.027 \lesssim \xi_e \lesssim +0.086$. Models with ξ_e outside this range are excluded.

Thanks a lot Gary.
Together with the BBN, CMB studies provided up to now the stronger evidence in favor of the current cosmological scenario. Most of them come from the COBE mission. We now have the opportunity of speaking with the Nobel Laureate John Mather, who won the Nobel Laureate together with George Smoot for the fundamental cosmological results of the COBE mission. Here, we will ask him to review the characteristics of COBE and the importance of its discoveries.

2.8 CMB Observations and Main Implications

2.8.1 The COBE Legacy

Dear John *(Mather)*, **COBE opened the so-called era of precision cosmology with the up-to-now best measure of the CMB spectrum and discovered the CMB large scale anisotropy. Can you tell us about the scientific adventure of COBE? Why such a project has been so relevant in the context of physical cosmology in the beginnings of 1990?**

The discovery of the expanding Universe and its modern details have a fascinating history lasting for about a century, almost entirely (except for Einstein's conceptual leaps into relativity) propelled by new technological advances that enabled new measurements.

In 1912, V.M. Slipher discovered that galaxies were moving away from us at high speeds, but it was not yet known that galaxies are made of stars. In 1916, Einstein gave us the theory of General Relativity, but in solving the equations, he assumed that the Universe could not be expanding, and added the Λ constant to permit this. In 1922, Friedmann's solution to Einstein's equations showed that the Universe could be expanding. Also in 1922, Edwin Hubble measured the first galactic distances by resolving Cepheid variables, stars which vary periodically in brightness and can serve as standard candles. In 1927, Lemaître argued from Einstein's equations that the Universe should be expanding, and described the "primeval atom". In 1929, Hubble showed the linear relationship of speed of galactic recession to the distance scale, and the expanding Universe was widely recognized, if not accepted. In 1946, Hoyle showed that the chemical elements could have been made in stars. In 1948, Gamow, Alpher, and Herman showed that the residual heat of the Big Bang should still exist, with a temperature of about 5 K. In 1965, the discovery of that radiation by Penzias and Wilson strongly supported the Big Bang theory, but continued difficulties in measuring its spectrum (brightness vs. wavelength) undermined confidence. In 1990, the COBE mission ended the uncertainty about the spectrum and later determined that the temperature is 2.725 K, and in 1992, the same team discovered the anisotropy (hot and cold spots in the map) of the background radiation. It is now believed that the anisotropy comes from cold DM, already beginning to cluster at the time of the decoupling of ordinary matter (when the Universe was about 3.89×10^5 years old, and 1/1090 as large as it is today). The DM concentrations caused gravitational redshifts (the Sachs–Wolfe effect) and so are detectable. Without them, gravitation would not have been able to stop the cosmic expansion locally, and there would be no galaxies.

2.8.1.1 Importance of the CMB

The discovery and detailed measurements of the CMB radiation have changed cosmology from extreme speculation to detailed calculation based on widely accepted concordance models of the early Universe. This extraordinary circumstance is due to the fact that the CMB is still by far the dominant radiation field in the Universe, both in energy content and in number of photons, and that there are wavelengths accessible for measurement at which the details can be observed with extraordinary precision. This scientific revolution has depended on the advent of cryogenic technology to achieve the needed sensitivity, and on access to high altitude balloons and outer space to avoid the interference of the Earth's atmosphere. It has also depended on a certain degree of luck, in that the model for the growth of density fluctuations from the primordial seeds has matched the observations extremely well. It is now generally thought that the distribution of matter through the Universe is the result of gravitation acting on three fluids: DM, ordinary matter, and photons, starting with nearly scale-free initial fluctuations produced by unknown processes in the Big Bang. Based on this success, it is now possible to consider searches for the tiniest details of the angular distribution, polarization, and energy spectrum of the background radiation. With additional luck, these details could reveal the nature of the forces acting in the earliest moments of the Big Bang, the history of energy release after the Big Bang, and the beginnings of disequilibrium processes based on atomic and molecular physics. All are critically important parts of our own history, the path that was followed to produce the only known example of intelligent life on a small planet around a small star in an ordinary galaxy.

2.8.1.2 Cosmology in 1974

In 1974, when (acting on recommendations from my postdoctoral advisor Pat Thaddeus) I organized a team to propose the COBE mission, there were only a few pieces of evidence about the Big Bang. These were the following: the abundances of the light elements (hydrogen and helium, with traces of lithium, beryllium, and boron), the existence of the cosmic microwave background radiation (apparently filling the Universe uniformly with heat at about 3 K), and the number of types of neutrinos determined from particle accelerator experiments (three, to match the three types of leptons: electrons, muons, and tau mesons). There were several great questions still open: Why is the Universe so uniform? What distributed the galaxies where they are? Was there some kind of energy release following the Big Bang? Was the Big Bang hot or cold? Was it possible that the Universe only looks like it came from a Big Bang, but is self-renewing with continuing matter creation, as in the Steady State theory?

The discovery of the CMB in 1965 almost ended the debate about the Steady State theory, which did not naturally have a place for a BB radiation field. In the Steady State theory, there would be a background radiation field, but it was produced by stars distributed out to immense distances through an infinite amount of time, and would not necessarily have a BB distribution with wavelength. The critical test of the Big Bang theory was therefore: does the CMB have the spectrum of a perfect BB? In the Big Bang story, the radiation field is a remnant of extremely hot and dense early times, in which there is plenty of time to achieve the equilibrium spectrum represented by a BB. In the Steady State theory, there is no such hot and dense period in which equilibrium would be achieved.

By the early 1970s, there was good evidence that the CMB had a roughly BB spectrum, but the measurements were difficult because the radiation is faint (relative to terrestrial sources), the atmosphere interferes, and the needed technology was new. As a result, almost every measurement seemed to suggest that there might be some problem with the CMB spectrum, at levels ranging from 3σ (fairly important) to factors of 50 (extremely important but very implausible). Historically, 3σ results are common in measurements, and so I was never worried that the Big Bang explanation was in trouble, but there was some room for doubt, and space for disbelievers to invent radically different stories.

Clearly, it was essential to get better data. Measurements of the CMB are typically done by differential instruments, designed to compare the sky with a reference object (or another piece of the sky). With differential instruments, it is possible to make very good comparisons even if the instrument sensitivity is not enough to do it quickly. If we could build an instrument package in space, we could get years of observing time to compensate for the faintness of the signal relative to the sensitivity of the instruments.

2.8.1.3 COBE

The COBE satellite was a NASA mission to measure the microwave and infrared background radiation of the Universe. It was launched at dawn by a Delta 2 rocket on 18 November 1989 from Vandenberg Air Force Base in Lompoc, California. It is (still) in a Sun-synchronous circular polar orbit at an altitude of 900 km, with an orbit plane 99° from the equator and approximately perpendicular to the Sun. The Equatorial bulge of the Earth causes the orbit plane to precess once per year. The COBE carried three instruments, the Far Infrared Absolute Spectrophotometer (FIRAS), the Differential Microwave Radiometers (DMR), and the Diffuse Infrared Background Experiment (DIRBE). The FIRAS and the DIRBE were cooled inside a liquid helium cryostat to about 1.5 K.

The COBE mission was proposed in 1974 by a team consisting of M. Hauser, J. Mather, D. Muehlner, P. Thaddeus, R. Silverberg, R. Weiss, D. Wilkinson, with a complement of four instruments. Instruments like the DMR were also proposed in 1974 by teams including L. Alvarez and G. Smoot from UC Berkeley, and S. Gulkis and M. Janssen from the Jet Propulsion Laboratory (JPL). In 1976, NASA chose a mission definition study team to work with Goddard Space Flight Center. Around 1979, NASA decided to plan to build the satellite in-house at Goddard. In 1982, construction was approved. In 1986, the Space Shuttle Challenger explosion required the COBE to be redesigned for launch on a Delta rocket. After launch in 1989, the mission operated with liquid helium for 10 months, and continued for total lifetime of 4 years. Data analysis was completed after 8 years.

2.8.1.4 Objectives of COBE

In 1974, there was no good theory for how much the CMB spectrum might differ from the perfect BB form. There were many radical ideas about the structure of the Universe, how there might have been explosive events, BHs, decaying elementary particles , and so forth. Similarly, there was no good theory for how much the CMB might differ in brightness from one direction to another. It was expected that there would be a small effect due to the motion of the Solar System relative to the rest of the Universe, but when that was finally measured, it was still a challenge to explain the measured velocity based on the gravitational forces acting over the age of the Universe.

As a result of all this uncertainty, there was no serious theory to guide the design of the COBE mission. We decided that our objective would be to measure as well as we are allowed to do by the local astrophysical environment: the interference from dust in the Solar System, the Milky Way, and other galaxies. As it turns out, this was achieved, but with difficulty. During the preparation of the COBE mission design, there were times when the project was delayed for budgetary reasons: the Infrared Astronomical Satellite (IRAS) satellite, a science and technology precursor, was over budget, and the HST was too. We had an opportunity, urged by team member Phil Lubin, to use newer and better microwave receiver technology. The new receivers were about twice as good as the old, and without this added sensitivity, we might never have made the discovery of the CMB fluctuations.

The COBE was designed with three main objectives: (1) Measure the spectrum of the CMB as precisely as possible, and measure any differences from the predicted BB form. (2) Make an all-sky map of the CMB and characterize any differences from uniformity. (3) Measure the cosmic infrared background radiation, presumably produced throughout the history of the Universe by stars and BHs, and absorbed by dust and re-radiated as infrared.

2.8.1.5 COBE Spacecraft

The design of the COBE spacecraft is illustrated in Fig. 2.12. The entire mission design is concentrated on protecting the instruments from the intense heat and light of the local spacecraft environment, and on determining the contributions of local astrophysical sources to the measured background radiation. Electrical power is provided by three deployable solar panel wings, and a deployable conical sunshield protects the instrument package from heat and radiation from the Sun and the Earth. The spacecraft spins around the symmetry axis approximately once per 72 s to scan the lines of sight of the DMR and DIRBE rapidly around the sky. The selected orbit plane is approximately perpendicular to the Sun and the instrument package points away from the Earth, and so the Sun never illuminates the instruments and the Earth peeks over the shield for only brief periods (up to 20 min per orbit) for 3 months per year, and not at all at other times. The orientation is sensed by Sun and Earth sensors and controlled by magnetic torquer bars, reaction wheels, and a large counter-rotating wheel to oppose the spin angular momentum of the spacecraft.

The FIRAS and DMR occupy the interior space of the cryostat, and the DMR instrument occupies three boxes around the outside of it.

2.8.1.6 Instruments On-Board COBE

FIRAS

The main purpose of the FIRAS was to compare the spectrum of the CMB radiation with a precise BB, and to measure any deviations. As the CMB is supposed to

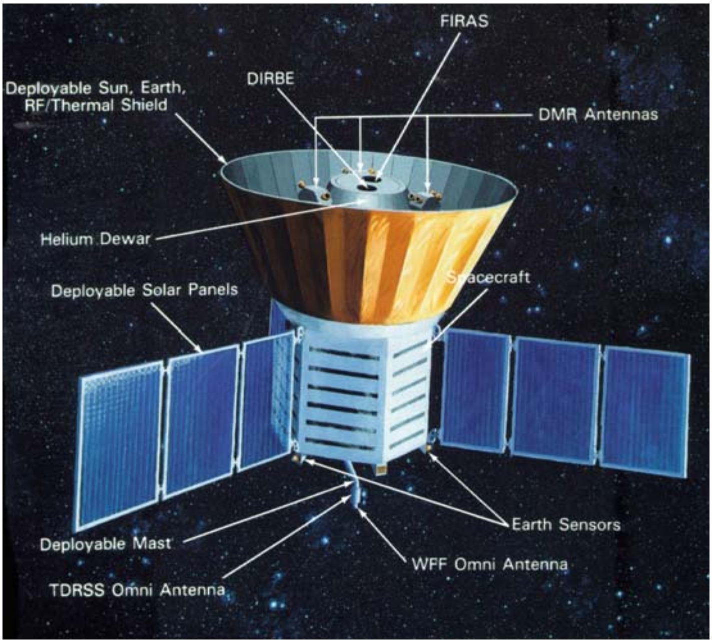

Fig. 2.12 COBE spacecraft in orbit 900 km above Earth, in an orbit inclined 99° to the Equator. The instruments are protected by a conical shield. From the COBE Science Working Group and Goddard Space Flight Center

equilibrate very rapidly to a BB form in the hot and dense conditions of the early Universe, and the expansion is supposed to preserve the BB form in the absence of major energy release after the explosion, the measurement is an essential test of the Big Bang theory.

The FIRAS covered the entire wavelength range from 100 μm to 1 cm in two bands, separated at 0.5 mm. It received light from a cone 7° in diameter, aligned with the COBE spin axis. To obtain precise spectra, it was constructed as a symmetrical, differential Michelson interferometer. With a BB filling the second input of the interferometer, roughly matched to the temperature of the CMB, the resulting interferograms were approximately nulled. In addition, a precise external calibrator was placed occasionally in the aperture. The calibrator was designed as a re-entrant cone, like a trumpet mute, and when inserted in the FIRAS aperture had a calculated error of a few parts per million deviation from BB form. Its temperature was measured by three germanium resistance thermometers with calibrations traceable to the National Bureau of Standards (now NIST). The optical concept is illustrated in Fig. 2.13.

Each side of the interferometer was provided with a pair of bolometric detectors, with the wavelengths separated by a dichroic filter. Cosmic rays incident on the detectors limited the sensitivity, but this effect has now been developed to a science

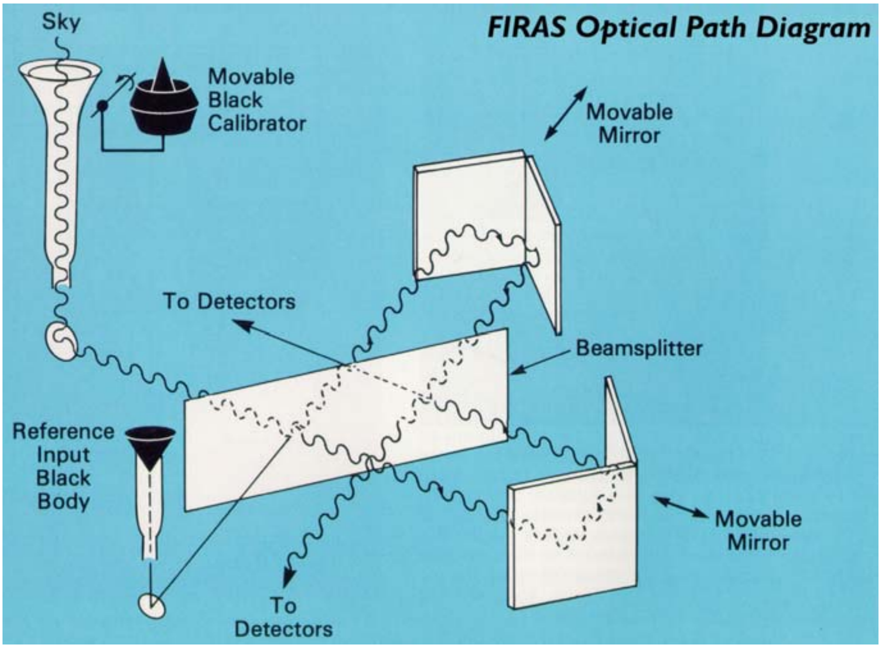

Fig. 2.13 The FIRAS was a differential Michelson interferometric spectrometer. Light from the sky was defined by a compound parabolic concentrator. Polarizers were used to split the beams and recombine them. Michelson won the Nobel Laureate in Physics in 1907 for his invention of the interferometer. From the COBE Science Working Group and Goddard Space Flight Center

itself, and bolometric detectors are now the best available X-ray detectors as well, combining high efficiency with direct measurement of the energy of each photon.

DMR

The purpose of the DMR was to map the CMB. The instrument operated at three frequencies, 31.4, 53, and 90 GHz, with two receivers at each frequency. Each receiver was fed by a ferrite Dicke switch operated at 100 Hz, alternately connecting two corrugated horn antennas to the receiver. The horns were pointed 60° apart and 30° off the spin axis. The low sidelobes of the horn antennas were critical to the success of the mission, as the Earth was not always far below the edge of the sunshield. Observations of hundreds of millions of differential measurements were combined in a least-squares fitting procedure to make maps of the sky at all three frequencies. Instrument errors were modeled in the fits, and included sensitivity to the Earth's magnetic field, the spacecraft's magnetic field, diffraction of the Earth and Sun over the sunshield, temperature variations, etc. The most important astrophysical challenge was to understand and compensate for emissions of electrons in our Galaxy, which emit by the free–free process (collisions with protons) and the synchrotron process (spiraling around the magnetic fields.) In the end, small cosmic anisotropies were found, with an amplitude of approximately a part in 10^5 of the total intensity, and a roughly scale-invariant angular power spectrum. These anisotropies are interpreted as gravitational redshifts in the early Universe, as cosmic background photons escape from potential wells of different depths (the Sachs–Wolfe effect). When these anisotropies were discovered, their significance was immediately clear, although there was no theory whatever at the time that the COBE was proposed. It is now commonly believed that the primordial density fluctuations tie very well to the cosmic structures now observed in the galaxy distributions, and that gravitational forces acting on the primordial fluctuations are sufficient to explain the entire panoply of large-scale structures of galaxies and their clusters.

DIRBE

The purpose of the DIRBE instrument was to measure the cosmic infrared background radiation from the collected emission of all the generations of stars and galaxies since the beginning. The DIRBE was a small telescope (20 cm aperture) observing at 10 wavelengths from 1.25 to 240 μm. The main challenge for the DIRBE team was to measure and understand the changing foreground emissions from the interplanetary dust. After many years, the result was obtained that there are both near IR and far IR cosmic background radiation fields, with a total brightness equal to that of all the known visible and near IR galaxies. These results are an important constraint on the history of star and galaxy formation, and are still not fully explained. For a review see [231].

2.8.1.7 The COBE Success

Cosmology in 1990

In 1981, Alan Guth announced his theory of cosmic inflation. Somehow, this announcement was taken more seriously than some of the earlier versions, outlined, for example, by Floyd Stecker, Demos Kazanas, Qaisar Shafi, and Katsuhiko Sato. My initial opinion was that it was a very nice theory but probably untestable. However, I am happy to be wrong. The theory gives an explanation for a huge cosmic mystery: why is the Universe so uniform? It also makes a specific prediction that the primordial inhomogeneities in the distribution of matter and energy should have no preferred physical size. This prediction had been made for different and very general reasons many years before, by Harrison and Zel'dovich in 1970 and 1972. Modern versions of the theory of inflation say that although there is no preferred size of the density fluctuations, there might be a slight preference for large over small, or viceversa.

In the 1980s, progress was being made in measuring these density fluctuations using telescopic surveys. The objective is to find all the galaxies in a certain region of sky and with certain properties like brightness, measure their distances (or estimate them from their velocities), make maps, and calculate correlation statistics. The results were initially very surprising: there are huge empty regions (cosmic voids), and huge concentrations of galaxies into clusters, and the clusters themselves are clustered into superclusters. There was a discovery of the Great Attractor (one of these huge concentrations) and the Great Wall (a huge sheet-like structure with many galaxies). Then, there was a theoretical effort to explain how galaxies could be located in patterns like these. If we could guess how matter was arranged in the Big Bang, we should be able to predict how it would move under the influence of gravity as billions of years pass by. Analysis was insufficient for this task, and supercomputer simulations have become the only reliable way to do calculations.

There were many uncertainties. First, we did not know the age of the Universe, or the density of matter in it. The age was uncertain because it is difficult to measure the Hubble constant, the expansion speed divided by the distance. We did not yet know for sure whether the Big Bang was the right theory. And, there was already evidence that the amount of matter was not right. We have excellent calculations of nuclear reaction rates in the first few minutes of the Big Bang, and to match the measured abundances of the elements, there cannot be very much ordinary matter. But there is clear evidence from the rotation of galaxies that there is much invisible matter, which has naturally been called DM. It turns out to be possible to explain the distribution of galaxies much better if one imagines that there are various types of DM (called cold and hot, according to whether the particles are massive or light).

It was also recognized that if there were a cosmic DE (modern name), the Universe would be accelerating outwards and that the whole calculation of the growth of cosmic structure would be affected. However, to most physicists this idea seemed an unnecessary complication, not favored by Occam's Razor. But astrophysical discoveries continue to be surprising: things are always more complicated the more closely they are examined. In this context, the scholars of the growth of the cosmic structure were not so surprised as others by the discovery of the DE.

By the time that the COBE was ready for launch in 1989, I felt very confident that the energy spectrum of the CMB would turn out to be very close to the perfect BB form. None of the proposed explanations for the previously measured spectrum distortions seemed plausible. The fundamental reason is simple: photons outnumber particles of matter by about a billion to one. There was just not a good story about how the matter particles could have a big influence on the photons. Only radical changes to the history and contents of the Universe would make a big change to the CMB spectrum. For instance, suppose that some primordial elementary particle was unstable and decayed away more than a year after the Big Bang. The energy release from such a particle would be eventually added to the CMB by way of hot electrons interacting with the CMB, and would produce a distorted spectrum. Or, suppose that some sort of BHs or other explosive events produced the distribution of matter. The energy associated with them would eventually be taken up by hot electrons, and some would be added to the CMB. In 1989, as the reason for cosmic structure was not yet known, such extraordinary explanations were still somewhat plausible.

By 1992, when the COBE team announced the discovery of cosmic anisotropies, theorists were becoming very confident that the anisotropies would have to be found, with about the same properties that were measured (e.g., [267]). These theorists based their calculations on the simulations of the formation of galaxies and their measured clustering properties. Without some initial density variations from the Big Bang, galaxies could not be formed. Nevertheless, the actual discovery was greeted warmly, and Stephen Hawking was quoted as saying it was the discovery of the century, if not of all time.

Thank you John. Indeed COBE has been a cornerstone mission for cosmology: the first detection of CMB anisotropies has been widely recognized as the cosmological discovery of the century and COBE marked the transition toward modern precision cosmology. In this respect, the WMAP mission with its wealth of information on a wide set of cosmological parameters likely represents the most successful current example. We will see why in the next interview with Charles Bennett.

2.8.2 WMAP

Dear Charles *(Bennett)***, after COBE, WMAP likely represents the most relevant cosmological space mission of the new century. Can you illustrate the scientific context asking for such a mission?**

The WMAP mission provided detailed temperature and polarization maps of the CMB radiation over the full sky. Five frequency bands allow for the separation of the CMB from astrophysical foreground emission. WMAP data (either alone or in combination with other cosmological data) provide tight constraints on the history, contents, and geometry of our Universe. With WMAP, the CMB became the premier baryometer of the Universe; the existence of DE was solidly confirmed and determined to be 74% of the mass-energy of the Universe; the age of the Universe was determined to be 13.7 billion years (within $\sim$1%, even with no external data); the epochs of matter–radiation equilibrium, decoupling, and reionization have been determined. The precisely determined coherence of the detected baryon acoustic oscillations across the sky, the lack of spatial curvature, and the spectral index of the spatial fluctuations all support the simplest predictions of the inflationary paradigm. WMAP provided a precision calibration of the baryon acoustic oscillations, which have been seen in galaxy redshift surveys, and can now serve as a standard ruler for DE measurements. As a result, WMAP has spawned a new generation of ground and space DE experiments. WMAP products are released through the Legacy Archive for Microwave Background Data Analysis (LAMBDA)[3].

The central organizing idea of our modern view of cosmology is the "Big Bang" expansion of the Universe. Despite the widespread misconception, the Big Bang theory is not a theory of the origin of the Universe. The theory contains no physics

[3] See LAMBDA in web page list

to explain the "initial singularity". Rather, the Big Bang theory posits that the early Universe was hot and dense and has expanded adiabatically ever since. Thermal equilibrium was established between the matter and electromagnetic radiation when the Universe was young. As the radiation was in thermal equilibrium with matter, the spectrum of the radiation equilibrated to the specific BB form. In the billions of years of the adiabatic expansion of the Universe the initially high-energy radiation of the infant Universe became low energy microwave radiation, with a temperature now <3 K. This radiation was predicted by Ralph Alpher and Robert Herman at the Johns Hopkins Applied Physics Laboratory just after World War II [7], and was accidentally discovered by Arno Penzias and Robert Wilson in 1965 at the Bell Telephone Laboratories in New Jersey [400].

Does the CMB radiation have the predicted BB spectrum expected from radiation that had been in thermal equilibrium with matter early in the history of the Universe? A small rocket experiment initially tested this in the late 1980s, with a payload designed to measure the spectrum of the CMB. The published results from this experiment indicated a significant deviation from a BB spectrum. This result was explained by unexpected processes (such as the decay of an elementary particle) in the early Universe, and by a variety of other physical processes that violate the assumption of an adiabatic expansion of the Universe, which is the fundamental postulate of the Big Bang theory.

Led by John Mather, the COBE mission was launched into a polar orbit in 1989 from the Vandenberg Air Force Base in California. In 1990, with only 9 min of flight data used, the COBE mission made a precise measurement of the CMB spectrum, and it turned out to be a BB as originally predicted by Ralph Alpher. With the full mission data set, the spectrum of the CMB was measured with stunning precision and accuracy. COBE found the radiation temperature to be 2.725 ± 0.001 K, and limited any deviations from a BB to 50 parts per million of the peak brightness. This marked the beginning of the age of "precision cosmology".

A vexing question about the CMB remained. This question centers on the uniformity of the CMB temperature across the sky. Penzias and Wilson measured the radiation to be uniform across the sky to $\sim$10%. Later measurements would show the radiation to be uniform across the sky to nearly a part in hundred thousand. There is one exception: at the level of a part of a thousand, a Doppler effect from our own proper motion through the cosmos produces a dipole temperature distribution across the sky. This was first established in the early 1970s by the measurements of Conklin [111] and Henry [234]. I will ignore this "local" dipole effect in the remainder of this discussion.

Throughout the 1980s, measurements of the CMB revealed that the Universe began with a very smooth and uniform distribution of matter and radiation. Yet, today we see a very lumpy Universe. How did this transformation happen? At what level is there nonuniformity (i.e., anisotropy") in the temperature of the CMB radiation? In the late 1960s, it was believed that small initial fluctuations in the Universe grew under the influence of gravity to become the lumpy Universe we see today. Sachs and Wolfe established [483] that the dimensionless temperature fluctuations, $\Delta T/T$, are linear with the primordial gravitational potential fluctuations $\Delta\Phi/c^2$

according to $\Delta T/T = -\frac{1}{3}\Delta\Phi/c^2$. Thus, a few stalwart experimentalists set out to try to detect the tiny temperature fluctuations expected from the gravitational potential fluctuations needed to seed the formation of structures in the Universe. Theorists predicted that these fluctuations might be at the level of 10%, or a part in a thousand. The experimentalists eventually ruled these out (other than the dipole). A cycle ensued where theorists would predict ever smaller levels of fluctuations and experimentalists would respond with null results at ever lower levels. The lack of temperature fluctuations called into question the viability of the idea that structures formed in the Universe by gravitational collapse. We were missing something.

Suppose there is a type of matter that does not interact with light at all. That is, it does not emit, absorb, or scatter light. We call this "nonbaryonic DM". The "non-baryonic" specification distinguishes this type of matter from protons and neutrons, which may be dark only in the sense that atoms may emit radiation that is experimentally difficult to detect. This is often called DM; however, it is fundamentally different from nonbaryonic DM that does not interact with light at all.

The extra gravity from a sufficient density of nonbaryonic DM allows low level initial fluctuations to lead to today's lumpy Universe. Nonbaryonic DM could move quickly (nearly at the speed of light), like neutrinos, or slowly (much less than the speed of light). Relativistic matter impedes the formation of lumpy structures, so slow-moving matter is needed, that is, "cold dark matter". Thus was born a new cosmological model, dubbed the standard CDM model, which assumed enough CDM to make the geometry of the Universe flat.

The standard CDM model came about well after the COBE mission was proposed. During the long period of time over which the COBE mission was developed, suborbital measurements progressed so effectively that it was not clear that COBE would be a significant improvement. I volunteered to lead an effort to make the COBE radiometers more sensitive by redesigning, rebuilding, and retesting them for operation at colder temperatures, and do it within the requisite tight schedule. With the support of the COBE Science Team, I prevailed upon Tony Kerr, of the National Radio Astronomy Observatory, to serve as a technical consultant for the critical mixer portion of the redesign effort. The effort was completed on schedule, and made a pivotal difference in the ability of the COBE mission to realize one of its major accomplishments.

The data analysis team that I led produced the full sky CMB anisotropy maps. Provided with these maps, Ned Wright analyzed the statistics to arrive at the tell-tale excess variance of 30 μK rms (10 parts per million) at a 10° angular scale. I led the effort to determine how much of the excess variance might be due to astrophysical foreground emission, and Alan Kogut led the effort to place quantitative limits on how much of this excess variance could arise from systematic measurement errors. In 1992, the COBE discovery of primordial temperature fluctuations at a part in 100,000 was presented in a suite of four papers [34, 287, 520, 593]. The notion that the structures in the Universe evolved gravitationally from primordial gravitational potential fluctuations was now secure, but only with a substantial cold DM component in the Universe.

COBE was sensitive only to angular scales of 7° and larger. Pairs of spots on the sky with these separation angles could not have had time to equilibrate unless they were in causal contact at very early times, as in Inflation theory. Thus, the COBE measurements were of the scalar primordial gravitational potential fluctuations.

In summary, the two key determinations from COBE were (1) the verification of the BB spectrum of the CMB radiation, and (2) the determination of the amplitude of the primordial gravitational fluctuations. These were major achievements, but still left many unanswered cosmological questions. How old is the Universe? How much of it is made of baryons (atoms) and how much of cold DM? What is the geometry of the Universe? What were the physical processes in the very early Universe? How fast is the Universe expanding and is that rate constant or changing? Is the energy of empty space (Einstein's cosmological constant) zero? Did the Universe begin with a period of intense inflation, and if so, which kind?

To answer these questions, we needed new measurements. We needed to measure the CMB radiation on angular scales much smaller than COBE's 7° scale. We needed to measure the sub-horizon non-primordial development of the CMB temperature fluctuations. In 1993, I led a team that proposed the Microwave Anisotropy Probe (MAP) to NASA to do that.

What have been the fundamental experimental guidelines and technical solutions originally adopted by the WMAP mission?

The WMAP science team required that the relative CMB temperature be measured accurately over the full sky. The overriding design requirement was to control systematic errors that would otherwise contaminate the measurements. It was a fundamental goal of the WMAP mission to produce a map of the sky with very small instrumental correlations between pixels, so that the statistics of the sky could be tested in a relatively straightforward manner. This requirement motivated the mission name, "MAP". To achieve this, WMAP uses differential microwave radiometers that measure temperature differences between pairs of pixels on the sky. The differential measurement approach was fundamentally important, as was the need for multiple interconnections between measurements on the sky. The interconnections should be established rapidly, before the instrument has time to change. In 1 h WMAP observes 30% of the sky.

The WMAP science team had additional priorities on their short list. Among these was to maintain sensitivity to polarized signals. No particular polarization sensitivity specifications were levied as we did not want this to drive the mission, and cause development delays and cost over-runs. However, we did not want to lose polarization information if it was avoidable. In the end, WMAP maintained polarization capability. The mission was not optimized for minimum polarization systematic errors. To do so would have meant a rapid modulation between polarization states. We felt that this was well beyond the scope of the mission. Rather, the mission was designed to minimize temperature systematic errors. The temperature signal was rapidly switched in a Dicke radiometer configuration, as well as with the spin ($\sim$1 min), precession ($\sim$1 h), and annual cycles in a complex sky pattern that multiply interconnected the temperature differences between spots on the sky.

To facilitate the separation of the CMB from foreground signals, WMAP uses polarization-sensitive radiometers at five separate frequency bands centered at 23, 33, 41, 61, and 94 GHz (wavelengths of 13, 9.1, 7.3, 4.9, and 3.2 mm). There are 4, 4, 8, 8, and 16 channels per frequency, respectively, with FWHM beam sizes of 0.81°, 0.62°, 0.48°, 0.33°, and 0.20°. The radiometers are rapidly modulated with a 2.5 kHz phase switch. Amplitude calibration relies on the in-flight modulation of the cosmic dipole, and beam calibration relies on in-flight observations of Jupiter.

The COBE data were adversely affected by the mission's low Earth orbit. A small amount of Earthshine was picked up by the experiment's feed horns. Also, the thermal environment of the spacecraft varied as a result of the low Earth orbit. Even the Earth's magnetic field environment contributed to COBE's sensitivity to systematic effects. Given that WMAP was to make far more sensitive measurements, it had to be far more immune than COBE to these effects. So the WMAP science team decided that it would be important to take the spacecraft out of low Earth orbit and place it in an orbit about the second Lagrange point, L2, of the Earth–Sun system. The observatory, seen in Fig. 2.14, is kept in continuous shadow by a large deployed sun-shield, which includes the solar panels that power WMAP. The thermal emission from the Sun, Earth, and Moon would be effectively blocked from the sensitive WMAP instrument, and the thermal environment of the spacecraft was designed to be highly stable.

It was an often repeated mantra of the WMAP science team that, "We do not want to do everything. We just want to make limited, but high-impact CMB measurements, and do it quickly and cheaply". In the WMAP proposal to NASA, the science team explicitly predicted that the first acoustic peak, discussed later, would be established before WMAP would have results (which turned out to be the case). The CMB power spectrum state-of-knowledge before the WMAP results is shown in the grey band in Fig. 2.15.

The WMAP science team also recognized that many other cosmological measurements of various kinds would progress rapidly. (This turned out to be the case with numerous experiments, including the SDSS galaxy redshift survey.) It was explicitly stated in the WMAP proposal that WMAP data would be most powerful when used in combination with other cosmological measurement results, such as from large galaxy redshifts surveys.

The WMAP mission was proposed to NASA in 1995, selected in April 1996, and confirmed for development in 1997. The satellite, with its single instrument, was built, tested, and launched in just 4 years. It was launched on a Delta II rocket on 30 June 2001, at 3:46 p.m. Eastern Daylight Time from the Cape Canaveral Air Force Station in Florida. Since August 2001, WMAP has continually surveyed the full sky, mapping out tiny differences in the temperature of the CMB. Originally called the Microwave Anisotropy Probe (MAP), the science team asked NASA to rename the mission to honor the memory and accomplishments of David T. Wilkinson, a member of the COBE and WMAP science teams and a pioneer in CMB studies. NASA agreed, and when the first results became available in February 2003, NASA announced the new mission name: the Wilkinson Microwave Anisotropy Probe (WMAP).

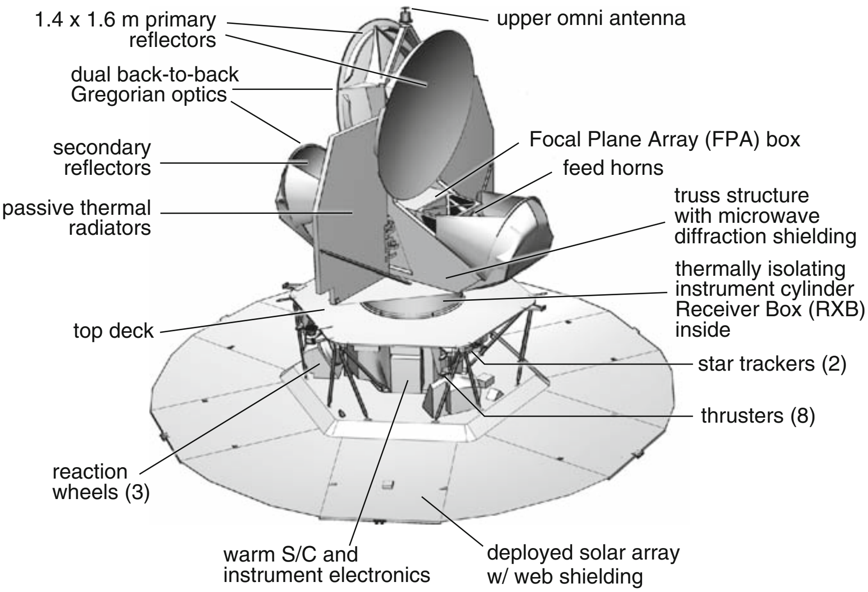

Fig. 2.14 Dual back-to-back Gregorian ($1.4 \times 1.6\,\mathrm{m}^2$) primary reflectors focus the microwave radiation from two spots on the sky roughly 140° apart and feed these signals to 10 separate differential receivers that are in an assembly directly underneath the primary optics. Large radiators, between the primary optics, passively cool the sensitive amplifiers in the receiver assembly below 90 K. The bottom half of the spacecraft provides the necessary avionics functions, such as command and data collection electronics, attitude (pointing) control and determination, power services, and a propulsion system. Attitude is controlled using two star trackers, two gyros, coarse, and fine Sun sensors, and three reaction wheels. The satellite spins at 0.464 rpm (∼2 min per spin) and precesses at 0.017 rpm (1 h per precession) about a 22.5° cone on WMAP-Sun line. Blow-down hydrazine propulsion with eight thrusters were used to achieve the L2 orbit and are also used for station-keeping. The spacecraft structure is made of carbon composite and aluminum materials. The total observatory mass at launch was 840 kg. Communications employ two omnidirectional antennas and a fixed medium gain antenna, which is used at 667 kbp for daily downlinks to the 70 m Deep Space Network. A $3.1\,\mathrm{m}^2$ GaAs/Ge solar array oriented 22.5° off the full Sun line, and a 23 A-hr NiH battery, provides the required 419 W. From [36]

WMAP provided a spectacular imaging of microwave sky components, both cosmological (i.e., CMB) and astrophysical (Galactic emissions, extragalactic foregrounds). How does the extraction of the CMB anisotropy depend on the subtraction of the various astrophysical foregrounds? How faithful is the WMAP measure of the cross correlation between temperature and polarization anisotropy and of the polarization itself at large/intermediate angular scales?

Foreground microwave radiation must be separated from the CMB. The fundamental basis for their separation is their differing spectra.

In the slice around the sky where the emission from the Milky Way galaxy is strongest, we do not have confidence in the precision of any model of microwave emission from the Galaxy. Instead, we simply omit this area of the sky from our

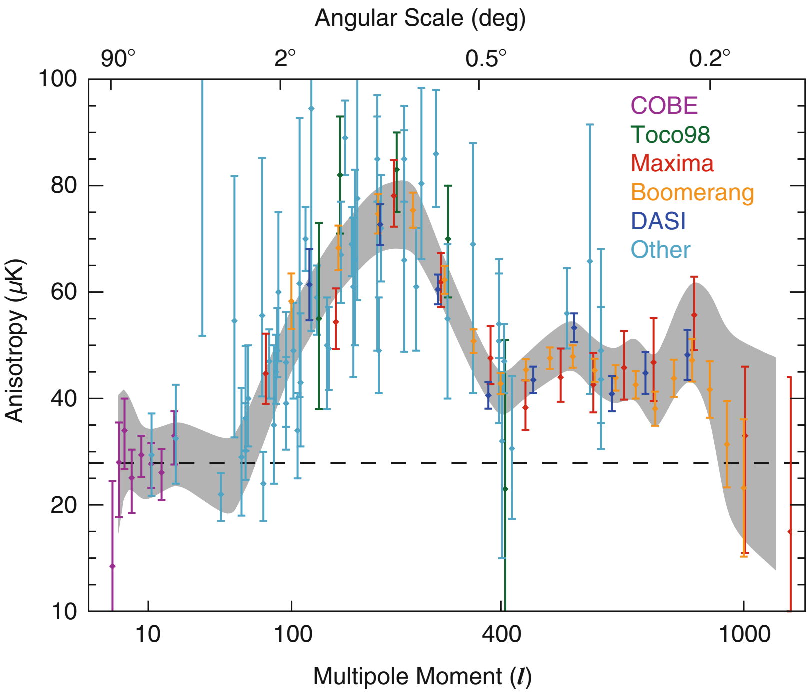

Fig. 2.15 Combining the existing measurements, the grey uncertainty band was the power spectrum state-of-knowledge before the WMAP results. From [36]

cosmological analyses. For the rest of the sky, we are fortunate that the CMB fluctuations dominate over the foregrounds, greatly reducing the required precision of foreground cleaning.

There are four approaches for handling of the foregrounds. First, we can mask out strong foreground regions. Second, we can physically model the individual foregrounds and subtract these models from the map. Third, we can attempt to cancel the foreground emission, regardless of the specific emission process. Fourth, we can fit multiple external foreground data templates to the WMAP data and subtract those signals. It is very useful to use more than one approach to compare and cross-check results.

We implemented a Maximum Entropy Method (MEM) to estimate each of the physical emission components using a combination of WMAP data and external data sets. This is used for understanding the individual foreground components, but is never used in the production of cosmological results.

By contrast, we combine the WMAP observations at different frequencies in such a way as to leave the CMB intact, while minimizing signals with spectra associated with foreground emission. This is, in a sense, a matched filter for signals with a CMB spectrum. We call this the Internal Linear Combination (ILC) method as it uses a linear combination of multifrequency data internal to the WMAP

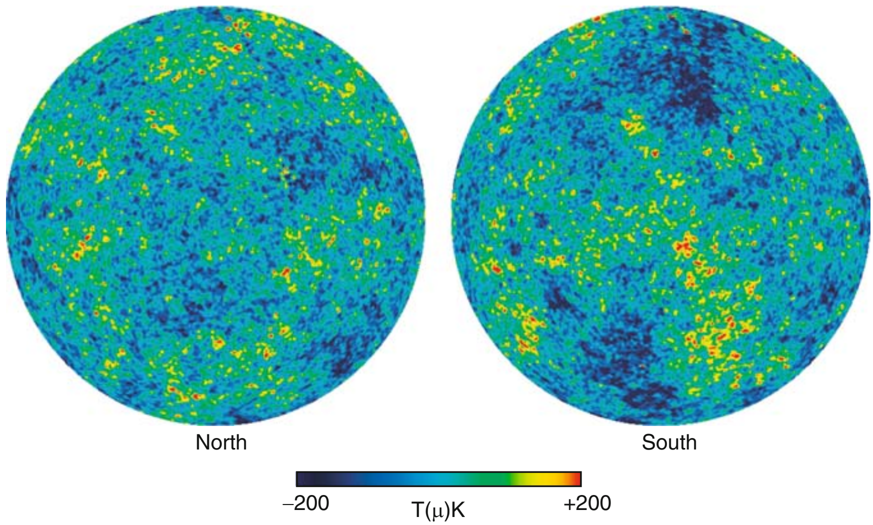

Fig. 2.16 The WMAP full sky map with foregrounds removed. The two views are centered on the north Galactic pole (*left*) and the south Galactic pole (*right*). From [241]

observations. The ILC map is shown in Fig. 2.16. The disadvantage of this technique is that the noise properties of the data become complex. This is not a problem for the largest angular scales as these data are limited by cosmic variance, but it is inappropriate for cosmological analyses on small angular scales.

A mainstay of WMAP analysis is foreground cleaning through the use of templates created from observations independent of WMAP data. For example, thermal dust emission is weak in the WMAP frequency range, but is very strong at much higher frequencies. Relative sky maps of dust emission from high-frequency surveys, where the emission is strong, can serve as a template for removing an estimate of dust emission from WMAP data. A set of templates, each with its own scaling amplitude, can be simultaneously fit to the WMAP data. The components are then subtracted at the appropriate scaling levels. This process is very effective, and does not complicate the noise properties of the WMAP data.

The problem of estimating, modeling, and/or cleaning the polarized foregrounds is far more difficult for three reasons: (1) there are no external high-quality polarization maps of the full sky to use as templates; (2) the polarization that WMAP observes generally has a low signal-to-noise ratio; and (3) the Galactic polarized emission dominates over the polarized CMB signal across much of the sky.

One solution we implement is to use the lowest frequency WMAP band as a polarization template for the higher frequency bands. We also use the limited data available in the polarization of starlight as a proxy for the polarization of thermal dust emission. We still mask out regions of the sky where the polarized foregrounds are too strong to correct with sufficient accuracy and precision.

After accounting for foregrounds, WMAP measures a small but significant CMB polarization signal on large angular scales. This signal has been used to specify the optical depth from the epoch of last scattering to the present. The results imply that the first stars formed and reionized the intergalactic gas ∼400 million years after Inflation. The polarization signal is still noise-dominated and will improve with additional data.

What kind of physical information is conveyed by these observables?

The WMAP full sky map of the fluctuations of the CMB temperature is largely a map of the sound wave pattern at the time of decoupling of photons and matter.

In the early Universe, CDM is attracted into primordial gravitational potential wells. Baryons are gravitationally attracted to the combination of the potential wells and the collapsing CDM; however, the enormous radiation pressure in the early Universe prevents the baryons from following the CDM. The tug-of-war between the gravitational attraction into potential wells and the restoring force of the radiation pressure sets up an oscillation in the photon–baryon fluid. This sound wave oscillation continues until the radiation pressure no longer restrains the gravitational collapse of the baryons. This happens precipitously at decoupling, when the temperature of the Universe drops to the point where free electrons bind to protons to form neutral hydrogen atoms. Thomson scattering that previously tied the photons to the electrons ceases, thereby removing radiation pressure from the matter. Without the fog of electrons, the light free-streams across the Universe, preserving a "snapshot" of the Universe at the time of decoupling.

Before the photons decoupled, the sound waves represented a series of compactions and rarefications of the gas. Photons from the hotter and denser compacted regions are slightly higher energy than photons from the rarefied regions.

The Universe is filled with sound waves between the end of inflation and decoupling. The smallest wavelength sound waves oscillate through many cycles during that 380,000 year interval. The fundamental sound wave is the one that goes from maximum to minimum compaction (or rarefaction) between the end of inflation and decoupling. Sound wave overtones of the fundamental have wavelengths that are an integer fraction of the fundamental.

The power spectrum of the CMB, in Fig. 2.17, reveals the sound waves. The largest peak, at about 0.6°, is the fundamental mode. The second peak is from a full cycle of maximum compaction to maximum compaction (or minimum to minimum). This pattern is seen all over the sky, in a coherent manner, which means the sound waves were triggered synchronously everywhere. This is now a powerful verification of a prediction of Inflation.

The reason that the harmonic peaks do not go on at equal amplitude to smaller scales is because they are damped on the scale of the mean-free-path of photon–electron Thomson scattering events.

As the maximum compaction and maximum rarefaction represent opposite extremes of the photon–baryon fluid, the ratio of the amplitudes of the first and second acoustic peaks is a direct measure of the baryon density of the Universe. With the precision of the WMAP data, this has become the premier baryometer of our

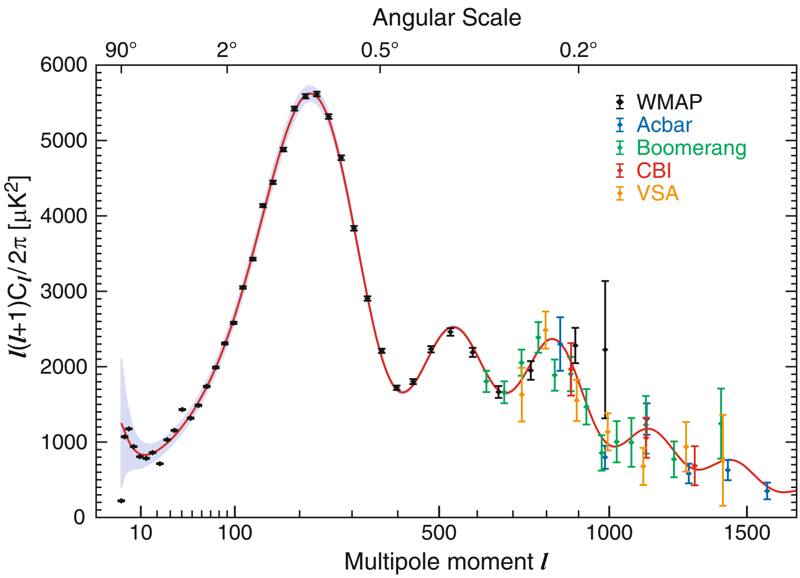

Fig. 2.17 The WMAP data are shown supplemented by other CMB data taken at higher angular resolution. Note the enormous improvement in the uncertainty between the pre-WMAP uncertainties in the grey band in Fig. 2.15 and the WMAP error bars in this figure. The best fit cosmological model is shown. We decompose the map into spherical harmonics and study the variance of the spherical harmonic coefficients as a function of multipole orders. The result shows a series of peaks, known as the Sakharov peaks, Doppler peaks, or baryon acoustic oscillations. This harmonic sequence as a function of wavenumber is the signature of a preferred angular scale. The preferred angular scale is a preferred length scale at the last-scattering surface: the distance that the sound waves traveled between the end of inflation and the epoch of decoupling. From [241]

Universe. The overall amplitude of all of the acoustic peaks measures the total matter density. The angular position of the acoustic peak pattern is a measure of the total mass-energy density in the Universe, as the image from decoupling is carried across the "lens" of curved space to reach us. The WMAP-measured peak location is consistent with no spatial curvature. This, along with multiple other measures of the cosmos, requires that a substantial amount of DE is needed.

The peaks are a manifestation of a preferred length scale. The angular size of the peaks therefore allows us to measure the distance to the last-scattering surface. The angular size depends very sensitively upon the spatial curvature of the Universe and less sensitively upon the details of DE. We conclude from the WMAP data that the radius of curvature of the Universe is much larger than the currently visible portion.

Thank you Charles. Now, before continuing the important discussion on the removal of foregrounds from the CMB signal, we briefly overview the results coming from balloon experiments. In fact, they allowed the first decisive probe of the geometry of the Universe, exploiting the degree and sub-degree pattern of CMB anisotropies. Balloon experiments are also crucial for the high resolution mapping *of the foreground at sub-millimeter and far-infrared wavelengths, a range uncovered by WMAP, and therefore for the separation of the foreground from the CMB signal. Juan Francisco Macias-Perez will discuss these aspects further.*

2.8.3 Balloon-borne Experiments

Dear Juan *(Francisco Macias-Perez)***, balloon-borne experiments allowed the discoveries of acoustic peaks in the CMB anisotropy temperature and polarization power spectra at degree and sub-degree scales and at the same time mapping the Galactic foreground produced by dust grains in both total intensity and polarization, observing at frequencies not accessible working at ground. Could you describe the most relevant observational results obtained on these lines?**

Balloon-borne experiments like the Far Infra-Red Survey (FIRS) [192], the Millimeter wavelength Anisotropy eXperiment (MAX) [8], Astrophysical Radiation with Ground-based Observatory (ARGO) [123], the Medium-Scale Anisotropy Measurements (MSAM) [97], the TopHat-Balloon-borne Anisotropy Measurement (TopHat-BAM) [560], and QMAP [128] were present from the beginning in the race for the discovery and measurement of the CMB temperature anisotropies at large and intermediate angular scales. However, the gold age of balloon-borne CMB experiments started in 1998 with the Balloon Observations Of Millimetric Extragalactic Radiation and Geophysics (BOOMERanG) [124] and the Millimeter Anisotropy eXperiment IMaging Array (MAXIMA) [228] flights. The BOOMERanG experiment [124] was based on high sensitive bolometer detectors cooled down to 300 mK and operating at 90, 150, 240 and 400 GHz, respectively. The 1998 BOOMERanG flight lasted for 10 days and lead to a high sensitivity CMB map covering about 2% of the sky at a resolution of 10 arcmin. MAXIMA was also based on high sensitive bolometers but cooled down to 100 mK and operating at 150, 240, and 410 GHz. The MAXIMA 1 flight just in few hours lead to a high sensitivity CMB map of 124 square degrees with a resolution of about 10 arcmin.

The first results on the measurement of the CMB temperature power spectrum came from the BOOMERanG team [125] and were confirmed shortly after by the MAXIMA team [228]. Both data sets showed a highly sensitive measurement of the CMB temperature power spectrum at degree and subdegree scales, showing a peak at about 1° and evidences for a second peak at smaller angular scales. These peaks were identified [125, 532] as the acoustic peaks expected in the case of inflationary cosmologies relegating topologically based cosmologies to a secondary role [67]. Furthermore, the position of the first acoustic peak at about 1° clearly favored a flat (Euclidean) universe [125] as predicted by inflationary cosmologies. To my opinion, these results are capital for modern cosmology as they have opened the era of inflationary cosmologies, which have proved very successful to date. The 1998, BOOMERanG and MAXIMA data sets have been deeply reanalysed [126, 314].

These analyses confirmed the first results and extended the measured CMB temperature power spectrum towards smaller angular scales, showing evidences of a third acoustic peak.

Although there were no doubt on the existence of the acoustic peaks in the CMB temperature power spectrum, their relative amplitude with respect to the large angular scale anisotropies measured by the COBE satellite [520] were matter of discussion. To solve this issue, the *Archeops* balloon-borne experiment was launched in 2001. *Archeops* [38] was also based on bolometers cooled down to 100 mK and operating at 143, 217, 353, and 545 GHz. In a 24 h flight, *Archeops* covered 30% of the sky, permitting the measurement of the CMB temperature power spectrum at large and intermediate angular scales, including the Sachs–Wolfe plateau and the first and second acoustic peaks [38, 559]. These results were of great importance as for the first time it was proved that large and intermediate angular scale experiments were properly intercalibrated. This clearly gave more confidence on the cosmological parameters estimated from CMB experiments [37].

The BOOMERanG [422], MAXIMA [489], and *Archeops* [117] CMB maps, because of their high quality, were also used to study the statistical properties of the CMB temperature anisotropies, which are expected to be Gaussian distributed in most inflationary models. These analyses have strongly constrained the nonlinearity parameter, f_{nl}, showing that the balloon-borne measured CMB temperature anisotropies are compatible with Gaussianity. These maps have been also used to study secondary sources of CMB anisotropies, as for example the Sunyaev-Zel'dovich Effect (SZE). A joint analysis [239] of the *Archeops* and WMAP data lead for the first time to a clear global detection of the SZE.

Finally, the 2003 BOOMERanG flight have greatly contributed to the measurement of the CMB polarization at intermediate and small angular scales. For this flight, BOOMERanG was equipped with eight polarization-sensitive bolometers at 150 GHz, allowing to measure the temperature and polarization power spectra, TT [261], TE [406], EE, and BB [364]. By comparison to previous WMAP results [385], the observed temperature and polarization correlation, and in particular, the phase shift between the temperature and polarization anisotropies, has been confirmed at smaller angular scales. This phase shift is compatible with that expected from inflationary cosmologies [328]. More importantly, the BOOMERanG team have obtained 4.8σ detection of the EE CMB modes and opened the race for the detection of the BB modes.

The contribution of balloon-borne experiments to the study of the dust foreground emission is not as important as their contribution to the study of the CMB anisotropies. This is due mainly to the fact that in most cases they observed clean regions on the sky where the dust contribution was subdominant and therefore the foreground contribution to the CMB signal minimal. Furthermore, most of the effort on the analysis side have been concentrated on the estimation of the CMB temperature and polarization anisotropies. However, as no other data at the millimeter regime are currently available, the analysis of their contribution on the understanding of dust emission is capital.

Clearly, the main result on dust emission from the BOOMERanG experiment (see [342]) is that, at high Galactic latitudes, the emission from dust in total intensity and polarization at 145, 245, and 345 GHz is lower than expected from the infrared data. Indeed, the study of the Spectral Energy Distribution (SED) of the dust emission have revealed a steeper spectral index. This is of course good news for future CMB experiments.

A more detailed study of the dust foreground emission have been performed with the *Archeops* data. In total intensity, the large angular scale dust emission at 353 and 545 GHz are consistent with the expected emission from the analysis of the infrared data (see [326] for details). In polarization, the analysis of the 353 GHz data (see [39]) have shown a significant Galactic large scale polarized emission coherent on the longitude ranges $100° \div 120°$ and $180° \div 200°$ with a degree of polarization at the level of $4 \div 5\%$, in agreement with expectations from starlight polarization measurements. Some regions in the Galactic plane (Gem OB1, Cassiopeia) showed an even stronger degree of polarization in the range $10 \div 20\%$. These results suggest a powerful grain alignment mechanism throughout the ISM and a coherent magnetic field coplanar to the Galactic plane. At high Galactic latitudes, the signal-to-noise ratio on the 353 GHz maps is too low for a detailed analysis; however, it is possible to perform statistical studies by computing the temperature and polarization angular power spectra (see [425] for details). From this analysis, the temperature angular power spectrum (APS) was shown to be compatible with the one expected from the extrapolation of the infrared data. At Galactic latitudes, $|b| > 5°$, the *Archeops* data show a correlation at 4σ level between the dust total intensity and polarization emissions at very large angular scales. This would indicate that the dust polarized emission constitutes a major foreground down to 100 GHz for the estimation of temperature and polarization CMB angular power spectra at very large angular scales. This is very important as the imprints of the physics of reionization and inflation are at these large angular scales.

The *Archeops* data have been also used to study compact dust sources [134], which can also be considered as foreground dust emission for CMB studies. Three Hundred and four sources have been found mainly in the Galactic plane where the signal-to-noise ratio was adequate. All sources show a dust-emission like modified BB emission spectrum with temperatures covering the range from 7 to 27 K. For the coldest sources ($T < 10\,$K), a steep ν^{β_d} emissivity law is found with a surprising $\beta_d \sim 3 \div 4$. This leads to an inverse relationship between the temperature and the emissivity index that cannot be explained by standard interstellar dust models and would require to invoke specific modifications of the optical properties of the dust. A similar inverse relationship was found by [143].

2.8.4 Far-IR Foreground

Dear Juan *(Francisco Macias-Perez)*, **we observe the Universe from our Galaxy. So both CMB experiments, cosmological extragalactic background, and distant galaxy observations are affected by emission and absorption by dust grains in the IGM. Can you physically discuss the main properties of this component?**

How it impacts cosmological observations in total intensity and in polarization? What kind of surveys are needed to precisely map this source of extinction and emission? What data analysis approaches are used to clean cosmological data from this spurious contribution?

Indeed, when we observe the Universe from earth or from space we look through the Galaxy and therefore both the Galactic emission and absorption of light must be taken into account. In a simplified vision, we can say that dust grains in the galaxy absorb photons at optical and ultraviolet wavelengths, then heat up and emit photons in the millimeter and submillimeter regimes.

The total absorbed intensity, I^{ab}, as a function of frequency, ν, can be expressed as follows

$$I_\nu^{\mathrm{ab}} = \alpha_\nu \times I_\nu^0, \tag{2.16}$$

where α_ν is the global absorption coefficient across the dust cloud and I_ν^0 is the initial source intensity. Notice that the value of the absorption coefficients depends on the composition of the dust grains that are typically formed of silicates or carbon compounds [175, 498]. The emitted intensity, I_ν^{em}, follows a modified BB spectrum of the form

$$I_\nu^{\mathrm{em}} \propto \nu^{\beta_{\mathrm{d}}} B_\nu (T_{\mathrm{d}}) \tag{2.17}$$

where B_ν is the Planck law. T_d and β_d are the dust temperature and emissivity indices, respectively. The dust temperature and emissivity index also vary with the composition of the dust grains.

Observations by the FIRAS instrument of the COBE satellite [107] and by the IRAS [251] satellite has revealed the emission of dust clouds on the whole celestial sphere. Using a combination of these data, Schlegel et al. [498] produced a map of the dust emission at 100 μm from which the temperature and emissivity indices of dust were derived by [175] for multicomponent models. The best-fit model consists of two main components with temperatures of 9.4 and 16.2 K and spectral indices of 1.67 and 2.7. The first and second components have been assimilated to silicates and carbonaceous components, respectively. However, no direct proof of this association has been presented to date and the physical pertinence of this model remains controversial.

2.8.4.1 Dust Extinction

If we observe an astrophysical source which is situated behind a dust cloud, a fraction of the source radiation is absorbed and therefore it appears fainter to us. This phenomenon is generally called extinction and is common in optical wavelength cosmology. As the observer is embedded in the Galaxy, for any position on the sky, in the path from the observer to the astrophysical source, we expect to find dust cirrus and therefore, an extinction of the flux of the source. In the context of cosmology, this problem is common to all large scale structure surveys (see, e.g., [505, 517]), which collect the emission from individual distant galaxies to infer the large scale structure of the Universe. It also affects SN search projects that measure the properties of distant type I SNe, which are considered to be standard candles and therefore can be used on cosmological tests (see [401] for review). The works of Schlegel et al. and Finkbeiner et al. [175, 498] are currently used to correct cosmological surveys from extinction.

In the case of cosmological surveys for which the redshift of the sources is in general needed, it is important to notice that the extinction by dust depends on frequency and therefore it is different for different spectral bands of observation (see [93] for details in the most commonly used spectral bands). As often, the redshift of cosmological sources is estimated by color (difference of flux between two given spectral bands) comparison; extinction by dust causes the reddening of the source and therefore an overestimation of the redshift of the source. To correct from this effect, the color correction maps derived from [175,498] are generally used. Notice that the reddening of the source does not affect the estimation of redshift using spectroscopy measurements, where the position of known absorption and emission lines is compared to a zero redshift reference or laboratory measurements.

2.8.4.2 Dust as a Foreground for CMB Studies

As the CMB temperature is 2.728 K [177], the frequency range of choice for observing the CMB emission is in the radio and millimeter regime from 10 to 800 GHz. At these frequencies, the emission from the Galaxy has three main components : free–free, synchrotron, and dust.

These foreground emissions contaminate the measurement of the CMB temperature anisotropies and need to be removed from the CMB maps before any cosmological analysis (see [66] for a review on the subject). On Fig. 2.18, we present the SED of the CMB and the Galactic foreground emissions outside the Galactic plane as a function of frequency in giga hertz. From the figure we can see that the Galactic free–free and synchrotron emissions dominate the CMB emission at low frequencies, roughly below 30 GHz, while the Galactic dust emission dominates at high frequencies, roughly above 300 GHz.

As discussed earlier, the vibrational dust emission follows a modified BB spectrum (see (2.17)) peaking around 100 μm. In brightness temperature units (we consider here the Raleygh–Jeans approximation of the BB spectrum), this spectrum transforms into a simple power law (see Fig. 2.18) in the radio and millimeter regimes where the spectral index has the same numerical value that the emissivity index defined earlier.

In first approximation, we can assume that the dust grains responsible for the millimeter and submillimeter emissions are the same as those responsible for the infrared emission. Therefore, an effective spectral index for this power law for each position on the sky can be obtained from the best-fit model proposed by [175]. Using such an effective spectral index and the 100 μm map of [498], we can model the expected dust radio and millimeter emissions (see [312] for an up to date version of this kind of models).

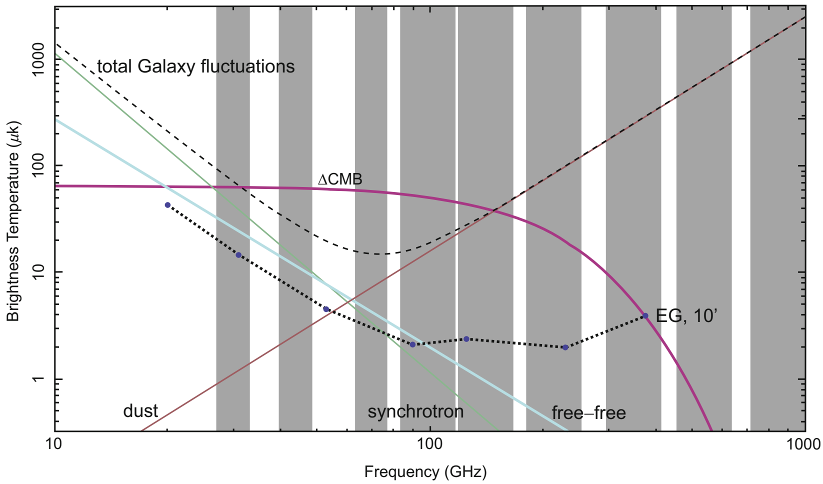

Fig. 2.18 Spectral energy distribution of the fluctuations of the CMB compared to that of Galactic and extragalactic foreground emissions (outside the Galactic plane) on an angular scale of 1° expressed in terms of antenna temperature as a function of frequency. Note that in this representation CMB fluctuations decrease with frequency sharply on the Wien side of the 2.7 K BB. The *Planck* frequency bands are shaded *grey*. The three highest frequencies will primarily measure dust emission

From above one might think that dust foreground emission is not a problem for CMB studies as we have good physical insights on dust and a model of the dust emission at the frequencies of interest. However, the problem is that we dispose few observations in the millimeter and submillimeter regimes. Furthermore, they cover only partially the sky as they come mainly from either balloon experiments like, for example, PRONAOS [506], *Archeops* [38], BOOMERanG [124], and MAXIMA [228] or Antartic ground based experiments, like, for example, BICEP [598] and QUaD [4].

In the radio domain, we are richer in observations mainly coming from the WMAP satellite [35], but the data are difficult to interpret as dust is not the dominant component. So, currently it is difficult to fully falsify or validate the dust emission model at the frequencies of interest for CMB. Furthermore, some of the above observations show already partial evidences of complexity on the size distribution and composition of dust grains.

BOOMERanG measurements (see [342]) at 145, 245, and 345 GHz of the high Galactic latitude sky, for which the dust emission is low, show evidences for a steeper dust spectral index. The *Archeops* maps [326] at 353 and 545 GHz, at low and intermediate Galactic latitudes, are in good agreement with the [175] model for the diffuse dust component, which dominates the signal. The overall dust brightness as obtained from the WMAP data foreground analysis (see, e.g., the recent 5 year results in [207]) seems to be largely consistent with the [175] best-fit model predictions, although the data prefer a spatial distribution somewhat less sharply peaked toward the Galactic center. For the high S/N regions, the dust spectral index is compatible with that expected from the model. A clear interpretation of these results is quite challenging. The BOOMERanG results are valid only in the case of very low dust emission. A quantitative analysis of the *Archeops* data specially at the lower frequencies, 143 and 217 GHz, is needed. The WMAP results are obtained by considering a multiple component model including synchrotron, free–free, dust and CMB emissions in a frequency range where the dust contribution is subdominant. As a summary, we could say that for frequencies above 100 GHz, the diffuse dust emission seems to be roughly consistent with the IRAS data, and from the point of view of foreground subtraction, the SED of the dust emission can be well approximated in units of brightness temperature by a power law of spectral index between 1.5 and 2.

Quite unexpectedly, the major surprise about foreground dust emission and how it relates to dust physics came from the radio frequency data. Kogut et al. [288] found a statistical correlation at high Galactic latitudes of COBE DMR observations at centimetric wavelengths with DIRBE dust emission at 140 μm. Similar results were obtained by [317] for radio observations at 14.5 and 32 GHz. This *anomalous* microwave emission was interpreted by [141] as electric dipole emission from small dust grains after ruling out on energy basis the hypothesis of free–free emission. In the [141] interpretation, commonly known as spinning dust model, the dust SED shows a peak at about 20 GHz (see dotted line on Fig. 2.19). The dust grains responsible for this emission, rotational dust, are not the same that the ones responsible for the vibrational dust emission discussed earlier and therefore this implies the existence of at least two distinct populations of dust grains in the IGM.

Further evidences for this *anomalous* microwave emission came from CMB experiments such as Saaskatoon, Tenerife, WMAP, and COSMOSOMAS (COSMOlogical Structures On Medium Angular Scales) (see [585] for a detailed review). In my opinion, the most conclusive evidences on spinning dust emission comes from the COSMOSOMAS observations of the Perseus molecular cloud by [585]. Figure 2.19 shows the SED of the Perseus molecular cloud from 0.4 to 3,000 GHz. In the figure, we clearly observe the vibrational dust emission at high frequency, which follows a modified BB spectrum of temperature 19 K and emissivity index 1.55, the free–free emission at low frequencies and an extra component in the range 10–30 GHz corresponding to the anomalous microwave emission. The latter is compatible with the spinning dust model proposed by [141], shown as dotted line in the figure. Although these results, local to the Perseus molecular cloud, are of great interest for the understanding of the physics of the anomalous microwave component, they cannot be generalized to the full sky and therefore, they cannot be directly used to model the spinning dust contribution to CMB maps. For this purpose, several analyses of the WMAP data have been carried out, but the results remain controversial. From the recent 5 years maps, [207] concludes to an excess of emission of 14% at 33 GHz, which could be either explained by steepening of the synchrotron spectral index or by an extra dust contribution in the form of spinning dust. Other authors, using alternative analyses like recently [361], have pointed out the fact that the extra emission is correlated with the vibrational dust emission. Although the

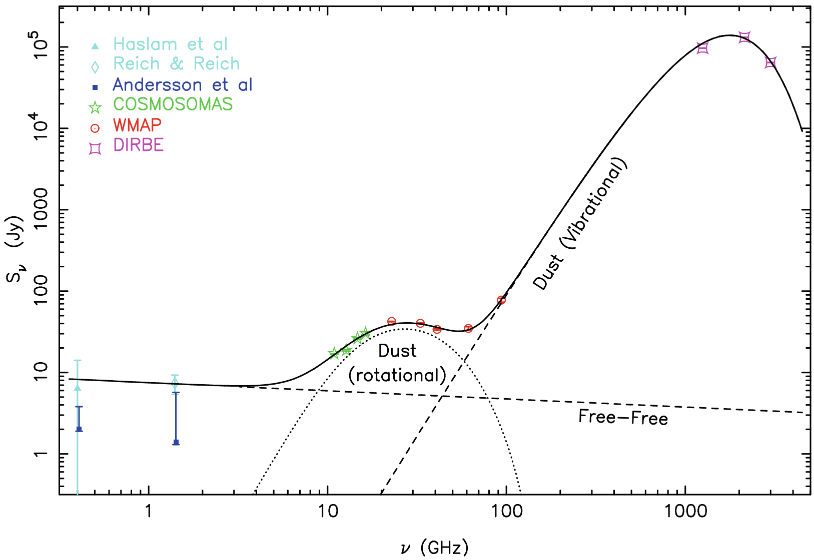

Fig. 2.19 Spectrum of G159.6–18.5. Points are shown for COSMOSOMAS, WMAP, DIRBE, and low-frequency (408 and 1,420 MHz) surveys convolved to give COSMOSOMAS-equivalent points. The size of the symbols corresponds to the sky-model uncertainty of 10%. Estimates are made of three foregrounds: free–free, rotational dust, and vibrational dust. Vibrational dust corresponds to a dust temperature of 19.0 K and an emissivity index of 1.55. A toy spinning dust model that consists of a linear combination of Draine and Lazarian [141] models is used. Figure courtesy of Watson et al. [585]

latter seems to validate the spinning dust hypothesis, we must remind that we also expect spurious correlation between dust and synchrotron as both follow the general morphology of the Galaxy. To really settle this issue, a better understanding of the Galactic synchrotron emission in the microwave regime is needed. So, although it may seem a paradox at first look, to improve our knowledge on the dust foreground emission, radio and microwave observations are needed.

2.8.4.3 Dust Polarized Emission

In the above discussion, we have voluntarily ignored the polarized foreground emission from Galactic dust. The main reason for this is that there is no dust polarization full sky survey neither in the infrared (the IRAS satellite did not include polarized detectors) nor in the submillimeter regimes. Currently, most of the information used to understand and model the dust polarization foreground emission comes from starlight polarization (see [233] for a detailed description of the available data).

Figure 2.20 shows a schematic representation of the IGM to illustrate how starlight polarization and the dust polarized foreground emissions are related. Dust

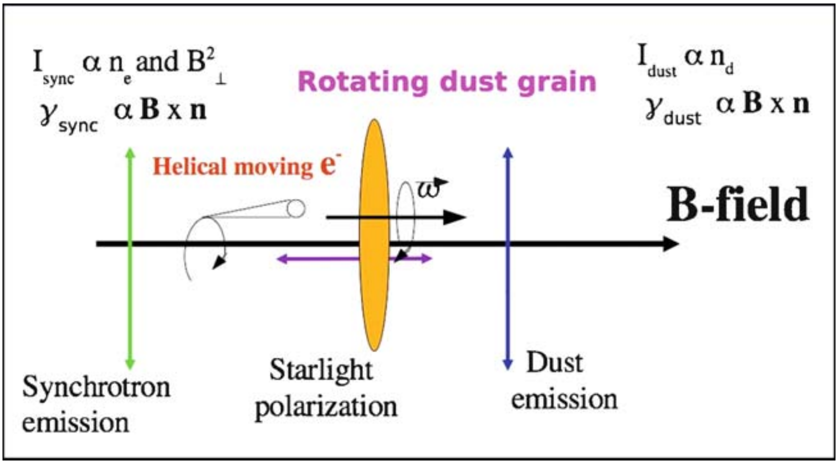

Fig. 2.20 The physics of the dust polarization emission

grains in the IGM are oblate, and in the presence of the Galactic magnetic field, they tend to align their larger axis perpendicular to the field lines and rotate so that their angular velocity, ω, is parallel to the field. Starlight is absorbed by the dust grain along its larger axis and therefore from the observer point of view it seems polarized parallel to the magnetic field direction. By contrast, the thermal dust emission occurs along the grain larger axis and therefore it is polarized perpendicularly to the magnetic field direction. As the grain rotates, the position of the observer modulates the polarization intensity. If the line of sight is parallel to the magnetic field line, the observed polarized intensity is zero.

On the contrary, if the line of sight is perpendicular to the magnetic field line, the polarization intensity is maximal. Between the two, the polarization intensity is modulated by a depolarization factor, which is a function of the angle between the magnetic field direction and the line of sight direction. This is a very simplified picture where we have assumed an optically thick dust grain and a mechanism for perfect grain alignment. As discussed earlier, Galactic dust grains do not verify the former and their opacity depend on their composition (see [498]). The currently proposed mechanism for grain alignment are not fully understood (see, e.g., [437]). However, under these assumptions, the measured starlight polarization suggest (see [181]) that at large angular scales the Galactic magnetic field lines follow the spiral arms as observed in other galaxies like M51. This has been independently confirmed using rotational measurements of pulsars [227]. Furthermore, the polarized *Archeops* maps at 353 GHz are compatible to first order with this interpretation (see [39]). For example, no polarization is observed in the Cygnus region for which the line of sight is parallel to the magnetic field lines that follow the spiral arm.

Based on the above discussion, the WMAP team [386] has constructed a simple model of the thermal dust polarized emission. They start from a model of the total intensity emission as the one discussed in the previous paragraphs and assume a global polarization factor. For each position on the sky, the polarization angle is assumed perpendicular to the starlight polarization angle. Then the polarized dust emission maps are corrected for a geometrical factor accounting for the expected depolarization. This approach seems to be sufficient to subtract most the dust polarized foreground emission on the WMAP maps that are noise dominated at high frequency. However, it is important to notice that this is a very rough correction as the starlight polarization data are spare and can only be used at very large angular scales and at high Galactic latitudes where the dust contamination at the WMAP frequencies is subdominant.

2.8.4.4 Component Separation Algorithms

Foreground removal and cleaning is crucial for CMB studies both in temperature and polarization. Many techniques and algorithms, commonly called component separation algorithms, have been developed for this purpose. A recent and detailed review of these methods in the framework of the analysis of the *Planck* data is presented in [312]. In general terms, we can separate those methods in two distinct categories: those based on foreground templates and those based on the statistical properties of foregrounds.

The foreground template-based techniques have been widely used on current and previous CMB experiments. Templates of the foreground emissions are generally constructed from observations at frequencies where the studied foreground emission is dominant. For example, in the case of vibrational dust emission, the [498] 100 μm map is currently used as a template when no high frequency data are available (see, e.g., [35]). When high frequency data are available (see, e.g., [326]), they are taken as template of the dust foreground emission. When dealing with diffuse Galactic foreground emissions, the templates are extrapolated to the CMB frequencies assuming either a simple multiplicative coefficient or a power law in units of brightness temperature. The multiplicative coefficient of the spectral index of the power law are allowed to vary spatially and are obtained by χ^2 minimization. Recently, bayesian sampling methods have been proposed (see, e.g., [157, 207]) to obtain more reliable estimates on the correlation between parameters and the error bars associated to them. So far, these algorithms are computationally expensive and are limited to low resolution data sets (angular scales of the order of the degree or larger). These foreground cleaning techniques are also used on polarization data as, for example, on [386] for which we have discussed earlier the way the dust polarization emission template is constructed.

The component separation algorithms based on the statistical properties of foregrounds aim at recovering from the data the spatial distribution of the different foreground emission components (no template is assumed). They can be divided into two main categories: those for which the SED of the foreground emissions are assumed known or partially known, typically called *nonblind* methods, and those for which the SEDs are also estimated from the data, typically called *blind* methods. In the latest years, there has been an explosion of the number of available methods (see [312]) mainly developed for the analysis of satellite data from WMAP and *Planck*. The main reason for this is that this kind of algorithms can work only in the case of observations at multiple frequency bands (5 or more). In here, we do not intend to give a detailed review of component separation algorithms, but rather to discuss their main characteristics. For this purpose, it is interesting to remind to the reader the Bayes theorem on which most of these techniques are based

$$P(M|D) = P(D|M) * P(M)/P(D), \tag{2.18}$$

where $P(M|D)$ represents the probability for model M, given the data D, also known as posterior probability, $P(D|M)$ is the probability of the data, given the model, also known as likelihood, and, $P(M)$ and $P(D)$ are the priors on the model and the data, respectively. In component separation algorithms, we search the model M that maximizes the posterior probability. Generally, we consider a flat prior for the data, $P(D) = 1$ and a Gaussian likelihood.

There exists two main *nonblind* algorithms: Wiener filtering and MEM ([242]). The former assume a Gaussian prior for the model while the latter assumes an entropic prior. The entropic prior permits a more general spatial distribution of the foreground emission and is more widely used both for intensity and polarization data. See, for examples, the analysis of WMAP data by [207] and of the future *Planck* data by [531]. For the *blind* case, most current techniques derive from the classical Fast Independent Component Analysis (FastICA) [334] and Spectral Matching Independent Component Analysis (SMICA) [131] algorithms. Both algorithms look for independent components in the data set and estimate simultaneously the noise in the data, the SED, and spatial distribution of the different physical components: CMB and foregrounds. FastICA works on the real space and search for independent components in the data that maximize the non-Gaussianity of the mixture. SMICA works on the spherical harmonic space and search for a fixed number of independent components in the data that maximizes the posterior probability described earlier. Both algorithms can be extended for polarization analysis, see, for example, [17].

2.8.5 Interstellar Medium

Dear Juan *(Francisco Macias-Perez)*, **the ISM consists of various gas and dust components, products of stellar and galaxy evolution. Can the study of the properties of the ISM in our Galaxy provide useful constraints on cosmological evolution? Can a comparative study of ISM in our Galaxy and in distant galaxies present synergies relevant for our comprehension of some aspects of cosmic evolution?**

By definition, the aim of cosmology is to study the Universe as a whole to determine globally its properties and evolution. As a consequence, most cosmological models present a stochastic nature and are not meant to exactly reproduce each of the observed structure but rather their statistical properties. For example, if we consider structure formation simulations, we do not expect them to match exactly with the observed structures at different epochs and positions. However, we expect, other than reproducing the observed statistical properties, to find within the simulations galaxies that look like the observed ones and in particular like the Milky Way.

At this respect, the study of the properties and time evolution of our Galaxy is of cosmological interest. Among many other issues, this kind of studies can help us to understand how disk galaxies are formed (see [513] for further readings), to identify the missing processes in the current physical models of galaxy formation and evolution (see, e.g., [112, 155]), to elucidate how and at which rate stars are formed (see, e.g., [272]), and to investigate the chemical evolution of galaxies [427]. Some of these questions will be observational, addressed by the future Global Astrometric Interferometer for Astrophysics (GAIA) satellite mission to be launched by ESA in 2011 (see [189]) and has been the main topics of a dedicated IAU symposium ([249]).

Gas and dust are the main components of the ISM and participate in the cycle of formation and destruction of stars in our galaxy and in other galaxies. The stars are formed by the gravitational collapse of the interstellar matter mainly dominated by hydrogen gas, they bright up for few billion years and the end of their lives they give away some of their matter in the form of dust. Therefore, the chemical composition of dust is closely related to that of the stars. So we can say that the study of the distribution and morphology of gas and dust in the Galaxy can help us to understand the morphology, dynamics, and SF history of the latter. As discussed before, these properties of the Galaxy are of cosmological interest. A cosmological scenario to be fully validated needs to be able to produce galaxies with the physical properties of the Milky Way. From this point of view, the study of gas and dust in the Galaxy, rather than constraining directly the cosmological parameters associated to a given particular model, can help to improve our knowledge on the physical processes involved in the evolution and formation of galaxies. In summary, we can conclude that the study of gas and dust in our galaxy provides indirect constraints on cosmological scenario of galaxy formation.

In the same lines, it seems natural to think about synergies between the Galaxy and other galaxies. For example, we expect external galaxies of the same type as ours to present similar physical and chemical properties. Being close to the ISM, we can perform detailed studies that are impossible for distant galaxies.

Thank you Juan. The precise understanding of the properties of the foreground is also extremely important in the radio domain, in particular, for the analysis of CMB data at centimeter and millimeter wavelengths. Wolfgang Reich will now address the problem of the contamination coming from synchrotron and free–free emissions and of the extrapolation from radio to microwave domain.

2.8.6 Radio Foregrounds

Dear Wolfgang *(Reich)***, both CMB experiments and extragalactic radio-source observations are masked by Galactic synchrotron and free–free emission that are well mapped by surveys at radio frequencies. After a brief presentation of the main properties of these Galactic emissions, can you discuss how they impact cosmological observations?**

The Solar system is located in the thin disk of the Galaxy at about 8 kpc distance from the Galactic Center. Our position inside the Galaxy becomes immediately evident when looking at the Milky Way in a clear night showing a band of diffuse light from unresolved stars extending from horizon to horizon. The advantage of being inside a galaxy is the possibility to study individual objects and structure details with much higher spatial resolution than possible in any other nearby spiral galaxy. The disadvantage is the difficulty to get a precise overview on its large-scale structure. In a remarkable paper, J.P. Vallée [566] made a statistical study of the results from numerous investigations on the properties of the Galactic spiral arms. Rather different methods were used and two, three, or four arms were reported. The percentage of studies where four arms were found increases from 60% in 1980 to about 85% nowadays, indicating the difficulty of determining basic large-scale parameters for the Milky Way. For most distant galaxies, such parameters can be easily taken from any image.

Sky surveys at radio frequencies show the Milky Way as a narrow ridge of bright emission stretching across the sky. The radio emission originates from the thin ISM between the stars, not from the stars themselves. Radio emission is not attenuated by dust as the light from more distant stars. Therefore, the observed radio emission includes all components along the entire line-of-sight through the Milky Way, except for very low frequencies, where thermal gas clouds become optically thick and absorb emission from behind. There is no way out: radio emission from the Milky Way adds to signals received from anything outside the Galaxy, thus also to signals from extragalactic radio sources or CMB anisotropies. However, for CMB observations of a not too large patch of sky, a proper planning will minimize the influence of Galactic emission in most cases.

The starting point to estimate the Galactic contamination is a look at radio continuum and polarization all-sky surveys, which were made at a number of frequencies. They were intended to understand the properties of the diffuse large-scale Galactic emission and also a number of large loop structures with diameters up to 100°, which are most likely very local objects. All ground-based all-sky radio surveys were made at frequencies up to 1.4 GHz, where synchrotron radiation is by far the dominating emission process out of the Galactic plane. The 1.4 GHz all-sky survey, which is a combined map from the northern and southern sky observations, is shown in Fig. 2.21 (Reich et al., in preparation). Synchrotron emission originates from relativistic electrons when being deflected in the Galactic magnetic field. Its steep power law spectrum results from the energy distribution of cosmic ray electrons. Diffuse thermal radiation as visible in Hα surveys extends up to high Galactic

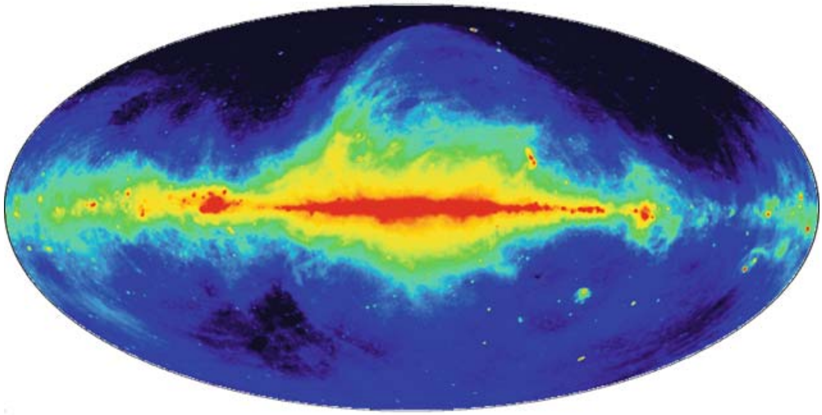

Fig. 2.21 All-sky total intensity distribution at 1.4 GHz shown in Galactic coordinates at an angular resolution of 36′

latitudes. However, its contribution at low frequencies is small, but increases towards higher frequencies because of its flat spectrum compared to that of Galactic synchrotron radiation. The synchrotron dominance is less for the quite intensive narrow emission ridge along the Galactic plane, where discrete thermal source complexes (HII-regions) were heated by massive stars. Also diffuse thermal emission makes a significant contribution in the Galactic plane even at low frequencies of a few hundred mega hertz.

All-sky surveys are difficult to perform in practice and it takes several years to complete them. This is the reason why only few have been made. They need suitable telescopes of about the same size in the northern and southern hemisphere. Survey observations are quite time consuming, and scheduling of large blocks of telescope time becomes difficult if they compete with projects where interesting results are obtained on short time scales. At higher frequencies, where sensitivities of a few milli-Kelvin are needed, atmospheric and environmental effects are larger than the receiver noise and change on shorter time scales than the duration of a single sweep of the telescope across the sky. These influences have to be properly taken into account during the data reduction process. Additional efforts have to be made to bring all-sky measurements on an absolute temperature level [442], which differs from the much more easy standard procedure, where for instance the edge areas of a map are defined as the local zero level. As a consequence of all these limitations, all-sky surveys have a relatively low angular resolution meaning about half a degree at best. Much higher angular resolutions, however, were achieved when mapping the narrow band of the intensive emission along the Galactic plane using synthesis telescopes with angular resolutions of about 1′ at 1.4 GHz. This allows to separate diffuse emission from discrete Galactic source structures with high accuracy and to study individual objects in detail.

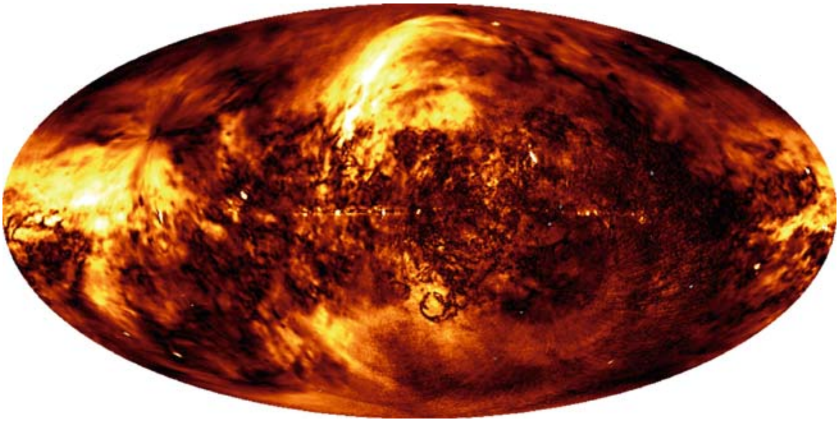

Fig. 2.22 All-sky polarized intensity distribution at 1.4 GHz corresponding to Fig. 2.21 (Reich et al. in preparation)

Also large-scale polarization surveys have been made. After the detection of diffuse polarized synchrotron emission in 1962 at Cambridge and Dwingeloo, numerous surveys mostly at 408 MHz were carried out. The highest frequency observed was at 1.4 GHz. These early data are limited in sensitivity and were severely undersampled. A new and so far unique ground-based all-sky survey at 1.4 GHz has just been recently obtained, which is shown in Fig. 2.22. Reich [441] has reviewed the status of Galactic polarization observations. The problem at 1.4 GHz are Faraday rotation effects in the ISM, which cause depolarization of the intrinsic polarized emission from the Milky Way. The amount of Faraday rotation is proportional to the thermal electron density and the magnetic field component along the line-of-sight. Faraday rotation depends on λ^2 and thus the influence of polarized synchrotron emission on polarized CMB fluctuations is even more difficult to estimate from low-frequency data than for total intensity emission.

CMB observations have to avoid the quite intense emission from the Galactic plane, where not only the high intensity level is of concern, but also the intensity fluctuations on small angular scales. It could be shown that a stripe of about $\pm 5°$ along the Galactic plane needs to be avoided at least. At high Galactic latitudes, thermal and nonthermal emission are weak enough to measure CMB total intensity anisotropies as shown by various ground-based and space missions. A complication may arise from the proposed GHz-emission originating from very small dust particles in an ambient magnetic field, what might result in another large-scale Galactic emission component. The emission spectrum from spinning dust particles was calculated to peak between 10 and 30 GHz and thus contaminates the lower frequency channels of WMAP and *Planck*. Nowadays, very little is known about the spinning dust component, reflecting the difficulty for observations in this frequency range. A few convincing observations have been made for individual Galactic sources, which prove that spinning dust emission exists and its spectrum agrees with theoretical predictions. However, for most HII-regions, their high-frequency spectrum

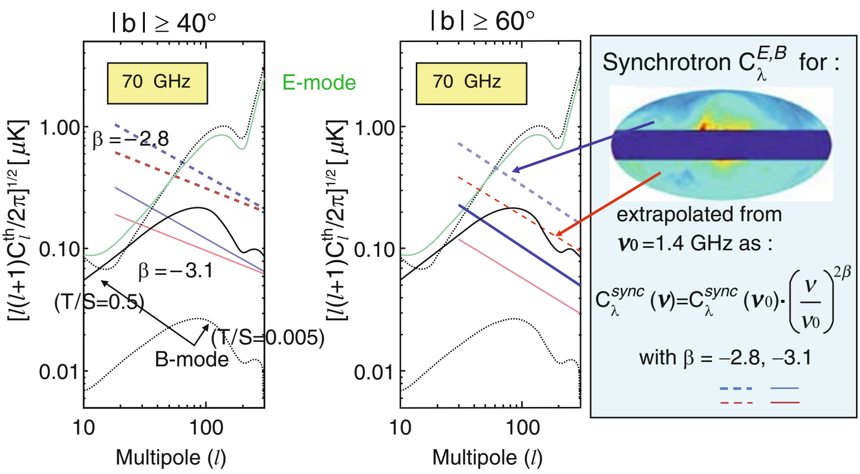

Fig. 2.23 Expected contamination of the E- and B-mode APS at 70 GHz by polarized Galactic emission at high latitudes for a range of temperature spectral indices. While the E-mode suffers little from Galactic foreground emission, the B-mode APS will be difficult to measure in case it is intrinsically weak. From [303]

is consistent with pure optically thin thermal emission. When adding all information about Galactic foreground components, the best frequency range for CMB anisotropy measurements, where the lowest foreground contamination is expected, seems to be around 70 GHz. Towards higher frequencies, thermal dust emission rises steeply.

70 GHz is also perfect for polarized CMB anisotropy observations, although a masking of Galactic large-scale emission for a stripe of about ±20° seems necessary for a clear detection of the E-mode anisotropy for angular scales smaller than about 4°. Detection of the B-mode spectrum will be extremely difficult, although its intensity level cannot be well predicted today. However, the most prominent peak in the B-mode anisotropy spectrum is expected for scales of the order of 2° and this peak might be detectable under favorable circumstances. A recent estimate on the contamination of polarized CMB fluctuations based on the 1.4 GHz map shown in Fig. 2.22 was made by La Porta [303] and is shown in Fig. 2.23.

What data analysis approaches are used to clean cosmological data from spurious foreground contributions?

The analysis of an observed radio map for signatures from the CMB is performed in terms of its APS, which describes the intensity fluctuations as a function of angular distance and can be compared with predictions made from cosmological models. While the APS from the CMB shows rather complex amplitude variations as a function of angular distance, the APS from Galactic emission is much more simple and can be fitted by some kind of power law plus eventually some additional curvature. However, some care must be taken to obtain a clean spectrum not being affected by discrete sources, beam size effects, noise contamination, and the limitations by the field size used for the analysis. Excessive strong emission along the Galactic plane and from nearby objects like the intense North Polar Spur, believed to be an evolved local SN remnant with a diameter of about 100° across the sky, have to be excluded. An estimate of the CMB anisotropy contamination by Galactic foreground emission needs an extrapolation of the APS derived from frequencies below 1.4 GHz, where Galactic survey observations are made, to frequencies higher than 20 GHz, where CMB observations are carried out. Galactic synchrotron emission declines with frequency and follows a steep power law spectrum. This power law extrapolation of the APS calculated at low frequencies is used for a prediction of the Galactic APS at high frequencies. However, the observed Galactic sky maps also contain some thermal emission, and attempts were made to separate both components from multifrequency data first and use these template maps to make predictions on their influence on CMB anisotropy measurements individually. These efforts led to "templates" of the sky for the various emission components, which have certainly some limitations, in particular when extrapolated to the high frequency range of CMB observations. Beside these "semi-blind" approaches also "blind" approaches were made, where the only assumption is that the CMB anisotropy field is the only component with a Gaussian distribution.

To what extent possible residuals may effect cosmological information?

One problem that limits the accuracy of an extrapolated Galactic APS in general is the uncertainty of the spectral index of the Galactic synchrotron emission at high frequencies. The spectrum of cosmic ray electrons becomes increasingly steeper with energy, which should be reflected in the observed radio spectrum. However, the spectral bend also depends on the strength of the magnetic field, which is not precisely known. Another problem is connected to unresolved extragalactic sources. The extrapolation of thermal emission is less problematic, as at least all diffuse emission can be assumed as optically thin in this frequency range and the spectral index is fairly accurately known. Visible sources have to be subtracted in any case, what is possible, although the accuracy may be limited in complex regions or unresolved double sources and some residuals may be left from this process in the map. The usually adopted Gaussian beam shape for the instrumental response of compact sources is a good although not perfect approximation. It has been shown that without source subtraction, a flattening of the APS on angular scales much larger than the angular resolution of the observations sets in. This is a severe limitation to extrapolate the diffuse Galactic emission APS at low frequencies towards high frequency CMB observations, in particular, when the angular resolution is low, as it is the case for all-sky surveys. The so-called "striping effects" are also of concern in the survey maps, which are always in the direction of the telescope movement and result from limited baseline stabilities of the long survey scans. However, as recently shown by [304], the overall APS is not really affected, as the resulting APS distortions are limited to a narrow range of angles.

How do Galactic depolarization effects impact CMB and extragalactic radio source observations?

Faraday effects in the Galaxy lead to depolarization of the intrinsically polarized emission. Synchrotron emission is polarized up to 75%, but much lower percentage polarizations are observed even at the highest frequencies, where all-sky maps from WMAP are available. The polarization maps from WMAP at 23 GHz and higher frequencies clearly show the intrinsic magnetic field direction. Observed Rotation Measure data from extragalactic sources shining through the entire Galactic disk indicate maximum deviations of the polarization angles of just a few degrees at 23 GHz due to Faraday rotation. Out of the Galactic disk, the polarization angle deviations are even smaller and typically below 1°. The low observed percentage polarization may result from a number of reasons. First of all the ratio of the random to regular (perpendicular) magnetic field component determines the observed percentage polarization. A small percentage polarization is thus expected in case the random magnetic field is significant or dominates, what seems to be the case throughout the Milky Way. The line-of-sight out of the Galaxy is always several kiloparsec close to the Galactic disk.

What can we learn from synchrotron emission about the magnetic field in our Galaxy?

The regular magnetic field direction seems to follow the spiral arms (Fig. 2.24) and thus changes direction with distance from the Sun. A reduced percentage polarization results by averaging the polarized synchrotron components along the line-of-sight. When comparing the recent 1.4 GHz polarization all-sky survey (Fig. 2.22) with that from WMAP at 23 GHz ([386]), significant depolarization in the direction

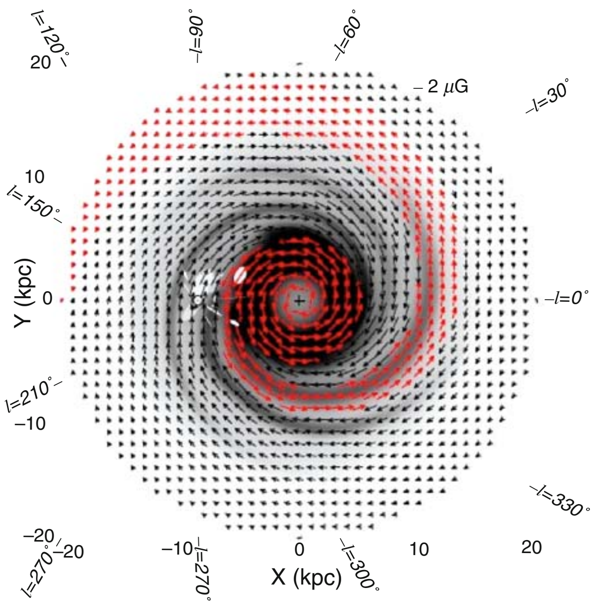

Fig. 2.24 Model of the regular Galactic magnetic field in the Galactic disk superimposed on the distribution of thermal gas (shown in *grey* scale). Some discrete thermal sources are indicated in addition. Inside the solar circle, a magnetic field reversal is needed to explain the observed Rotation Measures from extragalactic sources. From [542]

of the inner Galaxy is visible for latitudes $\pm 30°$. These areas cannot be used to make Galactic foreground predictions. At high Galactic latitudes, the spectrum of the polarized emission is close to that of the total intensities and depolarization effects are quite small, so that the Galactic influence on polarized CMB anisotropies can be estimated there (see Fig. 2.23).

Thank you Wolfgang.
Since its discovery, the standard interpretation of the microwave diffuse emission has not been fully accepted in the scientific community. We recall, that in the past, attempts had been pursued to explain the microwave background not as a residual of the hot Big Bang but in terms of integrated contributions by distant stars, possibly thermalized by dust. This, and other interpretations, different from the standard picture had trouble explaining the CMB properties as accumulated during the years following the discovery.

Some scientists hold unconventional views about the cosmological data and its interpretation. This is certainly the case of the next contribution by Pierre-Marie Robitaille, who in the interests of balance has been invited to take part to the debate. In the next interview, Pierre-Marie Robitaille will explain to us his radically different point of view in the interpretation of the CMB data.

2.8.7 A Radically Different Point of View on the CMB

Dear Pierre-Marie *(Robitaille)***, you have recently published a number of works posing several criticisms on the technique of data analysis used by the team of the COBE and WMAP satellites and advancing strong implications for the origin of the CMB. Would you please to summarize here your opinions?**

Recently, I have advanced serious concerns regarding the microwave background. My uneasiness, however, does not rest solely with satellite data. Rather, it originates with the measurement of Penzias and Wilson itself [400] and is based on a review of the initial formulations for the laws of thermal emission [276, 415, 416, 527, 589]. A deeper understanding of these laws is invoked in my work [456, 459]. When Penzias and Wilson reported a $\sim$3 K signal independent of seasonal variation or observational angle [400], they made a crucial link to thermal physics [276, 415, 416, 527, 589]. Their assumption that the signal [400] was thermal in origin [276, 415, 416, 527, 589] is now without dispute [107, 178]. At the same time, I believe that the outstanding signal-to-noise of the COBE FIRAS data [107, 178] implies a proximal source [454, 455, 458, 463–465, 469], not a powerful distant source [136]. I question the validity of the 3 K assignment [136, 400] not simply using this logic, but rather, by revisiting the laws of physics [456, 459]. I also believe that there are problems with the satellite data. I have objected [458, 465] to the manner in which the COBE [107] and WMAP [590] results were processed and presented. These objections have received some support [63, 64, 433]. I will now discuss my specific concerns beginning with thermal radiation and the assignment of a 3 K temperature [400].

2.8.7.1 Thermal Radiation

To evaluate when a temperature can be ascertained with confidence from an emission profile, it is necessary to consider the formulation of the laws of thermal emission [276,415,416,527,589]. Cosmology believes that these laws are universal: they can be applied independent of the nature of the emitting object. That is the basis of Kirchhoff's law of thermal emission, as was advanced in the mid-1800s [276]. Today, the universality of BB radiation constitutes one of the building blocks of astrophysics. Without universality, astronomy no longer knows the temperatures of all the stars [307, 308] and the absolute value of the Penzias and Wilson temperature [400] is not established. I have argued that a problem exists relative to the formulation of Kirchhoff's law [456]. By reviewing carefully what Kirchhoff observed and the basis for universality, one can demonstrate that Penzias and Wilson erred in setting their temperature [400].

2.8.7.2 Kirchhoff's Experiments (Literary Style Based on Kirchhoff's Own Actions and Those of His Contemporaries)

In the 1850s, Kirchhoff made little experimental progress in his efforts to characterize the radiation of various substances. He observed that the spectral profiles of substances were different [555] and the nature of emission with increasing temperature was highly variable. Some substances increased their total emission with temperature while, for others, total emission actually decreased [466]. For Kirchhoff, nothing appeared systematic, with the exception of graphite. Here, a rather smooth spectrum was detected. Moreover, when graphite was subjected to higher temperatures, a tremendous increase in emission was noted, which is given to us today by Stefan's Law [456, 466, 527]. In hindsight, it seems that Kirchhoff used graphite to extract the temperature of objects that lacked a predictable thermal behavior [416, 456, 467, 468]. Because graphite was understood, he used it to make a box wherein he enclosed the other experimental objects. Graphite was a perfect absorber. Any radiation incident upon its surface was absorbed in a manner independent of wavelength and angle of incidence. Therefore, when another object was placed within this box, the emission profile of the object was transformed to that of the graphite. Kirchhoff only had to wait for thermal equilibrium and then, by sampling from a small hole in the box, he could gather the temperature of the entire system, including the object of interest. This experiment critically depended on graphite. Had the walls been made of copper, for instance, a graphitic spectrum would never have been obtained [555]. Kirchhoff invoked graphite for a very specific reason: he required a perfect absorber. In turn, the graphite made up the heart of Kirchhoff's new light transformer [456]. Ironically, Kirchhoff wanted to move away from the graphite walls. He searched for far-reaching conclusions. As a theoretical physicist, he desired to formulate a mathematical law that could summarize his discoveries. His interest was in this law and not in the promotion of the graphite light-transformer [456]. Kirchhoff decided to go to the other extreme. Herein arose the next experiments and the false birth of universality [416, 456, 467, 468]. As graphite could be viewed as a perfect absorber, Kirchhoff wanted to build a second box from a perfect reflector, a polished metal, something like silver. He repeated his experiments with his new box. Each time, he placed an object in the box, waiting for equilibrium and examining the spectrum. But, no matter how long he waited, the spectrum within the box always failed to convert to that associated with graphite [416, 456, 467, 468]. Eventually, Kirchhoff had the idea of placing a small piece of graphite within the perfectly reflecting box [416, 456, 467, 468]. He considered the graphite particle to be the equivalent of a catalyst [416, 456, 467, 468], even though, in 1860, he could have had no idea how catalysis really worked. *Planck* would later affirm Kirchhoff's belief that simple catalysis was involved (see Sect. 52 in [416]). When Kirchhoff added the small piece of graphite to the perfectly reflecting box, the spectrum immediately converted to the thermal spectrum of graphite, and the concept of universality was born [416]. Kirchhoff argued that he could mathematically remove the graphite particle from the second box, as catalysts only affect the time required to reach equilibrium, but are not central participants. This is how Kirchhoff's law was formulated. BB radiation became universal and completely independent of the nature of the walls. Only enclosure was required. As long as enclosure was maintained, the nature of the enclosing object was irrelevant [416, 456, 467, 468].

2.8.7.3 Universality

Today, few would argue that the graphite particle was simply a catalyst. I believe that when Kirchhoff placed his graphite particle in the second box, it was as if he had lined the entire box with graphite [456]. After all, the photons within the box were moving near the speed of light and even the small piece of graphite was perfectly absorbing. Surely, it would not take long for all of the photons released by the object to strike the graphite particle. The object in the box was not perfectly absorbing. This was a characteristic exhibited solely by graphite and soot. As a result, the object was unable to reabsorb many of the photons re-emitted by graphite. Consequently, even a small piece of graphite could dominate the entire problem. As such, there can be no justification for universality on an experimental basis, as Kirchhoff, by placing the graphite particle within the perfect reflector, had essentially returned to the first box made of graphite [456]. Graphite was strictly required. Thus, experimental thermal radiation, outside the confines of the graphite box, becomes dependent on the physical nature of the emitting object [456]. Along the same line, Einstein's derivation of the Planckian equation [149, 150], based on his coefficients, depends on the presence of a Wien's field [589]. This type of field can be found only in a graphite box. As a result, far from affirming universality, Einstein's deviation actually refutes it, as I have highlighted [459]. There is no universality and if astrophysics is facing difficulties today, it is directly because of this erroneous concept. The simplest proof of a lack of universality can be found through examination of thermal emissivity tables [555]. Not a single natural BB has been discovered in the laboratory, and even many forms of graphite fail to exhibit the desired behavior over large ranges of frequencies. It is improper for astrophysics to set a temperature to the Universe for the simple reason that enclosure with a graphite box does not exist. In fact, irrespective of the reformulation of Kirchhoff's law [456, 467, 468], astrophysics still failed to meet the conditions demanded by both Kirchhoff and Planck, by taking temperatures outside the confines of a known enclosure. If this rule had been followed, science would never have assigned the microwave background to the Universe, as the validity of the 3 K temperature would have immediately been brought into question. Kirchhoff's law, even in its erroneous formulation [276, 416], required enclosure [456]. Now, even more is required [456]. The enclosure cannot be simply perfectly reflecting or adiabatic, as we all initially believed [306, 416, 455]. Rather, it must be a perfect absorber as best represented by graphite or soot [456]. The requirements for thermal equilibrium and enclosure can never be met when setting a temperature to the entire Universe. In addition, it is important to recognize that a spectrum may well appear to be thermal, and, indeed, it could be, but that does not make the resulting temperature valid. It is for this reason that enclosure is required. An object that sustains convection currents cannot be in thermal equilibrium [418]. Consequently, the proper elucidation of temperatures by thermal methods requires enclosure. If this requirement is not met, an alternative means of obtaining the temperature must be sought. It is improper, for instance, to take a Wien's displacement temperature when dealing with the liquid state because of convection. I have argued that the temperature of the Universe is derived from an oceanic signal [454, 455, 458, 463–465, 469] as first detected on Earth by Penzias and Wilson [400]. It is likely that the responsible physical species is the hydrogen bond between water molecules [463]. In a sense, one can think of a system of oscillators wherein the water molecules are the weights, and the springs are provided by this bond. As the hydrogen bond is so weak, this oscillator system generates a spectrum in the microwave region and not in the infrared. As for the lack of seasonal variation, it occurs because the energy levels in this oscillator system are completely occupied at Earthly temperatures. As a direct result, the temperature reported is not real, only apparent. It is a feature of an oscillatory system that constitutes a fleeting lattice in the liquid state.

In this respect, astrophysics has been misled because it failed to recognize the importance of the physical causes of thermal radiation as related to both solids and liquids.

2.8.7.4 The Physical Nature of Thermal Radiation

As the physical causes of radiation are being ignored, cosmology assumes that a thermal photon can be produced in the absence of condensed matter. This is because Kirchhoff advanced an idea [276], later adopted by Planck [416, 419] and all of physics [306], which was never re-evaluated, nor corrected. Planck writes (see Sect. 52 in [416]:

> It is therefore possible to change a perfectly arbitrary radiation, which exist at the start in the evacuated cavity with perfectly reflecting walls under consideration, into BB radiation by the introduction of a minute particle of carbon.

Planck then minimizes the importance of the particle of carbon, stating that it was simply a catalyst (see Sect. 52 in [416]). As a direct result of this idea, Planck's BB formula [415] has remained unlinked to physical reality to this day [456]. As I have mentioned, this is a significant oversight in modern physics [457]. One cannot achieve thermal equilibrium in the absence of matter. This, however, is the scenario which is invoked when the importance of graphite is minimized (see Sect. 52 in [416]). The second law of thermodynamics was developed as a result of studying matter, not simply isolated radiation. It is the matter that seeks thermal equilibrium. It does so by emitting radiation, but the radiation has no means of changing its nature in the absence of any further contact with matter. Radiation in the absence of matter cannot be subject to the second law of thermodynamics as it cannot reach an altered state. As such, when Kirchhoff and Planck minimize the effect of graphite, they are removing the matter that was vital to the achievement of the equilibrium they sought [416]. Radiation, in the absence of matter, knows no preferred state. When photons are produced, they can no longer monitor the temperature of the source. Yet cosmology is currently requiring all the photons in the Universe to do just that, irrespective of a known physical mechanism. This being said, Planck's equation must be linked to a direct physical cause [417]. Physics should provide for BB radiation the following: the physical setting, the nature of the energy levels, and the nature of the transition species [456, 457]. I have stated that BB radiation is absolutely dependent on the presence of a physical lattice as found only in condensed matter – either solid or liquid [456]. It is the vibration of atomic nuclei within the confines of the lattice that is responsible for the observed spectrum. Condensed matter physics may choose to advance an alternative explanation. However, now that graphite (or soot) is required, it is the nature of the material producing the emission that is important. Note that the re-examination of Kirchhoff's experiment [456] makes it clear that thermal radiation is dependent, not independent, on the nature of the enclosure. As a group, physicists should now try to understand what is happening within graphite itself to produce a thermal spectrum. Once condensed matter physics answers this question (and it is likely to involve the vibration of nuclei within the confines of a lattice field), then astrophysics will be bound by the same requirements.

If a Planckian spectrum is detected, it must be associated with a condensed state and, therefore, a lattice. Herein lays the primary reason why the Penzias and Wilson signal cannot originate from the remnants of a primordial Universe. Condensed matter, and the existence of a known vibrational lattice, is required. In addition, photons once produced cannot change their frequency in the absence of further contact with matter.

Planck approaches these ideas [417]:

> the production of radiant heat is a consequence of the act of emission, and its destruction is the result of absorption. Both processes, emission and absorption, have their origin only in material particles, atoms or electrons, not at the geometrical bounding surface; although one frequently says, for the sake of brevity that a surface element emits or absorbs

Further (see Sect. 144 in [416]):

> For among all conceivable distributions of energy the normal one, that is, the one peculiar to black radiation, is characterized by the fact that in it the rays of all frequencies have

> the same temperature. But the temperature of a radiation cannot be determined unless it be brought into thermodynamic equilibrium with a system of molecules or oscillators, the temperature of which is known from other sources. For if we did not consider any emitting and absorbing matter there would be no possibility of defining the entropy and temperature of the radiation, and the simple propagation of free radiation would be a reversible process, in which the entropy and temperature of separate pencils would not undergo any change.

Planck requires at the minimum a physical species, an oscillator. This is completely absent in the generation of the microwave background from a primordial source. In the case of Planckian spectra, it is the presence of a lattice and of nuclei which is important as I have noted [456].

This step takes us beyond Planck, in that, though he invoked oscillators, he never realized formally the crucial importance of linking his work to the lattice. In large measure, this was the direct result of the erroneous formulation of universality by Kirchhoff. Planck could not link to a direct physical species, as universality depends on independence from the nature of the walls. The erroneous basis for universality had very far reaching consequences, which I am trying to address today.

2.8.7.5 Thermal Radiation and the Photosphere

I have stated, for the same reasons, that the photosphere must be viewed as condensed matter [460, 461]. The thermal nature of the solar spectrum has long been established. However, astrophysics erred when it set this temperature [307, 308], in the same manner as cosmology has erred relative to setting the temperature of the microwave background [400]. Planck comments relative to the Sun (see Sect. 101 in [416]):

> Now the apparent temperature of the Sun is obviously nothing but the temperature of the solar rays, depending entirely on the nature of the rays, and hence a property of the rays and not a property of the Sun itself. Therefore it would be not only more convenient, but also more correct, to apply this notation directly, instead of speaking of a fictitious temperature of the Sun, which can be made to have a meaning only by the introduction of an assumption that does not hold in reality.

The Sun is not enclosed and therefore cannot meet the requirements for setting a proper temperature. The presence of convection currents on its surface also destroys any chance of maintaining local thermal equilibrium. These are the facts to which Planck is alluding. I have stated that the true temperature of the photosphere is not $\sim$6,000 K, but rather millions of degrees [455, 460, 461]. Given that the photosphere exhibits a thermal spectrum, a physical lattice must be present. The nuclei that make up the Sun cannot simply be arranged randomly, but rather, they must manifest spatial order, at least over short ranges. Short range spatial order is known to exist, for instance, in liquid water. Fleeting lattices are often used to describe such systems. The lattice is continuously forming and breaking down in association with flow. As the photosphere supports convection currents, then it must also have a fleeting lattice. In any event, spatial order is required for the production of a thermal spectrum. The existence of a photospheric lattice is a direct consequence of the physical mechanism required to generate thermal radiation. This line of reasoning leads to a photosphere in the condensed, not gaseous, state. The aforementioned discussion on the Sun has a direct implication relative to the microwave background [400]. The microwave background has a temperature extracted from a spectrum. This temperature, however, has no more meaning relative to the actual source temperature, than the erroneous temperature for the Sun, as Planck so ably reminds us.

2.8.7.6 Analysis of Satellite Data

WMAP: Given the eventual reassignment of the microwave background to the Earth, cosmology will find little to salvage in the WMAP satellite results and anisotropy maps [590]. Already, radiological standards relative to image processing have been applied to WMAP data and the findings were troubling [458]. I have made the point that the analysis of this data falls short of standard practices in imaging science [458]. The following five points summarize my findings relative to the WMAP data analysis. First, it is impossible to remove the Galactic foreground from these images without either ability to control the signal at the source or without a priori knowledge. This data analysis exactly parallels the principles of water suppression in the biological Nuclear Magnetic Resonance (NMR) laboratory [458]. In biological proton NMR, biophysicists are faced with the problem of removing the signal from a water line in aqueous samples to monitor weak signals from the molecules of interest. While this experimental setting may appear far removed from cosmology on the surface, it holds a beautiful example for science. Water suppression in biological NMR helps scientists to understand how to deal with weak signals in the presence of powerful dynamic range problems and overlapping signal contamination [458]. This is because the biological sample often contains water ($\sim$110 M in protons) and molecules of interest with proton concentrations in the mM range. The powerful water line in aqueous proton NMR can completely overshadow the weak signals of interest that resonate near this frequency precisely because it is 1,000 times more intense. Water suppression requirements are not unlike the problem of a strong Galactic signal at the same frequency of cosmological signals lying behind the Galaxy. In the case of WMAP, the proper elimination of strong overlapping foreground signals requires detailed a priori knowledge about the nature of our Galaxy and its exact emission profile in isolation of all other effects, as I have stated [458]. Such a priori knowledge will remain forever inaccessible. It is noteworthy that biophysicists have much more control and understanding of their experimental systems (i.e., the small NMR tubes filled with aqueous biological samples). Even so, biophysicists remain completely incapable of eliminating the water line using a priori knowledge alone. Rather, they achieve water suppression by manipulating the sample (or the spins) at the source. Sample manipulation is something that cannot be achieved with the galaxy. Biological water suppression has a direct implication in signal acquisition and processing across all disciplines. Astrophysics may not argue that it can achieve with the galaxy what biophysicists have failed to achieve with a 5 ml aqueous sample a man can hold in his hand. Relative to WMAP, the act of adding or subtracting component data sets in the region of the Galactic plane to eliminate the foreground signal is without any experimental justification other than the achievement of a null point [458]. Second, in removing the foreground and in obtaining "cleaned maps", the WMAP team is manipulating a large signal from the foreground, while trying to preserve very weak signals of interest even beyond the Galactic plane. This is one of the greatest problems in image processing. I have noted that it is impossible to remove powerful overlapping signals mathematically through "cleaning", without introducing error into the resulting maps [458]. In this regard, the WMAP team must not generate anisotropy through signal manipulation. Nonetheless, they have no means to ensure that this, in fact, has not occurred. Third, the WMAP team is also not justified in the dramatic yearly changes in the coefficients used in map reconstruction [458]. It is proper to correct for changes in receiver performance. A map could not be generated without accounting for these effects. However, the WMAP team alters the coefficients for map reconstruction on a year-by-year (or map-by-map) basis in a manner that cannot be solely attributed to changes in receiver sensitivity [458]. If cosmological meaning is to be excised from the anisotropy maps, the choice of coefficients used in map generation must not vary from one data collection period to another [458]. Fourth, on a pixel-by-pixel basis, the WMAP data set lacks the stability required for cosmological signals [458]. The multipole analysis does appear remarkably stable. However, this does not constitute a test of cosmological stability, as the multipole analysis results from signal averaging the contribution of every single pixel in the underlying maps. In data processing, it is known that signal averaging can mask signal instability. To ascertain signal stability, the behavior of the individual pixels in the anisotropy maps should be examined. These tests should compare the maps generated for each data acquisition period, while having no recourse to maps that average many years of observation. Using this simple analysis of the anisotropy maps on a pixel-by-pixel basis, it should become apparent that the WMAP images are too unstable to link this data to cosmological causes. Note, in this regard, that instability at the level of just a few pixels per year is significant for the cosmological community. This is because it is claimed that the signals must represent effects on time scales in the billions of years. If 100 pixels vary over a few years, than thousands of pixels must be inferred to vary over longer time periods. Over 100 years, a period of time irrelevant to cosmology, virtually every pixel in the anisotropy map could vary. Consequently, the anisotropy maps will never be able to display the stability required of an image with true cosmological value [458]. Fifth, the WMAP anisotropy maps are generated using a linear combination of the underlying frequency maps [590]. Yet, there are literally an infinite number of linear (or even nonlinear) combinations that could be used to generate anisotropy maps [458, 590]. As a result, there are an infinite number of possible maps. Astrophysics may choose to apply mathematical methods to extract a most probable solution. Nonetheless, in the end, there are no means of extracting the "definitive map". That is because other mathematical methods could easily be advanced rendering another map more probable. As a definitive map does not exist, then no map exists. We are left with an exercise in mathematical methods with no means of extracting the "true" anisotropy map of the sky. For instance, the map generated by [552] does not agree with the maps released by the WMAP team on a pixel by pixel basis [590]. Each group may therefore claim a solution, but there are no means of confirming the validity of the claim. This is a major difference relative to radiological imaging where the correct answer can be ascertained by different groups using other imaging modalities or even autopsy [460].

In summary, measurements of the microwave background have always been complicated [388]. Nonetheless, it is noteworthy that the assignment of a real temperature to this background constitutes of overextension of the laws of thermal emission [276, 415, 416, 456, 459, 527, 589]. There are three fundamental conclusions: (1) there is no universality [456, 459], (2) thermal emission must be linked to a discrete physical cause [457], and (3) the assignment of temperatures based on thermal emission requires both enclosure with a perfect absorber and local thermal equilibrium that have not been met. There remain numerous problems in generating anisotropy maps most notably: (1) the inability to properly remove the foreground, (2) a lack of signal stability on a cosmological timescale as reflected both by variations in coefficients used in map generation and pixel-by-pixel analysis, and (3) the absence of a unique solution. In the end, the assignment of the microwave background to the Earth should prevail. Occam's razor favors simple physical mechanisms for photon production (i.e., the hydrogen bond within water) and reasonable proximal sources (i.e., the oceans) over theoretical mechanisms (i.e., the production of thermal radiation without condensed matter) and strong distant sources (i.e., the Big Bang).

How could you explain in a common framework, alternative to the standard one, the consistency of the CMB absolute temperature with the level of dipole signature introduced by the observer motion and of its modulation introduced by the Earth motion around the Sun? Up to now, all CMB experiments at different frequencies and with different instrumentations give results consistent with COBE and WMAP. How do you explain this fact?

I have mentioned that the same dipole and its yearly modulation can be explained if an alternative common framework is used where two conditions are present: (1) the Earth is generating the monopole and is moving, along with the Sun, through the local group, and (2) another weak microwave field exists [63, 64, 433, 464, 465]. Presence of this weak field can be deduced both from COBE data, as will be illustrated later, and by the presence of a dipole in Relikt [279] and WMAP [590] data. The Sun, the Earth, COBE [107], and WMAP [590] are moving through this weak field. I have proposed that the weak field is not characterized by a single monopole. Rather, it is a noisy microwave background best characterized by the sum of perhaps many monopoles. It does not posses the spectral signature of a Planckian field. Nonetheless, motion through this field can generate a dipole as the magnitude of the latter is dependent only on the velocity and direction of the species moving through the field. Using these ideas, all the results obtained on Earth, above ground, and in space can be reconciled. In any event, the consistency of the CMB absolute temperature is not as firmly established as commonly believed.

2.8.7.7 Observations on Earth and in Low-Earth Orbit

First, the Earth cannot be modeled as a BB emitter near 300 K as the COBE team has assumed [107]. As the Earth is covered by water, its behavior in the microwave region is not the expected 300 K BB as I have stated [465]. This simply furthers my belief that the microwave background has not been properly evaluated in light of our knowledge of the Earth and of thermal radiation. Second, the atmosphere holds tremendous microwave power [121]. Microwave interference from the atmosphere also increases substantially with frequency [121]. This crucial observation has been relatively ignored in analyzing the microwave background. Any interference from the atmosphere was simply subtracted out. Yet, elastic scattering exists in the atmosphere. This scattering is vital in the generation of an isotropic monopole signal from the Earth, given that our planet is anisotropic [469]. Third, as the Earth is moving in the local group, if it generates a field, then a dipole must be present. The Earth, as the source of the microwave background following atmospheric scattering, is generally isotropic [469]. Consequently, it might be expected that the spectrum it emits would be relatively devoid of quadrupolar and octapolar contributions. This also helps to explain why the quadrapoles and octapole values observed in the microwave background are so low. But, a powerful dipole is present. This dipole, when viewed from the Earth, is the first derivative of the monopole generated by the oceans. Consider my assumption that the Earth is emitting the monopole and a resulting dipole. COBE, whose shields are most effective against unscatterred IR radiation from the Earth, senses the monopole emitted from the oceans. At the same time, the Earth is moving through the weak microwave field discussed earlier [63, 64, 433, 464, 465]. COBE reports anisotropy due to motion, but it cannot easily distinguish the effects. Only a small error attributed to a systematic problem is detected. COBE first reports a systematic error in obtaining the CMB absolute temperature from the dipole [178]. This was the first indication that a second field was present [63, 64, 433, 464, 465]. Indeed, although the COBE team eventually tried to account for the systematic errors they discovered [178, 345], questions remain in my opinion [465]. Given the signal-to-noise of the monopole as measured by COBE [107, 178], it is unusual that any systematic error should be found in this data [465]. The introduction of a second weak field reconciles this problem [63, 64, 433, 464, 465]. The systematic error obtained in setting the absolute CMB temperature from the dipole originates from the interaction of the monopole field, produced by the Earth, and the weak field through which the Earth moves [464, 465]. My idea that two field are interacting, namely the monopole field generated by the Earth and the weak field, accounts for the systematic error. This concept has received some support from Rabounski and Borissova [63, 64]. Rabounski has demonstrated that the presence of two fields can indeed account for a lower than expected monopole temperature when using dipole data [433]. In addition, Borissova and Rabounski [63, 64], have demonstrated, using general relativity, that the magnitude of the anisotropy does not change with altitude, unlike the magnitude of the monopole that drops rapidly as one moves away from the Earth. Nonetheless, when viewed from COBE [107], the dipole appears to fit the first derivative of a Planckian function, because its primary contribution is derived from the monopole generated by the Earth. The consistency of the CMB absolute temperature is not as firmly established as commonly believed.

2.8.7.8 General Remarks on Ground-based Measurements

Consider the steady state nature of the microwave background [463, 469]. If this background did originate from the Universe, it would strike the atmosphere of the Earth from all possible directions in a continuous manner. As steady state exists along with the presence of elastic processes, there can be no effective scattering of the signal, as there is a continuous influx of photons from all possible directions. Moreover, there can be no net absorption [463, 469]. This is because, any photon, which is initially absorbed, must be re-emitted, eventually. Any scattering associated with re-emission becomes irrelevant, because the source is not a point source. Thus, given sufficient time, averaging will cause both scattering and absorption–emission processes to become unimportant, given a cosmic origin. There can be no means of signal attenuation at elevated frequency [463, 469]. It could be argued that nonelastic processes are occurring in the atmosphere, with net energy absorption, and perhaps with the creation of signals at other frequencies. However, it should be noted that the photons involved carry microwave energies, not ultra-violet energies. Therefore, it is extremely unlikely that they possess sufficient energies to result in nonelastic interactions in the atmosphere. Rather, it is reasonable to believe that when a microwave photon of the proper energy is absorbed (for instance by a water molecule), the molecule makes a reversible transition to a higher vibrational–rotational state. When relaxation finally occurs, the photon is re-emitted, with scattering as the only indication that absorption has occurred. Yet, even the effect of this scattering becomes unimportant as the initial source is not localized. Since the microwave background is a steady state signal approaching, according to cosmology, from all directions in space, then there should have been no signal attenuation [463, 469].

2.8.7.9 Observations at the Second Lagrange Point

At the second Lagrange point (L2), the microwave field produced by the Earth is much too weak to be observed [63, 64, 433, 464, 465]. Nonetheless, as a weak microwave field exists at L2 [63, 64, 433, 464, 465], the WMAP satellite moves through this field in the same manner as COBE experienced on Earth. Consequently, as highlighted by Borissova and Rabounski using general relativity [63, 64], the same dipole (in magnitude and direction) will be sensed, including its yearly modulation. The differential instruments on WMAP are incapable of giving absolute temperature. Everything appears to fit, because the satellite is unable to directly assess the presence of the monopole [590]. The anisotropy maps generated at L2 resemble those acquired by COBE because these are indeed largely dominated by the Galactic foreground of the Milky Way and other galaxies, or point sources, in the Universe. Nothing in this data indicates a cosmological significance as I have mentioned in the previous section.

To Summarize

This discussion constitutes a reasonably sound basis for the belief that the monopole 3 K signal originates from the Earth [454, 455, 458, 463–465, 469]. The following points hold: (1) the monopole 3 K signal is powerful as monitored from COBE [107], (2) the systematic errors in determining the monopole temperature from dipole measurements [178] are easily addressed through the introduction of a second much weaker field [63, 64, 433, 464, 465], (3) the lack of quadrupolar and octapolar power in COBE [107] and WMAP data [590] point to a spherical source such as the Earth [107], (4) atmospheric scattering can convert anisotropic oceanic emissions into an isotropic signal [469], (5) the steady state nature of the microwave background also makes attenuation improbable if the background did arise from the Universe [463, 469], (6) the anisotropy maps generated by COBE and WMAP are expected to generally agree even if the monopole arises from the Earth, and (7) astrophysics has failed to fully consider that the Earth and its oceans cannot be treated as a 300 K BB emitters [465].

Certainly, the high precision derivation of the CMB temperature fluctuation in each sky pixel and its full cosmological exploitation calls for an experiment more precise than WMAP (like *Planck* for example) and improved data analysis techniques to subtract foregrounds and correct for systematic effects, objects of dedicated studies. On the other hand, the main cosmological information of WMAP relies only on the derived CMB APS. Even considering a single frequency map at about 60 or 90 GHz (where foregrounds are minimum), its shape can be obtained to first order by simply excluding the Galactic plane and it is in agreement with that expected by the Concordance Model (except for some anomalies) and, in particular, with an absolute CMB temperature of about 2.7 K. How can you explain this fact in your alternative framework?

2.8.7.10 The Excellent Agreement with the Concordance Model

Relative to the extraordinary fits obtained between the modern Concordance Model and the CMB APS, the result must be discounted. In large measure, I reject the premise of the question. I have already mentioned that the underlying anisotropy maps have serious problems, including the lack of stability required of data with true cosmological value. In any event, if the signal is reassigned, then the models will also be abandoned and any perceived agreement will be quickly forgotten. The issue at hand must remain focused on the experimental data and on the source of the microwave background itself. Nonetheless, for the time being, I concede that my framework cannot explain this apparent agreement.

2.8.7.11 The Value of More Satellite Measurements

The satellite missions continue to be important, but not for the expected reasons. Cosmology believes that the *Planck* mission, for instance, will provide data of increased resolution and quality relative to anisotropy [420]. Through its Low and High Frequency Instruments (LFI at 30, 44, 70 GHz and HFI at 100, 143, 217, 353, 545, 857 GHz), the *Planck* satellite [420] will also cover a much greater number of frequency bands than WMAP (23, 33, 41, 61 and 94 GHz) [590]. However, relative to anisotropy, *Planck* will be subject to the same shortcomings that plagued the WMAP mission [590]: inability to properly remove the foreground, lack of stability, and failure to offer a unique solution [458]. Rather than the five frequency bands found with WMAP, any analysis of the *Planck* LFI and HFI will have to contend with the linear combination of more frequency bands. Yet, cosmology cannot provide a unique solution for the combination of these frequency bands. The problems cannot be solved with improved data analysis techniques, as the question implies. Fundamentally, it is a problem of insufficient information. WMAP may have provided only five frequency bands, but *Planck*'s increased number of bands, far from providing more information, simply provide an even greater range of possible combinations to ascertain a final anisotropy map. More information will lead to even more variability, not to a definitive map. *Planck* is also promoted as being able to provide greater insight into the SZE. The analysis of the SZE requires excellent radiometer stability and signal-to-noise. However, the *Planck* LFI will not deliver the expected performance as I have highlighted [464]. The *Planck* LFI, while able to operate in absolute mode, are designed to function optimally as differential spectrometers, wherein the sky signal is constantly compared to two 4 K reference signals. It is critical for the *Planck* satellite that the sky signal has the expected amplitude. If not, the radiometer performance will become suboptimal [464]. Because the monopole is being produced by the Earth, the *Planck* LFI will be confronted with a significantly different environment than that for which it was designed. The satellite is expecting an effective sky temperature of $\sim$3 K, but will likely experience a much lower effective temperature. As a result, the knee frequencies of its low frequency instruments will rise, and the resulting frequency maps will be contaminated by significant stripes [464]. Though these can be removed to some extent with processing, the *Planck* LFI are likely to have signal-to-noise that remains inferior, not superior, to that obtained by WMAP. It is therefore unlikely that *Planck* will be able to make any contributions relative to the SZE. Nonetheless, as the LFI on *Planck* can operate in absolute mode relative to the 4 K references, *Planck* should be able to finally ascertain whether the Penzias and Wilson monopole signal [400] is present, or not, at L2 [63, 64, 433, 464, 465]. This will be the key finding from the *Planck* mission. In this respect, *Planck* will have tremendous value. With *Planck*, we will understand that there is no universality [456, 459]. The need to link the Planckian equation to a direct physical cause will also become evident. Physics will turn its attention slowly from the stars to the lattice and the liquid [456]. In this sense, *Planck* will become one of the most important instruments ever launched into space. It will be as if Max Planck himself [462] will return to tell us that something was actually missing relative to our understanding of BB radiation. The *Planck* satellite data will finally urge us to link Max Planck's equation to a physical cause. At the same time, it will remain interesting that such drastic methods were required to finally bring the needed corrections into the laws of thermal emission [276, 415, 416, 456, 459, 467, 468, 527, 589]. In this regard, a condensed photosphere will be the first product of restructuring these laws [460,461]. All these ideas depend on the 3 K monopole not being present at L2. This is something which, relative to currently planned experiments, only the *Planck* satellite can reveal.

The CMB absolute temperature is also indirectly probed through the analysis of various molecular lines in interstellar clouds. Also, the amplitude of SZE towards many clusters of galaxies is consistent with gas temperatures and densities derived from X-ray and optical observations assuming a CMB temperature of about 2.7 K. Assuming a completely different CMB absolute temperature would tremendously affect these arguments. How can you explain these evidences in your alternative framework?

Molecular Methods: Temperature analysis using molecular methods report local effects. It is not appropriate to extrapolate these to the entire Universe. Nonetheless, attempts have been made to obtain temperature readings, using molecular lines, from interstellar clouds [305,324,351,388,471,477,515,526]. Ideally, these methods depend on the presence of Local Thermal Equilibrium (LTE), wherein the species of interest interacts with an enclosed Planckian radiation, as detailed in Einstein's classic work [149], and discussed in astrophysical texts [471]. However, the assumption that LTE is maintained within interstellar clouds is uncertain at best. Astrophysics understands that, in the absence of LTE, the temperature obtained from such methods is a radiation temperature that may have little to do with the real temperature of the system [471]. Departures from the Planckian lineshape occur in the absence of LTE. It is also known that if collisions dominate the problem, an excitation temperature is actually obtained, which will be the mean between a kinetic and a radiation temperature [471]. In low density situations, the excitation temperature does approach the value of the radiation temperature. However, this requires knowledge of densities that are difficult to ascertain, although this has been reported [526]. Finally, I have also urged caution relative to universality [459] in direct connection to Einstein's analysis [149] from which all these methods are derived. Consequently, it will always be true that, unlike direct measurements from satellites, molecular methods generate highly inferred results. I have proposed that microwave power exists throughout space [465], but that it is not characterized by a unique temperature. A weak microwave field does exist at L2 [63, 64, 433, 464, 465] and this field may extend well beyond the galaxy. As hypothesized, the weak field constitutes a noisy spectrum made up of a very large number of apparent temperatures. As molecular species have specific vibrational–rotational energy states, it is well possible that a molecule in deep space reports the presence of a particular temperature close to 3 K. With cyanogen, in the ideal case, it is only the intensity of a few separate frequencies that have been sampled. In reality, a complete sampling might indicate that a much larger group of underlying apparent temperatures exist. Photons, after all, report only energies, not temperatures. It is the scientist who must interpret the possible temperatures based on spectra. In the presence of appropriate photons, a given temperature may well be extracted, but true temperature requires a complete understanding of the entire frequency spectrum in the region of interest. This is something that molecular methods cannot achieve as the presence of a Planckian spectrum cannot be deduced from just a few lines. At the same time, local sources within the interstellar clouds may produce the apparent 2.7 K temperature reported, for instance, by cyanogen [351]. Clearly, interstellar clouds do not contain only cyanogen. If the hydrogen bond between water molecules is indeed responsible for the 2.7 K signature obtained on Earth [463, 469], it is not unreasonable that water is present, even as ice crystals, in interstellar clouds. Microwave emissions from such particles might produce a local source for the 2.7 K field observed. Cyanogen in this case would be reporting not only an apparent temperature of 2.7 K, but rather, the presence of water aggregates within the interstellar cloud. It remains the case that molecular methods are fascinating, but inconclusive. Thus, the 2.7 K microwave background may well be reassigned to the Earth, even if molecular methods report an apparent 2.7 K temperature in interstellar clouds. In this case, it will be impossible to dissect whether experimental and theoretical problems were responsible for the perceived 2.7 K temperature in interstellar clouds or whether, indeed, water was also present. Given this scenario, the existence of a 2.7 K signal within interstellar clouds may always remain a mystery.

The Sunyaev–Zel'dovich effect: The evidence relative to the SZE effect requires tremendous signal-to-noise and resolution. Yet, when scientists examined WMAP data for this effect [320], it was remarkably absent from many of the possible regions that were candidates for the SZE. Consequently, arguments based solely on the presence of an SZE will continue to fall short, given the current state of astrophysical data. As explained earlier, *Planck* will not clarify this issue, because in the absence of a monopole signal at L2, its radiometers will not achieve the expected signal-to-noise and stability.

In summary, if the 3 K signal is absent at L2, there could be consequences for astrophysics that extend well beyond cosmology [460, 461]. Physics, and much of science, would have to take a necessary pause. Serious questions could be raised relative to the extent that the Universe can be monitored in its wonders and vastness. It would not be improper to humbly concede that there are subjects beyond human understanding. On this anniversary of Galileo, let us reflect not only on the marvels of science and discovery, but also on the lack of knowledge that still plagues mankind as the advancements spawned by the renaissance.

Thank you very much Pierre-Marie. We do not agree with many of your considerations. Without entering into specific discussions about the raised issues, and in particular on those regarding the Kirchoff's law that would imply to be engaged in endless discussions on fundamental physics, that has by now been derived from first principles and likely tested in thousands of experiments, we briefly comment on some aspects of the proposed view most critical for cosmology.

First, concerning WMAP data releases, all-sky maps are available separately for each year of observation. Each interested reader could verify the level of their *consistency, within instrumental sensitivity, with public available tools. Regarding the SZE towards clusters of galaxies, decreasing the CMB absolute temperature by three orders of magnitude, as proposed earlier, would imply the necessity of increasing the product of the cluster size with the density and temperature of electrons by a corresponding factor to keep the amplitude of the SZE at the same level. This is not compatible with the present observational view derived from data taken at various frequencies (see, e.g., the contribution by Isabella Gioia in this volume and references therein). We have also to remind that such a change in the CMB absolute temperature would dramatically affect BBN predictions, discussed in previous sections, and this would require a different theory to explain the observed abundances of light elements.*

Finally, it is remarkable how, other than for its extraordinary sensitivity to CMB anisotropies, the Planck *mission will represent also a further fundamental probe for the basic interpretation of the CMB in the cosmological context through an independent cross-check of CMB absolute temperature far from the Earth.* Planck *will be widely discussed by Charles Lawrence in Chap. 5.*

The CMB is the only background emission of genuinely diffuse origin. On the other hand, background emissions are observed at all wavelengths as results of the integrated contributions by astronomical objects. Their study provide interesting information on the evolution of the objects responsible for the considered background.

A particularly intriguing background emission is observed in the X-rays. The study of this emission has a direct impact in the cosmological context for several reasons. The next interview with Günther Hasinger will introduce us to the advantages of X-ray astronomy for observational cosmology.

2.9 Cosmological Information from X-Ray Astronomy

Dear Günther *(Hasinger)*, **X-ray astronomy opened a new window also for cosmological studies. Today the X-ray window is very important for the present day cosmology. Can you explain why?**

In my opinion, there are two areas where X-ray astronomy is making important contributions to our understanding of cosmology and cosmological evolution. The first area is the study of baryons and DM distributed in the LSS in the Universe, in particular in clusters and groups of galaxies, and their relevance to the determination of cosmological parameters and ultimately the nature of DE. The second area is the study of the Cosmic X-ray background and in particular the role of BHs throughout the history of the Universe.

2.9.1 Evolution of LSS and Nucleosynthesis

About 96% of the energy density of the Universe exists in the form of DM and DE, which govern the structure and evolution of the Universe on the largest possible scales. Clusters of galaxies are the largest collapsed objects in the Universe. Their formation and evolution is dominated by gravity, that is, DM, while their large scale distribution and number density depends on the geometry of the Universe, which in turn is affected by DE. Clusters are filled with hot baryonic gas, which is enriched with elements by star formation and stellar explosions, and is preferably detected by its high energy radiation. X-ray observations of clusters provide information on the DM and DE content of the Universe, on the amplitude of primordial density fluctuations, on the complex physics governing the formation and evolution of structures in the Universe, and on the history of metal synthesis. Nearby clusters of galaxies have been studied at great detail with existing X-ray satellites. These observations have shown a fascinating rich and complex spectrum of phenomena, for example, substructures of clusters and shocks induced by mergers and also the effects of the feedback of the central BHs, which require to include more complicated physics and to perform elaborate hydrodynamical simulations.

Nevertheless, clusters of galaxies have successfully been used to constrain cosmological parameters. Sizeable samples of clusters have been selected from the ROentgen SATellite (ROSAT) All-Sky Survey and have been utilized to determine cosmological parameter, in particular σ_8, the normalization of the primordial power spectrum[4], and Ω_m, the matter density of the Universe.

Figure 2.25 shows a recent compilation of constraints on the energy density of the Universe [51], based on measurements of harmonic fluctuations in the microwave background (WMAP), geometry measurements using SN Ia explosions as standard candles, and the determination of the power spectrum of large scale structure utilizing the ROSAT cluster sample. The fact that several independent sets of measurements gave consistent constraints of the energy density of matter (Ω_m) and of the mysterious DE (Ω_Λ) gave a lot of credibility to the existence of DE. To highlight the important role of X-ray astronomy in this context, it is important to note that the measurements of σ_8 originally disagreed significantly between different sets of measurements. While X-ray astronomy results, both based on the large scale structure [503] and also on the baryon content of clusters as a function of redshift [6] always found a relatively small value of σ_8 (around 0.7–0.8), observations based on gravitational lensing and on the microwave background originally found a rather higher value ($\sigma_8 = 1 \div 1.2$). However, after the analysis of the WMAP 3-year data, the CMB data are consistent with the original X-ray determination and also the lensing results start to get compatible results.

These analyses are based on rather small samples of relatively local clusters selected in X-rays ($\sim$1,000). With the upcoming eROSITA survey, aiming at the detection of $\sim 10^5$ clusters, these studies can be done with unprecedented statistical and systematic quality. This survey will be able to also constrain the equation of state of DE and its temporal evolution.

While the hot gas in clusters and groups of galaxies contains a large fractions of the total baryons in the local Universe, little is known about the fate of almost 50%

[4] σ_8 describes the amount of structure in the Universe and is represented as the rms matter fluctuations in spheres of $8\,h^{-1}$ Mpc.

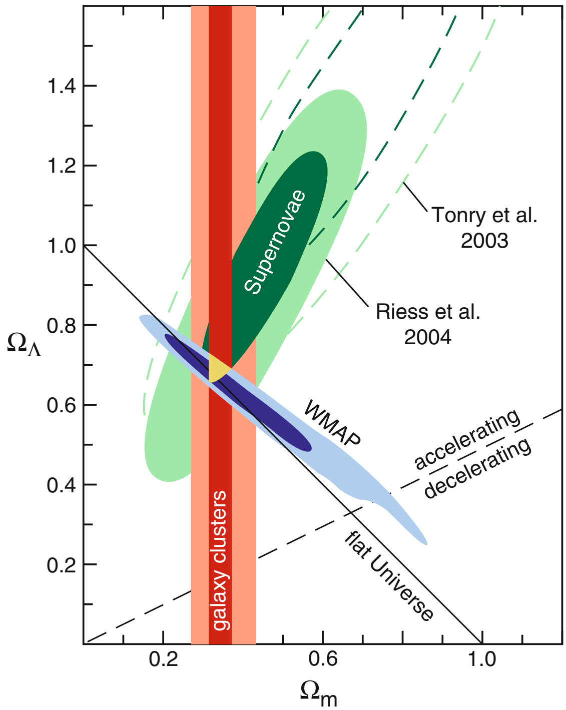

Fig. 2.25 Constraints on the cosmological parameters Ω_m and Ω_Λ from three different cosmological tests: SN type Ia explosions (*green*), analysis of fluctuations in the cosmic microwave background (*blue*), and abundance as well as spatial clustering analysis of clusters of galaxies detected in X-rays with ROSAT. From [51]

of the baryons, which are believed to reside in warm/hot filamentary structures, which are best observable with X-ray absorption spectroscopy. To study the genesis of groups and clusters of galaxies and the cosmic web at high redshifts ($z < 2$) and the evolution of the physical state and chemical abundances of the IGM, an X-ray telescope combining a very large collecting power with excellent energy resolution and good spatial resolution is necessary. This will be the aim of the XEUS mission[5].

2.9.2 *Coeval Evolution of Galaxies and Their Supermassive BHs*

The first stars and galaxies formed where gravity overpowered the pressure of the ambient baryons. Ultimately, gravity dominated and the first stellar mass BHs were formed, very likely in gamma-ray burst explosions. Supermassive BHs can grow in cataclysmic feeding events. The highest redshift accreting BHs known are around $z = 6.5$. The WMAP studies of the microwave background show that the first light must have ionized the Universe already as early as $z = 10 \div 20$. The fact

[5] In the meantime, the ESA/JAXA XEUS project and the NASA Con-X project have been merged into the International X-ray Observatory (IXO) project.

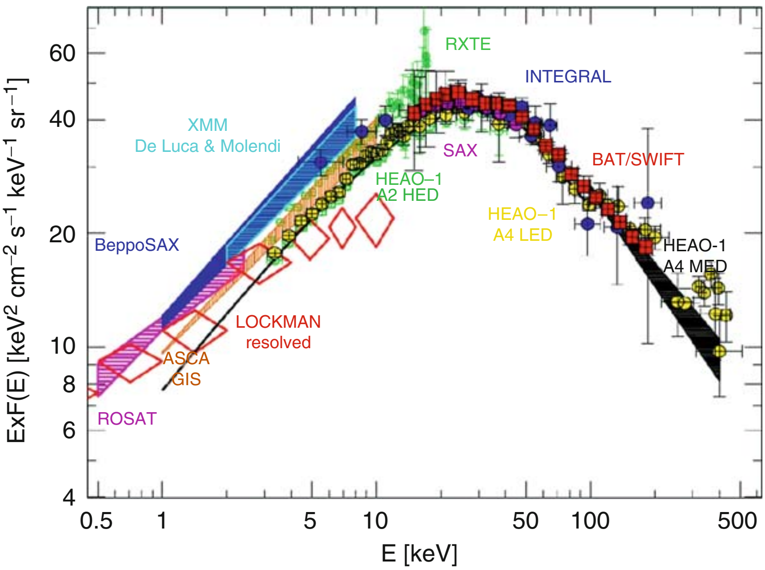

Fig. 2.26 Compilation of measurements of the X-ray background spectrum from many different observatories over the last 25 years. Modern measurements from collimated instruments aboard BeppoSAX, Integral/IBIS, and SWIFT/BAT confirm the shape of the background originally obtained from HEAO-1, but indicate a ∼10% higher background background intensity. Most recent determinations from imaging experiments below 10 keV indicate an even higher normalization; however, systematic effects may still exist there. The *diamonds* indicate the flux which has been resolved into discrete sources by the deepest XMM-*Newton* observation. Figure courtesy of M. Brusa and M. Ajello

that practically all galaxy bulges in the local Universe contain supermassive BHs, with a tight relation between BH mass and the stellar velocity dispersion, indicates a co-existence and co-evolution of stars and central BHs early in the Universe. Supermassive BHs must thus be an important constituent of the evolving Universe. Indeed, the X-ray background radiation, the first extragalactic background component to be discovered (see Fig. 2.26), can be interpreted as the echo of the building and growth phase of supermassive BHs at the centers of galaxies throughout the history of the Universe. The peak of the X-ray background observed at 30 keV is interpreted in terms of intrinsic absorption of the majority of AGN, which are shrouded by gas and dust clouds most likely from the accreted material. Indeed, folding the luminosity function and cosmic evolution of AGN (e.g., [230]) and an appropriate distribution of absorption column densities, one obtains excellent fits to the X-ray background and many other constraints, like, for example, the observed redshift distribution and number counts [195].

Only recently has the importance of feedback of stellar explosions and accreting BHs into the IGM and ISM, and thus their role for star and galaxy formation been realized. The major BH growth mode in the early Universe is believed to be triggered by major mergers of gas-rich galaxies, which can efficiently drive gas into the center of the merging galaxy and thus facilitates the accretion onto the BHs. Even before the final merger, the two BHs thus gain mass and radiate efficiently. The feedback of the photometric and kinetic energy output of the merging BHs is believed to be sufficient to drive out the gas from the forming galactic bulge and thus suppress further star formation. This mechanism has been made responsible for the tight correlation between BH mass and host galaxy properties in the local Universe and also for the fact that there is an upper mass cutoff in the galaxy population (see, e.g., [135]). The highest redshift quasars can be post-dicted by ab-initio hydrodynamical DM simulations of groups with several massively starforming protogalaxies merging successively in the early Universe [318]. The study of the birth and growth of supermassive BHs at $z \sim 10$ requires an unprecedented combination of large spectral throughput, high angular resolving power, and large field of view in the X-ray regime – namely XEUS – matching those of future optical and radio telescopes.

Could you discuss the complementarity between the cosmological information contained in the X-ray background and that derived from the accurate analysis of the various populations of extragalactic X-ray sources?

The study of the X-ray background and its constituents gives integrated constraints on the average population properties and their cosmological evolution. However, by definition, the objects contributing to the X-ray background are rather faint and typically at substantial redshifts ($z = 1 \div 2$). Therefore, very little detailed information on individual properties of objects can be derived. Also, X-ray surveys are always flux limited and therefore only provide a biased selection of objects. Therefore, X-ray background studies (as any other statistical cosmological study) have to be guided by population synthesis models, where the statistical properties of the underlying population of objects as well as their cosmological evolution have to be included to make predictions about observable parameters and compare these with the real observations. In the X-ray background models, these are in particular the typical spectral properties of the source emission mechanisms, the properties and geometry of the obscuring medium and the luminosity function of AGN and, most importantly, constraints on the cosmological evolution of all these quantities. These have to be derived using X-ray surveys, ideally in different X-ray bands and over a wide range of flux limits and solid angles. The limiting factor in these surveys is most often the tedious optical and NIR follow-up work that requires spectroscopy and precision photometry from large telescopes.

As an example of some of this work, Fig. 2.27 shows the evolution of the space density of ~1,000 AGN selected in the soft X-ray band (0.5–2 keV) in different luminosity intervals, which is one of the most comprehensive recent results after many years of work. This diagram shows in detail the so-called "downsizing effect" (sometimes also called "anti-hierarchical evolution") in AGN. The highest luminosity, and thus presumably most massive objects, for example, powerful quasars, have the maximum of their space density at high redshifts ($z = 2 - 3$), as is well known from studies of QSO in the radio and optical bands. On the other hand, lower luminosity objects appear significantly later in the history of the Universe, with the peak

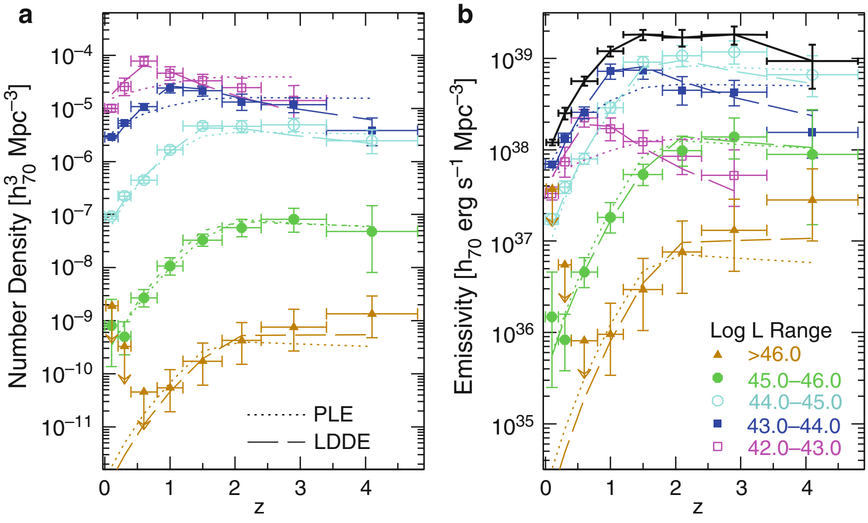

Fig. 2.27 (**a**) The space density of AGN as a function of redshift in different luminosity classes and the sum over all luminosities with log $L_X > 42$. (**b**) The same as (**a**), except that the soft X-ray emissivities are plotted instead of number densities. The uppermost curve (*black*) shows the sum of emissivities in all luminosity classes plotted. From [230]

of the lowest luminosity AGN (typically Seyfert galaxies) appearing at redshifts below unity, and thus much closer to our current epoch.

On the other hand, detailed studies of individual nearby sources are particularly important to gain a deeper understanding into the AGN emission processes, the geometry of the sources and their absorbing medium, and the relation of their properties to the environment, in which they live. As a particularly important example, I would like to discuss the by now famous galaxy NGC6240. This galaxy is one of the most exotic objects in the nearby Universe. It is a member of the class of ULIRGs. Such galaxies emit most of their total energy output in the far-infrared, indicating prodigious star formation, several hundred solar masses per year. NGC6240 is in the middle of a major merger and the two nuclei of the previously separate galaxies are still easily visible in the center of the "train wreck" of this collision. Neither the optical nor the infrared studies of this galaxy originally revealed the presence of an active BH in this galaxy; however, observations in the hard X-ray band (10–100 keV) with the BeppoSAX satellite clearly showed the presence of a heavily obscured, luminous AGN in this galaxy. Interestingly, the spectral energy distribution of NGC6240 from the radio to the hard X-ray band has a very similar shape as that of the overall cosmic background in the infrared, optical, and X-ray region. It contains all the ingredients we need to explain the extragalactic background, although they still need to be mixed in the right proportions.

The real revelation of the importance of NGC6240 for our understanding of the processes likely going on in the early Universe came from high resolution observations with the *Chandra* satellite. If two galaxies merge, and each of them contains a nuclear supermassive BH, could we catch them in a situation where both of the BHs are active? *Chandra* observations by [291] indeed revealed the presence of two active supermassive BHs in the two nuclei of NGC6240. Both of them are heavily obscured and can only be detected at hard X-ray energies. But both X-ray sources can be clearly identified as active BHs, in particular, from the strong fluorescent iron lines in their X-ray spectrum. Most importantly, although they are hardly visible at any wavelength other than hard X-rays, both nuclei have quasar-type X-ray luminosities, indicating a substantial BH growth already well before the final merger of the two BHs.

It is possible that we witness here in the local vicinity a rare case of what is going on regularly in the early Universe. A tight connection between AGN and star formation is expected to exist in merging systems, where both the BHs and the starburst have a common source of fuel driven into the center by the merger. Archibald et al. [9] have interpreted these phenomena as a natural sequence of galaxy formation. In their scenario, each galaxy forms around a seed BH of relatively low mass ($10 \div 1000\,M_\odot$). At first, starlight dominates the total output of the galaxy, because the seed BH still has to grow. The hole does so exponentially by swallowing material at the Eddington rate as fast as it can. After about 500 million years, the BH is massive enough – about a billion solar masses – that infalling material outshines the stars. A quasar is born. After a while, this quasar has eaten all the available fuel and falls asleep until new gas falls into the center, waking it up. Li et al. [318] have shown that Archibald's assumption about continuous Eddington accretion over a long period of time can be justified if the accreted fuel is provided by a sequence of mergers of smaller protogalactic clumps.

Thank you Günther.
We enter now in the nonlinear phase of the evolution of the Universe, resulting in the formation of the first structures. In the next section, Piero Madau will shed some light on these dark ages.

2.10 First Structures

Dear Piero *(Madau)***, the epoch of first stars, galaxies, SNe, and BHs is certainly one of the most intriguing problems of current astrophysics. Would you like to focus on the most important aspects connected with that dawn age in which the Universe emerged from the dark era? Which observations may help solving the actual controversies? Can you discuss the uncertainties in extending the Madau's plot to higher redshifts, closer to the epoch cosmic reionization?**

The development of primordial inhomogeneities into the nonlinear regime and the formation of the first astrophysical objects within DM halos mark the transition from a simple, neutral, cooling Universe – described by just a few parameters – to a messy ionized one – the realm of radiative, hydrodynamic, and star formation processes. The WMAP polarization data show that this transition must have begun quite early, and that the Universe was fully reionized some 350 million years after the Big Bang. It is a young generation of extremely metal-poor massive stars and/or "seed" accreting BHs in subgalactic halos that may have generated the ultraviolet radiation and mechanical energy that reheated and reionized most of the hydrogen in the cosmos. The detailed thermal, ionization, and chemical enrichment history of the Universe during the crucial formative stages around redshift 10 depends on the power-spectrum of density fluctuations on small scales, the stellar IMF and star formation efficiency, a complex network of poorly understood "feedback" mechanisms, and remains one of the crucial missing links in galaxy formation and evolution studies.

2.10.1 Preamble

Hydrogen in the Universe recombined about half a million years after the Big Bang, and cooled down to a temperature of a few kelvins until the first non-linearities developed, and evolved into stars, galaxies, and BHs that lit up the Universe again. In currently popular cold DM flat cosmologies (ΛCDM), some time beyond a redshift of 10, the gas within halos with virial temperatures $T_{\rm vir} \gtrsim 10^4$ K – or, equivalently, with masses $M \gtrsim 10^8\,[(1+z)/10]^{-3/2}$ $M_\odot$ – cooled rapidly due to the excitation of hydrogen Lyα and fragmented. Massive stars formed with some IMF, synthesized heavy elements, and exploded as type II SNe after a few $\times 10^7$ years, enriching the surrounding medium: these subgalactic stellar systems, aided perhaps by an early population of accreting BHs in their nuclei, generated the ultraviolet radiation and mechanical energy that contributed to the reheating and reionization of the cosmos. It is widely believed that collisional excitation of molecular hydrogen may have allowed gas in even smaller systems – virial temperatures of a thousand K, corresponding to masses around $5 \times 10^5[(1+z)/10]^{-3/2}$ $M_\odot$ – to cool and form stars at even earlier times [2, 76, 551]. Throughout the epoch of structure formation, the all-pervading IGM, which contains most of the ordinary baryonic material left over from the Big Bang, becomes clumpy under the influence of gravity, and acts as a source for the gas that gets accreted, cools, and forms stars within subgalactic fragments, and as a sink for the metal enriched material, energy, and radiation which they eject. The well-established existence of heavy elements like carbon and silicon in the Lyα forest clouds at $z = 2$–6 [482, 522] may be indirect evidence for such an early episode of pregalactic star formation. The recently released 5-year *WMAP* data require the Universe to be fully reionized by redshift 11 ± 1.4 [142], another indication that significant star-formation activity started at very early cosmic times.

The last decade has witnessed great advances in our understanding of the high redshift Universe, thanks to breakthroughs achieved with satellites, 8–10 m class telescopes, and CMB experiments. Large surveys such as the SDSS, together with the use of novel instruments and observational techniques have led to the discovery of galaxies and quasars at redshifts in excess of 6. At the time of writing, nine quasars have already been found with $z > 6$ [163], and one actively star-forming has been spectroscopically confirmed at $z = 6.96$ [253]. These sources probe an epoch when the Universe was <7% of its current age. *Keck* and *VLT* (Very Large Telescope) observations of redshifted H I Lyα ("forest") absorption have been shown to be sensitive probe of the distribution of gaseous matter in the Universe (e.g., [438]). Gamma-ray bursts have recently displayed their potential to replace quasars as the preferred probe of early star formation and chemical enrichment: GRB050904, the most distant event known to date, is at $z = 6.39$ [224]. The underlying goal of all these efforts is to understand the growth of cosmic structures, the properties of galaxies and their evolution, and ultimately to map the transition from the cosmic "dark age" to an ionized Universe populated with luminous sources.

Progress has been equally significant on the theoretical side. The key idea of currently popular cosmological scenarios, that primordial density fluctuations grow by gravitational instability driven by collisionless CDM, has been elaborated upon and explored in detail through large-scale numerical simulations on supercomputers, leading to a hierarchical ("bottom-up") scenario of structure formation. In this model, the first objects to form are on subgalactic scales, and merge to make progressively bigger structures ("hierarchical clustering").

Ordinary matter in the Universe follows the dynamics dictated by the DM until radiative, hydrodynamic, and star formation processes take over. According to these calculations, a truly inter- and proto-galactic medium (the main repository of baryons at high redshift) collapses under the influence of DM gravity into flattened and filamentary structures, which are seen in absorption against background QSOs. Gas condensation in the first baryonic objects is possible through the formation of H_2 molecules, which cool via roto-vibrational transitions down to temperatures of a few hundred kelvins. In the absence of a UV photodissociating flux and of ionizing X-ray radiation, three-dimensional simulations of early structure formation show that the fraction of cold, dense gas available for star formation and accretion onto seed BHs exceeds 20% for halos more massive than 10^6 $M_\odot$ already at redshifts 20 [299, 325].

In spite of some significant achievements in our understanding of the formation of cosmic structures, there are still many challenges facing hierarchical clustering theories. While quite successful in matching the observed large-scale density distribution (like, e.g., the properties of galaxy clusters, galaxy clustering, and the statistics of the Lyα forest), CDM simulations appear to produce halos that are too centrally concentrated compared to the mass distribution inferred from the rotation curves of (DM-dominated) dwarf galaxies, and to predict too many DM subhalos compared to the number of dwarf satellites observed within the Local Group [280, 331, 365]. Another perceived difficulty (arguably connected with the "missing satellites problem", see, e.g., [78]) is our inability to predict when and how the Universe was reheated and reionized, that is, to understand the initial conditions of the galaxy formation process and the basic building blocks of today's massive baryonic structures. We know that at least some galaxies and quasars had already formed when the Universe was less than 10^9 year old, and have made great progress in mapping the cosmic star formation history of bright galaxies up to redshift 7 or so [69, 329] (see Fig. 2.28). But when did the first luminous systems form,

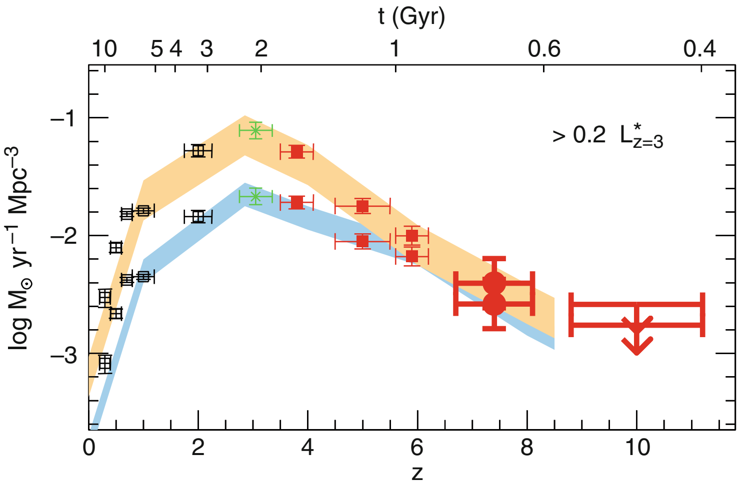

Fig. 2.28 Estimated SFR density as a function of redshift (integrated down to 0.2 $L^*_{z=3}$). The lower set of points give the SFR density without a correction for dust extinction, and the upper set of points give the SFR density with such a correction. This is also indicated with the shaded regions, respectively, where the width of these regions show the approximate uncertainties. From [69]

was star formation efficient in objects below the atomic cooling mass, and what was the impact of the first stars on the surrounding intergalactic gas? The crucial processes of star formation and "feedback" (e.g., the effect of the energy input from the earliest generations of sources on later ones) in the nuclei of galaxies are still poorly understood. Accreting BHs can release large amounts of energy to their surroundings, and may play a role in regulating the thermodynamics of the ISM, ICM, and IGM (e.g., [135]). The detailed astrophysics of these processes is, however, unclear. Although we may have a sketchy history of the production of the chemical elements in the Universe, we know little about how and where exactly they were produced and how they are distributed in the IGM and in the ICM . Finally, where are the first stars and their remnants now, and why do the hundreds of early-forming, massive satellites predicted to survive today in the Milky Way halo remain dark?

2.10.2 The Dark Age and the Emergence of Cosmic Structure

2.10.2.1 The Dark Age

The Universe became optically thin to Thomson scattering at redshift 1,100, and entered a "dark age". At this epoch the electron fraction dropped below 15%, and the primordial radiation cooled below 3,000 K, shifting first into the infrared and then into the radio. We understand the micro-physics of the post-recombination Universe well. The fractional ionization froze out to the value $\sim 3 \times 10^{-4}$: these residual

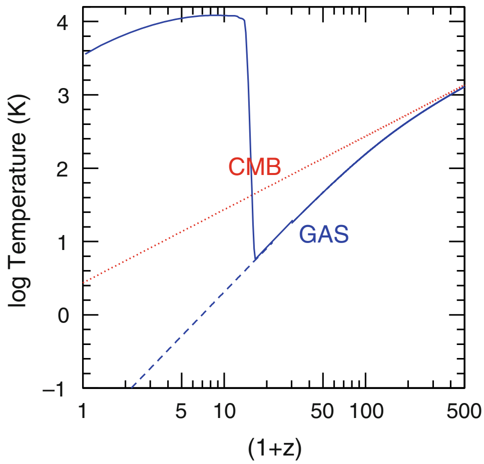

Fig. 2.29 Evolution of the CMB radiation (*long-dashed line*) and IGM gas (*solid line*) temperatures after recombination. The Universe is assumed to be reionized by ultraviolet radiation at $z \sim 10$. The *short-dashed line* is the extrapolated gas temperature in the absence of any reheating mechanism

electrons were enough to keep the matter in thermal equilibrium with the radiation via Compton scattering until a redshift of $z_t \sim 150$, that is, well after the Universe became transparent. Thereafter, the matter temperature decreased as $(1+z)^2$ due to adiabatic expansion (Fig. 2.29) until primordial inhomogeneities in the density field evolved into the nonlinear regime.

The minimum mass scale for the gravitational aggregation of CDM particles is negligibly small. One of the most popular CDM candidates is the neutralino: in neutralino CDM, collisional damping and free streaming smear out all power of primordial density inhomogeneities only below Earth-mass scales. Baryons, however, respond to pressure gradients and do not fall into DM clumps below the cosmological Jeans mass (in linear theory this is the minimum mass-scale of a perturbation where gravity overcomes pressure),

$$M_{\rm J} = 4\pi \frac{\rho}{3} \left(\frac{5\pi k_{\rm B} T_{\rm e}}{12 G \rho m_{\rm p} \mu} \right)^{3/2} \approx 9 \times 10^4 \ {\rm M}_\odot (a T_{\rm e}/\mu)^{3/2}. \qquad (2.19)$$

Here $a = (1+z)^{-1}$ is the scale factor, ρ the total mass density including DM, μ the mean molecular weight, and $T_{\rm e}$ the gas temperature. In the post-recombination Universe, the baryon-electron gas remained thermally coupled to the CMB, $T_{\rm e} \propto a^{-1}$, and the Jeans mass was independent of redshift and comparable to the mass of globular clusters, $M_{\rm J} \approx 10^6 \ {\rm M}_\odot$. At $z < z_{\rm t}$, the temperature of the baryons dropped as $T_{\rm e} \propto a^{-2}$, and the Jeans mass decreased with time, $M_{\rm J} \propto a^{-3/2}$. This trend was reversed by the reheating of the IGM. The energy released by the first collapsed objects drove the Jeans mass up to galaxy scales: baryonic density perturbations stopped growing as their mass dropped below the new Jeans mass. Photo-ionization by the ultraviolet radiation from the first stars and quasars heated the IGM to temperatures of $\approx 10^4$ K (corresponding to a Jeans mass $M_{\rm J} \lesssim 10^{10} \ {\rm M}_\odot$ at $z \simeq 11$), suppressing gas infall into low mass halos and preventing new (dwarf) galaxies from forming (see Fig. 2.30).

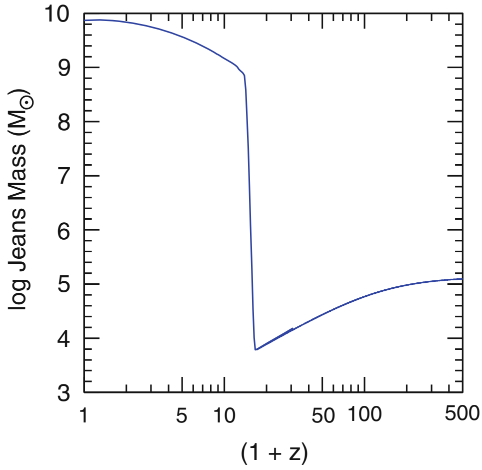

Fig. 2.30 Cosmological (gas + DM) Jeans mass for a Universe reionized by ultraviolet radiation

2.10.2.2 The Emergence of Cosmic Structure

As mentioned earlier, some shortcomings on galactic and sub-galactic scales of the currently favored model of hierarchical galaxy formation in a Universe dominated by CDM have appeared in the last decade. The significance of these discrepancies is still debated, and "gastrophysical" solutions involving feedback mechanisms may offer a possible way out. Other models have attempted to solve the apparent small-scale problems of CDM at a more fundamental level, that is, by reducing small-scale power. Although the "standard" ΛCDM model for structure formation assumes a scale-invariant initial power spectrum of density fluctuations, $P(k) \propto k^n$ with $n = 1$, there is strong evidence in the WMAP data for a departure from scale invariance, with a best fit value $n = 0.963^{+0.014}_{-0.015}$ [142]. The one and three-year *WMAP* data showed some preference for a scale-dependent index, $\mathrm{d}n/\mathrm{d}\ln k < 0$, that is, for a model in which the spectral index varies as a function of wavenumber k [523].

The 5-year WMAP data do not significantly prefer such a "running index". Models with either $n < 1$ or $\mathrm{d}n/\mathrm{d}\ln k < 0$ predict a significantly lower amplitude of fluctuations on small scales than standard ΛCDM. The suppression of small-scale power has the advantage of reducing the amount of substructure in galactic halos and makes small halos form later (when the Universe was less dense) hence less concentrated [601]. But it makes early reionization a challenge.

Figure 2.31 shows the linearly extrapolated (to $z = 0$) variance of the mass-density field for different power spectra. In the CDM paradigm, structure formation proceeds "bottom-up", that is, the smallest objects collapse first, and subsequently merge together to form larger objects. It then follows that the loss of small-scale power modifies structure formation most severely at the highest redshifts, significantly reducing the number of self-gravitating objects then. This, of course, will make it more difficult to reionize the Universe early enough. It has been argued, for example, that one popular modification of the CDM paradigm, warm DM (WDM),

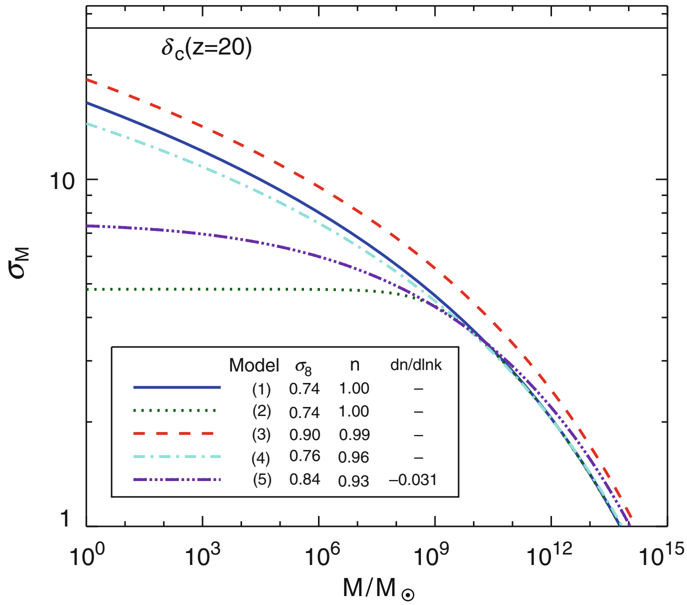

Fig. 2.31 The variance of the matter-density field vs. mass M for different power spectra. All models assume a concordance cosmology with parameters $(\Omega_m, \Omega_\Lambda, \Omega_b, h) = (0.29, 0.71, 0.045, 0.7)$. *Solid curve:* standard ΛCDM with no tilt, cluster normalized. *Dotted curve:* ΛWDM with a particle mass $m_X = 2\,\text{keV}$, cluster normalized, no tilt. *Dashed curve:* tilted *WMAP* model, *WMAP* data only. *Dash-dotted curve:* tilted *WMAP* model, including 2dFGRS and Lyα data. *Dash-triple dotted curve:* running spectral index *WMAP* model, including 2dFGRS and Lyα data. Here n refers to the spectral index at $k = 0.05\,\text{Mpc}^{-1}$. The horizontal line at the top of the figure shows the value of the extrapolated collapse overdensity $\delta_c(z)$ at $z = 20$

has so little structure at high redshift that it is unable to explain the WMAP observations of an early epoch of reionization [523]. The running-index model may suffer from a similar problem. A look at Fig. 2.31 shows that $10^6\,M_\odot$ halos will collapse at $z = 20$ from 2.9σ fluctuations in a tilted ΛCDM model with $n = 0.99$ and $\sigma_8 = 0.9$, from 4.6σ fluctuations in a running-index model, and from $5.7\,\sigma$ fluctuations in a WDM cosmology. The problem is that scenarios with increasingly rarer halos at early times require even more extreme assumptions (i.e., higher star formation efficiencies and UV photon production rates) in order to be able to reionize the Universe suitably early [94, 103, 223, 521, 594].

The study of the nonlinear regime for the baryons is far more complicated than that of the DM because of the need to take into account pressure gradients and radiative processes. As a DM halo grows and virializes above the cosmological Jeans mass through merging and accretion, baryonic material will be shock heated to the effective virial temperature of the host and compressed to the same fractional overdensity as the DM. The subsequent behavior of gas in a DM halo depends on the efficiency with which it can cool. It is useful here to identify two mass scales for the host halos: (1) a *molecular cooling mass* M_{H_2} above which gas can cool

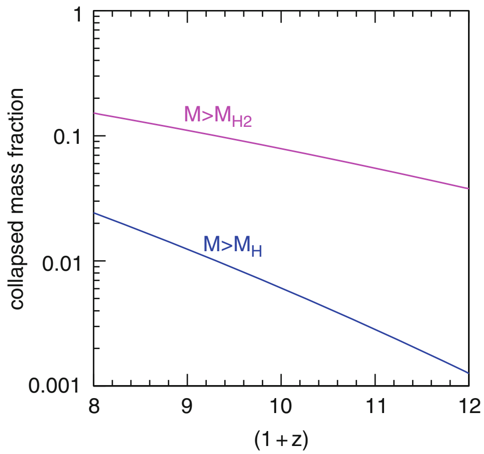

Fig. 2.32 *Solid lines:* Total mass fraction in all collapsed DM halos above the molecular cooling and the atomic cooling masses, M_{H_2} and M_{H}, as a function of redshift, for standard ΛCDM cosmology

via roto-vibrational levels of H_2 and contract, $M_{\mathrm{H}_2} \approx 10^5\,[(1+z)/10]^{-3/2}\ \mathrm{M}_\odot$ (virial temperature above 200 K); and (2) *an atomic cooling mass* M_{H} above which gas can cool efficiently and fragment via excitation of hydrogen Lyα, $M_{\mathrm{H}} \approx 10^8\,[(1+z)/10]^{-3/2}\ \mathrm{M}_\odot$ (virial temperature above 10^4 K). Figure 2.32 shows the fraction of the total mass in the Universe that is in collapsed DM halos with masses greater than M_{H_2} and M_{H} at different epochs.

2.10.2.3 The Epoch of Reionization

As hierarchical clustering theories provide a well-defined framework in which the history of baryonic material can be tracked through cosmic time, probing the reionization epoch may then help constrain competing models for the formation of cosmic structures. Quite apart from uncertainties in the primordial power spectrum on small scales, however, it is the astrophysics of baryons that makes us unable to predict when reionization actually occurred. Consider the following illustrative example:

Hydrogen photo-ionization requires more than one photon above 13.6 eV per hydrogen atom: of order $t/\bar{t}_{\mathrm{rec}} \sim 10$ (where $\bar{t}_{\mathrm{rec}}$ is the volume-averaged hydrogen recombination timescale) extra photons appear to be needed to keep the gas in overdense regions and filaments ionized against radiative recombinations [206, 330]. A "typical" stellar population produces during its lifetime about 4,000 Lyman continuum (ionizing) photons per stellar proton. A fraction $f \sim 0.25\%$ of cosmic baryons must then condense into stars to supply the requisite ultraviolet flux. This estimate assumes a standard (Salpeter) IMF, which determines the relative abundances of hot, high mass stars vs. cold, low mass ones.

The very first generation of stars ("Population III") must have formed, however, out of unmagnetized metal-free gas: numerical simulations of the fragmentation of pure H and He molecular clouds [2, 76] have shown that these characteristics likely led to a "top-heavy" IMF biased towards very massive stars (i.e., stars a few hundred times more massive than the Sun), quite different from the present-day Galactic case. Metal-free very massive stars emit about 10^5 Lyman continuum photons per stellar baryon [75], approximately 25 times more than a standard stellar population. A corresponding smaller fraction of cosmic baryons would have to collapse then into very massive stars to reionize the Universe, $f \sim 10^{-4}$. There are of course further complications. As, at zero metallicity, mass loss through radiatively driven stellar winds is expected to be negligible [298], Population III stars may actually die losing only a small fraction of their mass. If they retain their large mass until death, very massive stars with masses $140 \lesssim m \lesssim 260$ $M_\odot$ will encounter the electron–positron pair instability and disappear in a giant nuclear-powered explosion [184], leaving no compact remnants and polluting the Universe with the first heavy elements. In still heavier stars, however, oxygen and silicon burning is unable to drive an explosion, and complete collapse to a BH will occur instead [56]. Thin disk accretion onto a Schwarzschild BH releases about 50 MeV per baryon. The conversion of a trace amount of the total baryonic mass into early BHs, $f \sim 3 \times 10^{-6}$, would then suffice to reionize the Universe.

Even if the IMF at early times were known, we still would remain uncertain about the fraction of cold gas that gets retained in protogalaxies after the formation of the first stars (this quantity affects the global efficiency of star formation at these epochs) and whether – in addition to ultraviolet radiation – an early input of mechanical energy may also play a role in determining the thermal and ionization state of the IGM on large scales. The same massive stars that emit ultraviolet light also explode as SNe, returning most of the metals to the ISM of pregalactic systems and injecting about 10^{51} ergs per event in kinetic energy. A complex network of feedback mechanisms is likely at work in these systems, as the gas in shallow potential is more easily blown away [130], thereby quenching star formation. Furthermore, as the blast waves produced by SN explosions – and possibly also by winds from "miniquasars" – sweep the surrounding intergalactic gas, they may inhibit the formation of nearby low-mass galaxies, and drive vast portions of the IGM to a significantly higher temperature than expected from photo-ionization, so as to "choke off" the collapse of further galaxy-scale systems. Note that this type of global feedback is fundamentally different from the "in situ" heat deposition commonly adopted in galaxy formation models, in which hot gas is produced by SNe within the parent galaxy.

2.10.3 High Redshift Quasars and BH Feedback

Dear Piero *(Madau)*, the last years have seen a big advance in our understanding of the galaxy formation process. Of particular importance has been the discovery of quasars at high redshifts (up to $z = 6$ and beyond). Can you explain how these studies are connected with the present cosmological scenario? Another theme that is now achieving increasing popularity is that of the BHs

feedback in the galaxy evolution process. Can you point out on the current status of these studies?

The strong link observed between the masses of supermassive BHs at the center of most galaxies and the gravitational potential wells that host them suggests a fundamental mechanism for assembling BHs and forming spheroids in galaxy halos. The $m_{\rm BH}$-σ relation [169, 193] implies a rough proportionality between supermassive BH mass and the mass of the baryonic component of the bulge. It is not yet understood whether this relation was set in primordial structures, and consequently how it is maintained throughout cosmic time with such a small dispersion, or indeed which physical processes established such a correlation in the first place (e.g., [514]).

In CDM-dominated cosmologies, galaxy halos experience multiple mergers during their lifetime, with those between comparable-mass systems ("major mergers") expected to result in the formation of elliptical galaxies [240]. Simple models in which supermassive BHs are also assumed to grow during major mergers and to be present in every galaxy at any redshift – while only a fraction of them is "active" at any given time – have been shown to explain many aspects of the observed evolution of quasars [269] (see Fig. 2.33). The coevolution of supermassive BHs and their host galaxies in hierarchical structure formation scenarios gives origin to a number of important questions, most notably the following:

- Did the first supermassive BHs form in subgalactic units far up in the merger hierarchy, well before the bulk of the stars observed today? The seeds of the $z \sim 6$ quasars discovered in the SDSS had to appear at very high redshift, $z \gtrsim 10$, if they are accreting no faster than the Eddington rate. In hierarchical cosmologies, the ubiquity of supermassive BHs in nearby luminous galaxies can arise even if only a small fraction of halos harbor supermassive BHs at very high redshift [356].

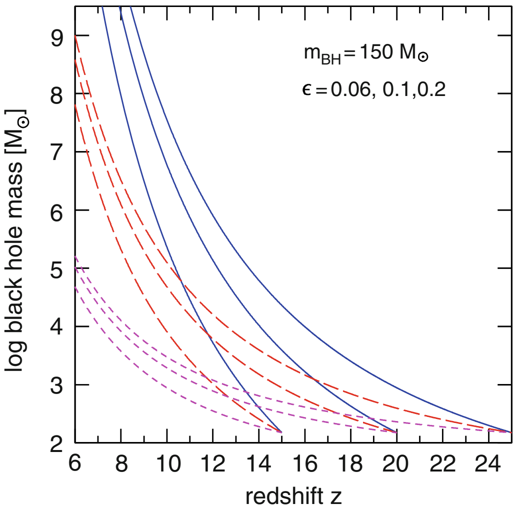

Fig. 2.33 Growth of supermassive BHs from early epochs down to $z = 6$, the redshift of the most distant SDSS quasars. The three sets of curves assume Eddington-limited accretion with radiative efficiency $\epsilon = 0.06$ (*solid lines*), 0.1 (*long-dashed lines*), and 0.2 (*short-dashed lines*). Gas accretion starts at $z = 15, 20, 25$ onto a seed black of mass $m_{\rm BH} = 150\ M_\odot$

- How massive were the initial seeds, and is there a population of relic pregalactic supermassive BHs lurking in present-day galaxy halos? A clue to these questions may lie in the numerous population of ultraluminous off-nuclear ("non-AGN") X-ray sources that have been detected in nearby galaxies. Assuming isotropic emission, the inferred masses of these "ULXs" may suggest intermediate-mass BHs with masses $\gtrsim$ a few hundred $M_\odot$ [262].
- Do supermassive BH binaries form and coalescence in large numbers? If supermassive BHs were common in the past (as implied by the notion that many distant galaxies harbor active nuclei for a short period of their life), and if their host galaxies undergo multiple mergers, then supermassive BH binaries will inevitably form in large numbers during cosmic history. Supermassive BH pairs that are able to coalesce in less than a Hubble time will give origin to the loudest gravitational wave events in the Universe [508].
- It was first proposed by [145] that the heating of the surrounding stars by a decaying supermassive BH pair would create a low-density stellar core out of a preexisting cuspy (e.g., $\rho_* \propto r^{-2}$) stellar density profile. If stellar dynamical processes can efficiently drive wide supermassive BH binaries to the gravitational wave emission stage, what is the cumulative dynamical effect of multiple BH mergers on galaxy stellar cusps?
- AGN powered by supermassive BHs keep the Universe ionized at $z \lesssim 4$, structure the IGM, and probably regulate star formation in their host galaxies [135]. Intermediate-mass holes accreting gas from the surrounding medium may shine as "miniquasars" at redshifts as high as $z \sim 20$. What is the thermodynamic effect of miniquasars on the IGM at early times?

Thank you Piero.
In the currently accepted cosmological scenario, larger structures formed from merging of smaller ones, in a hierarchical scheme. Clusters of galaxies are the largest observed self-gravitating structures of the Universe. The following interviews with Alan Dressler and Isabella Gioia will give an overview of clusters and their implications for cosmology. While Alan will focus on the properties of galaxies contained in clusters, Isabella will touch on the use of clusters as cosmological tools. Let us start with Alan, who will now describe to us the kind of information that can be derived from the morphology–density relation and the scaling relationship of galaxies.

2.11 Galaxy Clusters, The Largest Self-gravitating Structures of the Universe

Dear Alan *(Dressler)*, **galaxy clusters are the largest self-gravitating structures of our observable Universe. Understanding their properties is really important from a cosmological point of view. In particular, we are interested to better understand the importance for our cosmology of the Morphology–Density**

relation, the well know relation you have discovered. Early-type galaxies seem indeed overabundant at the center of clusters. How this observed fact can be useful for present day cosmological models? Could this relation be understood invoking the evolution of cosmic structures?

In modern cosmological simulations, rich clusters of galaxies are identified as intersections of filaments and/or sheets, which characterize the distribution of the DM on scales of tens of megaparsecs. In, for example, the Millennium Simulation, these one- and two-dimensional structures contain most of the mass and, as the Universe expands, DM – and the baryons that accompany it – move along the filaments and sheets towards the regions of highest density. It is natural to associate the large bound structures – rich clusters of galaxies – with these "nodes" in the DM distribution. Even though rich clusters are identified primarily by their baryonic component, that is, galaxies and hot gas, they obviously contain huge amounts of DM. The mass spectrum of clusters of galaxies, determined either through dynamical measurements (galaxy motions) or total luminosity under the assumption of an average mass-to-light ratio, has been compared to the expectations of the N-body simulations in a ΛCDM Universe and found to be in excellent agreement. In fact, the ability of the models to faithfully predict the power spectrum (volume density) of DM structures from the largest scales, as measured by WMAP or ground-based CMB experiments, is stunning, considering the small number of free parameters, many of which are also constrained by other astronomical measurements. Thus, it seems indisputable that the rich clusters of galaxies do represent the nodal structures seen in the simulations, and the evolution of their DM content, through steaming along the structures and mergers, provides a sound base upon which the evolution of rich clusters of galaxies can be investigated. The extraordinary discovery of "the bullet cluster" is an excellent example of how the DM simulations have succeeded in providing this foundation.

As remarkable a success story as the modeling of DM has been, so great a disappointment have been the attempts to add baryons, hydrodynamics, and star formation, at least if the aim is to predict the evolving properties of galaxies – on the baryonic side of the equation. Despite enormous effort, the modeling has demonstrated no predictive ability whatsoever. Although many results of theory papers describe a compatibility of theory and observations, such as the global history of star formation, or the formation of an elliptical galaxy through mergers, this success comes from adjusting and manipulating the models in order to match existing observations. I agree with Ostriker when he says that these hardly merit the label "prediction" – "postdiction" is a better description. This is not to say that the ability to match simulations with observed properties of galaxies and clusters is easy and useless – after all, there are constraints on physics that make the processes and tuning either reasonable or not, but matching the properties of real galaxies is a long way from demonstrating that this is the actual physics that is responsible for this or that phenomena.

I believe that it is fair to say, then, that the relationship between galaxy morphology and local galaxy density that I reported in 1980 has yet to provide meaningful constraints on models of structure formation. This is because the morphology–density relation is fundamentally about galaxy structure, morphology, and star formation history, none of which can as yet serve as true constraints in ΛCDM models plus baryons, not until they can do more than "match" observations. I should add, as an anecdote, that for many years after my 1980 paper, I was sought out by one theorist after another, each of whom gleefully reported that their own model of galaxy formation produced the morphology–density relation, sometimes approximately, sometimes exactly – but all were successful. A selection effect, to be sure no one rushes to report failure – but after a while it was clear that the morphology–density relation must be rather easy to manufacture – it must be basic, so much so that it might not be much help in ruling in and out models of the structural and stellar-population evolution of galaxies.

There are, however, clear lessons to be drawn from the morphology–density relation that provide some guidance. First, there is the "nature–nurture" debate, about which I believe the morphology–density relation has much to say. The striking thing is not that such relations exist – that ellipticals and S0s become more prominent, in higher-density regions, while the proportion of spirals and irregulars declines – but that these effects are such a slow function of local density – logarithmic, in fact. By almost any measure, one would expect that, if the morphology–density relation is the result of galaxy interactions with each other, or with the potential well of the cluster, or with the hot intracluster gas, these effects would be much stronger – at least linear with density if not stronger still. The fact that they are in fact so much weaker than linear has always suggested to me that the morphology–density relationship must have been established at much earlier times, before the density-contrast was amplified by gravity. If this is true, then the morphology–density relation must be a "legacy relation" that points more to "nature" – that is, initial conditions – than to late processes happening over the whole of cosmic time.

What this means for the models, of course, is that they must be able to produce all the major types of galaxies and have each type inhabit the full range of environments. Elliptical and S0 galaxies are found even in the "field" – the loose groups and isolated galaxies – although much less common, they are far from rare. This is an expression of the "slow" dependence of the morphology–density relation – spiral galaxies may be truly absent from the very cores of rich clusters, but elliptical and S0 galaxies in the field actually outnumber those in rich clusters, once you take into account how many more field galaxies there are compared to cluster galaxies. Therefore, it seems clear that galaxy morphology must be a result of early processes, when the difference between field and cluster density was small and both proto-field and proto-cluster contained strong density peaks that could become elliptical or bulge-dominated disk galaxies. A model that relies on bashing latter-day spirals together to make ellipticals, or stripping them in the cores of rich clusters – as early semi-analytical models from the Durham Group did – is clearly wrong.

But, more recently it has become apparent that there is something even more fundamental about the morphology–density relationship that escaped all of us 25 years ago. The Hubble sequence is better correlated with *mass* than it is with anything else, including local density. I should explain that. In the 1970s, a lot of energy went into comparing the luminosity functions of galaxies in clusters and the field – they are almost identical. Some even used the universality of the luminosity function as evidence that galaxy formation took place in the nearly uniform conditions of the young Universe (and all differences with environment developed much later), the opposite of what I have just argued. What escaped me and (as far as I know) everyone else is that, when the mass-to-light ratios of those populations are folded in, considering that the field is mostly low M/L spirals and the cores of clusters are dominated by high M/L ellipticals, there is a strong trend: high density regions contain more massive galaxies! The similarity of luminosity functions turns out to have been accidental and a red herring. It seems to me that this makes the existence of the morphology–density relation all the more natural: any sensible hierarchical model will result in more massive galaxies in denser environments, from halo mergers, gas infall, stellar mergers – whatever. This points directly to initial conditions, or early evolution, if you will – nature over nurture – in determining a galaxy's morphological type. It also emphasizes why, if ellipticals are the result of mergers, they are not mergers of "lightweight" spirals galaxies, but mergers in the first few Gyr of more massive gas rich building blocks, that may or may not have born any resemblance to today's spirals. If the N-body simulations can move beyond matching observations to develop rigorous recipes of star formation, cooling, feedback, and AGN influences, then the morphology-density relation should fall right out. This will include the "legacy" ellipticals in low-density environments, which must have developed from high-density peaks in the primordial spectrum that were *not* atop large swells that would become rich groups and clusters.

A final thought: as the Hubble observations of rich clusters have extended out to $z \sim 1$, the evolution of the morphology–density relation has been traced over the last half of a Hubble time. As expected, the elliptical galaxies are there, in their present-epoch numbers, even at $z \sim 1$. This agrees with the "nature" model of their formation, but many, perhaps most of the S0s seem to have been arrived later. Clearly, spiral galaxies are involved in the production of S0 galaxies, but in particular it seems that big-bulge spirals, being more massive and more common in dense regions, are the precursors of many of today's S0s. A successful cosmological model of structure grown should include a prescription for how this transformation took place, integrated into the building of the clusters themselves through infall.

Large scale structure studies may constrain, as CMB studies do, the cosmological parameters. Constraints can be found also for the mass of neutrino and for the ratio Ω_ν/Ω_m. There are evidences, however, that on scales smaller than 10 h^{-1} Mpc, different galaxy types are clustered differently. The "bias" in general may introduce systematic effects on the determination of the cosmological parameters from redshift surveys . What is the real progress that can be achieved from a better knowledge of galaxy clustering?

As I mentioned in my discussion earlier, the morphology–density relation looks somewhat different when viewed through the prism of galaxy mass. This is important because mass is the most fundamental parameter used to describe a galaxy. Galaxy morphology seems to be strongly dependent on mass, which is something that successful theoretical work must account for. About the goal of measuring cosmological parameters with galaxies and clusters, I am less confident about this approach as compared to using type Ia SNe to measure the equation-of-state and its evolution, simply because I believe that evolutionary effects are small or can be controlled for SNe-based studies. In a word, the evolution of large scale structure does not feed back into the evolution of the SNe themselves, and that makes them a good tool for the job. It is not so obvious that the same is true for galaxies or galaxy clusters, where there is likely some crosstalk between structure growth and the properties of galaxies.

The idea of bias – that different morphological types traced the distribution of underlying mass differently – arose in the 1980s when observations of large-scale flows of galaxies were compared to expectations from the large-scale distribution of light and the CMB dipole. Compared to the distribution of light, and assuming that mass is traced by light, early-type galaxies were found to be biased tracers – more concentrated than the mass. The bluer the wavelength band in which the galaxies were selected, the greater this effect. As galaxy surveys moved to longer wavelengths, the effect of bias – the difference in how the late- and early-type galaxies trace mass – becomes smaller. Now that we have galaxy surveys in the infrared, I would expect that some of these statistics – if recalculated – would reduce to no bias at all. In other words, a mass-selected sample of galaxies is likely to be a good tracer of the mass distribution of the large-scale structure. This is what is important for a robust determination of cosmological parameters, for example, using galaxy redshift surveys to map the "baryon oscillations" that are an imprint from the Big Bang and thus provide constraints on the equation of state.

If the goal is to study the *cause* for the bias of galaxies as a function of morphological type, the approach is different. Here we would like to accentuate as much as possible the differences between the way galaxy light traces underlying (dark) mass, in order to try to understand the elements of galaxy formation and evolution that might be responsible. In other words, the bias is just another constraint on models of large-scale structure formation, one that cannot be thoroughly investigated until the models can reproduce the morphology of galaxies from (nearly) first principles of physics. It is my opinion that the question of whether light traces mass on all scales smaller than 50 Mpc has never been fully resolved – the community lost interest in the subject during the time when N-body simulations were so successfully reproducing the gross properties of observed large-scale structure. I myself have been trying to carry on a program that would make better measurements of galaxy non-Hubble motions to obtain better constraints. I believe that real progress – improving significantly on what has already been done – requires imaging from the HST. Unfortunately, I and many others interested in this program have found Hubble time-allocation committees unenthusiastic.

I should mention that, although I do not think using galaxies or clusters as probes of cosmological parameters is as robust as the CMB or SNe techniques, the question is so important that attempts to make independent checks on the evolution of the acceleration are well justified. In this respect, clusters of galaxies seem to have an increasing role to play, now that physical parameters such as mass and density can be more accurately measured. When these quantities were estimated from galaxy motions and distributions, the cosmic scatter was discouraging, but using gravitational lensing and constructing mass-limited samples using the Sunyaev-Zel'dovich effect, the prospects are much improved. If the physical parameters of galaxy clusters at cosmological distances can be well determined, an important additional measurement of the history of the cosmic expansion will be forthcoming.

The observed properties of the scaling relations, such as the Fundamental Plane of Early-type galaxies and the Tully–Fisher relation for spirals, seem to point toward the "monolithic" scenario of galaxy formation, while numerical simulations are in agreement with the "hierarchical merging" scenario. Do you want to express your point on view on this?

Hierarchical models of the formation of cosmic structure became popular in the 1980s, and their dominance ever since is well justified. People tend to forget the dilemma confronting the cosmological models of the day, which were dominated by baryons. CMB experiments achieved ever-more-sensitive temperature measurements, yet the fluctuations which were needed to account for today's structure failed to emerge. The introduction of a CDM constituent, interacting only through gravity and slowly streaming, rescued astrophysics from this existential dilemma. CDM offered a robust prediction for the detection of the fluctuations, at a $\Delta T/T \sim 10^{-5}$ K, where they were subsequently found by the COBE satellite, in one of the highpoints of twentieth century physics. As the field of CMB observations has flourished, an astonishing agreement of the theoretical and observed power spectrum has emerged, over orders of magnitude of scale – this in itself leaves only a little room for neutrinos or other kinds of warm or hot DM. Thus, it seems that the N-body simulations based on CDM have from the start been on a firm footing.

From the point of view of an observer, however, the stunningly successful predictions of large scale structure came with a price. Built-in was a galaxy-formation picture in which the largest galaxies were expected to have formed *last*, with the latest and greatest amount of star formation – a picture that was in screaming disagreement with observations from early on. My spectroscopy of distant cluster and field galaxies with Jim Gunn in the 1980s, and the many photometric and spectroscopic studies that followed, showed a Universe in which the most massive galaxies – the ellipticals – appeared to be old even 5 Gyr ago. This was true not only for the most massive cD and radio galaxies, which had been studied as far back as the 1960s, but also for a factor of 10 down the mass function to quite average elliptical galaxies. Contrary to early predictions of the N-body simulations, the Universe at $z \sim 1$ was found to be remarkably similar to what we see around us today. In fact, evolution was observed to be least for the massive galaxies but most for the smallest galaxies observable at $z \sim 1$, $M \sim 0.1\,M^*$. In a hierarchical Universe, these small galaxies were expected to have formed first and served as building blocks to make the larger galaxies as the Universe aged.

Of course, we are all older and wiser now. Before there were hydrodynamic components to the N-body simulations, and some attempt to model star formation and other baryon-related physics, it was foolish to relate anything about actual observed galaxies to the evolution of DM halos gleaned from the simulations, which is what we were doing. It is clear now that the 0.1 M* galaxies that are observed to be rapidly evolving at $z \sim 1$ are much different than the higher-density objects of similar mass that were the building blocks of galaxy building at $z \geq 2$. And, just because the DM halos of massive galaxies were still being assembled as late as $z \sim 1$, we recognize now that the history of star formation may have followed a quite different course.

That being said, observers like myself continue to point to an apparently non-hierarchical character of the evolution of luminous galaxies. Thanks to Hubble Deep Fields (HDFs) and 8-m class ground-based telescopes, we can say with some certainty that the most massive galaxies were the first to form: they have, on average, the oldest stars, and there is good evidence from K-band luminosity functions that they were assembled early as well. Researchers who are building the N-body simulations have finally addressed this question head on, recognizing that – at the very least – the models should account for the fact that the most massive galaxies in the densest regions – the cores of rich clusters – have had very little if any star formation as $z \sim 2$. These efforts have focused on understanding the effect on star formation of feedback that must accompany large energy release through star formation and accretion onto massive central BHs – AGN. These are about the only tools a ΛCDM modeler has to push baryons around and prevent them from forming stars. The AGN mechanism, where copious gravitational energy is released through accretion onto a central, massive BH, seems to be getting the most attention, partly because of the popularity of the idea of co-evolution of massive BHs and a galaxy's overall stellar content revealed in the "$M_{BH} - \sigma$ relation". The energy released in these events are sufficient to have a significant effect on star formation, but proving that this feedback mechanism couples efficiently to gas on scales of many kiloparsecs, is a difficult problem, especially as dust appears to play a substantial role in the evolution of young galaxies. It is also unclear that the AGN mechanism would work equally well for galaxies located at the centers of the deep potential wells of groups and clusters (where hot gas might be rapidly cooling) as well as for galaxies that are not so environmentally privileged. This may be a problem because from a stellar population perspective, all such galaxies share the signatures of early, completed star formation.

But, from my point of view, the challenge for the most massive galaxies is symptomatic of a general difficulty that the simulations have in general with producing galaxies whose evolution will lead to the Milky Way and its neighbors. For quite a while it has also been recognized that the hierarchical nature of the structure building with CDM means that the models "have issues" with disk galaxies. Large spiral and S0 galaxies, which dominate the Universe, are more easily formed in some kind of monolithic collapse of a much larger gaseous structure spun by tidal torques from its neighbors (and more likely to survive if they do not suffer bombardment from significant satellites). It is easy to see that a hierarchical merger of small building blocks, ever increasing in mass and scale, will not lead simply to a galaxy with large angular momentum – this is, of course, one of the attractions of the idea of forming ellipticals through mergers. This discrepancy between models and observations for disk galaxies is sometimes referred to as the "Tully–Fisher" problem – the rotation rate for a given mass, a remarkably tight relation for real galaxies, is hard to reproduce in a hierarchical simulation. It is not just the overall rotation, but the small scatter in the relation, that provides a natural challenge to the theory. Once again, the appeal is to feedback to blunt the natural features of hierarchical clustering, this time not from accretion onto massive BHs but from the vast energy release of early star formation over the full extent of the galaxy. Our understanding of feedback is still rudimentary, and star formation and what controls it is still mostly hand-waving (but this may be changing!), so it is difficult to know how believable are these patches and add-ons that enable the models to deliver large, high-angular momentum disk galaxies.

For me, this is the greatest frustration in attempting to compare models and observations. I'm afraid I know of no case where simulations have reported properties of simulated galaxies, as a function of environment or cosmic age, for example, where those properties were previously unknown, that is, a true prediction. Rather, the masters of the simulations find ways to adjust the models, within the bounds of nonexotic physics, to match the observations, and there is no clarity as to whether "the fix" to this problem is compatible to the one you heard about last month to resolve a different discrepancy. There is no sense of observation–leading–theory–leading–observation–leading–theory that has been a hallmark of the field of astrophysics. The problems of star formation and feedback, laced with difficult issues like turbulence and magnetic fields, are not likely to be "solved" any time soon, so this makes it very hard for observers like me to be impressed with what the simulators have accomplished, even though we surely should be. When I look back at what we understood about galaxy formation and the large-scale structure of the Universe when I was in graduate school, I am more than amazed at how far we have come – I never would have predicted such a thing. Both observation and theory have – one way or another – ambled, stumbled, and surged along, and together they have produced spectacular progress. Maybe if there is still a note of discontent with the simulations not being sufficiently "predictive", maybe it is just from being spoiled!

Thanks a lot Alan. Clusters are important for cosmology also for the properties of their ICM. In the following interview, Isabella Gioia will review the basic ideas and the multifrequency approaches adopted in these studies. She will also illustrate how some relationships observed in clusters can be used to infer cosmological information.

2.12 A Multifrequency View of Galaxy Clusters

Dear Isabella *(Gioia)*, **clusters of galaxies are powerful X-ray emitters that can be easily detected out to high redshifts, and therefore are very important tools for cosmologists. In particular, the relation between the X-ray luminosity and the temperature, and temperature and mass of the ICM offers a way to convert a luminosity function into a mass function with obvious consequences for present day cosmology. Can you comment on the state-of-the-art of these studies? How can they constrain cosmological parameters?**

To answer such a question I need to give a little introduction on why clusters of galaxies have always been a preferred tool of cosmologists. I will start with a short description of what clusters are, give a quick background on the different wavelengths, and describe why X-ray clusters play such an important role in astronomy. Finally, I will give some description of the data that can be collected at the different wavelengths and discuss the advantages of the multi-wavelength approach for cosmology, and what needs to be done to improve our understanding of galaxy clusters.

2.12.1 Clusters of Galaxies: An Introduction

About 30 years ago, all clusters of galaxies were selected at optical wavelengths, as the easiest way to identify a cluster is to search for an overdensity in the projected distribution of galaxies in optical images. The pioneering work of Abell [3] and later the catalogs by Zwicky and collaborators [605] made astronomers aware of how many concentrations of galaxies were present in the nearby Universe (that I define here as objects with a redshift $z \leq 0.15$). However, the most visible part of galaxy clusters, all of the stars in all of the galaxies that make up the cluster, contributes only a small fraction of the mass of the cluster. Clusters host manifold components, such as individual galaxies and hot gas (the baryonic component), invisible DM, and what are commonly referred to as "nonthermal components". It is well known that a fraction of clusters (about 40% among rich, hot clusters [199]) shows large scale synchrotron radio emission with no obvious connection to the cluster galaxies, and therefore associated with the ICM (see [165] for a review on the subject). Such extended radio sources are a direct and clear probe of the existence of cluster scale magnetic fields and relativistic particles spread over the same large volume. The composite images in Figs. 2.34 and 2.35 are an illustration of the different features

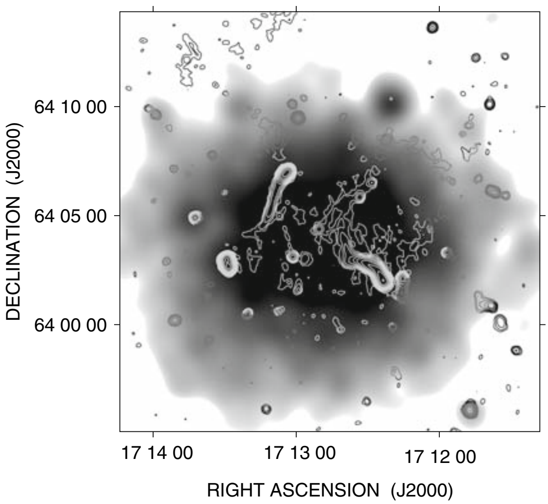

Fig. 2.34 This composite image contains two views of the cluster Abell 2255. Superimposed onto the *ROSAT*-PSPC X-ray emission [167] (in shades of *grey*) are the VLA 1.4 GHz radio emission [212] represented as iso-contours. Courtesy of F. Govoni and M. Murgia

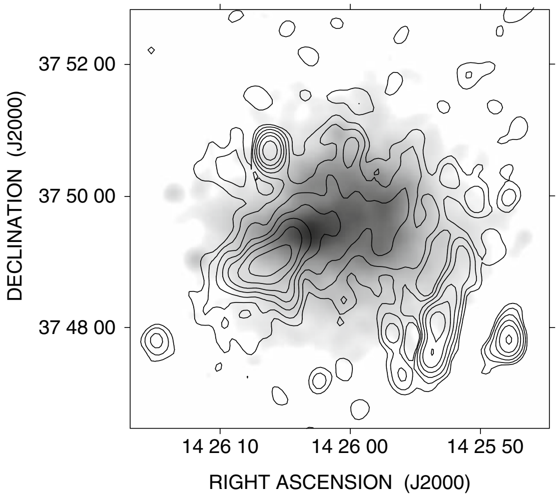

Fig. 2.35 This composite image contains two views of the cluster Abell 1914. Superimposed onto the *Chandra* X-ray emission [211] (in shades of *grey*) are the VLA 1.4 GHz radio emission [18] represented as iso-contours. Courtesy of F. Govoni and M. Murgia

seen in radio with the Very Large Array (VLA) and in X-rays with the *ROSAT*-PSPC and with *Chandra*, for two clusters, Abell 2255 and Abell 1914.

The total mass of a cluster is dominated by the non-baryonic component (the invisible DM) that we know exists because of its gravitational pull on the luminous matter. While the baryonic component of a cluster can be directly observed at optical and X-ray wavelengths, the invisible DM can only be measured through the effect of gravitational lensing on the background galaxies or observing other dynamical manifestations of the clusters. Roughly, it is estimated that the composition of a cluster is 3% galaxies, 17% ICM, and 80% DM. Thus the total mass of a cluster, which is the property we need to know to use clusters as cosmological tools, is dominated by the invisible, collisionless DM.

2.12.2 Clusters of Galaxies in X-Rays

Observations of galaxy clusters in the X-ray band trace the intracluster gas, and hence provide an efficient and physically meaningful method for the identification and selection of galaxy clusters. Over the past decade, studies based on the current generation of X-ray satellites (*Chandra* and XMM-*Newton*) have completely changed our X-ray view of galaxy clusters. The large collecting area of XMM-*Newton*, combined with the very fine angular resolution of *Chandra*, have contributed to unveiling the complex structure and physics of the hot ICM.

The physics of X-ray emission from clusters of galaxies is pretty straightforward. Simple gravitational processes dominate cluster formation and evolution and imply that clusters are still forming today. The evolution of clusters is simple, being driven by the gravity of the underlying mass density field of the Universe and of a collisionless collapse of the cluster DM. These same formation processes also heat gas trapped by the cluster potential, which then produces optically thin thermal radiation. The evolution of cluster X-ray emission can be more reliably calculated compared to that of other objects visible at cosmological distances, such as galaxies and quasars, and the cluster evolution calculations may be verified by direct observations of nearby objects. Thus observations of the X-ray evolution of clusters provide a robust measure of the evolution of cosmic structure and therefore constrain the cosmology of the Universe.

The advent of X-ray imaging in the 1980s revealed that clusters are extended and powerful sources, with luminosities up to $10^{45}\,\mathrm{erg\,s^{-1}}$ for the richest clusters, that emit by optically thin thermal bremsstrahlung from hot ($\sim 10^8$ K), low-density ($\sim 10^{-3}$ atoms cm^{-3}) gas. Their total masses are in the range from a few times $10^{13}\,\mathrm{M_\odot}$ for the poorest groups to more than $10^{15}\,\mathrm{M_\odot}$ for the most massive clusters. In the X-ray, sky clusters appear as high contrast objects, given the dependence of the X-ray emission on the square of the gas density, and can be seen up to high redshift. In addition, the X-ray luminosity, $\mathrm{L_X}$, correlates well with the cluster mass, the cluster property most directly related to cosmological parameters (even though, as I will discuss later, $\mathrm{L_X}$ is not the most accurate of all proposed X-ray indicators for the total mass of a cluster).

Since the early 1990s searches for clusters in X-ray surveys discovered many bound systems out to cosmologically interesting distances (see the pioneering work by Gioia and collaborators with the *Einstein Observatory* [196, 197, 238], and the many X-ray surveys that came out later with the *ROSAT*-PSPC detector (cf. among others [84, 198, 474, 574]). X-ray selection has the unique advantage of revealing physical objects, deep potential wells in the case of clusters, thus avoiding the problem of contamination by foreground galaxies and stars as can happen with optical selection. This is a fundamental point, especially when one deals with very distant clusters, which are the main players in cosmological studies. An additional fundamental advantage of X-ray selection is the ability to define flux-limited samples with well understood selection functions that allow one to evaluate the volume surveyed and thus lead to a straightforward computation of comoving number densities. Figure 2.36 illustrates the sky coverage of several X-ray surveys carried out over the last two decades. Completeness is an important quantity in observational cosmology. A well defined and complete sample is designed to detect all objects with luminosity (or any other cluster quantity) above a given value and within a given redshift, and thus it can be reliably used for cosmological studies.

However, the most important cluster parameter, its mass, is not directly observable. So observers generally proceed by using some other observable like X-ray luminosity or temperature as a surrogate for cluster mass and linking that observable with mass through a simple scaling relation. Numerical simulations of cluster formation indicate that these relations can be quite accurate (e.g., [160] and [77] show in simulations without radiative cooling or star formation, that cluster temperature tracks cluster mass to within about 15%; see among others also [159, 294, 367] for simulations with cooling and star formation). Several proxies of the total cluster mass have been proposed based on cluster observables such as galaxy velocity

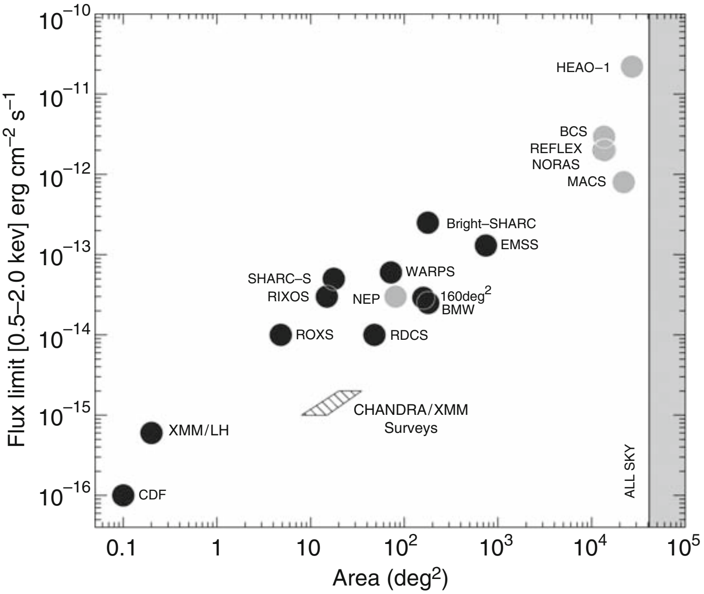

Fig. 2.36 Solid angles and flux-limits of X-ray surveys carried out over the last two decades. *Dark filled circles* represent serendipitous surveys constructed from a collection of pointed observations. *Light shaded circles* represent surveys covering contiguous area. The hatched region is a predicted locus of serendipitous surveys with *Chandra* and XMM-*Newton*. From [475]

dispersion [200], optical light [204, 426], mass of the ICM (many papers by many authors), Sunyaev-Ze'ldovich decrement [90], gravitational lensing [180, 264, 563] (see [580] for description of all these methods). Two easy-to-obtain X-ray observables are X-ray luminosity and X-ray gas temperature, which are both found to correlate more tightly with the cluster virial mass than other cluster properties, like for instance optical richness.

The cluster X-ray luminosity is the most straightforward mass indicator to measure observationally as a minimum number of X-ray photons is required. However, as most of the luminosity comes from the dense central region of the clusters (the radius of the core is much smaller than the virial radius), L_X is the least accurate of all proposed X-ray indicators for the total mass, given the large scatter and deviations of the slope of the luminosity-mass, L_X-M, relation [443] from self-similar model predictions[6]. One way to calibrate the L_X-M relation is to combine the M–T_X relation (whose scatter has been found to be considerably smaller, see, for example, [11, 176, 576]) with the observed L_X-T_X relation. Observational studies

[6] Those ICM models whose physics is based on the assumption that only gravity determines the thermodynamical properties of the hot diffuse gas are called self-similar models [263]. In such models clusters of different sizes are expected to be scaled version of each other as gravity does not have a preferred scale.

have found that the slope of the L_X-T_X relation is steeper than self-similar predictions (e.g., [10, 339]), and the entropy in cluster cores is higher than that predicted (e.g., [423]), indicating important non-gravitational effects (such as cooling, mergers, etc.) on the energy budget of clusters. One has to pay attention also to the evolution of the mass-observable relations. For instance, Branchesi and collaborators [71, 72] find a significant evolution in the L_X-T of a sample of 40 archival *Chandra* and XMM-*Newton* clusters, similar or stronger than the self-similar model, from $z = 0$ to $z \leq 0.3$, followed by a much weaker, if any, evolution at higher redshifts (see also [350]). The higher-z weaker evolution seems compatible with an increasing importance at high redshift of non-gravitational effects in the structure formation process (e.g. [579, 580]).

The X-ray temperature of the ICM [236, 237, 384] is another common indicator for mass. The X-ray temperature is closely related to the depth of a cluster potential well and can be observed with current X-ray detectors up to $z \sim 1$ and beyond. Under the assumptions of hydrostatic equilibrium and isothermality (simplifying assumptions that are not necessarily true in reality), one can derive the total mass in X-rays by knowing the baryon density from the X-ray surface brightness and the temperature of the hot gas. These two quantities are readily available today with the detectors onboard *Chandra* and XMM-*Newton* satellites, which can measure both simultaneously. The masses obtained in this way are very close to those obtained through the virial theorem namely $T \propto M^{2/3}$. It is worth mentioning here that the very accurate temperature profiles out to large radii now provided by the current X-ray telescopes have actually allowed Vikhlinin and collaborators [576] to relax the assumption of isothermality. They have used the best available *Chandra* observations for 13 low-redshift clusters and made direct hydrostatic mass estimates from the gas temperature profiles.

In the recent past, several authors have used the cluster baryon mass as a proxy for the total mass, thus avoiding all the uncertainties of the M–T_X and M_{tot}–L_X relations [6, 158, 575, 578]. The advantage is that it can be measured from X-ray imaging alone and is a robust and complementary indicator to the others for constraining cosmological parameters. An additional recently proposed [295] mass indicator is defined as the product of the X-ray derived gas mass and the average temperature, $Y_X = M_g T_X$, that strongly correlates with cluster mass with only $5 \div 8\%$ intrinsic scatter. However, non-gravitational processes can potentially alter the mass–temperature relation, the baryon-to-DM ratio of clusters, and the redshift evolution of both these quantities. Maughan [349] followed up on this and found from the $L_X - Y_X$ relation for 115 *Chandra* clusters that the X-ray luminosity is a robust, low-scatter mass proxy.

To wrap up this part, I would say that X-ray is a fundamental band to identify and characterize galaxy clusters. Current X-ray telescopes show us the very detailed fine structure of cluster emission up to distant redshifts (see for instance Fig. 2.37 for an XMM image of one of the most distant serendipitously selected X-ray clusters at $z = 1.4$, [370]) something unthinkable until a decade ago. The many X-ray surveys from previous missions, either serendipitous or all-sky surveys, have been demonstrated to be promising tools for the characterization of the properties of galaxy

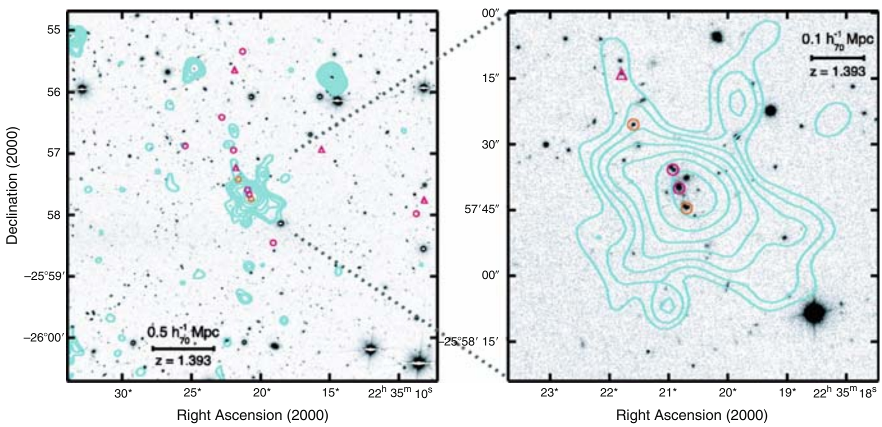

Fig. 2.37 Galaxy cluster XMMU J2235.32557, at $z = 1.393$. *Left*: VLT FORS2 R-band image overlaid with X-ray contours from a 45 ks XMM-*Newton* observation. *Right*: VLT-ISAAC Ks image overlaid with the same X-ray contours. Spectroscopically confirmed members ($1.38 < z < 1.4$) are indicated as *circles* or *triangles*. From [370]

clusters. I believe that we can get even more stringent constraints on cosmological parameters as more sensitive and statistically significant surveys made with the current telescopes become available.

2.12.3 Clusters of Galaxies as Cosmological Tools

Clusters of galaxies are the highest peaks in a cosmic terrain driven by gravitational clustering and represent the largest scale of fully collapsed structures in the Universe [391, 394]. Thus they offer a unique insight into the formation of structures and into the parameters governing their evolution. The internal mix of components within clusters, as well as the space density and temperature distribution function of the most distant and massive clusters, can be used to determine fundamental cosmological parameters. Other cluster measurements useful for cosmological studies include the power spectrum of the three-dimensional distribution of clusters, and the baryon fraction and its evolution. These studies have been carried out by a number of authors over the years; among them see, for instance, [6, 61, 62, 152, 158, 235, 236, 407, 475, 575, 578], which is not a complete list. These works have used the mass function as given by [428] or [512] or by Jenkins [257]. The values of the mean mass density, Ω_m, and DE density, Ω_Λ, of the Universe are fundamental data for cosmological theories. These quantities are conveniently parametrized in terms of the critical density, $\rho_0 = 3H_0^2/(8\pi G)$.

The growth rate of the density perturbations depends primarily on Ω_m and, to a lesser extent, on Ω_Λ at least out to $z \approx 1$, where we can study clusters observationally. The abundance of rich clusters of galaxies is extremely sensitive to the

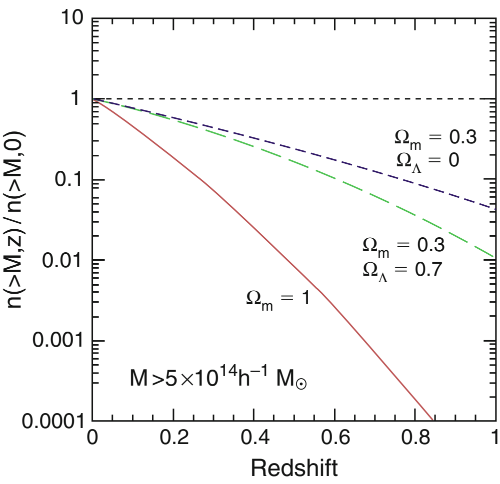

Fig. 2.38 Evolution of $n(>\mathrm{M}, z)$ for $\mathrm{M} > 5 \times 10^{14}$ h^{-1} $\mathrm{M}_\odot$ for three cosmologies (*solid line*, $\Omega_m = 1$; *long-dashed line*, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$; *short-dashed line*, $\Omega_m = 0.3$, $\Omega_\Lambda = 0$) with $\sigma_8 = 0.5$ for the $\Omega_m = 1$ case and $\sigma_8 = 0.8$ for the low-density models. From [475]

amplitude of the mass density fluctuations on a scale of 8 h^{-1} Mpc, or σ_8, while the evolution of the abundance is extremely sensitive to Ω_m and to a lesser extent to Ω_Λ. An additional parameter is the DE parameter w [89], the ratio between the pressure and energy density in the equation of state of the DE component[7]. The value of w is less constrained by clusters.

Figure 2.38 shows the sensitivity of the cluster mass function to cosmological models. Both the X-ray Luminosity Function (XLF) (the number density of galaxy clusters having a given X-ray luminosity) and the X-ray Temperature Function (XTF) (the number density of galaxy clusters having a given temperature) for both nearby and distant clusters have been used as a proxy for the mass function by a number of authors. When only local cluster data are used, there is a degeneracy between σ_8 and Ω_m. See discussion in [61, 62] on how the resulting constraints on the σ_8–Ω_m plane vary by changing the parameters that define the M-$\mathrm{L_X}$ relation. To break this degeneracy, one can either use the evolution of the XLF with redshift, or consider measurements at other spatial scales, such as the fluctuations in the CMB with appropriate assumptions. Many X-ray surveys have shown that the comoving number density of clusters at a given luminosity from $z \sim 0.8$ to the present changes very little for $\mathrm{L_X} \leq 10^{44}$ erg s^{-1}. Evolution is seen only for clusters with $\mathrm{L_X} \geq 10^{45}$ erg s^{-1} (see, among others, e.g., [197, 369, 475] and Fig. 2.39 for a compilation of high-redshift XLFs that highlights evolution). The situation becomes worse when one wants to investigate the DE parameter w. In that case, investigators combine constraints from both SNe and clusters, or weak lensing, the cosmic microwave background and clusters to improve the constraints.

The degeneracy between σ_8 and Ω_m may also be broken by measuring the evolution of the cluster temperature function. The first cosmological measurement using

[7] If $w = -1$, then the DE is the cosmological constant, if $-1 < w < 0$, then it is called "quintessence" or Q component [89].

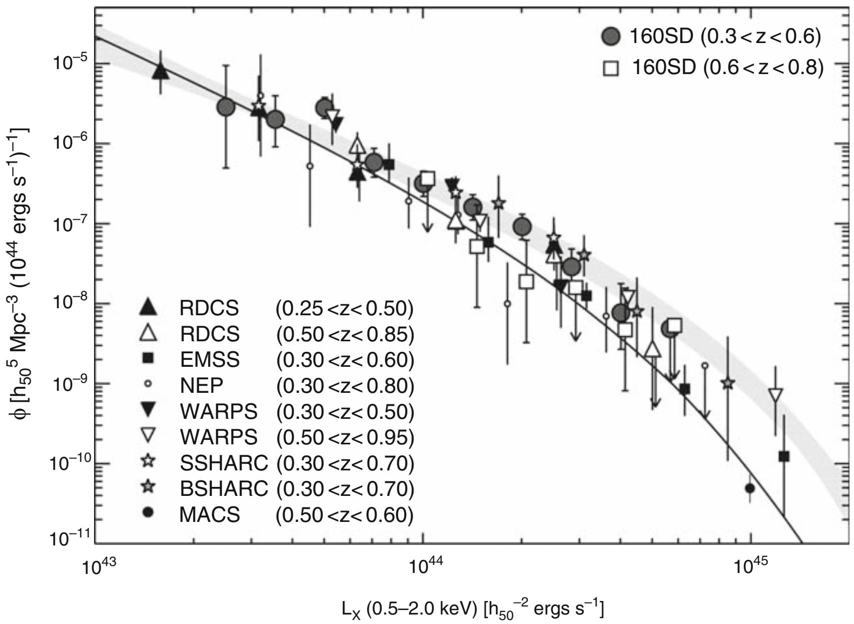

Fig. 2.39 Compilation of high-redshift XLFs as measured by eight independent X-ray flux limited surveys. The *shaded region* delineates the regime of the local XLF, whereas the *solid line* is an evolving model XLF. From [369]

the evolution of the XTF at redshift greater than zero was done by Henry [235] who derived $\Omega_m \approx 0.5 \pm 0.15$. Many updates on both theoretical and observational side followed among others, see [47, 62, 139, 152]. Henry [236] measured the X-ray temperature with ASCA of all but one cluster in the *Einstein* Extended Medium-Sensitivity Survey [196] high-redshift ($z \geq 0.3$) sample and compared the data with a complete sample of low-redshift clusters that also had temperature measurements [237]. Using constraints provided by the SNIa Hubble diagram and the cosmic microwave background fluctuations, he found that all three bands (clusters, SN, CMB) intersect at $\Omega_m \approx 0.3$ and $\Omega_\Lambda \approx 0.7$, with the quintessence equation of state $w = -0.42 \pm 0.21$ and $\sigma_8 = 0.66 \pm 0.16$. The last determination by the same author (Henry, in preparation) can be considered as the state-of-the-art in the field. Figure 2.40 shows the intersect in the Ω_m–σ_8 plane of three bands representing three different clusters analyses. The cluster constraints (dotted line) define a narrow band in the Ω_m–σ_8 plane, which intersects with constraints from the WMAP 5-year data [142] (solid line) and weak lensing data [33] (dashed line). Allen et al. [6] use *Chandra* measurements of the X-ray mass gas fraction for 42 clusters in the range $0.05 < z < 1.1$ to constrain the mean matter density, the DE, and DE parameter w. Combining the X-ray gas fraction f_{gas} with constraints from SN and WMAP 3-year studies and for a flat cosmology, they obtain a tight $\Omega_m = 0.253 \pm 0.021$ and $w = -0.98 \pm 0.07$. Mantz and collaborators [338] derive a precise determination of

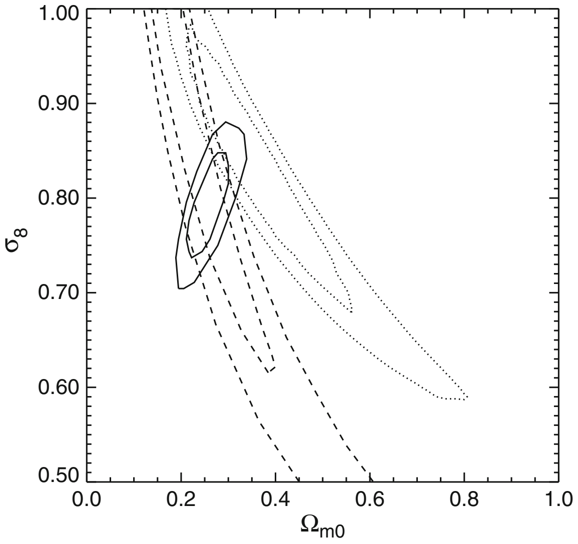

Fig. 2.40 68% and 95% confidence contours for two parameters for three analyses: WMAP 5-year data (*solid line*, [142]), cluster data (*dotted line*; Henry et al., in preparation), and weak-lensing data (*dashed line*; [33]). All constraints intersect at σ_8=0.85, at $\Omega_{m0} = \Omega_m = 0.3$

the DE equation of state combining the X-ray luminosity function data of the most luminous clusters out to $z = 0.7$ with SN, WMAP 3-year, and cluster gas fraction data. They find $\Omega_m = 0.269 \pm 0.016$, $\sigma_8 = 0.82 \pm 0.03$, and $w = -1.02 \pm 0.06$, in agreement with earlier galaxy cluster studies.

This demonstrates that we have already enough information from cluster samples to also constrain the DE content of the Universe, one of the most ambitious targets of modern cosmology. Thus we understand why cosmologists love to work with galaxy clusters. The reason is simple: they are tools for precision cosmology through the evolution of their mass function.

The properties of clusters are investigated through a multifrequency approach: from radio emission and SZE, to IR-Opt-UV mapping, to X-ray emission. For example, a cluster typically appears more extended when mapped through the SZE than through its X-ray emission. Can you discuss the wealth of astrophysical information achievable through the comparison among data at various bands? Which advantages for cosmology come from such multiwavelength approach?

A multi-wavelength approach in any branch of astronomy is of importance as different bands highlight different properties of the emission mechanisms or detect different components of the astronomical objects that contribute to our understanding of their physics, formation, and evolution. The composite image of Abell 520 in Fig. 2.41 highlights the usefulness of the multi-wavelength approach to detect different emission sources from clusters of galaxies [333]. The different waveband measurements are complementary in setting more stringent cosmological

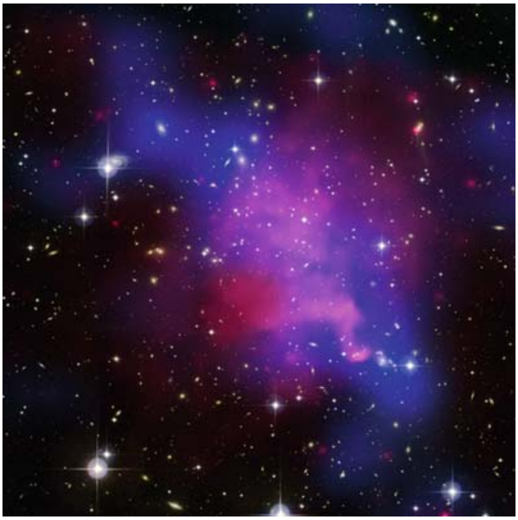

Fig. 2.41 This composite image (which appears in the *Chandra* online Photo Album; see CHANDRA-phot in web page list) contains three views of the cluster Abell 520. The hot gas as detected by *Chandra* is colored *red*. Optical data from the Canada–France–Hawaii and *Subaru* telescopes show the starlight from the individual galaxies (*yellow* and *orange*). The location of most of the matter in the cluster (*blue*) was also found using these telescopes, by means of weak gravitational lensing of the distant galaxies by the intervening matter. From NASA, CXC, CFHT, and University of Victoria. Courtesy of A. Mahadavi and CXC

constraints. X-ray clusters can be used alone to constrain cosmological parameters or can be combined with independent methods (weak-lensing, CMB anisotropies, SZE, SNe, to name a few) and different wavelength data (optical, radio, infrared etc.). As we have seen, the cluster mass is a parameter of great value in observational cosmology. A combination of several, independent cluster mass estimates is likely to provide the most accurate results.

In the optical, the mass-to-light ratio or the mass-richness relation, as well as mass estimates based on the dynamics of member galaxies, have been used by a number of authors with some success [46, 200, 201, 205]. I would like to mention here some of the optical cluster surveys that have overcome the problem of projection effects. I am referring to the work of Gladders and Yee [204] (the Red-Sequence Cluster Survey) who demonstrated that two filter imaging is sufficient to perform a clean cluster search using the cluster red sequence of early-type galaxies, even when probing deeply into the mass function. Zaritsky and collaborators (see Las Campanas Distant Cluster Survey [208]) adopted a different method where clusters are detected as positive surface brightness fluctuations in the background sky.

An unquestionable unique tool to study the matter distribution of the Universe is the use of the weak gravitational lensing of distant galaxies by intervening matter. We have seen in the previous section how the use of weak-lensing coupled with CMB and X-ray data has led to much more stringent constraints on Ω_m and σ_8. Weak lensing has benefitted from the excellent optical surveys currently available with multi-color data and superb image quality over wide areas [33, 106, 185]. The larger areas enable the measurement of the lensing signal out to much larger radii, thus improving the reliability of the results [243]. See Hoekstra and collaborators [244] for a review on weak gravitational lensing.

In the radio band, the pioneering work of Feretti and collaborators [165, 166, 168, 203, 210, 211] have unveiled large diffuse cluster components in the ICM due to synchrotron radio emission not directly related to the cluster galaxies. The study of these sources (called radio halos, relics, and mini-halos according to their size, shape, and location with respect to the cluster center) is very important as they are large scale features that are related to other cluster properties in the optical and X-ray domain, and are thus directly connected to the cluster history and evolution. The radio halos are indicators of cluster mergers, probes of the ICM magnetic fields. They will eventually allow us to constrain models of decaying/annihilating DM species. The radio relics are likely tracers of shock waves during the structure formation. The radio mini-halos are found in the center of clusters with cooling cores and will allow us to investigate the interaction between the relativistic plasma and the thermal plasma at the cluster centers. The future SKA[8] will dramatically improve the knowledge of these sources, thanks to the detection of new objects, and to the detailed studies of their spectra and polarized emission. See also [85] for the contribution of SKA to future CMB spectrum space experiments.

I would like to mention that the nonthermal component in clusters with radio halos has been detected in the hard (above 25 keV) X-ray band due to inverse Compton scattering by relativistic electrons of the CMB photons. Fusco-Femiano and collaborators [187] found it in Coma, A2256, and A754, among other clusters. This is another manifestation of the same relativistic electrons that emit by synchrotron in the radio band. The detection of the nonthermal hard X-rays has enjoyed healthy debate up until now among the different observers.

Another powerful observational tool for cosmology is the SZE (see review by [45, 90]), which is a distortion in the CMB spectrum caused by the CMB photons passing through the hot ICM and inverse Compton scattering off the energetic electrons. The effect is insensitive to the redshift of the clusters, thus making the method well suited for studies of clusters at high redshift where the abundance of galaxy clusters critically depends on the underlying cosmology. While the thermal SZE is a function of the electron number density, n_e, the X-ray emission scales as n_e^2. Thus clusters are more extended when mapped in SZE than in X-rays (see Fig. 2.42 for X-ray and SZE maps of three distant clusters). The different dependence on the gas density enables a determination of the direct distance to the galaxy cluster that is independent of the extragalactic distance ladder, up to high-z clusters. The great merit of SZE is that combined with other observational diagnostics of clusters (X-ray emission, weak and strong lensing, galaxy velocity dispersion measurements) can provide a measure of the basic cosmological parameters like for instance the Hubble constant.

[8] The radio telescope will have an effective collecting area more than 30 times greater than the largest current telescope. "Galaxy evolution, cosmology and DE" is one of five projects identified by the radio astronomy community as being the key science drivers for the SKA.

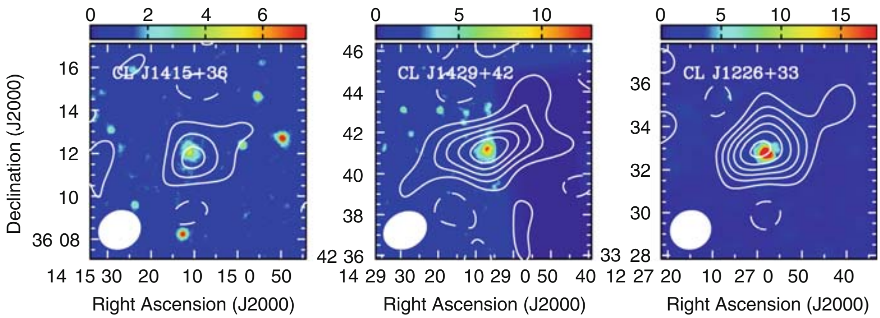

Fig. 2.42 SZE effect measurements (contours) overlaid on XMM-*Newton* X-ray images of Cl J1415.1+3612 ($z = 1.03$), Cl J1429.0+4241 ($z = 0.92$), and Cl J1226.9+3332 ($z = 0.89$) (from left to right). The SZE observations were obtained at 30 GHz during the commissioning period of the new, eight-element interferometer SZA. In each panel, the FWHM of the synthesized beam of the SZE observations is shown by the filled ellipse in the bottom left corner. From [368]

Recently, Bonamente et al. [52] used 38 clusters with *Chandra* and SZE data to find a value for $H_0 = 76.9^{+3.9}_{-3.4}$ (68% confidence level) with $\Omega_m = 0.3$ and $\Omega_\lambda = 0.7$ cosmology in agreement with result obtained by the HST for clusters at low redshift, and with the ΛCDM Concordance Model.

The cosmic soft excess is a phenomenon exhibited by a fraction of clusters ($\approx 30 \div 40\%$) in the extreme ultraviolet or in the soft (1 keV) X-ray band [319]. Since its discovery, the properties and origin of the cosmic soft excess have been subject of debate. The cosmic soft excess has been detected by the Extreme Ultraviolet Explorer and by several X-ray telescopes, including the current XMM-*Newton*. Both the thermal and nonthermal interpretation on the cosmic soft excess origin have been considered and the issue is still under study.

In conclusion, the wealth of astrophysical information currently available on galaxy clusters can give us a deeper understanding of the Universe we live in. We have entered in a promising era for cosmology with clusters. Today, scientists are adopting the multi-frequency approach to carry out cosmological studies as each wavelength contributes a little piece of information that makes sense once the whole puzzle is assembled. We have come a long way since when astronomers were looking for overdensities of galaxies to discover clusters! X-ray observations, optical and infrared observations of the cluster member galaxies and weak lensing of background galaxies by the deep cluster potential are complementary probes of high-redshift clusters. Measurements of the SZE have been used to determine cluster properties such as the gas and total masses, electron temperatures, as well as to constrain the cosmological distance scale [52, 221, 309].

The interested reader can also refer to [59, 60, 396, 556] for further discussions on cosmology with clusters.

Thank you Isabella.

We will now enter in another big open question of modern cosmology: the DM. After more than 20 years of researches, we have not been able to identify the culprit *of the missing mass, although many suspects do exist now for the so-called weak interacting particles. A lot of hopes have been put in the LHC, which is expected to provide fundamental results in a couple of years.*

Now Simon White will enter this subject mainly from a historical perspective, telling us his point of view on DM.

2.13 Dark Matter in Modern Cosmology

Dear Simon *(White)*, **DM is still a mystery after 20 years of astronomical research. The most convincing evidence of its existence comes from the observations that luminous objects (stars, gas clouds, globular clusters, galaxies) move faster that they should if they were only felt by Newtonian gravity. Would you like to express your point of view on this fact, by reviewing briefly this subject?**

The need for very substantial amounts of unseen, apparently DM was first realized more than 75 years ago by the Swiss astronomer, Fritz Zwicky. In his 1933 paper, Zwicky pointed out that the motions of galaxies inside the well-known and very populous Coma cluster of galaxies are so large that the mass required to confine the galaxies within the observed system is many times larger than the total mass of all the visible stars. This deduction has stood the test of time, and present observations suggest that stars only account for a few percent of the mass in clusters according to either Newton's or Einstein's theory of gravity.

X-ray telescopes showed in the 1970s and 1980s that a significant part of the "missing mass" is in the form of a diffuse hot plasma that fills intergalactic space in clusters like Coma. This X-ray emitting IGM contains several times as much mass as the galaxies themselves, but according to current data these two observable "baryonic components" (i.e. components built out of protons, neutrons, and electrons like all familiar materials) together account for only about 15% of the total mass of the cluster. Thus 85% remains unseen.

In the 1970s, it gradually became accepted that individual galaxies like our own are surrounded by massive halos at least ten times the size of the visible galaxy and containing at least ten times as much mass. Such dark halos are needed to explain the orbital motions of dwarf galaxies and gas clouds, and seem to link naturally to the more massive dark halos in which galaxy clusters are embedded. By the end of the decade, massive dark halos had become part of the standard model for the growth of galaxies and larger structures out of a near-uniform Big Bang. At that time, most authors considered optically faint stars as the most plausible objects to make up these dark halos.

A new idea became dominant in the early 1980s as people investigated the possibility that the DM might be made of a new kind of matter: free, weakly interacting, electrically neutral elementary particles. The most plausible candidate is a neutrino, and recent experiments have demonstrated that at least one of the currently known neutrinos has an astrophysically significant mass. Already by 1983, however, numerical simulations had demonstrated that such neutrinos could be at most a small fraction of the DM. The problem is that normal neutrinos emerge from the hot Big Bang with relatively large thermal motions. These prevent such hot DM from condensing into objects less massive than the largest galaxy clusters, and so give rise to a pattern of structure that is inconsistent with observation.

This set-back led astrophysicists to suggest that the DM might consist of a previously unknown kind of elementary particle that emerges from the early Universe with much smaller thermal motions than neutrinos. Simulations soon demonstrated that such CDM could plausibly give rise to dark halos and large-scale structure similar to those observed, and indeed the first sets of simulations in the early 1980s already included models with DM, DE, and inflationary initial conditions, all the elements of our current concordance cosmology.

The older dynamical evidence for DM has been supplemented in recent years by two new and powerful observational probes. One is gravitational lensing, which measures how the images of distant objects are distorted by the gravitational field of the objects they pass on the way to our telescopes. Lensing is now able to measure the statistical properties of the cosmic mass distribution to high precision, and current results are consistent with the predictions of simulations of structure formation in a CDM Universe. The second probe is based on the statistical properties of fluctuations in the CMB. The measured fluctuations are very weak, so weak, in fact, that they are only consistent with the emergence of galaxies at low redshift if the mass fluctuations from which galaxies and galaxy clusters grew are not fully reflected in the CMB maps. This is indeed the case if the DM is *non-baryonic* (i.e., not made out of quarks), but not otherwise.

Dynamical, lensing, and CMB probes now provide detailed measurements of the cosmic mass distribution over a wide range of scales and cosmic epochs. Generally these agree with our standard Concordance Model and in some cases they are starting to give rather precise tests of it. There are some open questions about structure on small scales; the central structure of small dwarf galaxies appears to differ from the simplest CDM predictions, and only a few of the most massive DM substructures predicted to surround the Milky Way have been detected as dwarf galaxies. To my mind, these are issues that are more likely to be resolved through a better understanding of how galaxies form, than by abandoning our working model for DM. Time will tell.

There is still a small but active community of people investigating the possibility that DM is a chimera, that the phenomenology it is invoked to explain is actually a consequence of a departure of the laws of gravity from those set out by Newton and Einstein. The proposed alternatives have grown substantially in sophistication over the last decade. They are still incomplete, in that they have not yet demonstrated that they can explain how the structure seen in the nearby Universe (through lensing and dynamics observations) grew from that seen at high redshift in the CMB, but they have made significant progress towards this goal. It is disturbing, that although such theories have been around for about 30 years, they have not yet been rigorously excluded.

I think that a truly convincing demonstration that the DM is some new type of free elementary particle can only come through detection of a nongravitational signature. Several possibilities currently look promising. Many plausible candidates are Majorana particles with a small but significant annihilation cross-section. The DM distribution might therefore "glow" at a level that may be detectable with upcoming gamma-ray observatories. The first detection would likely be of the Milky Way's own dark halo, which would produce emission across the whole sky with surface brightness dropping smoothly with increasing distance from the Galactic Center. A number of suggested candidates would also be directly detectable in experiments on Earth, and indeed one Italian team is already claiming a detection which has yet to be confirmed by other groups.

Thank you Simon.
In the following interview, Matthias Steinmetz will review the problems of the CDM scenario, focusing on the more controversial aspects recently discussed in the literature.

2.13.1 Issues of the CDM Scenario

Dear Matthias *(Steinmetz)*, **in spite of the successful matching of CDM simulations with the observed large scale density distribution, there are still problems for this theory at small scales. They may be summarized by the following items:**

- **Cusps: is there too much DM in halo galaxy centres? Do we need DM around elliptical galaxies?**
- **Halo substructure issues: the number of sub-halos greatly exceed the observed number of Milky Way satellites.**
- **Angular momentum issues: the baryonic matter has an angular momentum distribution very different from that predicted for DM halos.**
- **Halo and galaxy merging history.**
- **The CDM theory predicts an older age for smaller galaxies contrary to what seems now established by observations.**

Can you discuss these problems? Could the CDM theory overcome such difficulties?

The past decade has witnessed tremendous advances in our understanding of the processes that seed structure in the Universe and shape its evolution in time. Thanks to a dramatic development in our technical capabilities in measuring and surveying the sky, we now have a solid "panchromatic" view of the Universe covering – still with considerable gaps – the epochs from recombination ($z \approx 1{,}100$) to the present day. This progress was accompanied by similar progress in developing concepts on how structure forms in the Universe. Guided by the idea that the mass content of the Universe is dominated by DM, cosmological models based on the paradigm of an inflationary Universe dominated by CDM have proved remarkably successful at accounting for observations. The free parameters of astrophysical relevance in this modeling are surprisingly few: the current rate of universal expansion, H_0; the mass density parameter, Ω_m; the value of the cosmological constant, Ω_Λ, or, alternatively, an equation of state characterizing a DE field; the primordial baryon abundance, Ω_b; and the overall normalization of the power spectrum of initial density fluctuations, typically characterized by σ_8, the present-day rms mass fluctuations on spheres of radius $8h^{-1}$.

As this model was introduced in the early 1980s, the values of these parameters have been revised and tuned to match an ever growing list of observational constraints. In its current form as the so-called Concordance Model, or ΛCDM for short, it is indeed consistent with an impressive array of well-established fundamental observations (as discussed in many contributions in this volume), and future observations are likely to further constrain these parameters. Note, however, that most the successes listed above are on large ($\gtrsim$1 Mpc) scales, that is, scales that are close to the linear regime, where theoretical predictions are straightforward and comparison with observation is more direct. Once the main parameters of the model are fixed, the amplitude and shape of fluctuations on smaller scales follows directly from the theory, a fact that allows for the validity of the scenario to be directly scrutinized on kpc, that is, galactic, scales, where a wealth of observational data exists. And indeed, these are the scales where controversy on the validity of the ΛCDM paradigm exist, as should be discussed in the following sections.

2.13.1.1 The Cusp Problem: Is There Too Much DM in Halo Galaxy Centers?

The flat rotation curves of observed spiral galaxies provide very strong evidence for substantial amounts of DM in the outer regions of galaxies. The total amount of DM amounts to at least 5–10 times the visible mass. Also the escape velocity of stars in the Milky Way, as recently measured by large kinematic surveys like RAVE and SEGUE [518, 595], is consistent with that conclusion. However, in a number of galaxies where detailed stellar and gas dynamical observations are possible, the actual DM contribution within the optical radius appears rather small: many disks, including, it seems, also the Milky Way, are "maximal". Depending on assumptions on the rotation velocity and the Sun's distance to the Galactic center, there may be actually very little room for substantial amounts of DM in our Galaxy within the solar circle ([44, 153, 373]).

For galaxies other than the Milky Way, the problem is echoed by notorious problems of semi-analytical models to simultaneously fit the luminosity function of galaxies and the Tully–Fisher relation. Possible solutions either requires halos considerably less concentrated than predicted by the ΛCDM model or – according to our current understanding – rather exotic solutions, for example, that dark halos do not contract while considerable amounts of baryons accumulate at their centers [144].

The long and controversial debate, whether the density profiles of DM halos in low surface density galaxies have cores or cusps is a related issue. Because of its negligible primordial velocity dispersion, cold DM particles can achieve enormous phase space densities. As a result, numerical simulations have consistently shown that near the centers of halos, the density profiles of virialized CDM halos diverge. These divergent profiles are at odds with the usual interpretation given to the "solid-body" HI rotation curves reported for some low surface brightness dwarf galaxies. Because low surface brightness galaxies are systems where the baryonic contribution to the gravitational potential is almost negligible, solid-body rotation in these systems rather implies a constant density core. While the most recent generation of high resolution simulations report density profiles of slope shallower than $n = -1$, they are still too steep to be consistent with the measured HI rotation curves.

However, the detailed comparison between observational data and model predictions has been plagued by systematic problems on the observational (e.g., beam smearing, estimate of the HI surface density), as well as on the simulation side (e.g., resolution effects), and, in particular, comparing the two against each other. While the systematic effects in observation and theoretical modeling could largely be quantified, uncertainties in the comparison persist, for example, to properly account for triaxial halos, or to incorporate the intrinsic scatter of DM halo profiles compared to the averaging fitting formula.

Among all the reported problems of the ΛCDM model on small scales, I see the apparent lack of DM (if compared with the ΛCDM predictions) near the centers of galaxies to be the most uncomfortable one. Comparisons based on extended 2D velocity maps (e.g., using integral field spectrographs) may provide an interesting route to settle remaining systematic uncertainties.

2.13.1.2 Halo Substructure Issues: The Number of Sub-halos Greatly Exceed the Observed Number of Milky Way Satellites

As of a few years ago, about 11 satellite systems of the Milky Way were known. High resolution N-body simulations indicate that, if the dark halo of our own Milky Way is made up of CDM, it should contain at least several hundred DM sub-condensations [270, 280, 365]. Should most of these dark sub-condensations harbor luminous galaxies with velocity dispersions comparable to those of their surrounding halos, this would result in a considerable, if not dramatic, mismatch.

However, owing to a number of developments, the halos substructure issue is looking less critical than a 10 years ago:

- By systematically surveying $\sim$7,500 square degrees of the night sky, the SDSS could identify 10 new satellites for the Milky Way, that is, almost double the number of known satellites. Extrapolated to the full celestial sphere, this implies $\sim$60 potential new satellites for the Milky Way, making the discrepancy between model prediction and observations far less severe.
- A critical issue seems to be to properly associate the stellar velocity dispersion, which is measured at the center of dwarf galaxies (typically within a few tens of pc), with the circular velocity of the surrounding DM halo, which is measured

at the virial radius (typically several to several tens of kpc). Owing to the shallower density profile at small radii, the velocity dispersion should be a factor 2–3 smaller than the virial velocity of the halo [232], resulting in a strongly reduced discrepancy.

- The existence of a photoionizing UV background after $z \sim 6$ is likely to be sufficiently strong to suppress star formation in late forming satellite, further alleviating the problem [78].

Even if some astrophysical mechanism like photoionization can prevent forming luminous galaxies within sub-halos, these may still manifest themselves by perturbing the dynamics of the cold, thin disk of the Galaxy. To quantify the damaging potential of this effect is still under investigation (see, e.g., [137, 179]).

Among all the reported problems of the ΛCDM model on small scales, I see the satellite problem as the least critical, as several promising solutions exist. Most critical in my opinion is to settle the question, to what extend DM substructure is really perturbing galactic disks.

2.13.1.3 Angular Momentum Issues: The Baryonic Matter has an Angular Momentum Distribution Very Different from that Predicted for DM Halos

A difficulty indirectly related to the substructure problem concerns the angular momentum of gaseous disks assembled in hierarchically clustering scenarios. On the global scale, the situation is quite promising: analytical models [162, 362] show that tidal torques exerted by the surrounding mass distribution are inducing just enough angular momentum during the phase before turn-around in order to explain the sizes of the present day population of disk galaxies.

However, this success depends on two critical assumptions, namely that (1) the initial angular momentum distribution of gas is identical to that of the DM and that (2) the subsequent dynamical evolution of the gaseous component proceeds under conservation of angular momentum.

While numerical simulation support the first assumption [510, 567], the second one seems much less obvious. In the absence of heating, most of the mass of a galactic disk formed within a CDM halo is accreted through mergers of proto galaxies whose own gas components have previously collapsed to form centrifugally supported disks. Numerical simulations show that most of the angular momentum of the gas is transferred to the surrounding halos due to dynamical friction during mergers (see Fig. 2.43). As a result, the spin of *gaseous* disks formed by hierarchical mergers are much lower than those of observed spirals. This result also holds, when the numerical resolution is substantially increased compared to earlier studies [284]. Indeed the principle effect that owing to dynamical friction angular momentum is transported from an infalling satellite to the DM halo can also be seen in N-body simulations of much higher resolution.

Energetic feedback due to late stages of stellar evolution ("SN feedback") has often been advocated as a means to prevent gas from collapsing early into the

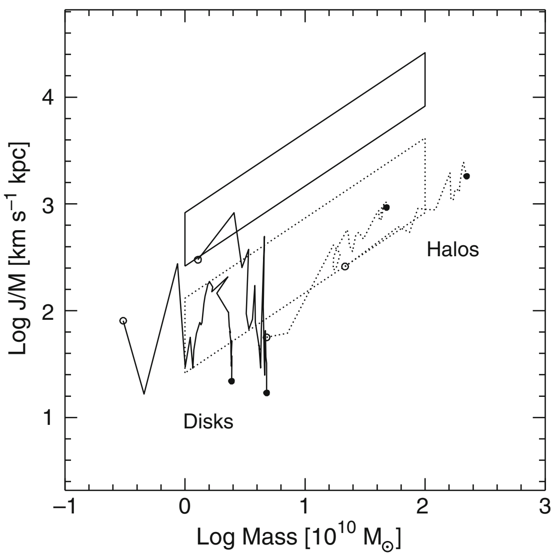

Fig. 2.43 Evolution of the dark halo and central gaseous disk in the J/M vs. M plane, from $z = 5$ (*open circles*) to $z = 0$ (*solid circles*). The evolution of two systems is shown. Note that at $z = 5$, the gas and DM appear to have the same specific angular momentum. The mass of the system grows steadily by mergers, which are accompanied by an increase in the spin of the halo and a decrease in the spin of the central disk. The latter results from angular momentum being transferred from the gas to the halo during mergers. From [372]

progenitor potential wells of present day galaxies, thus preventing large angular momentum losses during the assembly of the galaxy disk. Although plausible, early simulations have not been able to demonstrate directly that this is a viable solution.

Considerable progress has been made, however, in the past few years, thanks to a number of effects (see, e.g., [1, 209, 453]):

- The currently favored ΛCDM model (as opposed to a high-Ω_m standard CDM model as it was used in the simulation work in the 1990s) helps in reproducing disks as it suppresses late time accretion and its accompanying damaging effect. Furthermore, owing to the higher baryon fraction, disks are more self-gravitating and therefore less susceptible to angular momentum transport.
- Many current numerical experiments focused on individual halos that are isolated and have not undergone a major merger within the past 5–10 Gyr. While such halos are not untypical in ΛCDM, they are still to some extent "special".
- Higher numerical resolutions allows more advanced and more realistic methods in comparing the properties of model disks with those of observed spiral galaxies.
- The most important effect is the inclusion of (realistic?) efficient feedback models that prevent gas from excessive cooling. The reliable implementation of these description seem to require also a sufficiently high numerical resolution.

At the moment there seems to be good promise that we can reasonably reproduce the properties of a Sb-type disk galaxy in a high-resolution numerical simulation with advanced feedback models. However, whether this also works to reproduce essentially bulge-less galaxies of Hubble-type Sc/Sd and that we are capable to reproduce the properties of the present day population of disk galaxies (i.e., reproduce the right mix of morphologies) still needs to be shown.

2.13.1.4 Halo and Galaxy Merging History. The CDM Theory Predicts an Older Age for Smaller Galaxies Contrary to What Seems Now Established by Observations

Indeed, the rest band luminosity of galaxies with the highest specific star formation rates seems to decline with redshift, an effect discovered by Cowie et al. [114] that has been reconfirmed by various studies in the recent years[9]. Essentially, it is equivalent to the well known fact that early type and more massive galaxies seem to have the peak of their star formation rate at earlier times than late type galaxies of lower mass. A similar trend seems also to be present in the accretion histories of supermassive BHs (see, e.g., [230]).

At first sight this seems to be contrary to the expectation of the hierarchical galaxy formation model, whose main prediction is that small structures form first, and larger structures form by merging of smaller structures. However, such a statement only holds if one identifies the present day generation of low mass objects with those low mass objects that formed at early epochs, a misconceptions that is contrary to the predictions of the hierarchical clustering paradigm! Indeed, the first generation of low mass objects are merged into larger objects and cease to exist. The first stars formed come from the highest density peaks and can nowadays be found at the centers of the most massive objects (see e.g. [529, 588]). The present generation of low mass objects are formed in lower density environments, that is, they stem from less pronounced peaks in the cosmic density field.

So is there no problem at all? Quite the opposite, but the involved physics is more settled: while the core of a growing high mass object contains a substantial mass in the form of very old stars, this object will continue to accrete matter. In the absence of strong feedback processes, accreted gas would eventually be transformed into stars, resulting in ellipticals who are much bluer than their observed counterpart. However, the mode of gas accretion is changing over cosmic epoch: gas accreted at early times falls in cold while gas accreted at later epochs is shock heated to the virial temperature (see, e.g., Figs. 3 and 4 in [372]). This effect can be translated into a critical, redshift-dependent mass below which gas comes in preferentially in the form of cold clumpy flows, likely subject to rapid star formation, while above that critical mass, gas is accreted as a tenuous hot medium and thus unlikely to form stars [129]. But this effect still seems not to be sufficient to suppress late star formation all together. Gas dynamical simulations reveal that even in the presence of moderate feedback owing to SN explosions, the resulting model elliptical galaxies are relatively blue (see, e.g., [358]). More recent work demonstrates that feedback owing to accretion onto the central supermassive BH [514] as the most promising cure to this problem, an idea that meanwhile has been in various semi-analytical

[9] The often used term "downsizing" for this phenomenon is not quite correct and may indeed be misleading: for galaxies we observe a systematic shift in the mass of actively star forming galaxies, that is, a trend that applies to the properties of the population over time, but not to the individual object. In economy (from which this euphemism is borrowed), however, downsizing refers to the reduction of workforce of a particular (existing) company rather than the trend to smaller sized companies.

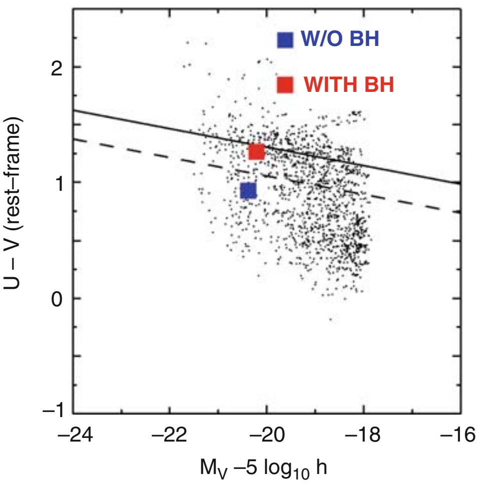

Fig. 2.44 The U-V color and the V-band magnitude of the central galaxy at $z = 0$ in the simulation without (*blue symbol*) and the simulation with (*red symbol*) AGN feedback (from [274]). The results of the simulations have been superimposed on a synthesized U-V vs. M_V Color–Magnitude Diagram (CMD) for a volume-weighted sample of $14.5 < r < 16.5$ SDSS galaxies, which we have reproduced from Bell et al. [29]

(see, e.g., [116]) and numerical approaches (see, e.g., [135]). Figure 2.44 ([274]) demonstrates how in a high resolution cosmological simulation of a forming elliptical galaxy the inclusion of AGN-feedback causes essentially the complete shut off of any late star formation, resulting in a "reddening" of the model galaxy by $\sim$0.3 mag in U-V at $z = 0$.

I see no critical discrepancy between the model predictions and observations. Key problem is to efficiently suppress late star formation in early type galaxies. The observed star formation history and the predicted mass accretion history in the ΛCDM model are fairly different, and so baryonic processes are required that are applicable to a wide range of physical conditions and that are efficiently operating. First models relying on AGN feedback are very promising; however, a far more thorough understanding on the physics of BH accretion and related feedback processes is needed.

Thank you Matthias.
The existence of DM can be proved also through the deflection it causes in the light rays coming from more distant sources. The last 10 years have seen the rapid expansion of surveys aimed at discovering such "lenses". This is because the lensing phenomenon may provide an unbiased indication of the underlying mass distribution of celestial objects. We will enter into this problem with the aid of Matthias Bartelmann, who will explain the ideas behind lensing and their utility for cosmology.

2.14 Lensing

Dear Matthias *(Bartelmann)*, **first of all, could you clarify the differences between weak and strong lensing? Are both effects useful for observational cosmology? If yes, would you explain in which way and what are the problems associated with such measurements?**

As Albert Einstein has told us in his General Theory of Relativity, the presence of all sorts of matter and energy modifies the structure of space–time. It becomes locally curved, in a way that is perhaps well described by the often used analogy with a rubber sheet. Imagine a rubber sheet stretched into a frame, and place a mass on it, such as a ball of steel. It will create a local sink. Now roll a much lighter marble over the rubber sheet, past the ball of steel: The marble's trajectory, which would have been straight in absence of the ball, will now appear curved towards it. Now imagine repeating this experiment with a hand-full of marbles rolled over the rubber sheet. Those coming closest to the ball will be deflected the most from their straight propagation, those far away will roll in an almost straight line. The marbles rolling past the right side of the ball will be deflected to the left, and reverse on the other side. In this way, the ball of steel will attract the marbles and focus their trajectories. This illustrates how General Relativity explains gravity: mass and energy create sinks in space–time, which curve trajectories of particles going through.

Gravity becomes a property of space–time in this way. This means that all kinds of particles will feel it and follow its pull, material particles, as well as photons. In General Relativity, light is thus deflected by lumps of mass or energy, in much the same way as ordinary, convex glass lenses act in geometrical optics. This illustrates the effect and the name of gravitational lensing: the deflection of light by matter.

In the limit of weak gravitational fields, the description of gravity by GR must approach Newtonian gravity, where the gravitational potential Φ is determined by its field equation, the Poisson equation. Now, in General Relativity, the gravitational potential quantifies the space–time curvature, if gravity is weak and its sources are slow. Specifically, this means that velocities must be very small compared to the speed-of-light c, and the Newtonian potential must be very much smaller than c^2. Then, we can proceed as in geometrical optics, as if the gravitational field was a medium with a refractive index, larger than unity by $2|\Phi|/c^2$. Everything necessary to understand the effects of gravitational lensing in almost all of its astrophysical and cosmological applications can now be derived from this refractive index and Fermat's principle, asserting that light chooses the path with the least or longest travel time.

The angle by which light is deflected derives from the gradient of the refractive index. Neighboring light rays passing a gravitationally lensing mass are deflected by somewhat different amounts. Imagine a circular source being lensed: its left edge will be deflected somewhat more or less than its right edge, typically causing the source to be distorted. Gravitational lensing is astigmatic. It creates a distorted image of the world in its background. What is an annoying feature of an imperfect glass lens is the central reason why lensing is so useful for astrophysics and cosmology. Its astigmatism can reveal the presence of the lensing matter, regardless of whether they shine or not. Lamps in a room seen through a door made of corrugated glass appear as irregular spots of light. Knowing something about how lamps would look through the open door, the structure of the glass can be recovered from the shape of the spots.

Applying lensing to astrophysics and cosmology, we are in much the same situation. We know that there are structures on a continuum of scales from stars and planets to the filaments surrounding the large voids in the Universe with their diameters of more than ten million light-years. The matter condensed in these structures is mostly dark, but must act as gravitational lens for the sources in its background. The astigmatism of gravitational lensing creates coherent distortion patterns around the lenses, which can be measured from the shapes of distant sources in their background.

A thin light bundle of parallel rays passing a mass can be changed by gravitational lensing in three different ways. Its cross section can be deformed, its rays can converge or diverge after the passage, and it can be twisted. The twist is typically very small and can usually be ignored. The deformation of the cross section is caused by the gravitational tidal field and described by a two-component quantity, the shear. Over- or underdense matter in the bundle makes it converge or diverge, and this is characterized by the suitably scaled mass density, called convergence. Shear and convergence are both linear combinations of second derivatives of the gravitational potential. They measure the curvature of the potential in two different ways, and are related to each other through the potential.

Strong and weak gravitational lensing can now be distinguished by the magnitude of these two quantities, convergence and shear. Both are dimensionless because the gravitational potential was scaled by c^2 before, and its derivatives are taken with respect to angular coordinates on the sky. If convergence and shear are small compared to unity, lensing is called weak, otherwise strong.

Weak lensing leads to a mild convergence or divergence of light bundles and distorts formerly circular into elliptical bundle cross sections. Strong lensing typically occurs when the bundle converges so strongly after the lensing event that its cross section shrinks to a point, or when it is distorted so strongly that its cross section is heavily deformed into all kinds of shapes. The bundle may even split into several different bundles. Imagine ideally circular sources. Weak lensing will make them appear as mildly elliptical, slightly enlarged or shrunk images. Strong lensing can create multiple images from individual sources, distort them strongly into extended, more or less curved images, so-called arcs, and magnify them considerably. Strong lensing gives rise to features that are immediately evident: arcs in galaxy clusters (see, e.g., Fig. 2.45), multiple images of single sources around galaxies (see, e.g., Fig. 2.46), ring-shaped or partially ring-like images occur and can usually be easily identified as effects of strong gravitational lensing.

Measurement problems related to strong lensing typically occur when the sources are faint. Strong lensing is a rare phenomenon because sources must appear projected very close to sufficiently massive lenses for it to occur. Strongly lensed sources are typically faint. Often, spectroscopy is necessary to uniquely identify multiple images with a single source. In extended, strongly distorted images of faint sources, it is often hard to identify where the images disappear in the sky background, and therefore their sizes, shapes, and total brightnesses are hard to measure.

Weak lensing occurs everywhere in the sky, but is typically hard to recognize. Most often, the effect is so weak that the ellipses formed from hypothetical circular background sources have major and minor axes differing by a few percent only, but sources are not circular. Rather, they intrinsically have elliptical or more

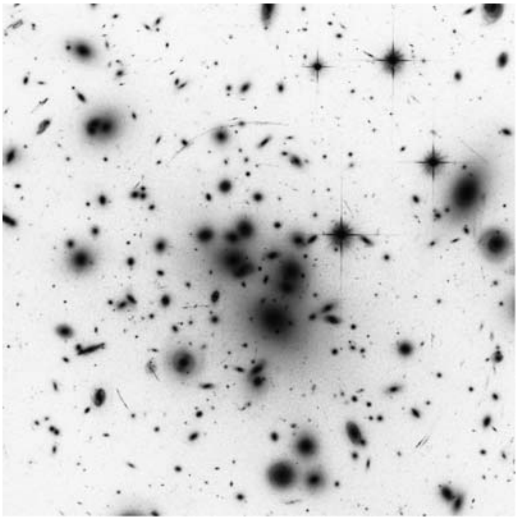

Fig. 2.45 Among the most massive galaxy clusters in our visible Universe is Abell 1689. The image shows many background galaxies distorted into large gravitational arcs, most prominently the one to the top left of the cluster centre. From HST Archive

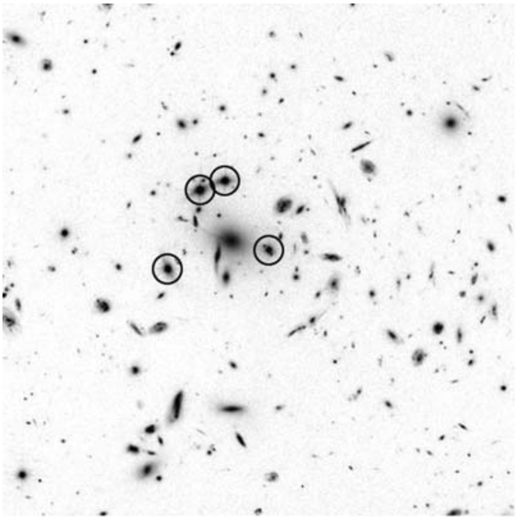

Fig. 2.46 The mass of a galaxy group splits the quasar SDSS J1004 + 4112 into five images, of which four (encircled) are clearly visible on this image. This is so far the most dramatic example of a multiple-image gravitational lens system, with a maximum image separation of 15 arcsec. From HST Archive

complicated shapes, which are then slightly modified by the weak gravitational lensing. Only if many images are superposed, their intrinsic shapes average to circles, and then the weak lensing can be identified as the distortion of the averaged image that is expected to approximate a circle in absence of lensing. Shapes of very many faint images need to be measured precisely, which is a very difficult task in detail.

Could you discuss the cosmological parameters that can be constrained by observations of lensing and how robust is their estimation?

Weak gravitational lensing mainly probes two cosmological parameters, the total matter density and the fluctuation amplitude of density perturbations. This is easy to see. Without matter fluctuations, there were no lensing effects at all. In a homogeneous universe, no lensing effects would be visible, and so fluctuations are needed for lensing effects to occur. The larger the fluctuation amplitude is, the more pronounced the lensing effects will become. As the fluctuation level is measured relative to the mean background density, the same relative fluctuation amplitude means higher fluctuations if the mean density increases. Therefore, the amplitude of cosmological weak lensing effects is proportional to the product of the relative fluctuation level and the mean matter density. This product is the primary cosmological parameter constrained by gravitational lensing by the large-scale structures in the Universe.

This statement is almost but not precisely correct. Like ordinary glass lenses, the efficiency of gravitational lenses depends on where they are put with respect to the source and the observer. They are least efficient when placed right in front of the observer or the source, and most efficient when placed approximately half-way between them. A gravitational lens, characterized by its fluctuation amplitude, will therefore create more or less visible effects depending on its distance. The distance, in turn, depends on the geometry of the Universe, which is characterized by the densities of all forms of matter and energy contained in the Universe. In addition to the simple proportionality to the mean matter density, there is also a weak geometrical dependence of cosmological weak lensing on the mean matter density and possible other forms of energy. Weak lensing can thus not constrain the matter density and the fluctuation level separately, but some combination of these two parameters. Additional knowledge is needed to constrain either one of them.

The dependence of weak lensing on the distance allows another important cosmological constraint. We know mainly from measurements of the temperature fluctuations in the CMB that matter contributes only approximately 30% of the total energy content of the Universe, and observations of a certain class of stellar explosions, the type Ia SNe, indicate that the missing 70% could be in the form of Einstein's cosmological constant, or something with similar behavior called the DE. We have at best speculative ideas as to what this DE may be. It may reveal itself by its influence on structure growth in the Universe. For this to be exploited, accurate measurements are required of how the amplitude of cosmic structures grows over time.

The distance dependence of gravitational lensing enables constraints on the cosmic structure growth. More distant sources are most efficiently lensed by more distant lenses. Separating sources by distance and analyzing their weak distortions separately reveals how the lensing efficiency of the intervening structures grew, which gives us indirect information on the behavior of the DE.

Strong lensing depends much more indirectly on the cosmological parameters, mainly because it is much more sensitive to the details of the mass distribution inside the lenses than weak lensing. Cosmological information from strong lensing is mainly derived from counting strong-lensing events. They come in essentially two flavors, multiply or strongly distorted images. It requires sufficiently many gravitational lenses in the right distance range, with enough mass concentrated strongly enough to reproduce the observed number of strongly gravitationally lensed objects.

The expected number of such suitable lensing objects depends first and foremost on the same parameters as weak cosmological lensing, namely on the mean matter density and the amplitude of the density fluctuations. But again, there is also an interesting dependence on structure growth and therefore on the behavior of the DE through its effect on how quickly structures can grow.

This type of measurement is promising in particular, because the number of gravitational lenses as massive as galaxy clusters depends exponentially on the matter-fluctuation amplitude. Changing it by 20% may change the number of strong-lensing events by an order of magnitude. On the other hand, strong lensing depends equally sensitively on the details of the internal structure of the lensing masses, its possible substructures, and its deviations from spherical symmetry. Also, the number of background sources needs to be accurately known for predictions of the expected number of strong-lensing events. And finally, since strongly lensed images are so rare, large portions of the sky need to be imaged with sufficiently long exposure times to obtain reliable number counts the theoretical predictions can be compared with.

While cosmological constraints from weak lensing have already been obtained with impressively small uncertainty, it remains to be shown what cosmological information we shall be able to infer from strong lensing. The capability of both, strong and weak lensing, to reveal the nature of the DE could not be exploited yet, for both observational and theoretical uncertainties. Nonetheless, they are very promising techniques, and much hope is put into them.

Gravitational lensing is now considered a robust method to derive masses. Would you like to summarize the basic idea, the cosmological implications, and the uncertainties associated with such measurements?

Being caused by the presence of massive objects, both variants of the gravitational lensing effect can be used to determine masses. The basic principle is simple. Imagine a strong gravitational lens, giving rise to either a multiply-imaged quasar or a large arc-like image. Now add mass to the lens. It is plausible that the separation between the split images will increase, or that the arc will move away from the center of the lens. Therefore, keeping all other parameters fixed, the angular separation of strongly lensed images from the lens' center must be a measure of the mass contained in the lens.

Geometry plays its role also here. If the lens is at an efficient position, approximately half-way between the source and the observer, then the same amount of mass will have more effect on the image splitting than if the lens is at a more unfavorable position. Therefore, if we wish to infer the mass, we need to know the geometry of the lens system before. This can be achieved by measuring the cosmological redshifts of both the source and the lens, from which we can calculate their distances if we assume a cosmological model. In other words, mass measurements through the image splitting caused by strong lensing require that a cosmological model be assumed and redshifts to the lens and the source be measured.

Suppose we have this information, a cosmological model and the necessary redshifts, and we trust it. Can we then infer the mass in a unique way? Unfortunately, not quite. The answer depends very much on how the lensing mass is distributed. To begin with a simple case, suppose the lens is axially symmetric, that is, its mass density depends only on the projected distance from its centre. Now, let us add some mass, also in an axially symmetric way, but at a radius *larger* than the separation of all lensed images from the lens' center. In other words, we add a ring of mass, concentric with the lens, but including all existing images. The images will remain unchanged because the additional ring of matter will not cause any net deflection in its interior. At best, therefore, strongly lensed images allow the determination of the mass enclosed by themselves, but not at all the total lensing mass. With strongly lensed images usually occurring arcseconds up to about an arcminute away from their lenses' centers, they typically measure a small fraction of the lensing mass.

Accepting that, let us consider the next difficulty. Quite obviously, the increase in the image splitting after adding mass to the lens must depend on where we deposit this additional mass. If we place it all right at the centre of the lens, its effect will be different than if we smoothly spread it out. Even for an ideal, axially symmetric lens, we need to know the radial profile of the mass density before we can convert the image splitting to a mass. We thus have to note that masses enclosed by strongly lensed images can be determined, provided we have reasons to accept axial symmetry and can reasonably guess the mass density profile.

Asymmetric lenses are more difficult to address. Suppose we start from an axially symmetric lens that we stretch somewhat in one direction perpendicular to the line-of-sight. The strongly lensed images will move to some degree, depending on where they were before the stretching. The asymmetry will increase the gravitational tidal field, or the shear, and this will allow less mass to create a similar image splitting as before, or the same mass to split images by a larger amount. Moreover, if the axial symmetry is broken, even mass surrounding the images will influence their locations and their appearance. Thus, only if we can reasonably assess the degree of asymmetry can we hope to infer masses correctly.

But we are not done yet. Lensing is sensitive to all matter along the line-of-sight towards a source and not only to the one dominating galaxy or galaxy cluster where the lensing effects appear. Lensing measures all mass in a cylinder reaching from the observer to the source, enclosed by the lensed images. Along this cylinder, mass is more or less efficient depending on geometry. An additional mass placed into the cylinder and shifted from the observer to the source will have its largest effect approximately half way from the bottom to the top of the cylinder. Therefore, lensing masses are always masses within cylinders, in which all contributions are summed according to their geometrical weight.

This sketches mass determinations based on strong lensing. Weak lensing provides another method because measured image distortions are caused by the shear. Its components are determined by linear combinations of second-order derivatives of the lensing potential, whose Laplacian is proportional to the surface-mass density of the lens. Therefore, measurable image distortions and the surface-mass density are related through the lensing potential. This suggests a conceptually simple and elegant procedure, by which the ellipticity measurements are inverted to obtain the surface-mass density.

This method has its own difficulties. First, the lens inversion returns the convergence rather than the physical surface-mass density. Converting the convergence to physical units requires the geometry of the lens system to be known, that is, the redshifts of lens and sources and the cosmological model. Even if these data are given or can reasonably be guessed, the lens inversion is not unique because an infinite class of transformations can be applied to the lens without changing its weak-lensing effects. Infinitesimal transformations of that sort correspond to replacing the lens by the lens placed on a sheet of constant matter density, hence this class of transformations has been called the mass-sheet degeneracy.

Even though the mass-sheet degeneracy prevents the unique measurement of lens masses, density profiles can be uniquely determined because they only need differential information from one radius to the next. However, the statement made above remains true that lensing can measure only total masses or mass densities projected within cylinders along the line-of-sight. How much of this mass can be attributed to an individual objects remains uncertain.

The interested reader can refer to [502] for a textbook, to [283,371,500,501,583] for lectures, and to [26, 354, 440] for reviews on weak gravitational lensing.

Thank you Matthias. We now turn around our discussion, with the aim of highlighting the cosmological importance of normal stars.
The evolution of stars is quite well understood from a theoretical point of view. This means that stars are natural clocks for measuring the evolution of the Universe. In the following interview, Cesare Chiosi will remind us these nice properties of stars related to their physical evolution and their utility to constrain the age of the Universe.

2.15 Constraints on the Universe Age from Stellar Evolution

Dear Cesare *(Chiosi)*, **we all learned in astrophysical courses that stars are powerful cosmic clocks. Indeed the Universe cannot be younger than the estimated ages of stars. So, stars in these terms provide very important tests for any cosmological hypothesis. Can you please summarize why stars are so important for cosmology? How stellar evolution can constrain the cosmological parameters that describe our Universe?**

To answer your questions, here I briefly discuss the problem of the age of the oldest stars in different astrophysical environments, going from the very first generations at high red-shift, the globular clusters, the open clusters, the Galactic bulge and the field. Most of the discussion refers to the situation in which Color-Magnitude Diagrams (CMDs) for the various stellar populations are available. I will also review the case in which the stellar populations cannot be resolved into single stars so that strategies based on integrated properties (colors, line strength indices, etc.) must be adopted to infer the age of the stellar content. The case of elliptical galaxies is examined at some extent in relation to the wider subject of galaxy formation in cosmological context.

2.15.1 The Very First Generation: PopIII

When the very first generation of stars (otherwise known as PopIII) took place out of the primordial gas (hydrogen, helium, traces of light elements, and no metals) is not firmly established nor the detection of these stars has been assessed (see [259] for a recent review). The major problem is with their mass spectrum. If for some reason only stars with mass in excess of about $0.8\,M_\odot$ could form, detecting them now is not possible. Their existence should be inferred by their remnants (BHs, neutron stars, and white dwarfs) and more important the kind of chemical enrichment they could have produced. If there were also stars less massive than the above limit some of them could be around. How to detect them? The expectation is that their spectrum should resemble pure BB emission. There are indeed a handful of objects with extremely low metal content that best candidate themselves as genuine PopIII stars whose traces of heavy elements are probably due to surface contamination by the interstellar medium. These PopIII stars are very old, the age however being confined between the age of the Universe based on WMAP data ([523]) of 13.75 ± 0.16 Gyr and the current estimates for Globular Clusters (see below).

2.15.2 Ages from Star Clusters

2.15.2.1 Galactic Globular Clusters

The revolution in modern photometric techniques and instrumentation dramatically improved the quality and richness of the CMDs of Galactic Globular Clusters (GGCs), thus allowing comparisons with theoretical models for low-mass stars of unprecedented sophistication. As an exhaustive referencing to the impressive list of high quality CMDs is beyond the scope of this review, we limit ourselves to quote the databases of CMDs of GGCs by [229,412,476,493]. There also exists an equally impressive list of theoretical studies for low-mass stars at varying basic parameters: mass, helium abundance Y, metallicity Z (this latter separated into the three components: [CNO/H], [α/H], and [Fe/H], and their relative proportions), mixing length in the outer convective layer, opacities, nuclear reactions rates, mixing process, diffusion processes, equation of state, neutrino energy losses, mass loss by stellar wind, etc. Most of these models are calculated all the way from the main sequence to the latest stages, thus making available homogeneous sets of evolutionary tracks and isochrones. Among the studies presenting extensive grids of stellar models, we recall [40, 42, 53, 113, 133, 140, 202, 408, 409, 570, 597].

To be safely used in the interpretation of the CMDs, the evolutionary models must be tested for accuracy in the input physics and adequacy of the physical assumptions, and finally calibrated using known reference objects (see [188]). In addition to this, accurate transformations from the theoretical HRD (M_{bol}, T_{eff}) into an observational CMD (in many different photometric systems) are required. The adequacy of the physical assumptions (e.g., type of mixing) can be tested by comparing model predictions (e.g., lifetimes) to their observational counterpart (star counts and luminosity functions). It is worth recalling that this comparison is meaningful only if the observations satisfy certain conditions on the minimum number of stars to be sampled, that is, *the fuel consumption theorem* [445].

Determining the age of GGCs is a complex game (see [100]), which requires a knowledge of many parameters, such as the helium content Y, metallicity [M/H], CNO abundance [CNO/H], distance modulus, and reddening together with important physical phenomena such as diffusion, sedimentation, and rotation, and finally the so-called second parameter driving the morphology of the Horizontal Branch (HB). For long time, the extension towards the blue of the HB was considered to depend on the metallicity: the HB gets bluer at decreasing [Fe/H]. Very soon many exceptions were found, thus suggesting that other parameters are at work.

Since GGC stars are too cool to allow direct spectroscopic measurements of the abundance by mass of helium Y, less direct methods are used. A general assumption is that helium abundance in Globular Clusters (GCs) reflects primordial nucleosynthesis as GGCs are among the oldest objects in the Universe. In general, the helium abundance is estimated from the big-bang nucleosynthesis [377, 378, 381, 382]. The classical value was $Y_{\rm p} = 0.235 \pm 0.005$. WMAP data yield $Y_{\rm p} = 0.2482 \pm 0.0003$ [524]. It is generally assumed to be constant throughout the halo clusters, even if helium abundance has been often considered a candidate for the second parameter (see below).

The metallicity is usually referred to the observed abundance [Fe/H] and the nowadays accepted metallicity scales come from Zinn and West [604] or Carretta and Gratton [92]. The majority of GGCs have [Fe/H] values from -1.0 to -2.3 dex with typical uncertainty of 0.15 dex. However, recalling that in metal-poor stars, the abundances of Ne, Mg, Si, and S are significantly enhanced with respect to [Fe/H] [215, 374], this latter alone is not fully representative of the real content of heavy elements.

It is well known that the morphology of the turnoff greatly depends on the abundance of oxygen. The controversy of the oxygen enhancement in cluster stars as measured by [O/Fe] is still far from being solved. Given [O/Fe] $= 0$ for the Sun by definition, the questions are whether [O/Fe] is different for the halo stars and whether it varies with [Fe/H]. Observing giant stars in GGCs, Pilachowski et al. [410] and Gratton and Ortolani [214] obtained [O/Fe] $= 0.25$ regardless of [Fe/H] and [O/Fe] $= 0.40$, respectively. However, the question arises whether this result holds for all the stars in a cluster, main sequence included, or whether it is limited to giants. An enhancement of the oxygen abundance in giant stars resulting from inner processes can be excluded on the basis of stellar evolution theory (see [215]).

The distance scale of GGCs is another topic of strong controversy. Most determinations of the distance modulus reduce to comparing the apparent magnitudes of the RR Lyrae stars or HB stars with the corresponding absolute visual magnitudes. This is a rather complicated subject that cannot be properly summarized here, which ultimately lead to assess whether or not a correlation between M_V(RR) and [Fe/H] exists and tray to fix the slope and zero point of this relation. Owing to the far-reaching implications of the $M_V(\mathrm{RR}) - [\mathrm{Fe/H}]$ relation on the age problem, this topic is a matter of debate. For more details on the subject, the reader should refer to the reviews [100, 446].

Over the years, the situation has much improved (in particular after the advent of Hipparcos). Recent compilations of GGC distances are by [88, 439].

Since there are many GGCs with low color excess ($E_{B-V} \leq 0.1$) spanning a broad range of metallicities (up to [Fe/H] $= -1$), reddening is not a serious problem in finding the intrinsic color of the turnoff.

Under the action of gravity, in low mass stars helium can sink inward relative to hydrogen. This process may affect the age in two ways. First the lower relative central abundance of hydrogen decreases the main sequence lifetime. Second a higher relative hydrogen abundance in the envelope results in a larger radius (lower T_{eff}) without changing the Red Giant Branch (RGB) position. The main sequence turnoff is redder, thus implying that a lower age is required to fit a given cluster (see [96]). The reduction in age is estimated to be about $10 \div 20\%$. Rotation does not seem to affect the age in a significant fashion (see [132]).

For long time, the observation of CMDs showing sequences of virtually undetectable width indicated a uniform abundance of heavy elements within the stars of a particular cluster (two exceptions were known to exist: ω Cen and M22, which showed star to star differences of about 0.5 dex), see [28, 387]. In recent years, even the GGCs that were considered the prototypes of single stellar populations have been found to possess a complex chemical structure [411]. Some of them indeed show at least two zero age main sequences, with the bluer one being more metal and helium rich [359, 413]. The estimated helium abundance can be as high as $\mathrm{Y} = 0.40$ or so. This implies an enormous enrichment rate $\Delta Y/\Delta Z$ and a complicated dynamical and star forming history.

Determining Ages. Given a good CMD, ages can be derived by means of the classical Isochrone Fitting (IF) method, the ΔV method, and the $\Delta(B - V)$ method, each of which suffering significant uncertainty.

In the *Isochrone Method*, all the parameters discussed above are necessary. Therefore, the ages obtained from isochrone fitting are by far the most uncertain (see [188, 446]). Most studies in the past have assumed solar [O/Fe] and have found ages going from $10 \div 12$ Gyr (e.g., [533]). If [O/Fe] varies with [Fe/H], this age range is less clear. If helium diffusion is included, an age reduction of 2 Gyr is possible, as estimated by [96].

The ΔV method rests on the fact that the turnoff magnitude becomes fainter as a cluster evolves, while the HB luminosity is virtually constant. ΔV is the magnitude difference between the turnoff and the HB at the turnoff color. This method is independent of reddening. Furthermore, the magnitude of RR Lyrae stars and turnoffs are likely scarcely dependent on [O/Fe] and helium diffusion. The disadvantage with this method is that not all GGCs possess RR Lyrae stars, and some HBs are not horizontal. Furthermore, the turnoff is almost vertical, which makes uncertain the definition of the turnoff magnitude as well. It requires an assumption for the helium abundance. Finally, there is the effect of the controversial relations M_V(RR)–[Fe/H] and [CNO]–[Fe/H] (see [100]). On the observational side, ΔV does not correlate with [Fe/H], shows some scatter at given metallicity, and the mean value is 3.54 [79]. An overview of the possible alternatives at varying the [CNO]–[Fe/H] relation and of the slope and zero points of the M_V(RR)–[Fe/H] relation is given by [188] to whom we refer. In brief, the solution goes from coevality and Age Metallicity Relationship (AMR) to large age differences (about 4 Gyr according to [486]) and a reasonable AMR (the oldest clusters are metal poor, the young ones metal rich), to even counter-intuition relationships between Y, metallicity, and age. Absolute ages are uncertain because they strongly depend on the zero-point of M_V(RR)–[Fe/H] relationship.

The $\Delta(B-V)$ method is based on the color difference between the turnoff and the base of the RGB [491, 569]. This color difference decreases as the cluster age increases. Assuming that the mixing length used in stellar models is calibrated, the method is independent of distance, reddening, photometric zero point, helium abundance, and, to first order it seems to be insensitive to variations in [Fe/H]. The major uncertainties are with the transformations from $T_{\rm eff}$ to colors, the degree of helium diffusion, and [O/Fe], all of these affecting the turnoff color. $\Delta(B-V)$ is reduced by an increased age, an enhancement in oxygen abundance, and helium diffusion. According to VandenBerg et al. [569], this method is particularly suited to determine relative ages. One of earliest age determinations with this method was by Sarajedini [490]. Using the Yale isochrones of Green and Demarque [217] and assuming [O/Fe] $= 0$, they estimated an age for the oldest clusters of about 18 Gyr and an age range of at least 2.5 Gyr. Enhancement of [O/Fe] and/or helium diffusion would reduce the age by about 2 Gyr (see [430]). As far as the age spread among clusters with similar metallicity is concerned, VandenBerg et al. [569] give the following indication. The most metal-poor clusters ([M/H] $= -2.1$) are uniform in age within 0.5 Gyr; clusters with [M/H] $= -1.6$ are also coeval, though some age spread cannot be excluded; finally the most metal-rich clusters, $[\mathrm{M/H}] \geq -1.3$, appear to encompass a significant range [486]. This indicates that the age spread increases with the metallicity as expected if the collapse of the halo was of prolonged rather than of short duration ($\leq$1 Gyr).

The age spread, and age-metallicity relation, if real, not only could be a solution to the problem of the second parameter controlling the morphology of the CMDs of GGCs, but also constrain the time scale and mechanism of halo formation. Long ago, Searle and Zinn [507] made the hypothesis that age is the second parameter driving the morphology of HBs (the metallicity is the first). Other second parameter candidates, such as Y, [CNO/Fe], or core rotation have been considered that could also account for the observed differences (see [313]) but to date only the age seems to provide an explanation compatible with both the standard theory of stellar evolution and the observed distribution of RR Lyrae stars. There are many pairs of GGCs with nearly identical metal content and different HBs that are ideal laboratories for testing the possibility that the age is the second parameter (see [100] for a review of the subject). In any case, the effect of other second parameters ought to be kept in mind.

Second Parameter. Because an important characteristic of the second parameter phenomenon is its systematic variation with the galactocentric distance, Searle and

Zinn and Lee [313, 507] sought for a global interpretation of the available information correlating [Fe/H], [CNO/Fe], HB type, galactocentric distance, and relative ages of GGCs. In the Lee ([313] and references therein) scenario, very likely the age is the second parameter that has the largest influence in determining the HB morphology, and the clusters in the inner halo ($R_G \leq 8$ kpc) are in the mean several Gyrs older than the outer halo clusters. At the same time, arguments are given that run counter to the hypothesis that helium abundance, core rotation, or [CNO/Fe] abundance are the second parameter. If this interpretation is correct, it lends support to the idea of prolonged phase of Halo formation, possibly involving mergers and accretion of large fragments with independent dynamical and nucleosynthetic histories [311]. It is worth recalling that, as pointed out by Sandage [485], a significant age spread among GCs does not contradict the picture of Galaxy formation suggested long ago by Eggen et al. [148].

Absolute Ages. Although absolute ages are less important from the point of view of interpreting the CMD of GGCs, they are very important in cosmology (set a lower limit to the age of the Universe) in understanding the mechanism of galaxy formation. The above discussion has clarified that the absolute age depends very strongly on the accuracy and adequacy of both observational parameters and stellar models, and explains why over the years the absolute ages have varied a lot. As nowadays, the absolute GC ages are best determined using the ΔV technique and its more recent follow-ups [122], in spite of the fact it requires a calibration of the absolute magnitude RR Lyrae stars. These ages depend upon the distance scale to GGCs. Therefore, the absolute ages are subject to change as soon as the basic parameters are improved; over the years they have changed a lot. For instance, (1) the ages estimated by Sarajedini and King [492] for a selected sample of GGCs show that their distribution peaks at about 16 ± 2 Gyr, with wings going down to 10 Gyr and up to 20 Gyr; (2) in the nineties the best estimate of the GGC absolute ages was $(13 \div 15) \pm 3$ Gyr (see [446]). The first uncertainty is due to stellar models, and the second one to the observational data (mostly the distance). This estimate is hard to reconcile with the cosmological value of the age of the Universe (which however was not yet firmly established at that time). (3) Chaboyer [95] reviewing the constraints on the absolute magnitude of RR Lyrae stars in metal poor stars combined with a detailed study of the uncertainties in the theoretical models lowered the age of the oldest GGCs to 13.2 ± 1.5 Gyr. (4) Recently, De Angeli et al. [122] using the ΔV-method find that all clusters with $[\mathrm{Fe/H}] < -1.7$ are old and coeval, with the possible exception of two objects, which are marginally younger. The age dispersion for the metal-poor clusters is 0.6 Gyr (rms), consistent with a null age dispersion. Intermediate-metallicity clusters ($-1.7 < [\mathrm{Fe/H}] < -0.8$ are on average 1.5 Gyr younger than the metal-poor ones, with an age dispersion of 1.0 Gyr (rms) and a total age range of $\sim$3 Gyr. About 15% of the intermediate-metallicity clusters are coeval with the oldest clusters. All the clusters with $[\mathrm{Fe/H}] > -0.8$ are 1 Gyr younger than the most metal-poor ones, with a relatively small age dispersion, although the metal-rich sample is still too small to allow firmer conclusions. There is no correlation of the cluster age with the galactocentric distance. The reference age is 11.2 Gyr with the Zinn and West [604] metallicity scale and 10.9 with the

Carretta and Gratton [92] scale. Concordance with cosmological determination of the age of Universe is nowadays possible: the current age of GCs is in the range $(10 \div 12) \pm 1$ Gyr.

2.15.2.2 Galactic Open Clusters

The Old Open Clusters (OGC), whose ages range from say 1 to 8–9 Gyr, trace most of the history of the Galactic Disk. Therefore, the correct ranking of OGCs as a function of age, chemical composition, and kinematical properties is of paramount importance to understanding the process of SF in the Galactic Disk. Furthermore, having turnoff masses between $1\,M_\odot$ and $2\,M_\odot$, they are probes of stellar structure in that mass range, in which the transition from radiative to convective cores on the main sequence, from p-p chain to CNO cycle for the core H-burning phase, and from very bright RGBs as in M67 to much less evident RGBs as in the Hyades occur. A review of the subject still valid today is given by [100]. A somewhat old determination of ages for a selected sample of OGCs is by [91]. The ages span from 0.9 Gyr for NGC 2477 to $8 \div 9$ Gyr for NGC 6791. Another attempt is by [109], who derived estimates of the age from the turn-off colors of all the clusters of the Lund list. The ages go from a 0.01 Gyr to a few Gyr, the age distribution peaks at about 0.4 and 6 Gyr and varies from region to region of the Galactic Disk. Another method to derive OGC ages makes use of the White Dwarfs cooling sequences (see [256, 581] and references). An interesting issue is whether there is an age gap between the youngest GGCs and the OGCs. Apparently there is no such gap, but the two populations do not overlap [30].

2.15.2.3 Galactic Bulge, Field Stars, and the Local Group

Galactic Bulge. The dominant stellar population of the central bulge of the Milky Way is old, with roughly solar metallicity [447]. The age is very similar to that of the old metal-rich Bulge GCs and to 47 Tucanae, which is an old object (about 12 Gyr). Stellar composition measurements (Keck/HIRES data) confirm that bulge stars are enhanced in Mg and Ti, which favors a short formation time scale (see below). Finally, HST/NICMOS data indicate that the stellar population within the central 100 pc of the Milky Way is very old (likely the first one formed in the Galaxy).

Galactic field stars. Pont and Eyer [424] studying the classic sample of Edvardsson et al. [146], who derived the AMR of 189 field dwarfs with precisely determined abundances, argued that much of the observed scatter in the AMR is caused by the interplay between the systematic biases affecting the traditional age determination, the color mismatch with the evolution models, and the presence of undetected binaries. Using new parallax, temperature, and metallicity data, the age determination for the same sample indicates that the intrinsic dispersion in the AMR is at most 0.15 dex and probably lower. In particular, they show that the presence of old, metal-rich objects ([Fe/H] ~ 0.0 dex, age > 6 Gyr) and young, metal-poor objects

([Fe/H] ~ -0.5 dex, age $\leq$ 6 Gyr) in many observed AMRs is an artifact of too simple a treatment of the age determination. The incompatibility of those AMRs with the notion of a well-mixed interstellar medium may therefore only be apparent.
Local Group. The galaxies of the Local Group can be easily resolved into stars so that good CMDs are available to decipher the past history of SF and chemical enrichment. But for the two spirals, the remaining dwarf galaxies of different morphological type, going from dwarf irregulars to dwarf ellipticals, show rather complicated SF histories (see [191, 216, 343, 344]) characterized by several bursts of activity. In any case, clusters and field stars as old as those in the Milky Way are found.

2.15.3 Ages from Integrated Properties

As we go beyond the Local Group, stellar populations in clusters, fields, and galaxies as a whole can be hardly resolved into single stars, so that CMDs are not available for their interpretation. We must address to integrated properties, magnitudes, colors, line strength indices, etc. Basically, two techniques to measure metallicity and age are to our disposal: the line strength indices introduced long ago by Faber et al. [161] and used by many authors (see [545, 546] and references), and the continuum magnitudes and colors (see [434, 436] and references).

Single Stellar Populations. Prototypes of Single Stellar Populations (SSPs) are the GCs in the classical view (the possibility that GCs themselves are composite systems is still at its infancy and will not be taken into account here).

Galactic Globular Clusters. Line strength indices in the Lick system have been calculated for GGCs (e.g., [431]). They have been analyzed with theoretical SSPs at varying age, metallicity, and degree of α enhancement (see [431, 548]). The overall agreement between theory and observation is satisfactory. Ages, metallicity, and α-enhancement inferred from the line strength indices are those typical of GCs. The comparison has been pushed to a higher level by deriving the ages, metallicities, and α-enhancement for a sample of GCs for which the same parameters are known from the CMDs [122, 429, 497]. Metallicities, degree of enhancement, and ages derived from indices fully agree with those obtained from the CMDs [548]. This holds true both for individual clusters and the mean values for the whole sample: $\langle \mathrm{Age} \rangle = 11.2 \pm 1.8$, $\langle [\mathrm{Z/H}] \rangle = -0.95 \pm 0.62$.

Extragalactic Globular Clusters. Systems of GCs have been discovered and studied in external galaxies ranging from dwarfs to giants spanning the whole Hubble sequence of morphological types. The situation has been recently reviewed by [74] to whom we refer. Most galaxies have bimodal color distributions of GCs reflecting two sub-populations: metal-poor and metal-rich. The characteristics of both populations are correlated with those of the parent galaxy. Most likely, the metal-poor GCs were formed very early on in low mass DM halos. The metal-rich ones were born in subsequent dissipational building up of their parent galaxy. Their age and metallicity indicate that most massive early-type galaxies formed the bulk of their stars at early stages. Therefore, the systems of GCs in external galaxies provide strong constraints on the dominant mechanism of galaxy formation (see below).

SSP Manifolds: The Case of Early Type Galaxies. One of the major challenges of modern astrophysics is to understand the origin and evolution of galaxies, the bright ellipticals in particular. In a Universe dominated by CDM and some kind of DE in the form of a nonzero Cosmological Constant Λ, and containing a suitable mix of baryons and photons, cosmic structures are formed by the gravitational collapse of DM. They are organized in a hierarchy of complexes (haloes) inside which baryonic matter dissipates its energy and collapses to form luminous systems. In this context, the formation of elliptical galaxies can be reduced to the following schemes, even if a sharp distinction among the various scenarios might not exist in reality [395]:

(1) The *monolithic scenario*, which predicts that all elliptical galaxies form at high redshift by rapid collapse and a single prominent SF episode ever since followed by quiescence (e.g., [310]). In favor of this scheme are the observational data that convincingly hint for old and homogeneous stellar populations (see [101] for more details). Along the same line of thought is *revised monolithic* proposed by Schade et al. [495] to account for some evidences of SF at $0.2 < z < 2$ inferred from emission line of [OII], and the nearly constant number frequency of elliptical galaxies up to $z \simeq 1$: a great deal of the stars in massive elliptical galaxies are formed very early at high redshift and the remaining few ones at lower redshift.

(2) The *hierarchical scenario*. The scrutiny of nearby elliptical galaxies shows a large variety of morphological and kinematical peculiarities (such as counter-rotating cores, small gaseous disks and shells), and the occurrence of SF in a recent past (see [321] and references). All this leads to the hierarchical picture in which elliptical galaxies are formed by mergers and/or accretion of smaller units over the Hubble time scale (see [372] and references). Therefore, at increasing the look back time, the density in comoving space of bright (massive) elliptical galaxies should decrease by a factor 2 to 3 (e.g. [270]). In favor of this view are (1) some observational evidences that the merger rate likely increases with $(1 + z)^3$ [390] together with some hint that the EG colors become bluer at increasing complexity of galaxy structure (perhaps tracing some SF associated to merger events); (2) many successful numerical simulations of galaxy encounters, mergers, and interactions (e.g., [24]). However, contrary to the expectation from this model, the number density of elliptical galaxies does not seem to decrease with the redshift, at least up to $z \simeq 1$ [250]. A variant of the hierarchical scheme is the *dry merger*, in which bright elliptical galaxies form by encounters of quiescent, no star forming galaxies [29], motivated by the fact the red peak of the color distribution of galaxies shows mild evolution in the B-band starting from $z \simeq 1$. As the red color of the peak suggests that no new stars are being formed in old elliptical galaxies in this time interval, the way out is to suppose that the hierarchical mass growth of elliptical galaxies is due either to galaxies in which SF is truncated by some physical process, or by "dry mergers" of smaller red, gas-poor objects.

Elliptical Galaxies: Old or Young? Diagnostic to Disposal. Independent of the scenario for the formation and evolution of elliptical galaxies, there are some primary observational constraints on the history of SF and the formation mechanism that models should meet [98]. So far as we know, broad-band colors, line strength indices, and chemical properties of bright (massive) elliptical galaxies are compatible with the notion that the bulk of the stars formed in the remote past, even though some secondary activity is possible. In contrast, the situation with fainter, dwarf (less massive) elliptical galaxies is more uncertain because they show a much wider scatter in their properties. This statement stems from several observational evidences and simple theoretical arguments:

(1) Chemical abundances. Elliptical galaxies are metal rich objects with mean metallicity in the range $-0.8 < [\mathrm{Fe/H}] < 0.3$ [73, 282]. They are also characterized by large [α/Fe] ratios (from 0.05 to 0.3 dex) in their nuclei [301]. This hints for a short duration of the SF activity (see below). Abundance gradients exist in elliptical galaxies with typical values $\Delta[\alpha/\mathrm{Fe}]/\Delta \log R \simeq -0.3$, which are best reproduced by outside-in models [414,547]. It is not clear whether there is a correlation between abundance gradients and galactic mass. This is predicted by the monolithic models of Larson [310] and Chiosi and Carraro [99].

(2) The Color–Magnitude Relation. On the average, elliptical galaxies get redder at increasing luminosity, that is, follow a mean (Color–Magnitude Relation) CMR, which is tight for cluster galaxies [70] and more dispersed for objects in small groups and in the field [504]. According to Bower et al. [70], the tightness of the cluster CMR is compatible with most of the stars in elliptical galaxies being formed at redshift $z > 2$ (look-back time of about 10 Gyr for a Universe with $q_0 = 0.5$ and $H_0 = 70$) and within the first 1–2 Gyr. The situation with field elliptical galaxies is perhaps compatible with SF spread over longer periods of time, that is, 2–3 Gyr according to [41]. The CMR is commonly explained by the SN-driven galactic wind model of Larson [310] as the consequence of a mass-mean metallicity relation, where the massive elliptical galaxies are more metal-rich. In the SN driven galactic wind model, massive elliptical galaxies, owing to their deeper gravitational potential, retain gas and form stars for longer periods of time than the low-mass ones. The alternative that the CMR is an age sequence, with bluer galaxies younger than red ones, has been proved not to be viable by Kodama and Arimoto [285].

(3) Fundamental Plane (FP). The tightness of the FP seen edge-on (M/L vs. M) of elliptical galaxies in Virgo and Coma clusters suggests a short duration for the SF activity in those elliptical galaxies [101].

(4) Line strength Indices. Over the years, many attempts have been made to infer ages, metallicities, and degree of enhancement in α-elements by means of line strength indices (see [544–546] and references). In a few cases, other broad-band colors such as (1,550-V), the signature of UV-excess in elliptical galaxies, and/or velocity dispersion σ have also been considered, for example, [73, 435]. Looking at the position of elliptical galaxies in various diagnostic planes, for example, the popular H_β vs. [$MgFe$], it was soon evident that (1) for field elliptical galaxies, there is a large scatter in H_β, commonly attributed to the stronger sensitivity of this index to SF with respect to other indicators. Similar scatter occurs also with other indices of the same type, for example, H_γ. (2) The field elliptical galaxies do not follow the relation expected in this diagram for objects matching the CMR (Bressan

et al. [73]). In contrast, cluster elliptical galaxies nicely follow it even though a significant scatter is seen [300, 301]. (3) While the simplest explanation would be that some galaxies are truly young objects, the reality could be more complex in the sense that *all elliptical galaxies are old but underwent different histories of SF*. Some of them completed their stellar activity in the distant past with no evidence of subsequent episodes. Others had a more prolonged SF history, perhaps in recurrent episodes of short duration and low intensity [73, 557, 558]. Bressan et al. [73] argued that *the total duration of SF should get longer at decreasing velocity dispersion σ (mass)*. See also [421] for similar conclusions. (4) Delayed or prolonged or even secondary SF seems to be more probable in field or loose group elliptical galaxies than in those belonging to compact groups and clusters [300–302]. Does all this imply that a galaxy's environment plays an important role in determining its evolution? (5) The situation with dwarf elliptical galaxies seems to be the same independently of the age (redshift) of the galaxy population. Indeed, the existence of blue, low-luminosity elliptical galaxies in Coma and Abell 851 detected in the far UV [87] supports the possibility that secondary activity of SF in these galaxies has randomly occurred all over their lifetime. In contrast, the bright elliptical galaxies once more seem to be very old objects undergoing passive evolution after the early, dominant SF. (6) Interacting and noninteracting elliptical galaxies (the former in low density environment and the latter in the field) have the same distribution in the H_β–[MgFe] plane, thus suggesting that dynamical interactions are not a necessary prerequisite for the occurrence of secondary activity, but internal causes are equally possible [321]. (7) The distribution in the above plane is much steeper than the path followed by aging SSPs. Adopting a simple scheme for SF reduced to a primary (P) and a secondary (S) event, each characterized by age and intensity (i.e., the relative fraction of mass turned into stars), T_i, I_i (where $i = P, S$) respectively, [321] noted that to recover the observed distribution (a) the age and duration of the primary event of SF must be old (age comparable to the Hubble time) and longer than about 2 Gyr (i.e., a sizable fraction of the Hubble time), otherwise a tight clump of low H_β galaxies accompanied by a tail of high H_β objects distributed along the locus of an aging SSP of suited composition would be seen; (b) a significant chemical enrichment must have taken place ($\Delta \log(Z)/\Delta \log(t) \simeq 0.7$); (c) finally, the intensity of secondary activity (if any) should not exceed say 2%, as otherwise too many galaxies with strong H_β would be expected. The same conclusions were also reached by [546] using more sophisticated Monte Carlo simulations. *Therefore, the observational data seem to suggest large age ranges for the bulk of the stars and substantial chemical enrichment (up to a metallicity about two times solar).* (8) Finally, limited to a handful of objects for which all necessary information is available, it seems that the SF process lasted longer in the center than in the external regions [73, 547].

(5) Broad band colors as tracers of SF. To what extent a past episode of SF would reflect onto broad-band colors, like $(B - V)$, of the host galaxy as we see it today? Answering this question could somehow constrain current models of galaxy formation in the monolithic or hierarchical scheme. To this aim Chiosi and Carraro [99] and Tantalo and Chiosi [546], using the same SF scheme of [321],

presented photometric simulations at varying age T_P, I_P, T_S, and I_S, that is, the age (in Gyr) and the intensity (percentage of total mass) of the primary and secondary star formation events. It turns out that for many combinations of the parameters, the resulting $(B - V)$ color would be too blue as compared to the typical colors of elliptical galaxies, $(B - V) = 0.95 \pm 0.025$. Secondary SF engaging 5–10% of the total mass and taking place as early as 5–6 Gyr ago would be detectable. The situation gets even worse for higher I_S and/or lower T_S. *This implies that only remote or minor star forming events are allowed.*

(6) Enhancement in α-elements. In elliptical galaxies, Mg_2 seems to increase faster than $\langle Fe \rangle$ (the same for similar indices). This is interpreted as the signature that elliptical galaxies are enhanced in α-elements. Furthermore, Mg_2 is known to increase with the velocity dispersion (and hence mass and luminosity) of the galaxy [86]. Standing on this body of data, the degree of enhancement in α-elements should increase from dwarf to massive elliptical galaxies (see [346] and references). The simplest and most widely accepted interpretation stands on the different duration of the SF period and the different contribution to α-elements and Fe by type II SNe from massive stars (mostly producing α-elements) and type Ia SNe from mass accreting White Dwarfs (mostly generating Fe). Type Ia occur later than type II SNe [220]. With the standard SN driven galactic wind model of Larson [310] and the classical IMF, to reproduce the observational [α/Fe]–mass relationship, the total duration of SF should decrease with the galaxy mass. *This is exactly the opposite of what is required by the same model to explain the CMR.*

(7) High red-shift data. Lyman-break and SCUBA galaxies at $z \geq 3$, where SF is as high as $40 \div 1{,}000 M_\odot$/yr, can be the young elliptical galaxies [127]. The existence of old fully assembled massive spheroidal galaxies already at $1.6 \leq z \leq 1.9$ [104] indicates an early formation of elliptical galaxies at least for $z \geq 2$. Recently, HST data has provided evidence of old massive spheroids at very high redshifts. Mobasher et al. [363] reported evidence for a massive post star-burst galaxy at $z \simeq 6.5$. Recent Spitzer-data (and Space InfraRed Telescope Facility: SIRTF) in the far infrared have revealed the existence of very massive galaxies already in place at redshift $z > 6$. Furthermore, evidences for mass "downsizing" and "top-down" assembly of elliptical galaxies ([105, 444]) arise from analyses of the rest-frame B-band Combo-17 and Deep2 luminosity functions and photometric studies of galaxies at $z = 1$ [286]. Finally, HST observations have also brought into evidence galaxies in place at $z > 7 \div 8$ [68].

(8) In favor of low redshift formation of elliptical galaxies, the following arguments are commonly invoked: (1) the relative high values of H_β observed in nearby elliptical galaxies; (2) the blue cores found in some elliptical galaxies in HDF [355] indicating recent SF; (3) The tight constancy of the FP with the redshift, which can be attributed to conspiracy of age and metallicity; more metal-rich galaxies are younger than the low metal ones [170]; (4) the apparent paucity of high luminosity elliptical galaxies at $z \simeq 1$ compared to now [271, 355, 602]. However, Yamada et al. [596] found that $60 \div 85\%$ of early type galaxies are already in place at $z = 1$.

To summarize, stars formed rather early on in the history of the Universe: PopIII objects appeared at about 13 Gyr ago. The bulk of stars in galaxies was also formed early on; according to modern determination of the age of GGCs, the age falls in the range $10 \div 12$ Gyr, the upper value likely holding for the metal-poor ones. The oldest extra-galactic GCs have similar ages. The history of cluster and field SF in individual galaxies seems to differ a lot depending on the morphological type, mass, and environmental conditions. In any case, as far as we can tell, in most galaxies very old stellar populations have been detected. Finally, the age of the Universe derived from WMAP data and that obtained from stars agree each other. Stars (and stellar models in turn) are indeed good clocks (perhaps the best one to our disposal).

Thank you Cesare.
At the end of this chapter, we decided to put a review of the cosmological parameter H_0, the Hubble constant. We have explained before the reason of our choice. Michael Rowan-Robinson reviews for us the most recent determination of this parameter.

2.16 The Distance Scale, A Road Towards Modern Cosmology

Dear Michael *(Rowan-Robinson)*, **how fundamental have distance measurements been for cosmology?**

Mankind's efforts to measure the distances of the planets, stars, and galaxies are closely bound up with the evolution of our ideas about the Universe we find ourselves in. This link stretches from classical times to today, with the very latest analysis of the fluctuations in the cosmic microwave background.

Aristotle (384-322 BC) was the first to estimate the size of the earth, using the angle of the shadow of a pole at noon at a location 100 miles south of the equator. Eratosthenes and Poseidonius later used a similar method. All these estimates are within about 10% of the modern value. In the second century BC, Hipparcos used an eclipse method to estimate the distance of the moon and deduced a value 59 Earth Radii, compared to the modern value of 60.3. Aristarcos tried to estimate the distance of the Sun using an eclipse method, but was out by a factor of 20. The Greeks also gave us Euclidean geometry (Euclid 300 BC), the idea of absolute, uniform time (Aristotle), and the idea of an infinite physical frame (the atomists, Epicurus). Interestingly, and contrary to the picture held by medieval thinkers, Aristotle believed that the stars were at a range of distances.

A discovery of Copernicus (1473–1543) that is less well-known than his heliocentric system is that he gave, for the first time, the correct relative distances of the Sun and planets. His values were within 5% of the modern values. The absolute scale of the Solar system was not determined accurately till the nineteenth century.

Galileo provided direct evidence for the Copernican picture and laid the foundations for Newtonian dynamics. The Copernican picture immediately implied a much greater distance for the "immovable" stars. Newton tried, unsuccessfully, to estimate the distances of stars through their brightness, but the first step on the distance ladder outside the Solar system was taken by Bessel in 1838 when he measured the

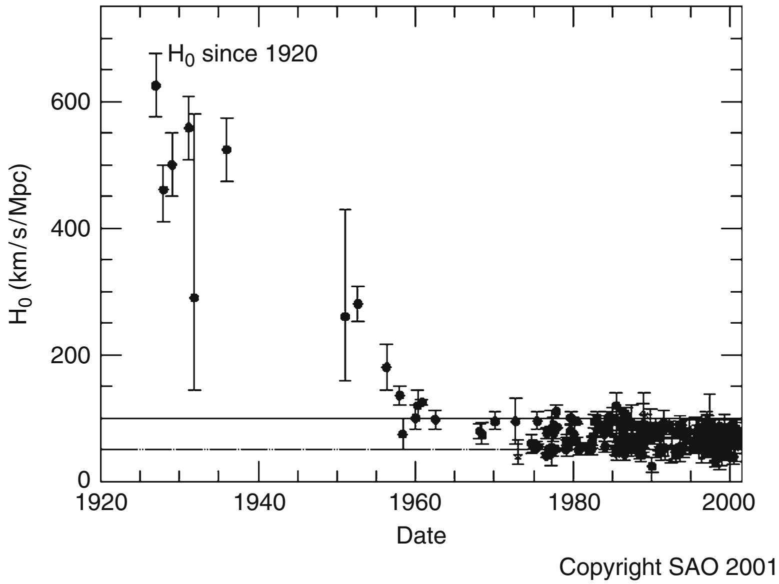

Fig. 2.47 Compilation by J. Huchra (Smithsonian Astrophysical Observatory) of measurements of the Hubble constant from 1927–2001. From [247]

parallax of 61 Cyg, its change in apparent direction on the sky due to the earth's orbit round the Sun. This was the final proof of the Copernican system. Bradley had discovered aberration, the elliptical motion of all stars on the sky due to the earth's motion, a century earlier.

The next crucial step on the distance ladder, still of prime importance today, was the discovery by Henrietta Leavitt in 1912, working at the Harvard Observatory, that the periods of Cepheid variable stars in the Small Magellanic Cloud are related to their luminosity, the period-luminosity relation. In 1924, Edwin Hubble used Leavitt's discovery to estimate the distance of M31, the Andromeda Nebula. It clearly lay far outside our Milky Way Galaxy, thus resolving the long-standing controversy about the spiral nebulae and opening up the Universe of galaxies. Five years later he announced, based on the distances of 18 galaxies, that the more distant a galaxy, the faster it is moving away from us (the Hubble Law):

$$\text{velocity}/\text{distance} = \text{constant} = H_0. \tag{2.20}$$

This is just what would be expected in an expanding Universe. Alexander Friedmann had shown in 1922 that expanding Universe models are what would be expected according to Einstein's General Theory of Relativity, if the Universe is (a) homogeneous (everyone sees the same picture) and (b) isotropic (the Universe looks the same in every direction). This unlikely assumption, the cosmological principle, had been introduced by Einstein in 1917 when he derived a static model of the Universe in which gravity is balanced by a new force, the cosmological repulsion.

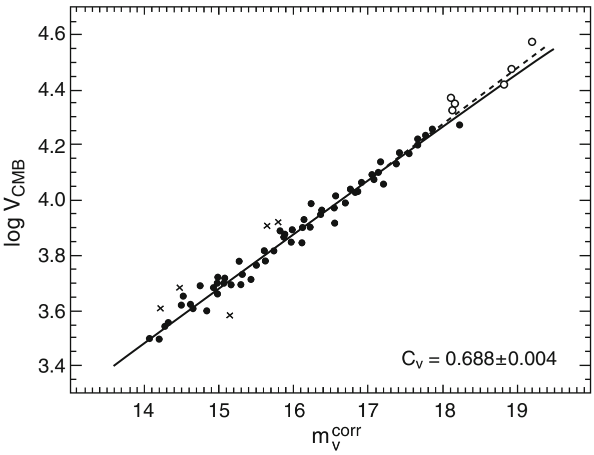

Fig. 2.48 Sandage et al.'s Hubble law based on type Ia SNe. (From [488]). The recession velocity is plotted against the corrected visual magnitude at maximum light

Einstein's inspired guess that the Universe must be very simple (homogeneous and isotropic) is confirmed to very high accuracy today.

But there has been continuing controversy over H_0?

Hubbles's estimate of H_0 was 500. Now H_0 has the dimensions of 1/time and so $1/H_0$ is the expansion age of the Universe, the age the Universe would have if no forces were acting. Hubble's value for H_0 implied an age of the Universe of 2 billion years and it was soon realized this was shorter than the age of the earth, derived from radioactive isotopes. From 1927 to 2001, the value of the Hubble constant was a matter of fierce controversy. Baade pointed out in 1952 that there were two different types of Cepheid, so Hubble's calibration had been incorrect. This reduced H_0 to 200. In 1958, Sandage recognized that objects that Hubble had thought were the brightest stars in some of his galaxies were in fact HII regions and arrived at the first recognizably modern value of 75. During the 1970s, there was an acute disagreement between Sandage and Tammann, on the one hand, favoring $H_0 = 50$, and de Vaucouleurs, on the other, favoring $H_0 = 100$. This disagreement stimulated me to write my monograph *The Cosmological Distance Ladder* [478], in which I set out to review all aspects of the distance ladder and to reconcile the systematic differences in distance estimates from different methods. With an objective weighting scheme based on quoted errors, and with higher weight for purely geometric distance methods (or those based on theoretical arguments), I concluded that there were systematic errors in the type Ia SN method (too high distances) and in the Tully–Fisher and HII region methods (too low) and that the best overall value for H_0 was 67 ± 12.

$H_0 = 67$ would give an expansion age for the Universe of 15.3 billion years (Gyr). In the simplest, Einstein de Sitter ($\Omega_m = 1$, $\Lambda = 0$) model, with only gravity acting to slow the expansion, the age of the Universe would be 10.2 billion years. This could be compared with ages of the oldest stars in globular clusters, 10–15 Gyr. (Chaboyer [96] estimated 12.6 ± 1.1 Gyr), and the age of the Galaxy derived from radioactive isotope abundances, also 10–15 Gyr. Was this already a headache for the Einstein de Sitter model?

2.16.1 HST Key Program

Following the launch of the HST in 1990, and the subsequent repair mission, substantial amounts of HST time were dedicated to measuring Cepheids in galaxies out to distances of 20 Mpc, to try to measure the Hubble constant accurately and to give the different distance methods a secure and consistent calibration. The HST Key Program soon split into two teams, one led by Wendy Freedman, Jeremy Mould, and Rob Kennicutt, and the other by Allan Sandage and Gustav Tammann. In 2001, Freedman et al. [182] announced their final result: $H_0 = 72 \pm 8$.

This as we shall see agreed extremely well with the first results from the WMAP CMB mission (72 ± 5, [523]). It gave an age of the Universe for an Einstein de Sitter model of 9.1 Gyr, which meant that a positive cosmological constant would be required for constancy with the age of the oldest stars. However, again, as we shall see later, evidence from type Ia SNe presented in 1998 was supporting the idea of a positive cosmological constant. However, it was still of interest to see whether there were any possible doubts about this HST Key project value for H_0. The uncertainties in this value are (1) the distance of the Large Magellanic Cloud, which remains uncertain by 10%, (2) the adopted Cepheid calibration, based on OGLE Cepheids, (3) corrections for the effects of dust extinction, (4) corrections for differences in metallicity between the LMC and the Cepheid host galaxies, (5) corrections for the local peculiar velocity flow. Using the Freedman et al. data, my own best estimates for these corrections and the weighting scheme of [478], I concluded [523] $H_0 = 63 \pm 6$.

2.16.2 Type Ia Supernovae

In 1998, two teams announced that using type Ia SNe as standard candles out to significant redshifts (~ 0.5) implied that the cosmological constant Λ had to be greater than zero [403, 450, 499]. There were issues with (1) the treatment of extinction by dust, and (2) the consistency of the assumed correlation of the luminosity at maximum light with the exponential decline rate after maximum raised by [315, 479]. I also raised two other issues: (3) inconsistencies with earlier type Ia SN data, (4) inappropriate use of SNe not observed before maximum light. A group which combined member of the high redshift SN team and the HST Key project team announced a value for H_0 from type Ia SNe of 68 ± 5 [194].

The SN data is clearly excellent and the latest results [16, 452], reaching out to redshift 1.5, are extremely impressive. A problem with the [194] analysis was that it used photographic magnitudes for some of the older SNe. Riess et al. [451] used new HST-ACS observations of Cepheids in galaxies with well-observed recent type Ia SNe and concluded that $H_0 = 73 \pm 6$. This analysis also demonstrated that some of the inconsistencies with earlier type Ia SNe can be attributed to systematic errors in the photographic magnitudes. The issue of the luminosity-decline rate relation has been addressed by [258, 375, 584]. There are still some unresolved inconsistencies in the derivation of extinction, which can only be resolved with the use of more photometric bands in future SN studies.

Do we have a consensus about H_0 today?

With the WMAP 3-year results yielding $H_0 = 73 \pm 3$ [524], it looks as though we have a consensus around $H_0 = 73$, $\Omega_m = 0.25$, $\Lambda = 0.75$, and an age of the Universe 13.7 Gyr. However, in 2006, Sandage, Tammann, and Saha [488] announced the results of their HST programme, with $H_0 = 62 \pm 5$.

This was based on a new extensive study of the Cepheid period–luminosity relation [543], and recognition that there is a difference between the P-L relation in the Galaxy and the LMC [487]. They used a new Cepheid calibration based on the Baade–Wesselink expanding photosphere method, and so do not incur the uncertainty in the LMC distance. And they give a new discussion of extinction in SNe. In my view, this is an analysis that has to be taken very seriously. I will discuss below whether this is inconsistent with the WMAP CMB estimate.

Mike Feast gave a recent review of work on H_0 [164]. New HST Cepheid distances by [31] and revised Hipparcos parallaxes result in a revision of Sandage et al.'s H_0 value from 62 to 69.6 [568]. The Freedman et al. value [182] is also increased. Macri et al. [327] have shown that the Cepheid distance to NGC 4258 is consistent with a geometrical estimate derived from maser emission. The latest H_0 estimates from the gravitational lens time delay method are 68 ± 10 [376], 72 ± 10 [488], and from the Sunyaev–Zel'dovich method for clusters of galaxies are 66 ± 14 [260] and 76 ± 10 [52].

The gravitational lens time delay and Sunyaev–Zel'dovich methods offered the prospect of completely independent geometrical methods, which could be applied at high redshift, thereby overcoming any uncertainty due to the peculiar velocities of local galaxies. The gravitation lens time delay method uses double images of quasars caused by gravitational lensing by an intervening galaxy. If the background quasar varies its light output, the two images will be seen to vary out of phase because of the different time it takes the light to arrive via the two different routes. The time delay can then be used to estimate the distance of the quasar. The Sunyaev–Zel'dovich method is based on the fact that very hot X-ray emitting gas in rich clusters of galaxies interacts with the photons of the cosmic microwave background to produce either a brightening or dimming at microwave wavelengths. A combination of the microwave and X-ray data allows the distance to be estimated if the gas cloud is assumed to be spherical and smooth.

Unfortunately, both methods appear to have irreducible systematic uncertainties. In the case of gravitationally lensed systems, we have to know the exact distribution of matter, including DM, in the foreground lensing galaxy. In the case of the S-Z method, we cannot be sure that the gas clouds are spherical and there is a strong possibility that the gas is clumpy.

How have measurements of CMB fluctuations impacted H_0?

Finally, I want to discuss the distance method that takes us to a redshift of 1,100, the angular scale of the first Doppler peak in the CMB fluctuations. If we think we know the physics of the Universe at the time of decoupling of matter from radiation (epoch of "recombination"), then we know the sound speed in the Universe at that time and hence the linear scale of the acoustic horizon. This will translate to the angular scale of the largest structures in the CMB fluctuations, the first Doppler peak, depending on H_0 and the cosmological model. Analysis of the CMB fluctuations usually proceeds by fitting the whole CMB fluctuation spectrum, with some assumptions about the primordial density fluctuation spectrum (usually that it is a power-law, sometimes with the further restriction that the power-law index has the Harrison–Zel'dovich scale-free value $n = -1$), the spatial curvature (often taken to be zero) and requiring consistency with other astrophysical data (type Ia SNe, large-scale structure). Results of some of these CMB analyses are summarized in Table 2.1.

Spergel et al. [523] show that with the assumption of a power-law spectrum, but no restriction to flat models, consistency with the WMAP (year 1) fluctuation spectrum can be achieved with a wide range of cosmological models and values for H_0. Priors on H_0 or the assumption of flatness then force us to the $\Omega_\Lambda = 0.75$ consensus model. However, if we also drop the assumption of a power-law primordial density fluctuation spectrum, which is not necessarily expected in a Universe that has been through a series of phase transitions, the possibilities are opened up even further. Blanchard et al. [48] showed that if we relax the assumption of a power-law to the simplest alternative, a broken power-law, then we can fit the CMB fluctuation spectrum just as well as the consensus model with an $\Omega_m = 1$, $\Omega_\Lambda = 0$ (Einstein de Sitter) model, provided $H_0 = 46$.

We can also get consistency with galaxy large-scale structure data, provided there are one or more neutrinos with a mass of a few keV, such that $\Omega \sim 0.2$ (a mixed DM

Table 2.1 CMB fluctuation results for H_0

Data set	Assumptions/other data	H_0	Reference
BOOMERanG, MAXIMA	Flat Universe	75 ± 10	Jaffe et al. [255]
WMAP first year		72 ± 5	Spergel et al. [523]
WMAP first year	SDSS LSS data	68 ± 10	Tegmark et al. [553]
WMAP first year	BAO data	65 ± 4.5	Eisenstein et al. [151]
WMAP 3-year data		73 ± 3	Spergel et al. [524]
WMAP 3-year data	LSS, BAO	$69 - 72$	Spergel et al. [524]
WMAP 5-year data	LSS, BAO	70.1 ± 1.3	Komatsu et al. [290]

model). However, this model is inconsistent with the type Ia SN data and $H_0 = 46$ is 3σ from the direct HST Key Project estimates. Recently, Shafieloo and Souradeep [509] have confirmed that the low-H_0, Einstein de Sitter model is as good a fit to the WMAP CMB fluctuation spectrum as the consensus model if the primordial density fluctuation spectrum is allowed to have a free form.

So the CMB fluctuations do not on their own determine H_0. An important advance is the discover of the BAO feature in the power spectrum of galaxy density fluctuations ([108, 151]). This feature is essentially the same acoustic horizon scale seen in the CMB fluctuations, but now seen in galaxy redshift surveys at $z \sim 0.35$. At this epoch, it has a linear scale of about 150 Mpc. Blanchard et al. [49] admit that this feature, if confirmed (it is about $2\div3\sigma$ significance at the moment), would be fatal for their low-H_0, Einstein de Sitter model. The combination of the CMB first Doppler peak and the baryon acoustic oscillation peak is the ultimate geometric measurement of H_0. Using the WMAP 5-year data combined with baryon acoustic oscillation and type Ia SN data, Komatsu et al. [290] conclude that $H_0 = 70.1 \pm 1.3$.

Figure 2.49 shows the result of applying these two tests, the angular scale of the first Doppler peak and of the baryon acoustic oscillation peak, as a pure diameter-distance test. The black line shows the locus of a zero-curvature Universe. The solid curves show the loci that give the same observed angular scale for the first Doppler peak, for $H_0 = 73, 65, 48$. We see that the $H_0 = 48$ curve intersects the zero-curvature line at the $\Omega_m = 1$, $\Omega_\Lambda = 0$ Einstein de Sitter model, consistent with the [48] and [509] claims.

The broken curves show loci for the same three values of H_0 for models which give the same observed angular diameter for the baryon acoustic oscillation peak at $z = 0.35$. The $H_0 = 65$ locus passes close where the corresponding first Doppler

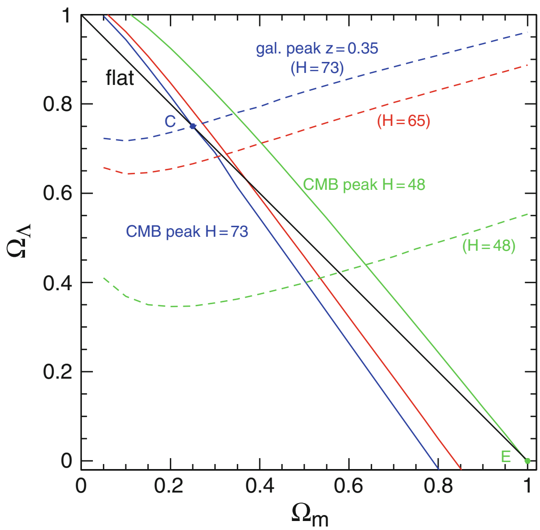

Fig. 2.49 Ω_Λ vs. Ω_m, showing results of CMB and BAO diameter–distance tests. *Black solid line*: zero curvature. *Solid curves*: loci that give the same observed angular scale for the first CMB Doppler peak, for $H_0 = 73$ (*blue*), 65 (*red*), 48 (*green*). *Broken curves*: models that give the same observed angular diameter for the baryon acoustic oscillation peak at $z = 0.35$ for $H_0 = 73, 65, 48$ (*same color code*)

peak locus intersects the zero curvature line. However, the $H_0 = 48$ locus lies nowhere near the Einstein de Sitter model. The conclusion is that $H_0 = 65$ is consistent with this combined test, but $H_0 = 48$ is ruled out.

To Summarize

(1) Local direct estimates of H_0 are in the range 62–72, with an uncertainty of 10%, and this is a big advance on the range 50–100 from the 1970s. My 1985 value of 67 still looks quite plausible.
(2) The CMB fluctuation estimates lie in the range 65–73, depending on the assumptions made, with an uncertainty of 4%.
(3) The angular scale of the baryonic acoustic oscillation peak, combined with the CMB first Doppler peak, is the ultimate geometrical measurement of H_0, and this reduces the claimed uncertainty to 2%.
(4) It is still worthwhile to improve the direct local estimates of H_0. If an accuracy of, say, 1%, could be achieved, then there is the prospect of learning about new physics beyond the Standard Model of Particle Physics, as tension between direct and CMB estimates would show that the underlying assumptions being used in the CMB physics were incorrect. This is a very challenging goal and might take a decade to achieve. Such an accuracy might be achieved through:

(a) Direct parallax measurements of the distance to the LMC, from the GAIA mission. This is the main uncertainty in most current local estimates, accounting for 10% uncertainty in H_0.

(b) Use of the Baade–Wesselink expanding atmosphere method for Cepheids and SNe. For Cepheids this would reduce the uncertainty in the absolute calibration. To generate accurate atmospheric models for type Ia SNe is challenging, but would enormously improve confidence in the method and would eliminate the need for ad hoc corrections for the luminosity dependence on decline rate.

(c) Use of multi-wavelength photometry, especially in the infrared, to control extinction and metallicity. Further work on estimating distances via Cepheids and SNe really does not seem worthwhile without this development.

(d) Much better mapping of the local density field would be needed to reduce the uncertainty in the peculiar velocities of galaxies.

In conclusion, progress in understanding the Universe has always been strongly connected with our ability to measure distances. Today we have distance measurements to redshift 1,100, the epoch when matter and radiation finally decoupled at the end of the hot Big Bang phase. Apparently, we have reached a precision of 2% in our measurements of the Hubble constant and a consensus model of the Universe with a dominant role for DE. Our inability to provide a motivation for this DE remains troubling and we should remain open to the possibility of new physics beyond the Standard Model, which might change our whole picture of the Universe.

Thank you very much Michael.

References

1. Abadi, M., et al.: Simulations of galaxy formation in a Λ cold dark matter universe. II. The fine structure of simulated galactic disks. Astrophys. J. **597**, 21 (2003)
2. Abel, T., Bryan, G.L., Norman, M.L.: The formation and fragmentation of primordial molecular clouds. Astrophys. J. **540**, 39 (2000)
3. Abell, G.O.: The distribution of rich clusters of galaxies. Astrophys. J. Suppl. **3**, 211–288 (1958)
4. Ade, P., et al.: First season QUaD CMB temperature and polarization power spectra. Astrophys. J. **674**, 22A (2008)
5. Agnoletto, I., et al.: SN 2006gy: Was it really extra-ordinary? Astrophys. J. **681**, 1348 (2009)
6. Allen, S.W., et al.: Improved constraints on dark energy from chandra X-ray observations of the largest relaxed galaxy clusters. Mon. Not. R. Astron. Soc. **383**, 879 (2008)
7. Alpher, R.A., Bethe, H.A., Gamow, G.: The origin of chemical elements. Phys. Rev. **73**, 803 (1948)
8. Alsop, D.C., et al.: A search for anisotropy in the cosmic microwave background on intermediate angular scales. Astrophys. J. **395**, 317–325 (1992)
9. Archibald, E.N., et al.: Coupled spheroid and black hole formation, and the multifrequency detectability of active galactic nuclei and submillimetre sources. Mon. Not. R. Astron. Soc. **336**, 353 (2003)
10. Arnaud, M., Evrard, A.E.: The $L_X - T$ relation and intracluster gas fractions of X-ray clusters. Mon. Not. R. Astron. Soc. **305**, 631–640 (1999)
11. Arnaud, M., Pointecouteau, E., Pratt, G.W.: The structural and scaling properties of nearby galaxy clusters. II. The $M - T$ relation. Astron. Astrophys. **441**, 893–903 (2005)
12. Arp, H.: NGC 7603, a galaxy connected to a companion of much larger redshift. Astrophys. Lett. **7**, 221 (1971)
13. Arp, H.: Quasars, redshifts and controversies. Interstellar Media, Berkeley, CA, p. 204 (1987)
14. Arp, H.: Seeing red : redshifts, cosmology and academic science. Apeiron, Montreal, Canada (1998)
15. Asplund, M., et al.: Lithium isotopic abundances in metal-poor halo stars. Astrophys. J. **644**, 229 (2006)
16. Astier, P., et al.: The supernova legacy survey: Measurement of Ω_M, Ω_λ and w from the first year data set. Astron. Astrophys. **447**, 31 (2006) [astro-ph/0510447]
17. Aumont, J., Macias-Perez, J.F.: Blind component separation for polarized observations of the cosmic microwave background. Mon. Not. R. Astron. Soc. **376**, 739A (2007)
18. Bacchi, M., et al.: Deep images of cluster radio halos. Astron. Astrophys. **400**, 465–476 (2003)
19. Bachev, R., et al.: Average ultraviolet quasar spectra in the context of eigenvector 1: A Baldwin effect governed by the eddington ratio? Astrophys. J. **617**, 171–183 (2004)
20. Bahcall, J.N., Hills, R.E.: The hubble diagram for the brightest quasars. Astrophys. J. **179**, 699–704 (1973)
21. Baldwin, J.A., et al.: Relative quasar luminosities determined from emission line strengths. Nature **273**, 431–435 (1978)
22. Baldwin, J.A., Wampler, E.J., Gaskell, C.M.: Emission-line properties of optically and radio-selected complete quasar samples. Astrophys. J. **338**, 630–653 (1989)
23. Bania, T.M., Rood, R.T., Balser, D.S.: The cosmological density of baryons from observations of $^3He^+$ in the milky way. Nature **415**, 54 (2002)
24. Barnes, J.E., Hernquist, L.: Transformations of galaxies. II. Gasdynamics in merging disk galaxies. Astrophys. J. **471**, 115 (1996)
25. Baron, E., et al.: Type IIP supernovae as cosmological robes: A spectral-fitting expanding atmosphere model distance to SN 1999em. Astrophys. J. **616**, L91 (2004)
26. Bartelmann, M., Schneider, P.: Weak gravitational lensing. Phys. Report **340**, 291 (2001) [astro-ph/9912508]

27. Baskin, A., Laor, A.: On the origin of the CIV Baldwin effect in active galactic nuclei. Mon. Not. R. Astron. Soc. **350**, L31–L35 (2004)
28. Bedin, L.R., et al.: The double main sequence of Omega Centauri. Mem. S.A.It. Suppl. **5**, 105 (2004)
29. Bell, E.F., et al.: Nearly 5000 distant early-type galaxies in COMBO-17: A red sequence and its evolution since $z \sim 1$. Astrophys. J. **608**, 752 (2004)
30. Bellazzini, M., et al.: Open clusters in the log age vs. MV plane. Mem. S.A.It. Suppl. **79**, 3 (2008)
31. Benedict, G.F.: Hubble space telescope fine guidance sensor parallaxes of galactic cepheid variable stars: Period-luminosity relations. Astron. J. **133**, 1810 (2007)
32. Benetti, S., et al.: The exceptionally bright type Ib supernova 1991D. Mon. Not. R. Astron. Soc. **336**, 91 (2002)
33. Benjamin, J., et al.: Cosmological constraints from the 100-deg^2 weak-lensing survey, Mon. Not. R. Astron. Soc. **381**, 702–712 (2007)
34. Bennett, C.L., et al.: Preliminary separation of galactic and cosmic microwave emission for the COBE differential microwave radiometer. Astrophys. J. Lett. **396**, L7 (1992)
35. Bennett, C.L., et al.: First-year Wilkinson microwave anisotropy probe (WMAP) observations: Preliminary maps and basic results. Astrophys. J. Suppl. **148**, 1–27 (2003)
36. Bennett, C.L., et al.: The microwave anisotropy probe mission. Astrophys. J. **583**, 1 (2003)
37. Benoit, A., et al.: Cosmological constraints from archeops. Astron. Astrophys. **399**, L25–L30 (2003)
38. Benoit, A., et al.: The cosmic microwave background anisotropy power spectrum measured by Archeops. Astron. Astrophys. **339L**, 19 (2003)
39. Benoit, A., et al.: First detection of polarization of the submillimetre diffuse galactic dust emission by Archeops. Astron. Astrophys. **424**, 571B (2004)
40. Bergbusch, P.A., Vandenberg, D.A.: Models for old, metal-poor stars with enhanced α-element abundances. III. Isochrones and isochrone population functions. Astrophys. J. **556**, 322 (2001)
41. Bernardi, M., et al.: Cluster versus field elliptical galaxies and clues on their formation. Astrophys. J. Lett. **508**, L143 (1998)
42. Bertelli, G., et al.: Isochrones for an extended region of the Z-Y plane including new asymptotic giant branch models. Astron. Astrophys. submitted
43. Bertola, F., Sulentic, J.W., Madore, B.: New ideas in astronomy. Cambridge University Press, Cambridge (1988)
44. Binney, J., Evans, N.W.: Cuspy dark matter haloes and the Galaxy. Mon. Not. R. Astron. Soc. **327**, 27 (2001)
45. Birkinshaw, M.: The Sunyaev-Zel'dovich effect. Phys. Rep. **310**, 97–195 (1999)
46. Biviano, A., et al.: On the efficiency and reliability of cluster mass estimates based on member galaxies. Astron. Astrophys. **456**, 23–36 (2006)
47. Blanchard, A., et al.: A new local temperature distribution function for X-ray clusters: Cosmological applications. Astron. Astrophys. **362**, 809–824 (2000)
48. Blanchard, A., et al.: An alternative to the cosmological concordance model. Astron. Astrophys. **412**, 35 (2003)
49. Blanchard, A., et al.: Large-scale galaxy correlations as a test for dark energy. Astron. Astrophys. **449**, 925 (2006)
50. Blondin, S., et al.: Using line profiles to test the fraternity of type Ia supernovae at high and low redshift. Astron. J. **131**, 1648 (2006)
51. Böhringer, H.: Cosmological tests with galaxy clusters. In: Relativistic astrophysics and cosmology – Einstein's legacy. Aschenbach, B., Burwitz, V., Hasinger, G., Leibundgut, B. (eds.), Springer, Heildelberg, p. 289 (2007)
52. Bonamente, M., et al.: Determination of the cosmic distance scale from Sunyaev-Zel'dovich effect and chandra X-ray measurements of high-redshift galaxy clusters. Astrophys. J. **647**, 25–54 (2006)
53. Bonatto, Ch., Bica, E., Girardi, L.: Theoretical isochrones compared to 2MASS observations: Open clusters at nearly solar metallicity. Astron. Astrophys. **415**, 571 (2004)

54. Bond, J.R., Efstathiou G.: Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter. Astrophys. J. Lett. **285**, L45 (1984)
55. Bond, J.R., Efstathiou G., Silk J.: Massive neutrinos and the large-scale structure of the universe. Phys. Rev. Lett. **45**, 1980 (1980)
56. Bond, J.R., Arnett, W.D., Carr, B.J.: The evolution and fate of very massive objects. Astrophys. J. **280**, 825 (1984)
57. Bonifacio, P.: On the lithium content of the globular cluster M 92. Astron. Astrophys. **395**, 515 (2002)
58. Bonifacio, P.: et al.: The lithium content of the globular cluster NGC 6397. Astron. Astrophys. **390**, 91 (2002)
59. Borgani, S.: Cosmology with clusters of galaxies. (2006) [astro-ph/0605575]
60. Borgani, S., Guzzo, L.: X-ray clusters of galaxies as tracers of structure in the universe. Nature **409**, 39 (2001)
61. Borgani, S., et al.: Cosmological constraints from the ROSAT deep cluster survey. Astrophys. J. **517**, 40–53 (1999)
62. Borgani, S., et al.: Measuring Ω_0 with the ROSAT deep cluster survey. Astrophys. J. **561**, 13–21 (2001)
63. Borissova, L., Rabounski, D.: On the nature of the microwave background at the lagrange 2 point. Part II. Prog. Phys. **4**, 84–95 (2007)
64. Borissova, L., Rabounski, D.: PLANCK, the satellite: A new experimental test of general relativity. Prog. Phys. **2**, 3–13 (2008)
65. Botticella, M.T., et al.: Supernova rates from the southern intermediate redshift ESO supernova search (STRESS). Astron. Astrophys. **479**, 49 (2008)
66. Bouchet, F.R., Gispert, R.: Foregrounds and CMB experiments I. Semi-analytical estimates of contamination. New Astron. Rev. **4**, 443 (1999)
67. Bouchet, F.R., et al.: Evidence against or for topological defects in the BOOMERanG data? Phys. Rev. **D 65**, (2002)
68. Bouwens, R.J., et al.: Galaxies at $z \simeq 7-8$: z_{850} – Dropouts in the hubble ultra deep field. Astrophys. J. Lett. **616**, L79 (2004)
69. Bouwens, R.J., et al.: $z \sim$ 7–10 galaxies in the HUDF and GOODS fields, and their UV luminosity functions. Astrophys. J. **686**, 230 (2008)
70. Bower, R.G., Lacey J.R., Ellis R.S.: Precision photometry of early type galaxies in the coma and virgo clusters - a test of the universality of the colour-magnitude relation - Part II - Analysis. Mon. Not. R. Astron. Soc. **254**, 601 (1992)
71. Branchesi, M., et al.: High redshift X-ray galaxy clusters. I. The impact of point sources on the cluster properties. Astron. Astrophys. **472**, 727–737 (2007)
72. Branchesi, M., et al.: High redshift X-ray galaxy clusters. II. The $L_X - T$ relationship revisited. Astron. Astrophys. **472**, 739–748 (2007)
73. Bressan, A., Chiosi, C., Tantalo, R.: Probing the age of elliptical galaxies. Astron. Astrophys. **311**, 425 (1996)
74. Brodie, J.P., Strader, J.: Extragalactic globular clusters and galaxy formation. Ann. Rev. Astron. Astrophys. **44**, 193 (2006)
75. Bromm, V., Kudritzki, R.P., Loeb, A.: Generic spectrum and ionization efficiency of a heavy initial mass function for the first stars. Astrophys. J. **552**, 464 (2001)
76. Bromm, V., Coppi, P.S., Larson, R.B.: The formation of the first stars. I. The primordial star-forming cloud. Astrophys. J. **564**, 23 (2002)
77. Bryan, G.L., Norman, M.L.: Statistical properties of X-ray clusters: Analytic and numerical comparisons. Astrophys. J. **495**, 80–99 (1998)
78. Bullock, J.S., Kravtsov, A.V., Weinberg, D.H.: Reionization and the abundance of galactic satellites. Astrophys. J. **539**, 517 (2000)
79. Buonanno, R., Corsi, C.E., Fusi-Pecci, F.: The ages of globular clusters and the sandage period-shift effect. Astron. Astrophys. **216**, 80 (1989)
80. Burbidge, G.: Noncosmological redshifts. Pub. Astron. Soc. Pacific **113**, 899 (2001)
81. Burbidge, G.: An accidental career. Ann. Rev. Astron. Astrophys. **45**, 1 (2007)

82. Burbidge, E., Burbidge, G.: The redshifts of galaxies and QSOs. In: Current issues in cosmology. Pecker, J.C., Narlikar, J. (eds.), Cambridge University Press, Cambridge, p. 17 (2006)
83. Burbidge, E., et al.: Synthesis of the elements in stars. Rev. Mod. Phys. **29**, 547 (1957)
84. Burenin, R.A., et al.: The 400 square degree ROSAT PSPC galaxy cluster survey: Catalog and statistical calibration. Astrophys. J. Suppl. **172**, 561–582 (2007)
85. Burigana, C., De Zotti, G., Feretti, L.: Sunyaev-Zel'dovich effects, free-free emission, and imprints on the cosmic microwve background. New Astron. Rev. **48**, 1107–1117 (2004)
86. Burstein, D., et al.: The far-ultraviolet spectra of early-type galaxies. Astrophys. J. **328**, 440 (1988)
87. Buson L., et al.: Ultraviolet imaging of the galaxy cluster CL0939+4713 (Abell 851) at $z = 0.41$. Astrophys. J. **531**, 684 (2000)
88. Cacciari, C.: Globular cluster distances. In: New horizons in globular cluster astronomy. Piotto, G., Meylan, G., Djorgovski, S.G., Riello, M. (eds.), ASP Conf. Proc. **296**, 329 (2003)
89. Caldwell, R.R., Dave, R., Steinhardt, P.J.: Cosmological imprint of an energy component with general equation of state. Phys. Rev. Lett. **80**, 1582–1585 (1998)
90. Carlstrom, J.E., Holder, G.P., Reese, E.D.: Cosmology with the Sunyaev-Zel'dovich effect. Ann. Rev. Astron. Astrophys. **40**, 643–680 (2002)
91. Carraro, G., Chiosi, C.: The galactic system of old open clusters: Age calibration and age-metallicity relation. Astron. Astrophys. **287**, 761 (1994)
92. Carretta, E., Gratton, R.G.: Abundances for globular cluster giants. I. Homogeneous metallicities for 24 clusters. Astron. Astrophys. Suppl. **121**, 95 (1997)
93. Carroll, B.W., Ostlie, D.A.: An introduction to modern astrophysics. Addison Wesley, Reading, MA (1996)
94. Cen, R.: The implications of Wilkinson microwave anisotropy probe observations for population III star formation processes. Astrophys. J. **591**, L5 (2003)
95. Chaboyer, B.: Globular clusters age dating. In: Astrophysical ages and time scales. ASP Conf. Ser. **245**, 162 (2001)
96. Chaboyer, B., et al.: The effect of helium diffusion on the ages of globular clusters. Astrophys. J. **388**, 372 (1992)
97. Cheng, E.S., et al.: Detection of cosmic microwave background anisotropy measurement. Astrophys. J. Lett. **488**, L59–L62 (1997)
98. Chiosi, C.: On the age of elliptical galaxies. In: Spectro-photometric Dating of Stars and Galaxies. Hubney, I., Heap, S.R., Cornett, R.H. (eds.), ASP Conf. Proc. **192**, 251 (1999)
99. Chiosi, C., Carraro, G.: Formation and evolution of elliptical galaxies. Mon. Not. R. Astron. Soc. **335**, 335 (2002)
100. Chiosi, C., Bertelli, G., Bressan, A.: New developments in understanding the HR diagram. Ann. Rev. Astron. Astrophys. **30**, 235 (1992)
101. Chiosi, C., et al.: A new scenario of galaxy evolution under a universal initial mass function. Astron. Astrophys. **339**, 355 (1998)
102. Ciardi, B., Ferrara, A.: The first cosmic structures and their effects. Space Sci. Rev. **116**, 625 (2005)
103. Ciardi, B., Ferrara, A., White, S.D.M.: Early reionization by the first galaxies. Mon. Not. R. Astron. Soc. **344**, L7 (2003)
104. Cimatti, A., et al.: Old galaxies in the young universe. Nature **430**, 184 (2004)
105. Cimatti, A., Daddi, E., Renzini, A.: Mass down-sizing and top-down assembly of early-type galaxies. Astron. Astrophys. **453**, L29 (2006)
106. Clowe, D., et al.: Weak lensing mass reconstructions of the ESO distant cluster survey. Astron. Astrophys. **451**, 395–408 (2006)
107. COBE website, http://lambda.gsfc.nasa.gov/product/cobe/
108. Cole, S., et al.: The 2dF galaxy redshift survey: Power-spectrum analysis of the final data set and cosmological implications. Mon. Not. R. Astron. Soc. **362**, 505 (2005)
109. Colegrove, P.T.: Open cluster ages and an apparent periodicity in the history of star formation in the solar neighborhood. PhD thesis. University of Nevada, Reno, USA (1999)
110. Collier, S., et al.: A new direct method for measuring the hubble constant from reverberating accretion discs in active galaxies. Mon. Not. R. Astron. Soc. **302**, L24–L28 (1999)

111. Conklin, E.K.: Velocity of the earth with respect to the cosmic background radiation. Nature **222**, 971 (1969)
112. Cooper, J.L., et al.: Three-dimensional simulations of a starburst-driven galactic wind. Astrophys. J. **674**, 157C (2008)
113. Cordier, D., et al.: A large stellar evolution database for population synthesis studies. III. Inclusion of the full asymptotic giant branch phase and web tools for stellar population analyses. Astron. J. **133**, 468 (2007)
114. Cowie, L.L., et al.: New insight on galaxy formation and evolution from keck spectroscopy of the hawaii deep fields. Astron. J. **112**, 839 (1996)
115. Croom, S.M., et al.: The correlation of line strength with luminosity and redshift from composite quasi-stellar object spectra. Mon. Not. R. Astron. Soc. **337**, 275–292 (2002)
116. Croton, D.J., et al.: The many lives of active galactic nuclei: Cooling flows, black holes and the luminosities and colours of galaxies. Mon. Not. R. Astron. Soc. **365**, 11 (2006)
117. Curto, A., et al.: Constraints on the non-linear coupling parameter f_{nl} with the archeops data. Astron. Astrophys. **486**, 383 (2008)
118. Cyburt, R.H., Fields, B.D., Olive, K.A.: The NACRE thermonuclear reaction compilation and big bang nucleosynthesis. New Astron. **6**, 215 (2001)
119. Cyburt, R.H., Fields, B.D., Olive, K.A.: Primordial nucleosynthesis with CMB inputs: Probing the early universe and light element astrophysics. Astropart. Phys. **17**, 87 (2002)
120. Daigne, F., et al.: Hierarchical growth and cosmic star formation: Enrichment, outflows and supernova rates. Astrophys. J. **647**, 773 (2006)
121. Danese, L., Partridge, R.B.: Atmospheric emission models: Confrontation between observational data and predictions in the 2.5–300 GHz frequency range. Astrophys. J. **342**, 604–615 (1989)
122. De Angeli, F., et al.: Galactic globular cluster relative ages. Astrophys. J. **130**, 116 (2005)
123. de Bernardis, P., et al.: Degree-scale observations of cosmic microwave background anisotropies. Astrophys. J. Lett. **422**, L33–L36 (1994)
124. de Bernardis, P., et al.: Mapping the CMB sky: THE BOOMERanG experiment. New Astron. Rev. **43**, 289 (1999)
125. de Bernardis, P., et al.: A flat universe from high-resolution maps of the cosmic microwave background radiation. Nature **404**, 955–959 (2000)
126. de Bernardis, P., et al.: Multiple peaks in the angular power spectrum of the cosmic microwave background. Astrophys. J. **564**, 559 (2002)
127. de Mello, D., Johansson, E.P.G., Markström, P.: A database of galaxies at high redshifts. Astrophys. Space Sci. **281**, 535 (2002)
128. de Oliveira-Costa, A., et al.: Mapping the cosmic microwave background anisotropy: Combined analysis of QMAP fligths. Astrophys. J. Lett. **509**, L77–L80 (1998)
129. Dekel, A., Birnboim, Y.: Galaxy bimodality due to cold flows and shock heating. Mon. Not. R. Astron. Soc. **368**, 2 (2006)
130. Dekel, A., Silk, J.: The origin of dwarf galaxies, cold dark matter, and biased galaxy formation. Astrophys. J. **303**, 39 (1986)
131. Delabrouille, J., et al.: Multidetector multicomponent spectral matching and applications for cosmic microwave background data analysis. Mon. Not. R. Astron. Soc. **346**, 1089D (2003)
132. Deliyannis, C.P., Demarque, P., Pinsonneault, M.H.: The ages of globular cluster stars – Effects of rotation on pre-main-sequence, main-sequence, and turnoff evolution. Astrophys. J. Lett. **347**, L73 (1989)
133. Demarque, P., et al.: Y^2 isochrones with an improved core overshoot treatment. Astrophys. J. Suppl. **155**, 667 (2004)
134. Desert, F.-X., et al.: Submillimetre point sources from the archeops experiment: Very cold clumps in the galactic plane. Astron. Astrophys. **481**, 411D (2008)
135. Di Matteo, T., Springel, V., Hernquist, L.: Energy input from quasars regulates the growth and activity of black holes and their host galaxies. Nature **433**, 604 (2005)
136. Dicke, R.H., et al.: Cosmic black-body radiation. Astrophys. J. **1**, 414–419 (1965)
137. Diemand, J., Kuhlen, M., Madau, P.: Formation and evolution of galaxy dark matter halos and their substructure. Astrophys. J. **667**, 859 (2007)

138. Dietrich, M., et al.: Elemental abundances in the broad emission line region of quasars at redshifts larger than 4. Astron. Astrophys. **398**, 891 (2003)
139. Donahue, M., Voit, G.M.: $Omega_m$ from the temperature-redshift distribution of EMSS clusters of galaxies. Astrophys. J. Lett. **523**, L137–L140 (1999)
140. Dotter, A., et al.: The ACS survey of galactic globular clusters.II. Stellar evolution tracks. Isochrones, luminosity functions, and synthetic horizontal-branch models. Astrophys. J. **134**, 376 (2007)
141. Draine, B.T., Lazarian, A.: Diffuse galactic emission from spinning dust grains. Astrophys. J. **494L**, 19D (1998)
142. Dunkley, J. et al.: Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from WMAP Data. Astrophys. J. Suppl. **180**, 306 (2009)
143. Dupac, X., et al.: Inverse temperature dependence of the dust submillimeter spectral index. Astron. Astrophys. **404**, L11–L15 (2003)
144. Dutton, A.A., van den Bosch, F.C., Courteau, S.: The Tully-Fisher zero point problem, ASPC **396**, 467 (2008)
145. Ebisuzaki, T., Makino J., Okumura, S.K.: Merging of two galaxies with central black holes. Nature **354**, 212 (1991)
146. Edvardsson, B., et al.: The chemical evolution of the galactic disk. I. Analysis and results. Astron. Astrophys. **275**, 101 (1993)
147. Efstathiou, G., Sutherland, W.J., Maddox, S.J.: The cosmological constant and cold dark matter. Nature **348**, 705 (1990)
148. Eggen, O.J., Lynden-Bell, D., Sandage, A.: Evidence from the motions of old stars that the Galaxy collapsed. Astrophys. J. **136**, 748 (1962)
149. Einstein, A.: Strahlungs-emission und absorption nach der Quantentheorie. Verhandlunger der Deutschen Physikalischen Gesellschaft **18**, 318 (1916)
150. Einstein, A.: Phys. Zs. **18**, 121 (1917). English Translation: On the quantum theory of radiation, by D. der Haar: The old quantum theory. Pergamon Press, New York, p. 167 (1967)
151. Eisenstein, D.J., et al.: Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies. Astrophys. J. **633**, 560 (2005)
152. Eke, V.R., et al.: Measuring Ω_0 using cluster evolution. Mon. Not. R. Astron. Soc. **298**, 1145–1158 (1998)
153. Eke, V.R., Navarro, J.F., Steinmetz, M.: The power spectrum dependence of dark matter halo concentrations. Astrophys. J. **554**, 114 (2001)
154. Elias–Rosa, N., et al.: Anomalous extinction behaviour towards the type Ia SN 2003cg. Mon. Not. R. Astron. Soc. **369**, 1880 (2006)
155. Elmegreen, B.G.: Variation of galactic bar length with amplitude and density as evidence for bar growth over a hubble Time. Astrophys. J. **672**, 1006 (2008)
156. Elvis, M., Karovska, M.: Quasar parallax: A method for determining direct geometrical distances to quasars. Astrophys. J. **581**, L67–L70 (2002)
157. Eriksen, H.K., et al.: Joint bayesian component separation and CMB power spectrum estimation. Astrophys. J. **676**, 10E (2008)
158. Ettori, S., Tozzi, P., Rosati, P.: Constraining the cosmological parameters with the gas mass fraction in local and $z > 0.7$ galaxy clusters. Astron. Astrophys. **398**, 879–890 (2003)
159. Ettori, S., et al.: Evolution at $z \geq 0.5$ of the X-ray properties of simulated galaxy clusters: Comparison with observational constraints. Mon. Not. R. Astron. Soc. **354**, 111–122 (2004)
160. Evrard, A.E., Metzler, C.A., Navarro, J.F.: Mass estimates of X-ray clusters. Astrophys. J. **469**, 494–507 (1996)
161. Faber, S.M., et al.: Old stellar populations. II - An analysis of K-giant spectra. Astrophys. J. Suppl. **57**, 711 (1985)
162. Fall, S.M., Efstathiou, G.: Formation and rotation of disc galaxies with haloes. Mon. Not. R. Astron. Soc. **193**, 189 (1980)
163. Fan, X.: Evolution of high-redshift quasars. New Astron. Rev. **60**, 665 (2006)
164. Feast, M.: From IRAS to Herschel/Planck: Cosmology with infrared surveys, London (2007)
165. Feretti, L.: Non-thermal emission from the intracluster medium. Adv. Space Res. **36**, 729–737 (2005)

166. Feretti, L., Giovannini, G.: Diffuse cluster radio sources. IAU Symp. **175**, 333–338 (1996)
167. Feretti, L., et al.: The radio and X-ray properties of Abell 2255. Astron. Astrophys. **317**, 432–440 (1997)
168. Feretti, L., Burigana, C., Ensslin, T.A.: Diffuse radio emission from the intracluster medium. New Astron. Rev. **48**, 1137–1144 (2004)
169. Ferrarese, L., Merritt, D.: A fundamental relation between supermassive black holes and their host galaxies. Astrophys. J. **539**, L9 (2000)
170. Ferreras, I., Charlot, S., Silk, J.: The age and metallicity range of early-type galaxies in clusters. Astrophys. J. **521**, 81 (1999)
171. Ferris, T.: The red limit. The search for the edge of the universe. William Morrow, New York (1977); Quill, New York (1983) 2nd ed. revised and updated
172. Fields, B.D., Olive, K.A.: The evolution of ^{6}Li in standard cosmic-ray nucleosynthesis. New Astron. **4**, 255 (1999)
173. Fields, B.D., Sarkar, S.: Big-bang nucleosynthesis. New J. Phys. **G 33**, 220 (2006)
174. Fields, B.D., Olive, K.A., Vangioni-Flam, A.: Implications of a new temperature scale for halo dwarfs on LiBeB and chemical evolution. Astrophys. J. **623**, 1083 (2005)
175. Finkbeiner, D.P., Davis, M., Schlegel, D.J.: Extrapolation of galactic dust emission at 100 microns to cosmic microwave background radiation frequencies using FIRAS. Astrophys. J. **524**, 867–886 (1999)
176. Finoguenov, A., Reiprich, T.H., Böhringer, H.: Details of the mass-temperature relation for clusters of galaxies. Astron. Astrophys. **368**, 749–759 (2001)
177. Fixsen, D.J., Mather, J.C.: The spectral results of the far-infrared absolute spectrophotometer instrument on COBE. Astrophys. J. **581**, 817–822 (2002)
178. Fixsen, D.J., et al.: The cosmic microwave background spectrum from the full COBE FIRAS data set. Astrophys. J. **473**, 576–587 (1996)
179. Font, A.S., et al.: Halo substructure and disk heating in a Λ cold dark matter universe. Astrophys. J. Lett. **563**, L1 (2001)
180. Fort, B., Mellier, Y.: Arc(let)s in clusters of galaxies. Astron. Astrophys. Rev. **5**, 239–292 (1994)
181. Fosalba, P., et al.: Statistical properties of galactic starlight polarization. Astrophys. J. **564**, 762F (2002)
182. Freedman, W.L., et al.: Final results from the hubble space telescope key project to measure the hubble constant. Astrophys. J. **553**, 47 (2001)
183. Frieman, J.A., Turner, M.S., Huterer, D.: Dark energy and the accelerating universe. Ann. Rev. Astron. Astrophys. **46**, 385 (2008)
184. Fryer, C.L., Woosley, S.E., Heger, A.: Pair-instability supernovae, gravity waves, and gamma-ray transients. Astrophys. J. **550**, 372 (2001)
185. Fu, L., et al.: Very weak lensing in the CFHTLS wide: Cosmology from cosmic shear in the linear regime. Astron. Astrophys. **479**, 9–25 (2008)
186. Fukugita, M., Kawasaki, M.: Primordial helium abundance: A reanalysis of the Izotov-Thuan spectroscopic sample. Astrophys. J. **646**, 691 (2006)
187. Fusco-Femiano, R., et al.: Hard X-ray radiation in the coma cluster spectrum. Astrophys. J. Lett. **513**, L21–L24 (1999)
188. Fusi-Pecci, F., Cacciari, C.: On the ages of galactic globular clusters. In: New windows to the universe. Sanchez, F., Vasquez, M. (eds.), Cambridge University Press, Cambridge, p. 364 (1990)
189. GAIA Science Advisory Group, GAIA: Composition, formation and evolution of the galaxy (2006), http://www.rssd.esa.int/GAIA
190. Galama, T.J., et al.: An unusual supernova in the error box of the γ-ray burst of 25 April 1998. Nature **395**, 670 (1998)
191. Gallart, C., Zoccali, M., Aparicio, A.: The adequacy of stellar evolution models for the interpretation of the color-magnitude diagrams of resolved stellar populations. Ann. Rev. Astron. Astrophys. **43**, 387 (2005)
192. Ganga, K., et al.: The amplitude and spectral index of the large angular scale anisotropy in the cosmic microwave background ration. Astrophys. J. Lett. **432**, L15–L18 (1994)

193. Gebhardt, K., et al.: Black hole mass estimates from reverberation mapping and from spatially resolved kinematics. Astrophys. J. **543**, L5 (2000)
194. Gibson, B.K., et al.: The hubble space telescope key project on the extragalactic distance scale. XXV. A recalibration of cepheid distances to type IA supernovae and the value of the hubble constant. Astrophys. J. **529**, 723 (2000)
195. Gilli, R., Comastri, A., Hasinger, G.: The synthesis of the cosmic X-ray background in the chandra and XMM-Newton era. Astron. Astrophys. **463**, 79 (2007)
196. Gioia, I.M., et al.: The extended medium sensitivity survey. I - X-ray data and analysis. Astrophys. J. Suppl. **72**, 567–619 (1990)
197. Gioia, I.M., et al.: The extended medium sensitivity survey distant cluster sample - X-ray cosmological evolution. Astrophys. J. Lett. **356**, L35–L38 (1990)
198. Gioia, I.M., et al.: The ROSAT north ecliptic pole survey: The optical identifications. Astrophys. J. Suppl. **149**, 29–51 (2003)
199. Giovannini, G., Tordi, M., Feretti, L.: Radio halo and relic candidates from the NRAO VLA Sky Survey. New Astron. Rev. **4**, 141–155 (1999)
200. Girardi, M., et al.: Optical mass estimates of galaxy clusters. Astrophys. J. **505**, 74–95 (1998)
201. Girardi, M., et al.: Optical luminosities and mass-to-light ratios of nearby galaxy clusters. Astrophys. J. **530**, 62–79 (2000)
202. Girardi, L., et al.: Theoretical isochrones in several photometric systems. I. Johnson-Cousins-Glass, HST/WFPC2, HST/NICMOS, Washington, and ESO imaging survey filter sets. Astron. Astrophys. **391**, 195 (2002)
203. Gitti, M., et al.: Discovery of diffuse radio emission at the center of the most X-ray-luminous cluster RX J1347.5-1145. Astron. Astrophys. **470**, L25–L28 (2007)
204. Gladders, M.D., Yee, H.K.C.: The red-sequence cluster survey. I. The survey and cluster catalogs for patches RCS 0926+37 and RCS 1327+29. Astrophys. J. Suppl. **157**, 1–29 (2005)
205. Gladders, M.D., et al.: Cosmological constraints from the red-sequence cluster survey. Astrophys. J. **655**, 128–134 (2007)
206. Gnedin, N.Y.: Effect of reionization on structure formation in the universe. Astrophys. J. **542**, 535 (2000)
207. Gold, B., et al.: Five-year Wilkinson microwave anisotropy probe (WMAP) observations: Galactic foreground emission (2008) [arXiv0803.0715G]
208. Gonzalez, A.H., et al: The Las Campanas distant cluster survey: The catalog. Astrophys. J. Suppl. **137**, 117–138 (2001)
209. Governato, F., et al.: Forming disc galaxies in ΛCDM simulations. Mon. Not. R. Astron. Soc. **374**, 1479 (2007)
210. Govoni, F., et al.: Radio and X-ray diffuse emission in six clusters of galaxies. Astron. Astrophys. **376**, 803–819 (2001)
211. Govoni, F., et al.: Chandra temperature maps for galaxy clusters with radio halos. Astrophys. J. **605**, 695–708 (2004)
212. Govoni, F., et al.: A2255: The first detection of filamentary polarized emission in a radio halo. Astron. Astrophys. **430**, L5–L8 (2005)
213. Granato, G.L., et al.: Joint formation of QSOs and spheroids: QSOs as clocks of star formation in spheroids. Mon. Not. R. Astron. Soc. **324**, 757 (2001)
214. Gratton, R.G., Ortolani, S.: Metal abundances in metal-poor globular clusters. Astron. Astrophys. **211**, 41 (1989)
215. Gratton, R.G., Sneden, C., Carretta, E.: Abundance variations within globular cluster. Ann. Rev. Astron. Astrophys. **42**, 385 (2004)
216. Grebel, E.K.: Stellar populations in the local group of galaxies. In: Mikolajewska, J., Olech, A. (eds.) Stellar astrophysics with the world's largest telescopes. AIP Conf. Proc. **752**, 161 (2005)
217. Green, E.M., Demarque, P.: Revised yale isochrones and luminosity functions. VizieR online data catalog VI/40. Yale University Observatory, New Haven, Connecticut (1987)
218. Green, P.J., Forster, K., Kuraszkiewicz, J.: Quasar evolution and the Baldwin effect in the large bright quasar survey. Astrophys. J. **556**, 727–737 (2001)

219. Greggio, L.: The rates of type Ia supernovae. I. Analytical formulations. Astron. Astrophys. **441**, 1055 (2005)
220. Greggio, L., Renzini, A.: The binary model for type I supernovae – theoretical rates. Astron. Astrophys. **118**, 217 (1983)
221. Grego, L., et al.: Galaxy cluster gas mass fractions from Sunyaev-Zel'dovich effect measurements: Constraints on Ω_M. Astrophys. J. **552**, 2–14 (2001)
222. Haehnelt, M., Davies, M.B., Rees, M.J.: Possible evidence for the ejection of a supermassive black hole from an ongoing merger of galaxies. Mon. Not. R. Astron. Soc. **366**, 22 (2006)
223. Haiman, Z., Holder, G.P.: The reionization history at high redshifts. I. Physical models and new constraints from cosmic microwave background polarization. Astrophys. J. **595**, 1 (2003)
224. Haislip, J., et al.: A photometric redshift of $z = 6.39 \pm 0.12$ for GRB-050904. Nature **440**, 181 (2006)
225. Hamann, F., Ferland, G.J.: Elemental abundances in quasistellar objects: Star formation and galactic nuclear evolution at high redshifts. Ann. Rev. Astr. Astrophys. **37**, 487 (1999)
226. Hamuy, M.: Observed and physical properties of core-collapse supernovae. Astrophys. J. **582**, 905 (2003)
227. Han, J.L, et al.: Pulsar rotation measures and the large-scale structure of the galactic magnetic field. Astrophys. J. **642**, 868H (2006)
228. Hanany, S. et al.: MAXIMA-1: A measurement of the cosmic microwave background anisotropy on angular scales of 10'-5 degrees. Astrophys. J. Lett. **545**, L1–L5 (2000)
229. Harris, W.E.: A catalog of parameters for globular clusters in the milky way. Astron. J. **112**, 1487 (1996)
230. Hasinger, G., Miyaji, T., Schmidt, M.: Luminosity-dependent evolution of soft X-ray selected AGN, New chandra and XMM-Newton surveys. Astron. Astrophys. **441**, 417 (2005)
231. Hauser, M.G., Dwek, E.: The cosmic infrared background: Measurements and implications. Ann. Rev. Astron. Astrophys. **39**, 249 (2001)
232. Hayashi, E., et al.: The structural evolution of substructure. Astrophys. J. **584**, 541 (2003)
233. Heiles, C. : 9286 stars: An agglomeration of stellar polarization catalogs. Astron. J. **119**, 923–927 (2000)
234. Henry, P.S.: Isotropy of the 3 K background. Nature **231**, 516 (1971)
235. Henry, J.P.: A measurement of the density parameter derived from the evolution of cluster X-ray temperatures. Astrophys. J. Lett. **489**, L1–L5 (1997)
236. Henry, J.P.: X-ray temperatures for the extended medium-sensitivity survey high-redshift cluster sample: Constraints on cosmology and the dark energy equation of state. Astrophys. J. **609**, 603–616 (2004)
237. Henry, J.P., Arnaud, K.A.: A measurement of the mass fluctuation spectrum from the cluster X-ray temperature function. Astrophys. J. **372**, 410–418 (1991)
238. Henry, J.P., et al.: The extended medium sensitivity survey distant cluster sample - X-ray Data dnd interpretation of the luminosity evolution. Astrophys. J. **386**, 408–419 (1992)
239. Hernández-Monteagudo, C., et al.: Spectral indications of thermal Sunyaev-Zel'dovich effect in Arheops and WMAP data. Astron. Astrophys. **449**, 41–48 (2006)
240. Hernquist, L.: Structure of merger remnants. I - Bulgeless progenitors. Astrophys. J. **400**, 460 (1992)
241. Hinshaw, G., et al.: Three-year Wilkinson microwave anisotropy probe (WMAP) observations: Temperature analysis. Astrophys. J. Suppl. **170**, 288 (2007) [astro-ph/0603451]
242. Hobson, M., et al.: Foreground separation methods for satellite observations of the cosmic microwave background. Mon. Not. R. Astron. Soc. **300**, 1H (1999)
243. Hoekstra, H.: A comparison of weak-lensing masses and X-ray properties of galaxy clusters. Mon. Not. R. Astron. Soc. **379**, 317–330 (2007)
244. Hoekstra, H., Yee, H.K.C., Gladders, M.D.: Current status of weak gravitational lensing. New Astron. Rev. **46**, 767–781 (2002)
245. Hopkins, A.M., Beacom, J.F.: On the normalization of the cosmic star formation history. Astrophys. J. **651**, 142 (2006)

246. Howell, D.A.: Predicted and observed evolution in the mean properties of type Ia supernovae with redshift. Astrophys. J. **667**, L37 (2007)
247. Huchre, J.: The Hubble Constant. Science **256**, 5055 (1992)
248. Hutchings, J.B.: QSO hosts: The broad perspective. New Astron. Rev. **50**, 685–688 (2006)
249. IAU254: The galaxy disk in cosmological context. (2008) http://iau254.nbi.dk/index.html
250. Im, M., et al.: Luminosity functions of elliptical galaxies at Z ≤ 1.2. Astrophys. J. Lett. **461**, L79 (1996)
251. IRAS satellite web site, http://irsa.ipac.caltech.edu/IRASdocs/iras.html
252. Iwamoto, K., et al.: A hypernova model for the supernova associated with the γ-ray burst of 25 April 1998. Nature **395**, 672 (1998)
253. Iye, M., et al.: A galaxy at a redshift $z = 6.96$. Nature **443**, 186 (2006)
254. Izotov, Y.I., Thuan, T.X.: Systematic Effects and a New Determination of the Primordial Abundance of ^4He and dY/dZ from Observations of Blue Compact Galaxies. Astrophys. J. **602** 200 (2004) [astro-ph/0310421]
255. Jaffe, A.H., et al.: Cosmology from MAXIMA-1, BOOMERANG, and COBE DMR cosmic microwave background observations. Phys. Rev. Lett. **86**, 3475 (2001)
256. Jeffery, E.J., et al.: New techniques to determine ages of open clusters using white dwarfs. In: 15th European workshop on white dwarfs. Napiwotzki, R., Burleigh, M.R. (eds.), ASP Conf. **372**, 97 (2007)
257. Jenkins, A., et al.: The mass function of dark matter haloes. Mon. Not. R. Astron. Soc. **321**, 372–384 (2001)
258. Jha, S., Riess, A.G., Kirshner, R.P.: Improved distances to type Ia supernovae with multicolor light-curve shapes: MLCS2k2. Astrophys. J. **659**, 122 (2007)
259. Johnson, J.L., Greif, T.H., Bromm, V.: The first stars (2008) [arXiv:0802.0207v1]Au: Please provide complete details for the reference [259].
260. Jones, M.E., et al.: H_0 from an orientation-unbiased sample of Sunyaev-Zel'dovich and X-ray clusters. Mon. Not. R. Astron. Soc. **357**, 518 (2005)
261. Jones, W.W., et al.: A measurement of the angular power spectrum of the CMB temperature anisotropy from the 2003 flight of boomerang. Astrophys. J. **647**, 823–832 (2006)
262. Kaaret, P., et al.: Chandra high-resolution camera observations of the luminous X-ray source in the starburst galaxy M82. Mon. Not. R. Astron. Soc. **321**, L29 (2001)
263. Kaiser, N.: Evolution and clustering of rich clusters. Mon. Not. R. Astron. Soc. **222**, 323–345 (1986)
264. Kaiser, N., Squires, G.: Mapping the dark matter with weak gravitational lensing. Astrophys. J. **404**, 441–450 (1993)
265. Kang, H.-S., Steigman, G.: Cosmological constraints on neutrino degeneracy. Nucl. Phys. **B 372**, 494 (1992)
266. Kanipe, J., Webb, D.: The arp atlas of peculiar galaxies: A chronicle and observers guide. Willman Bell (2006)
267. Kashlinsky, A.: Microwave background anisotropies implied by large-scale galaxy correlations – The minimum of C(0) and cosmological parameters. Astrophys. J. Lett. **383**, L1–L5 (1991)
268. Kaspi, S., et al.: Reverberation mapping of high-luminosity quasars: First results. Astrophys. J. **659**, 997–1007 (2007)
269. Kauffmann, G., Haehnelt, M.G.: A unified model for the evolution of galaxies and quasars. Mon. Not. R. Astron. Soc. **311**, 576 (2000)
270. Kauffmann, G., White, S.D.M., Guiderdoni, B.: The formation and evolution of galaxies within merging dark matter haloes. Mon. Not. R. Astron. Soc. **264**, 201 (1993)
271. Kauffmann, G., Charlot, S., White, S.D.M.: Detection of strong evolution in the population of early-type galaxies. Mon. Not. R. Astron. Soc. **283**, L117 (1996)
272. Kawamura, A.: Molecular clouds and star formation in the magellanic clouds and the milky way. Astrophys. Space Sci. **313**, 145–151 (2008)
273. Kennicutt, R.C.: Star formation in galaxies along the hubble sequence. Ann. Rev. Astron. Astrophys. **36**, 189 (1998)

274. Khalatyan, A., et al.: Is AGN feedback necessary to form red elliptical galaxies? Mon. Not. R. Astron. Soc. **387**, 13 (2008)
275. Khokhlov, A., Mueller, E., Hoeflich, P.: Light curves of type IA supernova models with different explosion mechanisms. Astron. Astrophys. **270**, 223 (1993)
276. Kirchhoff, G.: Über das Verhältnis zwischen dem Emissionsvermögen und dem Absorptionsvermögen der Körper für Wärme und Licht. Annalen der Physik, **109**, 275–301 (1860)
277. Kirkman, D., et al.: The cosmological baryon density from the deuterium-to-hydrogen ratio in QSO absorption systems: D/H toward Q1243+3047. Astrophys. J. Suppl. **149**, 1 (2003)
278. Kirshner, R.P., Kwan, J.: Distances to extragalactic supernovae. Astrophys. J. **193**, 27 (1974)
279. Klypin, A.A., Strukov, I.A., Skulachev, D.P.: The relikt missions: Results and prospects for detection of the microwave background anisotropy. Mon. Not. Astr. Soc. **258**, 71–81 (1992)
280. Klypin, A., et al.: Where are the missing galactic satellites? Astrophys. J. **522**, 82 (1999)
281. Kneller, J.P., Steigman, G.: BBN for pedestrians. New J. Phys. **6**, 117 (2004)
282. Kobayashi, C., Arimoto, N.: Gradients of absorption-line strengths in elliptical galaxies. Astrophys. J. **527**, 573 (1999)
283. Kochanek, C.S.: Strong gravitational lensing. In: Gravitational lensing: Strong, weak and micro. Saas-Fee Advanced Course 33. Meylan, G., Jetzer, P., North, P., Schneider, P., Kochanek, C.S., Wambsganss, J. (eds.), p. 91 (2006)
284. Köckert, F., Steinmetz, M.: Simulating disk galaxies: First results of a systematical study. In: IAU Symp. 235, p. 114 (2007)
285. Kodama, T., Arimoto, N.: Origin of the colour-magnitude relation of elliptical galaxies. Astron. Astrophys. **320**, 41 (1997)
286. Kodama, T., et al.: Down-sizing in galaxy formation at $z \sim 1$ in the Subaru/XMM-Newton deep survey (SXDS). Mon. Not. R. Astron. Soc. **350**, 1005 (2004)
287. Kogut, A., et al.: COBE differential microwave radiometers - preliminary systematic error analysis. Astrophys. J. **401**, 1 (1992)
288. Kogut, A., et al.: Microwave emission at high galactic latitudes in the four-year DMR sky maps. Astrophys. J. **464L**, 5K (1996)
289. Kollatschny, W., Bischoff, K., Dietrich, M.: Strong spectral variability in NGC 7603 over 20 years. Astron. Astrophys. **361**, 901 (2000)
290. Komatsu, E., et al.: Five-year Wilkinson microwave anisotropy probe (WMAP) observations: Cosmological interpretation. APJS **180**, 330 (2009)
291. Komossa, S., et al.: Discovery of a binary active galactic nucleus in the ultraluminous infrared galaxy NGC 6240 using chandra. Astrophys. J. **582**, L15 (2003)
292. Korn, A.J., et al.: A probable stellar solution to the cosmological lithium discrepancy. Nature **442**, 657 (2006)
293. Kowalski, M., et al.: Improved cosmological constraints from new, old and combined supernova datasets. APJ **686**, 749 (2008)
294. Kravtsov, A.V., Nagai, D., Vikhlinin, A.: Effects of cooling and star formation on the baryon fractions in clusters. Astrophys. J. **625**, 588–598 (2005)
295. Kravtsov, A.V., Vikhlinin, A., Nagai, D.: A new robust low-scatter X-ray mass indicator for clusters of galaxies. Astrophys. J. **650**, 128–136 (2006)
296. Krisciunas, K., Phillips, M.M., Suntzeff, N.B.: Hubble diagrams of type Ia supernovae in the near-infrared. Astrophys. J. Lett. **602**, L81 (2004)
297. Krisciunas, K., et al.: The type Ia supernova 2004S, a clone of SN 2001el, and the optimal photometric bands for extinction estimation. Astron. J. **133**, 58 (2006)
298. Kudritzki, R.P.: Wind models and ionizing fluxes of massive stars at very low metallicity. In: The first stars. Weiss, A., Abel, T., Hill, V. (eds.), Springer, Heidelberg, p. 127 (2000)
299. Kuhlen, M., Madau, P.: The first miniquasar. Mon. Not. R. Astron. Soc. **363**, 1069 (2005)
300. Kuntschner, H.: The stellar populations of early-type galaxies in the Fornax cluster. Mon. Not. R. Astron. Soc. **315**, 184 (2000)
301. Kuntschner, H.: The stellar populations of early-type galaxies in the fornax cluster. Astrophys. Space Sci. **276**, 885 (2001)
302. Kuntschner, H., Davies, R.L.: The ages and metallicities of early-type galaxies in the fornax cluster. Mon. Not. R. Astron. Soc. **295**, L29 (1998)

303. La Porta, L.: The impact of galactic synchrotron emission on CMB anisotropy measurements. PhD thesis. Universities of Bonn and Cologne (2007)
304. La Porta, L., et al.: The impact of Galactic synchrotron emission on CMB anisotropy measurements. I. Angular power spectrum analysis of total intensity all-sky surveys. Astron. Astrophys. **479**, 641 (2008)
305. Lachieze-Rey, M., Gunzig, E.: The cosmological background radiation: Echo of the early universe. Cambridge University Press, pp. 158–161 (1999)
306. Landsberg, P.T.: Thermodynamics, interscience, NY, pp. 250–286 (1961)
307. Langley, S.P.: Experimental determination of wave-lengths in the invisible spectrum. Mem. Natl. Acad. Sci. **2**, 147–162 (1882)
308. Langley, S.P.: On hitherto unrecognized wave-lengths. Phil. Mag. **22**, 149–173 (1886)
309. LaRoque, S.J., et al.: X-ray and Sunyaev-Zel'dovich effect measurements of the gas mass fraction in galaxy clusters. Astrophys. J. **652**, 917–936 (2006)
310. Larson, R.B.: Dynamical models for the formation and evolution of spherical galaxies. Mon. Not. R. Astron. Soc. **166**, 585 (1974)
311. Larson, R.B.: Galaxy building. Pub. Amer. Soc. Pacific **102**, 709 (1990)
312. Leach, S., et al.: Component separation methods for the planck mission. Astron. Astrophys. **491**, 597 (2008)
313. Lee, Y.W.: Stellar evolution and period changes in RR Lyrae stars. Astrophys. J. **367**, 524 (1991)
314. Lee, A.T., et al.: A high resolution analysis of the MAXIMA-1 cosmic microwave background anisotropy data. Astrophys. J. **561**, L1–L6 (2001)
315. Leibundgut, B.: Cosmological implications from observations of type Ia supernovae. Ann. Rev. Astron. Astrophys. **39**, 67 (2001)
316. Leibundgut, B.: Supernovae and cosmology. Gen. Relat. Gravit. **40**, 221 (2008)
317. Leitch, E.M., et al.: An anomalous component of galactic emission. Astrophys. J. **486L**, 23L (1997)
318. Li, Y., et al.: Formation of $z \sim 6$ quasars from hierarchical galaxy mergers. Astrophys. J. **665**, 187 (2007)
319. Lieu, R., Mittaz J.: Observational constraints on models for the cluster soft excess emission. In: Soft X-ray emission from clusters of galaxies and related phenomena. Kluwer, Dordrecht, The Netherlands, p. 155 (2004)
320. Lieu, M., Shuang-Nan, Z.: The Sunyaev-Zel'dovich effect in a sample of 31 clusters: A comparison between the X-ray predicted and WMAP observed decrement. Astrophys. J.**648**, 176 (2006); comment available at: Big Bang's Afterglow Fails an Intergalactic Shadow Test, Physorg.com; http://www.physorg.com/news76314500.html
321. Longhetti, M., et al.: Star formation history of early-type galaxies in low density environments. IV. What do we learn from nuclear line-strength indices? Astron. Astrophys. **353**, 917 (2000)
322. Lopez-Corredoira, M., Gutierrez, C.M.: Two emission line objects with $z > 0.2$ in the optical filament apparently connecting the Seyfert galaxy NGC 7603 to its companion. Astron. Astrophys. **390**, L15 (2002)
323. Lopez-Corredoira, M., Gutierrez, C.M.: The field surrounding NGC 7603: Cosmological or non-cosmological redshifts? Astron. Astrophys. **421**, 407 (2004)
324. LoSecco, J.M., Mathews, G.J., Wang, Y.: Prospects of constraining cosmology with the extragalactic cosmic microwave background temperature. Phys. Rev. **D 64**, 123002 (2001)
325. Machacek, M.E., Bryan, G.L., Abel, T.: Effects of a soft X-ray background on structure formation at high redshift. Mon. Not. R. Astron. Soc. **338**, 273 (2003)
326. Macías-Pérez, J.F., et al.: Archeops in-flight performance, data processing, and map making. Astron. Astrophys. **467**, 1313 (2007)
327. Macri, L.M., et al.: A new cepheid distance to the maser-host galaxy NGC 4258 and its implications for the hubble constant. Astrophys. J. **652**, 1133 (2006)
328. MacTavish, C.J.: Cosmological parameters from the 2003 flight of Boomerang. Astrophys. J. **647**, 799–812 (2006)

329. Madau, P., et al.: High-redshift galaxies in the hubble deep field: Colour selection and star formation history to $z \sim 4$. Mon. Not. R. Astron. Soc. **283**, 1388 (1996)
330. Madau, P., Haardt, F., Rees, M.J.: Radiative transfer in a clumpy universe. III. The nature of cosmological ionizing sources. Astrophys. J. **514**, 648 (1999)
331. Madau, P., Diemand, J., Kuhlen, M.: Dark matter subhalos and the dwarf satellites of the milky way. Astrophys. J. **679**, 1260 (2008) [arXiv0802.2265]
332. Magain, P.: Discovery of a bright quasar without a massive host galaxy. Nature **437**, 381 (2005)
333. Mahdavi, A., et al.: A dark core in Abell 520. Astrophys. J. **668**, 806–814 (2007)
334. Maino, D., et al.: All-sky astrophysical component separation with fast independent component analysis (FASTICA). Mon. Not. R. Astron. Soc. **334**, 53 (2002)
335. Maiolino, R., et al.: Metals and dust in high redshift AGNs. Mem. Soc. Astron. It **77**, 643 (2006)
336. Mannucci, F., et al.: The supernova rate per unit mass. Astron. Astrophys. **433**, 807 (2005)
337. Mannucci, F., Della Valle, M., Panagia, N.: Two populations of progenitors for type Ia supernovae? Mon. Not. R. Astron. Soc. **370**, 773 (2006)
338. Mantz, A., et al.: New constraints on dark energy from the observed growth of the most X-ray luminous galaxy clusters. Mou. Not. Astr. Soc. **387**, 1179 (2008)
339. Markevitch, M.: The $L_X - T$ relation and temperature function for nearby clusters revisited. Astrophys. J. **504**, 27–34 (1998)
340. Marziani, P., et al.: Using quasars for cosmology. Mem. Soc. Astron. It. Suppl. **3**, 218–211 (2003)
341. Marziani, P., Dultzin-Hacyan, D., Sulentic, J.W.: Accretion onto supermassive black holes in quasars: Learning from optical/UV observations. In: New developments in black hole research. P.V. Kreitler (ed.), Nova Science Publisher, New York, pp. 123–183 (2006)
342. Masi, S., et al.: Instrument, method, brightness, and polarization maps from the 2003 flight of BOOMERanG. Astron. Astrophys. **458**, 687 (2006)
343. Mateo, M.L.: Dwarf galaxies of the local group. Ann. Rev. Astron. Astrophys. **36**, 435 (1998)
344. Mateo, M.: The stellar populations in the local group. Extragalactic binaries. 25th meeting of the IAU, Joint Discussion **13**, 13 (2003)
345. Mather, J.C., et al.: Calibrator design for the COBE far-infrared absolute spectrometer (FIRAS). Astrophys. J. **512**, 511–520 (1999)
346. Matteucci, F.: Galaxy evolution. Fundam. Cosmic Phys. **17**, 283 (1997)
347. Matteucci, F., Padovani, P.: Chemical evolution of galaxies and quasar metallicities. Astrophys. J. **419**, 485 (1993)
348. Matteucci, F., Recchi, S.: On the typical timescale for the chemical enrichment from type Ia supernovae in galaxies. Astrophys. J. **558**, 351 (2001)
349. Maughan, B.J.: The $L_X - Y_X$ relation: Using galaxy cluster X-ray luminosity as a robust, low-scatter mass proxy. Astrophys. J. **668**, 772–780 (2007)
350. Maughan, B.J., et al.: Images, structural properties, and metal abundances of galaxy clusters observed with chandra ACIS-I at $0.1 < z < 1.3$. Astrophys. J. Suppl. **174**, 117–135 (2008)
351. McKellar, A.: The problems of possible molecular identification for interstellar lines. Pub. Americ. Soc. Pacific **53**, 233 (1941)
352. McLure, R.J., Jarvis, M.J.: Measuring the black hole masses of high-redshift quasars. Mon. Not. R. Astron. Soc. **337**, 109–116 (2002)
353. Melendez, J., Ramirez, I.: Reappraising the spite lithium plateau: Extremely thin and marginally consistent with WMAP. Astrophys. J. Lett. **615**, L33 (2004)
354. Mellier, Y.: Probing the universe with weak lensing. Ann. Rev. Astron. Astrophys. **37**, 127 (1999)
355. Menanteau, F., Jimenez, R., Matteucci, F.: The origin of blue cores in hubble deep field E/S0 galaxies. Astrophys. J. Lett. **562**, L23 (2001)
356. Menou, K., Haiman, Z., Narayanan, V.K.: The merger history of supermassive black holes in galaxies. Astrophys. J. **558**, 535 (2001)
357. Meyers, R.: The cosmology quest: A floating world films. http://www.universe-film.com

358. Meza, A., et al.: Simulations of galaxy formation in a ΛCDM universe. III. The dissipative formation of an elliptical galaxy. Astrophys. J. **590**, 619 (2003)
359. Milone, A.P., et al.: The ACS survey of galactic globular clusters. III. The double subgiant branch of NGC 1851. Astrophys. J., **673**, 241 (2008)
360. Mirabel, I.: Phenomenological analogies in black hole systems of all masses. Prog. of Theor. Phys. Suppl. **155**, 71–82 (2004)
361. Miville-Deschenes, M.-A., et al.: Separation of anomalous and synchrotron emissions using WMAP polarization data. Astron. Astrophys. **490**, 1093 (2008)
362. Mo, H.J., Mao, S., White, S.D.M.: The formation of galactic discs. Mon. Not. R. Astron. Soc. **295**, 319 (1998)
363. Mobasher, B., et al.: Evidence for a massive poststarburst galaxy at $z \sim 6.5$. Astrophys. J., **635**, 832 (2005)
364. Montroy, T.E., et al.: A measurement of the CMB (EE) spectrum from the 2003 flight of boomerang. Astrophys. J. **647**, 813–822 (2006)
365. Moore, B., et al.: Dark matter substructure within galactic halos. Astrophys. J. **524**, L19 (1999)
366. Moran, E.C., et al.: The nuclear spectral energy distribution of NGC 4395, the least luminous type 1 seyfert galaxy. Pub. Astron. Soc. Pacific **111**, 801–808 (1999)
367. Muanwong, O., et al.: The effect of cooling and preheating on the X-ray properties of clusters of galaxies. Mon. Not. R. Astron. Soc. **336**, 527–540 (2002)
368. Muchovej, S., et al.: Observations of high-redshift X-ray selected clusters with the Sunyaev-Zel'dovich array. Astrophys. J. **663**, 708 (2007)
369. Mullis, C.R., et al.: Evolution of the cluster X-ray luminosity function. Astrophys. J. **607**, 175–189 (2004)
370. Mullis, C.R., et al.: Discovery of an X-ray-luminous galaxy cluster at $z = 1.4$. Astrophys. J. Lett. **623**, L85–L88 (2005)
371. Narayan, R., Bartelmann, M.: Gravitational lensing. In: Formation of structure in the niverse. Dekel, A., Ostriker, J.P. (eds.), p. 360 (1999)
372. Navarro, J.F., Steinmetz, M.: The effects of a photoionizing ultraviolet background on the formation of disk galaxies. Astrophys. J. **478**, 13 (1997)
373. Navarro, J.F., Steinmetz M.: The core density of dark matter halos: A critical challenge to the ΛCDM paradigm? Astrophys. J. **528**, 607 (2000)
374. Nissen, P.E., Edvardsson, B., Gustafsson, B.: Oxygen and α-element abundances in galactic disk stars as a function of stellar age. In: Production and distribution of C, N and O elements. Danziger, I.J., Matteucci, F., Kjar, K. (eds.), p. 131, ESO, Garching (1985)
375. Nobili, S., et al.: Restframe I-band hubble diagram for type Ia supernovae up to redshift $z \sim 0.5$. Astron. Astrophys. **437**, 789 (2005)
376. Oguri, M.: Gravitational lens time delays: A statistical assessment of lens model dependences and implications for the global hubble constant. Astrophys. J. **660**, 1 (2007)
377. Olive, K.A.: Big bang nucleosynthesis. Nucl. Phys. Suppl. **B 80**, 79 (2000)
378. Olive, K.A.: Big bang nucleosynthesis in the post-WMAP era. In: The new cosmology. Allen, R.E., Nanopoulos, D.V., Pope, C.N. (eds.), Conference on Strings and Cosmology, AIPC, **743**, p. 190 (2004)
379. Olive, K.A., Skillman, E.D.: On the determination of the He abundance in extragalactic H II regions. New Astron. **6**, 119 (2001)
380. Olive, K.A., Skillman, E.D.: A realistic determination of the error on the primordial helium abundance: Steps toward nonparametric nebular helium abundances. Astrophys. J. **617**, 29 (2004)
381. Olive, K.A., Steigman, G., Skillman, E.D.: The primordial abundance of 4He: An update. Astrophys. J. **483**, 788 (1997)
382. Olive, K.A., Steigman, G., Walker, T.P.: Primordial nucleosynthesis: Theory and observations. Phys. Report **333** 389 (2000) [astro-ph/9905320]
383. O'Meara, J.M., et al.: The deuterium to hydrogen abundance ratio towards the QSO SDSS1558-0031. Astrophys. J. Lett. **649**, L61 (2006)

384. Oukbir, J., Blanchard, A.: X-ray clusters in open universes. Astrophys. J. Lett. **262**, L21–L24 (1992)
385. Page, L., et al.: First year Wilkinson microwave anisotropy probe (WMAP) observations: Interpretation of the TT and TE angular power spectrum peaks. Astrophys. J. Suppl. **148**, 233 (2003)
386. Page, L., et al.: Three-year Wilkinson microwave anisotropy probe (WMAP) observations: Polarization analysis. Astrophys. J. Suppl. **170**, 335 (2007) [astro-ph/0603450]
387. Pancino, E., et al.: The multiple stellar population in ω Centauri: Spatial distribution and structural properties. Mon. Not. R. Astron. Soc. **345**, 683 (2003)
388. Partridge, R.B.: 3K: The cosmic microwave background radiation. Cambridge University Press, Cambridge, pp. 103–160, (1995)
389. Pastorello, A. et al.: Low-luminosity type II supernovae: Spectroscopic and photometric evolution. Mon. Not. R. Astron. Soc. **347**, 74 (2004)
390. Patton, D.R., et al.: Close pairs of field galaxies in the CNOC1 redshift survey. Astrophys. J. **475**, 29 (1997)
391. Peacock, J.A.: Cosmological physics, Cambridge University Press, Cambridge (1999)
392. Peebles, P.J.E.: The large-scale structure of the universe. Princeton University Press, Princeton (1980)
393. Peebles, P.J.E.: Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. Astrophys. J. Lett. **263**, L1 (1982)
394. Peebles, P.J.E.: Physical cosmology. Princeton University Press, Princeton, NJ (1993)
395. Peebles, P.J.E.: When did the large elliptical galaxies form?. In: A new era in cosmology. Metcalfe, N., Shanks, T. (eds.), ASPC, **283**, 351 (2002)
396. Peebles, P.J.E., Ratra, B.: The cosmological constant and dark energy. Rev. Mod. Phys, **75**, 569–606 (2003)
397. Peimbert, M., Peimbert, A., Ruiz, M.T.: The chemical composition of the small magellanic cloud H II region NGC 346 and the primordial helium abundance. Astrophys. J. **541**, 688 (2000)
398. Peimbert, A., Peimbert, M., Luridiana, V.: Temperature bias and the primordial helium abundance determination. Astrophys.J. **565**, 668 (2002)
399. Peimbert, M., Luridiana, V., Peimbert, A.: Revised primordial helium abundance based on new atomic data. Astrophys. J. **666**, 636 (2007)
400. Penzias, A.A., Wilson, R.W.: A measurement of excess antenna temperature at 4080 Mc/s. Astrophys. J. **142**, 419 (1965)
401. Perlmutter, S., Schmidt, B.: Measuring cosmology with supernovae. In: Supernovae and gamma ray bursts. K. Weiler (ed.). Lecture Note in Physics **598**, 195 (2003)
402. Perlmutter, S., et al.: Measurements of the cosmological parameters omega and lambda from the first seven supernovae at $z \geq 0.35$. Astrophys. J. **483**, 565 (1997)
403. Perlmutter, S.J., et al.: Measurements of Ω and Λ from 42 high-redshift supernovae. Astrophys. J. **517**, 565–586 (1999)
404. Pettini, M., et al.: Deuterium abundance in the most metal-poor damped Lyα system. Mou. Not. R. Astron. Soc. **391**, 1499 (2008)
405. Phillips, M.M.: The absolute magnitudes of type Ia SNe. Astrophys. J. Lett. **413**, L105 (1993)
406. Piacentini, F., et al.: A measurement of the polarization-temperature angular cross-power spectrum of the Cosmic Microwave Background from the 2003 flight of Boomerang. Astrophys. J. **647**, 833–839 (2006)
407. Pierpaoli, E., et al.: On determining the cluster abundance normalization. Mon. Not. R. Astron. Soc. **342**, 163–175 (2003)
408. Pietrinferni, A., et al.: A large stellar evolution database for population synthesis studies. I. Scaled solar models and isochrones. Astrophys. J., **612**, 168 (2004)
409. Pietrinferni, A., et al.: A large stellar evolution database for population synthesis studies. II. Stellar models and isochrones for an α-enhanced metal distribution. Astrophys. J. **642**, 797 (2006)
410. Pilachowski, C.A., Sneden, C., Wallerstein, G.: The chemical composition of stars in globular clusters. Astrophys. J. Suppl. **52**, 241 (1983)

411. Piotto, G.: Observational evidence of multiple stellar populations in globular clusters. IAU Symp. **246**, 141 (2008)
412. Piotto, G., et al.: HST color-magnitude diagrams of 74 galactic globular clusters in the HST F439W and F555W bands. Astron. Astrophys. **391**, 945 (2002)
413. Piotto, G., et al.: A triple main sequence in the globular cluster NGC 2808. Astrophys. J. Lett. **661**, L53 (2007)
414. Pipino, A., Matteucci, F.: Photochemical evolution of elliptical galaxies - I. The high-redshift formation scenario. Mon. Not. R. Astron. Soc. **347**, 968 (2004)
415. Planck, M.: Ueber das gesetz der energieverteilung in normalspectrum. Annalen der Physik, vol. 4, 553–563 (1901)
416. Planck, M.: The theory of heat radiation. P. Blakiston's Son, Philadelphia, PA (1914)
417. Planck, M.: Eight lectures on theoretical physics – delivered at Columbia university in 1909. Columbia University Press, New York, p. 72 (1915)
418. Planck, M.: Treatise on thermodynamics. Dover Publications, New York (1945)
419. Planck, M.: Scientific autobiography, philosophical library, New York (1949)
420. PLANCK website, http://www.rssd.esa.int/index.php?project=PLANCK&page=index
421. Poggianti, B.M., et al.: A photometric and spectroscopic study of dwarf and giant galaxies in the coma cluster. III. Spectral ages and metallicities. Astrophys. J. **562**, 689 (2001)
422. Polenta, G., et al.: Search for non-gaussian signals in the BOOMERANG maps: Pixel-space analysis. Astrophys. J. Lett. **572**, L27 (2002)
423. Ponman, T.J., Sanderson, A.J.R., Finoguenov, A.: The Birmingham-CfA cluster scaling project - III. Entropy and similarity in galaxy systems. Mon. Not. R. Astron. Soc. **343**, 331–342 (2003)
424. Pont, F., Eyer, L.: Isochrone ages for field stars: Method and application to the age-metallicity relation. Mon. Not. R. Astron. Soc. **351**, 487 (2004)
425. Ponthieu, N.: Temperature and polarization angular power spectra of galactic dust radiation at 353 GHz as measured by archeops. Astron. Astrophys. **444**, 327P (2005)
426. Postman, M., et al.: The palomar distant clusters survey. I. The cluster catalog. Astron. J. **111**, 615–641 (1996)
427. Prantzos, N.: An introduction to galactic chemical evolution. EAS Pub. Ser. **32**, 311 (2008)
428. Press, W.H., Schechter, P.: Formation of galaxies and clusters of galaxies by self-similar gravitational condensation. Astrophys. J. **187**, 425–438 (1974)
429. Pritzl, B.J., Venn, K.A., Irwin, M.J.: A comparison of elemental abundance ratios in globular clusters, field stars, and dwarf spheroidal galaxies. Astron. J. **130**, 2140 (2005)
430. Proffitt, C. R., VandenBerg, D.A.: Implications of helium diffusion for globular cluster isochrones and luminosity functions. Astrophys. J. Suppl. **77**, 473 (1991)
431. Puzia, T.H., et al.: Integrated spectroscopy of bulge globular clusters and fields. I. The data base and comparison of individual Lick indices in clusters and bulge. Astron. Astrophys. **395**, 45 (2002)
432. Quimby, R.M., et al.: SN 2005ap: A most brilliant explosion. Astrophys. J. Lett. **668**, L99 (2007)
433. Rabounski, D.: The relativistic effect of the deviation between the CMB temperatures obtained by the COBE satellite. Prog. Phys. **1**, 24–26 (2007)
434. Rakos, K., Schombert, J.: Age and metallicity estimation of globular clusters from Strömgren photometry. Pub. Astron. Soc. Pacific **117**, 245 (2005)
435. Rakos, K., et al.: Ages and metallicities of fornax dwarf elliptical galaxies. Astron. J. **121**, 1974 (2001)
436. Rakos, K., Schombert, J., Odell, A.: The age of cluster galaxies from continuum colors (2008) [arXiv:0801.3665v1]
437. Rao, R., et al.: High-resolution millimeter-wave mapping of linearly polarized dust emission: Magnetic field structure in Orion. Astrophys. J. **502L**, 75R (1998)
438. Rauch, M.: The lyman alpha forest in the spectra of QSOs. Ann. Rev. Astron. Astrophys. **36**, 267 (1998)
439. Recio-Blanco, A., et al.: Distance of 72 galactic globular clusters. VizieR Online Data Catalog **343**, 20851 (2006)

440. Refregier, A.: Weak gravitational lensing by large-scale structure. Ann. Rev. Astron. Astrophys. **41**, 645 (2003)
441. Reich, W.: Galactic polarization surveys. In: Cosmic polarization. Fabbri, R. (ed.). Research Signpost, Kerala, p. 91 (2006) [astro-ph/0603465]
442. Reich, P., Reich, W., Testori, J.C.: Spectral index variations of galactic emission. In: The magnetized interstellar medium, Uyanıker, B., Reich, W., Wielebinski, R. (eds.). Copernicus GmbH, Katlenburg-Lindau, p. 63 (2004)
443. Reiprich, T.H., Böhringer, H.: The mass function of an X-ray flux-limited sample of galaxy clusters. Astrophys. J. **567** 716–740 (2002)
444. Renzini, A.: Stellar population diagnostics of elliptical galaxy formation. Ann. Rev. Astron. Astrophys. **44**, 141 (2006)
445. Renzini, A., Buzzoni, A.: Global properties of stellar populations and the spectral evolution of galaxies. In: Spectral evolution of galaxies. Chiosi, C., Renzini, A. (eds.), p. 195. Reidel, Dordrecht (1986)
446. Renzini, A., Fusi-Pecci, F.: Tests of evolutionary sequences using color-magnitude diagrams of globular clusters. Ann. Rev. Astron. Astrophys., **26**, 199 (1988)
447. Rich, M.: Age of the bulge. In: Astrophysical ages and time scales. von Hippel, T., Simpson, C., Manset, N. (eds.), ASP, San francisco, p. 216 (2001)
448. Richards, G.T., et al.: The sloan digital sky survey quasar survey: Quasar luminosity function from data release 3. Astron. J. **131**, 2766–2787 (2006)
449. Riess,A.G., Press, W.M., Kirshner, R.P.: Determining the motion of the local group using type Ia SNe light curve shapes. Astrophys. J. Lett. **445**, L91 (1995)
450. Riess, A.G., et al.: Observational evidence from supernovae for an accelerating universe and a cosmological constant. Astron. J. **116**, 1009–1038 (1998)
451. Riess, A.G., et al.: Cepheid Calibrations from the Hubble Space Telescope of the Luminosity of Two Recent Type Ia Supernovae and a Redetermination of the Hubble Constant. Astrophys. J. **627**, 579 (2005)
452. Riess, A.G., et al.: New Hubble space telescope discoveries of type Ia supernovae at $z > 1$: Narrowing constraints on the early behavior of dark energy. Astrophys. J. **659**, 98–121 (2007)
453. Robertson, B., et al.: Disk galaxy formation in a Λ Cold Dark Matter Universe. Astrophys. J. **606**, 32 (2004)
454. Robitaille, P.M.L.: NMR and the age of the Universe. American Physical Society Centenial Meeting, BC19.14, March 21, (1999)
455. Robitaille, P.M.L.: The collapse of the Big Bang and the gaseous Sun. New York Times, March 17, (2002)
456. Robitaille, P.M.L.: On the validity of Kirchhoff's law of thermal emission. IEEE Trans. Plasma Sci. **31(6)**, 1263–1267 (2003)
457. Robitaille, P.M.L.: Comment to the NRC Committee on Condensed Matter and Material Physics (2005). [http://www7.nationalacademies.org/bpa/CMMP2010_Robitaille.pdf]
458. Robitaille, P.M.L.: WMAP: A radiological analysis II. American Phys. Soc., 8th Annual Meeting of the APS Northwest Section, May 18–20, (2006)
459. Robitaille, P.M.L.: An analysis of universality in blackbody radiation. Prog. in Phys. **2**, 22–23 (2006)
460. Robitaille, P.M.L.: The solar photosphere: Evidence for condensed matter. Prog. in Phys. **2**, 17–21 (2006)
461. Robitaille, P.M.L.: A high temperature liquid plasma model of the Sun. Prog. in Phys. **1**, 70–81 (2007)
462. Robitaille, P.M.L.: Max Karl Ernst Ludwig Planck: (1858–1947). Prog. Phys. **4**, 117–120 (2007)
463. Robitaille, P.M.L.: On the earth microwave background: Absorption and scattering by the atmosphere. Prog. in Phys. **3**, 3–4 (2007)
464. Robitaille, P.M.L.: On the nature of the microwave background at the Lagrange 2 point. Part I. Prog. Phys. **4**, 74–83 (2007)
465. Robitaille, P.M.L.: On the origins of the CMB: Insight from the COBE, WMAP, and Relikt-1 Satellites. Prog. in Phys. **1**, 19–23 (2007)

466. Robitaille, P.M.L.: The little heat engine: Heat transfer in solids, liquids and gases. Prog. Phys. **4**, 25–33 (2007)
467. Robitaille, P.M.L.: A ritical analysis of universality and Kirchhoff's law: A return to Stewart's law of thermal emission. Prog. Phys. **3**, 30–35 (2008)
468. Robitaille, P.M.L.: Blackbody radiation and the carbon particle. Prog. Phys. **3**, 36–55 (2008)
469. Robitaille, P.M.L.: The Earth microwave background (EMB), atmospheric scattering and the generation of isotropy. Prog. in Phys. **2**, L7–L8 (2008)
470. Rodrigues-Williams, L., Hogan, C.: Statistical association of QSO's with foreground galaxy clusters. Astron. J. **107**, 451 (1994)
471. Rohlfs, K., Wilson, T.L.: Tools of radio astronomy. Springer, New York, p. 284–392 (1996)
472. Romano, D., et al.: Chemical evolution in a model for the joint formation of quasars and spheroids. Mon. Not. R. Astron. Soc. **334**, 444 (2002)
473. Romer, A.K., et al.: Cluster Survey with XMM: Expected Catalog Properties and Scientific Applications. Astrophys. J. **547**, 594–608 (2001)
474. Rosati, P., et al.: A first determination of the surface density of galaxy clusters at very low x-ray fluxes. Astrophys. J. Lett. **445**, L11–L14 (1995)
475. Rosati, P., Borgani, S., Norman, C.: The evolution of X-ray clusters of galaxies. Ann. Rev. Astron. Astrophys. **40**, 539–577 (2002)
476. Rosenberg, A., et al.: Photometric catalog of nearby globular clusters. I. A large homogeneous (V,I) color-magnitude diagram data-base. Astron. Astrophys. Suppl. **144**, 5 (2000)
477. Roth, K.C, Meyer, D.M., Hawkins, I.: Interstellar cyanogen and the temperature of the cosmic background radiation. Astrophys. J. **413**, L67–L71 (1993)
478. Rowan-Robinson, M.: The cosmological distance ladder. W.H. Freeman, San Francisco (1985)
479. Rowan-Robinson, M.: Do Type Ia supernovae prove $\Lambda > 0$? Mon. Not. R. Astron. Soc. **332**, 352 (2002)
480. Ryan, S.G., Norris, J.E., Beers, T.C.: The spite lithium plateau: Ultrathin but postprimordial. Astrophys. J. **523**, 654 (1999)
481. Ryan, S.G., et al.: Primordial lithium and Big Bang nucleosynthesis. Astrophys. J. Lett. **530**, L57 (2000)
482. Ryan-Weber, E.V., Pettini, M., Madau, P.: Intergalactic CIV absorption at redshifts 5.4–6. Mon. Not. R. Astron. Soc. **371**, L78 (2006)
483. Sachs, R.K., Wolfe, A. M.: Perturbations of a cosmological model and angular variations of the microwave background. Astrophys. J. **147**, 73 (1967)
484. Salpeter, E.E.: The luminosity function and stellar evolution. Astrophys. J. **121**, 161 (1955)
485. Sandage, A.: On the formation and age of the Galaxy. J. R. Astron. Soc. Canada **84/2**, 70 (1990)
486. Sandage, A., Cacciari, C.: The absolute magnitudes of RR Lyrae stars and the age of the galactic globular cluster system. Astrophys. J. **350**, 645 (1990)
487. Sandage, A., Tammann, G.A., Reindl, B.: New period-luminosity and period-color relations of classical Cepheids. II. Cepheids in LMC. Astron. Astrophys. **424**, 43 (2004)
488. Sandage, A., et al.: The Hubble constant: A summary of the Hubble space telescope program for the luminosity calibration of type Ia supernovae by means of cepheids. Astrophys. J. **653**, 843 (2006)
489. Santos, M.G., et al.: Multiple methods for estimating the bispectrum of the cosmic microwave background with application to the MAXIMA Data. Mon. Not. R. Astron. Soc. **341**, 623 (2003)
490. Sarajedini, A.: Globular cluster photometry near the turn-off relative ages and the horizontal branch. In: Precision photometry: Astrophysics of the galaxy. Philip A.G.D., Upgren, A.R., Janes, K.A. (eds.), L. Davies., Schenectady, p. 55. (1991)
491. Sarajedini, A., Demarque, P.: A new age diagnostic applied to the globular clusters NGC 288 and NGC 362. Astrophys. J. **365**, 219 (1990)
492. Sarajedini, A., King, C.R.: Evidence for an age spread among the Galactic globular clusters. Astron. J. **98**, 1624 (1989)

493. Sarajedini, A., et al.: The ACS survey of galactic globular clusters. I. Overview and clusters without previous Hubble space telescope photometry. Astron J. **133**, 1658 (2007)
494. Saslaw, W., Valtonen, M. Aarseth, S.: The gravitational slingshot and the structure of extragalactic radio sources. Astrophys. J. **190**, 253 (1974)
495. Schade, D., et al.: Hubble Space telescope imaging of the CFRS and LDSS redshift surveys. III. Field elliptical galaxies at $0.2 \leq z \leq 1.0$. Astrophys. J. **525**, 31 (1999)
496. Schechter, P.: The Hubble constant from gravitatioal lens time delays. In: Gravitational lensing impact on cosmology, IAU Symp. **225**. Mellier, Y., Meylan, G. (eds.). Cambridge University Press, Cambridge, p. 281–296 (2005)
497. Schiavon, R.P., et al.: A library of integrated spectra of galactic globular clusters. Astrophys. J. Suppl. **160**, 163 (2005)
498. Schlegel, D.J., Finkbeiner, D.P., Davis, M.: Maps of dust infrared emission for use in estimation of reddening and cosmic microwave background radiation foregrounds. Astrophys. J. **500**, 525–553 (1998)
499. Schmidt, B., et al.: The high-Z supernova search: Measuring cosmic deceleration and global curvature of the Universe using type IA supernovae. Astrophys. J. **507**, 46 (1998)
500. Schneider, P.: Introduction to gravitational lensing and cosmology. In: Gravitational lensing: Strong, weak and micro. Meylan, G., Jetzer, P., North, P., Schneider, P., Kochanek, C.S., Wambsganss, J. (eds.), Saas-Fee Advanced Course 33. p. 1 (2006)
501. Schneider, P.: Weak gravitational lensing. In: Gravitational lensing: Strong, weak and micro. Meylan, G., Jetzer, P., North, P., Schneider, P., Kochanek, C.S., Wambsganss, J. (eds.), Saas-Fee Advanced Course 33. p. 269 (2006)
502. Schneider, P., Ehlers J., Falco, E.E.: Gravitational lenses. In: Gravitational lenses, XIV, 560, Springer Berlin Heidelberg, New York, pp. 112. (1992)
503. Schuecker, P., et al.: Observational constraints on general relativistic energy conditions, cosmic matter density and dark energy from X-ray clusters of galaxies and type-Ia supernovae. Astron. Astrophys. **398**, 867 (2003)
504. Schweizer, F., Seitzer, P.: Correlations between UBV colors and fine structure in E and S0 galaxies – A first attempt at dating ancient merger events. Astron. J. **104**, 1039 (1992)
505. 2dF survey web site http://www.aao.gov.au/2df/
506. Serra, G., et al.: PRONAOS : A two meter submillimeter balloon borne telescope. Adv. Space Res. **30**, 1297 (2002)
507. Searle, L., Zinn, R.J.: Compositions of halo clusters and the formation of the galactic halo. Astrophys. J. **225**, 357 (1978)
508. Sesana, A., et al.: Low-frequency gravitational radiation from coalescing massive black hole binaries in hierarchical cosmologies. Astrophys. J. **611**, 623 (2004)
509. Shafieloo, A., Souradeep, T.: Estimation of primordial spectrum with post-WMAP 3 year data. (2007) [astro-ph/0709.1944]
510. Sharma, S., Steinmetz, M.: The angular momentum distribution of gas and dark matter in galactic halos. Astrophys. J. **628**, 21 (2005)
511. Sharp, N.: Anomalous redshift companion galaxies - NGC 7603. Astrophys. J. **302**, 245 (1986)
512. Sheth, R.K., Tormen, G.: Large-scale bias and the peak background split. Mon. Not. R. Astron. Soc. **308**, 119–126 (1999)
513. Silk, J.: Formation and evolution of disk galaxies. J. Astrophys. Space Science **284**, 663–674 (2003)
514. Silk, J., Rees, M.J.: Quasars and galaxy formation. Astron. Astrophys. **331**, L1 (1998)
515. Silva, A.I., Viegas, S.M.: Physical conditions in QSO absorbers from fine-structure absorption lines. Mon. Not. R. Astron. Soc. **329**, 135–148 (2002)
516. Sivron, R., Tsuruta, S.: Dense clouds near the central engine of active galactic nuclei. Astrophys. J. **402**, 420 (1993)
517. Sloan Digital Sky Survey web site http://www.sdss.org/
518. Smith, M., et al.: The RAVE survey: Constraining the local galactic escape speed. Mon. Not. R. Astron. Soc. **379**, 755 (2007)

519. Smith, N., et al.: SN 2006gy: Discovery of the most luminous supernova ever recorded, powered by the death of an extremely massive star like η Carinae. Astrophys. J. **666**, 1116 (2007)
520. Smoot, G.F., et al.: Structure in the COBE differential microwave radiometer first-year maps. Astrophys. J. Lett. **396**, L1–L5 (1992)
521. Somerville, R.S., Bullock, J.S., Livio, M.: The Epoch of reionization in models with reduced small-scale power. Astrophys. J. **593**, 616 (2003)
522. Songaila, A.: The minimum universal metal density between redshifts of 1.5 and 5.5. Astrophys. J. **561**, 153 (2001)
523. Spergel, D.N., et al.: First-year Wilkinson microwave anisotropy probe (WMAP) observations: Determination of cosmological parameters. Astrophys. J. Suppl. **148**, 175–194 (2003)
524. Spergel, D.N., et al.: Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. Astrophys. J. Suppl. **170**, 377 (2007)
525. Spite, F., Spite, M.: Abundance of lithium in unevolved halo stars and old disk stars – Interpretation and consequences. Astron. Astrophys. **115**, 357 (1982)
526. Srianand, R., Petitjean, P., Ledoux, C.: The cosmic microwave background radiation temperature at a redshift of 2.34. Nature **408**, 931–935 (2000)
527. Stefan, J.: Über die beziehung zwischen der warmestrahlung und der temperature. Wein. Akad. Sitzber. **79**, 391–428 (1879)
528. Steigman, G.: Primordial nucleosynthesis in the precision cosmology era. Ann. Rev. Nucl. Part. Sci. **57**, 463 (2007)
529. Steinmetz, M., Müller, E.: The formation of disk galaxies in a cosmological context: Populations, metallicities and metallicity gradients. Astron. Astrophys. **281**, L97 (1994)
530. Stockton, A., Wyckoff, S., Wehinger, P.A.: On the distance to Markarian 205. Astrophys. J. **231**, 673 (1979)
531. Stolyarov, V., et al.: All-sky component separation for the Planck mission. Mon. Not. R. Astron. Soc. **336**, 97S (2002)
532. Stompor, R., et al.: Cosmological implications of the MAXIMA-1 high-resolution cosmic microwave background anisotropy measurement. Astrophys. J. Lett. **561**, L7–L10 (2001)
533. Straniero, O., Chieffi, A.: Isochrones for H-burning globular cluster stars. II – The metallicity range [Fe/H] = -2.3 to -0.5. Astrophys. J. Suppl. **76**, 525 (1991)
534. Sulentic, J.: On the association of quasars with bright galaxies. Astrophys. J. Lett. **244**, L53 (1981)
535. Sulentic, J.: Confirmation of the luminous connection between NGC 4319 and Markarian 205. Astrophys. J. Lett. **265**, L49 (1983)
536. Sulentic, J., Marziani, P., Dultzin-Hacyan, D.: Phenomenology of broad emission lines in active galactic nuclei. Ann. Rev. Astron. Astrophys. **38**, 521–571 (2000)
537. Sulentic, J., et al.: Radio-loud active galactic nuclei in the context of the eigenvector 1 parameter space. Astrophys. J. Lett. **597**, L17-L20 (2003)
538. Sulentic, J. et al.: VLT/ISAAC spectra of the Hβ; region in intermediate redshift quasars. Astron. Astrophys. **423**, 121–132 (2004)
539. Sulentic, J., et al.: VLT/ISAAC spectra of the Hβ; region in intermediate-redshift quasars. II. Black hole mass and Eddington ratio. Astron. Astrophys. **456**, 929–939 (2006)
540. Sulentic, J., et al.: C IVλ1549 as an eigenvector 1 parameter for active galactic nuclei. Astrophys. J. **666**, 757–777 (2007)
541. Sullivan, M., et al.: The Hubble diagram of type Ia supernovae as a function of host galaxy morphology. Mon. Not. R. Astron. Soc. **340**, 1057 (2003)
542. Sun, X., et al.: Radio observational constraints on Galactic 3D-emission models. Astron. Astrophys. **477**, 573 (2008)
543. Tammann, G.A., Sandage, A., Reidl, R.: New Period-Luminosity and Period-Color relations of classical Cepheids: I. Cepheids in the Galaxy. Astron. Astrophys. **404**, 423 (2003)
544. Tantalo, R., Chiosi, C.: Enhancement of α-elements in dynamical models of elliptical galaxies. Astron. Astrophys. **388**, 396 (2002)
545. Tantalo, R., Chiosi, C.: Measuring age, metallicity, abundance ratios from absorption-line indices. Mon. Not. R. Astron. Soc. **353**, 917 (2004)

546. Tantalo, R., Chiosi, C.: Star formation history in early-type galaxies. I. The line absorption indices diagnostics. Mon. Not. R. Astron. Soc. **353**, 405 (2004)
547. Tantalo, R., Chiosi, C., Bressan, A.: Ages and metallicities in elliptical galaxies from the H_β, $\langle Fe \rangle$, and Mg_2 diagnostics. Astron. Astrophys. **333**, 419 (1998)
548. Tantalo, R., Chiosi, C., Piovan, L.: New response functions for absorption-line indices from high resolution spectra. Astron. Astrophys. **481**, 462 (2007)
549. Taubenberger, S., et al.: SN 2004aw: confirming diversity of Type Ic supernovae. Mon. Not. R. Astron. Soc. **371**, 1459 (2006)
550. Teerikorpi, P.: The distance scale and Eddington efficiency of luminous quasars. (2005) [astro-ph/0510382]
551. Tegmark, M., et al.: How small were the first cosmological objects? Astrophys. J. **474**, 1 (1997)
552. Tegmark, M., de Oliveira-Costa, A., Hamilton, A.J.S.: A high resolution foreground cleaned CMB map from WMAP. Phys. Rev. **D 68(12)**, 123523 (2003)
553. Tegmark, M., et al.: Cosmological parameters from SDSS and WMAP. Phys. Rev. **D 69**, 103501 (2004)
554. Tifft, W.: Redshift periodicities, the galaxy-quasar connection. Astrophys. Space Sci. **285**, 429 (2003)
555. Touloukian, Y.S., Ho, C.Y.: Thermophysical properties of matter. New York, Plenum, v. 1–8, (1970)
556. Tozzi, P.: Cosmological parameters from galaxy clusters: An introduction. Lect. Notes in Phys. **720**, 125–156 (2007)
557. Trager, S.C.: The stellar population histories of early-type galaxies. II. Controlling parameters of the stellar populations. Astron. J. **120**, 165 (2000)
558. Trager, S.C., et al.: The stellar population histories of local early-type galaxies. I. Population parameters. Astron. J. **119**, 1645 (2000)
559. Tristram, M., et al.: The CMB temperature power spectrum from an improved analysis of the Archeops data. Astron. Astrophys. **436**, 785–797 (2005)
560. Tucker, G.S., et al.: Anisotropy in the microwave sky: Results from the first flight of the balloon-borne anisotropy measurement (BAM). Astrophysical J. Lett. **475**, L73–L76 (1997)
561. Turatto, M.: Classification of supernovae. In: Supernovae and gamma-ray bursters. K. Weiler. (ed.), Lecture Notes in Physics **598**, 21 (2003)
562. Turatto, M., Benetti, S., Pastorello, A.: Supernovae classes and subclasses. In: Supernova 1987A: 20 years after: Supernovae and gamma-ray bursters. AIP Conf. Ser. **937**, 187 (2007)
563. Tyson, J.A., Wenk, R.A., Valdes, F.: Detection of systematic gravitational lens galaxy image alignments – Mapping dark matter in galaxy clusters. Astrophys. J. Lett. **349** L1–L4 (1990)
564. Unwin, S.C., et al.: Taking the measure of the Universe: Precision astrometry with SIM planetQuest. Pub. Astron. Soc. Pacific **120**, 38–88 (2008)
565. Urry, C.: AGN unification: An update. AGN physics with the sloan digital sky survey. ASP Conf. Ser. **311**, 49–60 (2004)
566. Vallée, J.P.: The spiral arms and interarm separation of the Milky Way: An updated statistical study. Astron. J. **130**, 569 (2005)
567. van den Bosch, F.C., et al.: The angular momentum of gas in protogalaxies. I. Implications for the formation of disk galaxies. Astrophys. J. **576**, 21 (2002)
568. van Leeuwen, F., et al.: Cepheid parallaxes and the Hubble constant. Mon. Not. R. Astron. Soc. **379**, 723 (2007) [astro-ph/0705.1592]
569. VandenBerg, D.A., Bolte, M., Stetson, P.B.: Measuring age differences among globular clusters having similar metallicities – A new method and first results. Astron. J. **100**, 445 (1990)
570. VandenBerg, D.A., Bergbusch, P.A., Dowler P.D.: The Victoria-Regina stellar models: Evolutionary tracks and isochrones for a wide range in mass and metallicity that allow for empirically constrained amounts of convective core overshooting. Astrophys. J. Suppl. **162**, 375 (2006)
571. Vangioni-Flam, E., et al.: Lithium-6: evolution from Big Bang to present. New Astron. **4**, 245 (1999)

572. Véron-Cetty, M.-P., Véron, P.: A catalogue of quasars and active nuclei, 12th edition. Astron. Astrophys. **455**, 773–777 (2006)
573. Vestergaard, M., et al.: Mass functions of the active black holes in distant quasars from the sloan digital sky survey data release 3. [arXiv:0801.0243] (2007)
574. Vikhlinin, A., et al.: A catalog of 200 galaxy clusters serendipitously detected in the ROSAT PSPC pointed observations. Astrophys. J. **502**, 558–581 (1998)
575. Vikhlinin, A., et al.: Cosmological constraints from the evolution of the cluster baryon mass function at z$\sim$0.5. Astrophys. J. **590**, 15–25 (2003)
576. Vikhlinin, A., et al.: Chandra sample of nearby relaxed galaxy clusters: Mass, gas fraction, and mass-temperature relation. Astrophys. J. **640**, 691–709 (2006)
577. Vittorio, N., Silk, J.: Fine-scale anisotropy of the cosmic microwave background in a universe dominated by cold dark matter. Astrophys. J. Lett. **285**, L39 (1984)
578. Voevodkin, A., Vikhlinin, A.: Constraining amplitude and slope of the mass fluctuation spectrum using a cluster baryon mass function. Astrophys. J. **601**, 610–620 (2004)
579. Voit, G. M.: Expectations for evolution of cluster scaling relations. Adv. Space Res. **36**, 701–705 (2005)
580. Voit, G.M.: Tracing comic evolution with clusters of galaxies. Rev. Mod. Phys. **77**, 207–258, (2005)
581. von Hippel, T.: From young and hot to old and cold: Comparing white dwarf cooling theory to main-sequence stellar evolution in open clusters. Astrophys. J. **622**, 565 (2005)
582. Walker, T.P., et al.: Primordial nucleosynthesis redux. Astrophys. J. **376**, 51 (1991)
583. Wambsganss, J.: Gravitational microlensing. In: Gravitational lensing: Strong, weak and micro. Meylan, G., Jetzer, P., North, P., Schneider, P., Kochanek, C.S., Wambsganss, J. (eds.), Saas-Fee Advanced Course 33. p. 453 (2006)
584. Wang, L., et al.: Nonlinear decline-rate dependence and intrinsic variation of type Ia supernova luminosities. Astrophys. J. **641**, 50 (2006)
585. Watson, R. A., et al.: Detection of anomalous microwave emission in the Perseus molecular cloud with the COSMOSOMAS experiment. Astrophys. J. **624**, L89–L92 (2005)
586. Weedman, D.: A high-redshift "seyfert nucleus" in the arms of a spiral galaxy. Astrophys. J. **161**, 113 (1970)
587. Weedman, D.: Quasar Astronomy, Cambridge University Press, Cambridge (1986)
588. White, S.D.M., Springel, V.: Where are the first stars now? In: The first stars. Proceedings of the MPA/ESO Workshop. Garching, Germany, p. 327 (2000)
589. Wien, W.: Über die energieverteilung in emissionspektrum eines schwarzen korpers. Ann. Phys. **58**, 662–669 (1896)
590. WMAP website, http://map.gsfc.nasa.gov/
591. Wood-Vasey, W.M., et al.: Observational constraints on the nature of dark energy: First cosmological results from the ESSENCE supernova survey. Astrophys. J. **666**, 694 (2007)
592. Woosley, S.E., Weaver, T.A.: Sub-Chandrasekhar mass models for Type IA supernovae. Astrophys. J. **423**, 371 (1994)
593. Wright, E.L., et al.: Interpretation of the cosmic microwave background radiation anisotropy detected by the COBE Differential microwave radiometer. Astrophys. J. Lett. **396**, L13 (1992)
594. Wyithe, J.S.B., Loeb, A.: Was the Universe reionized by massive metal-free stars? Astrophys. J. **588**, L69 (2003)
595. Xue, X.-X., et al.: The Milky Way's circular velocity curve to 60 kpc and an estimate of the dark matter halo mass from kinematics of $\sim$2500 SDSS blue horizontal branch stars. Astrophys. J. **684**, 1143 (2008)
596. Yamada, T., et al.: The number density of old passively evolving galaxies at $z = 1$ in the Subaru/XMM-Newton deep survey field. Astrophys. J. **634**, 861 (2005)
597. Yi, S.K., Kim, Y.C., Demarque, P.: The Y^2 stellar evolutionary tracks. Astrophys. J. Suppl. **144**, 259 (2003)
598. Yoon, K.W., et al.: The Robinson gravitational wave background telescope (BICEP): A bolometric large angular scale CMB polarimeter, millimeter and submillimeter detectors and instrumentation for astronomy III. Zmuidzinas, Jonas, Holland, Wayne S., Withington, Stafford, Duncan, William D (eds.). Proceedings of SPIE **6275**, 62751K (2006)

599. Zamanov, R., Marziani, P.: Searching for the physical drivers of eigenvector 1: From quasars to nanoquasars. Astrophys. J. Lett. **571**, L77–L80 (2002)
600. Zehavi, I., et al.: A local Hubble bubble from type IA supernovae? Astrophys. J. **503**, 483 (1998)
601. Zentner, A.R., Bullock, J.S.: Inflation, cold dark matter, and the central density problem. Phys. Rev. **D 66/4**, 043003 (2002)
602. Zepf, S.E.: Formation of elliptical galaxies at moderate redshifts. Nature **390**, 377 (1997)
603. Zhu, X.-F., Chu, Y.-Q.: The association between quasars and the galaxies of the Virgo cluster. Astron. Astrophys. **297**, 300 (1995)
604. Zinn, R.J., West, M.J.: The globular cluster system of the galaxy. III – Measurements of radial velocity and metallicity for 60 clusters and a compilation of metallicities for 121 clusters. Astrophys. J. Suppl. **55**, 45 (1984)
605. Zwicky, F. et al.: Catalogue of galaxies and of clusters of galaxies. Calif. Inst. of Techn., Pasadena, 6 vols. (1961–1968)

Chapter 3
Astrophysical Cosmology

Contributions by Amedeo Balbi, Charles L. Bennett, Martin Bucher, Carlo Burigana, Peter Coles, Mauro D'Onofrio, Ruth Durrer, John Mather, Pavel Naselsky, Francesca Perrotta, Lucia A. Popa, David Spergel, Kandaswamy Subramanian, and Nicola Vittorio

3.1 Outline of the Chapter

The previous chapter was devoted to the main observational evidence on the basis of our comprehension of the Universe and to their main theoretical implications. Some theoretical and phenomenological topics of particular relevance for current cosmology with clear observational counterparts are rediscussed here, with a better attention to the corresponding concepts of fundamental physics and, to a certain extent, to the mathematical formalism at the basis of their formulation.

The presentation of the various themes is organized almost according to their relevance at increasing cosmic time, although the physical processes considered and the propagation of the generated photons to the observer occurred during various cosmic epochs.

We start with the presentation of the inflationary scenario, the current framework at the basis of the explanation on many of the fundamental properties of the observed Universe, described in Sect. 3.2 by Francesca Perrotta. While its link with the fundamental physics in the early Universe is far to be firmly understood, it is difficult to escape the main intuitions and predictions of this scenario.

A tight coupling to the fundamental physics of the early Universe, that is, to symmetry breaking phase transitions, is based on the topological defect models presented in Sect. 3.3 by Ruth Durrer, who discusses the fundamental theoretical aspects, the imprints on the cosmic microwave background (CMB), and the non-Gaussian signatures.

Note in passing that the energy scales considered in inflationary and topological defect models are few orders of magnitude higher than those achievable for particle physics experiments conducted directly through accelerators. Therefore, the Universe is likely our best high energy physics laboratory.

The contributions by Francesca and Ruth underline the relevance of the inhomogeneities in the Universe as a way to discriminate among different cosmological and fundamental physics scenarios. It is then natural to focus more on the nature of primordial perturbations, on their classification, and on the possibility to distinguish among different kinds of perturbations. The interview with Martin Bucher in Sect. 3.4 will clarify these aspects.

M. D'Onofrio and C. Burigana (eds.), *Questions of Modern Cosmology: Galileo's Legacy*, DOI 10.1007/978-3-642-00792-7_3,

Since its discovery, the CMB has been recognized to be one of the most crucial probe of the standard cosmological model and a promising laboratory for testing various properties of the Universe (e.g., geometry, thermal history, early epochs, seeds of structure formation, etc., some of which already discussed in the earlier sections). CMB is in fact the observed radiation field closest to the Big Bang, at least until the discovery of the cosmological gravitational background, and the bulk of the information it contains refers to cosmic epochs when nonlinearities were small and complex astrophysical phenomenologies of little importance.

In Chap. 2, the high precision measurements by the Cosmic Background Explorer (COBE), the Wilkinson Microwave Anisotropy Probe (WMAP), and balloon experiments were presented. Here we will focus on the most accepted theoretical framework that explains such data. The interview to John Mather in Sect. 3.5.1 is devoted to the main implications of the very stringent observational constraints on possible deviations from a Planckian spectrum. We then move to the relevance of CMB anisotropies. The interview to Nicola Vittorio illustrates the wealth of information that can be extracted through the analysis of the Angular Power Spectrum (APS) (Sect. 3.5.2.1) and on its link with the physics at the basis of CMB anisotropies. The angular power spectrum is a compressed statistical estimator of the CMB fluctuation pattern. While Nicola's presentation addresses, in particular, the significance of the power at different angular scales, the interview to Peter Coles in Sect. 3.5.2.2 is dedicated to the information contained in the phases of the full spherical harmonic expansion of the CMB anisotropies. After the presentation of these theoretical methods for the analysis of CMB data, the interview with David Spergel in Sect. 3.5.3 will review the information contained in the high accuracy all-sky data from WMAP with attention to the determination of cosmological parameters. In the last section, dedicated to CMB (Sect. 3.5.4), Peter Coles discusses the crucial problem of the geometry of the Universe, with emphasis on possible deviations from large scale uniformity and isotropy.

Following the expansion of the Universe, it is important to understand the various phases of the matter ionization, from the epoch of recombination, when electrons and nuclei formed neutral atoms, and the CMB photons could start to travel almost freely, to the subsequent phase of reionization associated with structure formation. These stages, previously discussed by Piero Madau in Chap. 2, mainly from an observational perspective, are revised in Sect. 3.6 in the interview with Pavel Naselsky, which emphasizes the key theoretical aspects (Sects. 3.6.1 and 3.6.2) and possible alternative ionization models (Sect. 3.6.3).

The large scale structure of the Universe constitutes another bench mark of cosmology. As pointed out by John Peacock in Chap. 2, the analysis of the Large Scale Structure (LSS) drove the change of paradigm from Cold Dark Matter (CDM) to Λ*CDM. In Sect. 3.7, Amedeo Balbi reviews the relevance of LSS studies, highlighting in particular the advantages of the power spectrum analysis and the link and synergy of LSS and CMB data. Then, Charles Bennett focuses on the imprints of Baryon Acoustic Oscillations (BAO) on the LSS power spectrum, detected thanks to the last generation of galaxy surveys (Sect. 3.7.1). Today the dynamics of structure formation and the LSS of the Universe are investigated not only with analytical* *methods but also through numerical simulations. In the last section dedicated to LSS (Sect. 3.7.2), Peter Coles presents the basic ideas and methods used in this research.*

The last two sections of this chapter are devoted to the discussion of two themes of high relevance in cosmology linked to fundamental physics. In Sect. 3.8, Lucia Popa outlines the state-of-the-art of neutrino physics, while in Sect. 3.9, Kandaswamy Subramanian addresses the important question of the role of cosmic magnetism. If, on the one hand, the physics of neutrinos is still poorly known but their effects for cosmology start to be understood, on the other hand, the fundamental physical laws of magnetism are established but the theoretical framework of their cosmological origin and astrophysical effects is only at its infancy.

Let us start with Francesca Perrotta, who will now highlight the epoch of inflation.

3.2 Inflation

Dear Francesca *(Perrotta)*, **inflation represents today a robust framework to explain many fundamental properties of the observed Universe, which, differently, could appear difficult to understand or, at limit, as paradoxes of modern cosmology, from the flatness and horizon problems to the generation of small perturbations, seeds of the current structures, from the dominance of matter over antimatter to the shape of primordial perturbation power spectrum, etc. To your opinion, what are the most crucial fundamental questions at the basis of inflation and the corresponding answers offered by this framework?**

The theory of inflation was born, thanks to a lucky and brilliant intuition by Guth in 1981 [109], at a time when the cosmological community was puzzled by the unsolved problems of the "standard model" (or "standard Big Bang cosmology"). The widely accepted framework was then based on General Relativity (GR), on the cosmological principle of large scale homogeneity and isotropy, and on the perfect-fluid representation of the Universe: such a description was able to predict the Hubble flow, the light elements abundance, the microwave background of radiation. However, despite its success, this picture was not able to explain some crucial feature of the visible Universe; in particular, the standard Big Bang model carries with itself the "flatness problem" (observations indicating that today $\Omega \approx 1$, which is an unstable solution of the standard model) and the "horizon problem" (the observed homogeneity between parts of the Universe which should never have been causally connected). The inflationary paradigm proposes an elegant escape to both. Let us see in more detail how the horizon problem manifests itself in the standard picture. In a Friedmann Universe, a region of physical length L becomes causally connected once it equals the horizon: only then, physical processes may eventually act to make the region homogeneous and isotropic. In a decelerated expansion, as the one produced, through GR, by ordinary matter or radiation, the effective horizon increases with time faster than any physical scale. Thus, any length scale will eventually enter the horizon at some time. The puzzling feature of this scenario shows when looking at the CMB: the CMB photons are a picture of the Universe as it was at recombination epoch, because, after decoupling, they simply freestream and, to a good approximation, do not interact with each other. The scale which is in horizon crossing at recombination epoch corresponds to $\sim 100\,\mathrm{h}^{-1}$ Mpc. The current distance to the last scattering surface is about $6{,}000\,\mathrm{h}^{-1}$ Mpc. So, the horizon at recombination subtends an angle $<1^\circ$ in the sky: two points on the microwave sky, separated by an angle larger than $\sim 1^\circ$, were not causally connected at recombination. On the other hand, the CMB appears to be isotropic up to one part over 10^5! The thermal equilibrium and homogeneity between regions of the last scattering surface, which are separated by distances large than the light could have traveled by the time they were imprinted on the sky, is the origin of the horizon problem.

The explanation proposed by the inflationary paradigm assumes that, during the very first epoch of the cosmological evolution, there is a period $(t_\mathrm{i}, t_\mathrm{f})$ ("inflation") in which the expansion is accelerated. During inflation, the superluminal evolution will redshift primordial fluctuations beyond the horizon: after that, the Universe expansion is decelerated again. In this case, a length L, previously causally connected, can become super-horizon sized during inflation, to reenter the horizon at a later time t: regions that we observe causally disconnected today were in causal contact before inflation. The (00) component of the Einstein equation and the trace component can be combined to obtain the second order equation for the scale factor evolution, in terms of the density and pressure of the cosmic fluid:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p)\,. \tag{3.1}$$

We can infer that, to provide accelerated expansion $\ddot{a} > 0$, the dominant component of the cosmic fluid during inflation should have an equation of state $p = w\rho c^2$, with $w < -1/3$. In most inflationary models, this condition is realized by mean of a scalar field, the "inflation". In particle physics, a scalar field is used to represent spin zero particles. It transforms like a scalar under coordinate transformations and, in a homogeneous Universe, it is a function of time alone. The effective energy density and pressure of a homogeneous scalar field φ are

$$\rho_\varphi = \frac{1}{2}\dot{\varphi}^2 + V(\varphi), \tag{3.2}$$

$$p_\varphi = \frac{1}{2}\dot{\varphi}^2 - V(\varphi)\,. \tag{3.3}$$

One can think the first term in each as a kinetic energy, and the second as a potential energy. This potential energy can be thought of as a form of binding energy. Like all systems, scalar fields try to minimize this energy: to allow inflation, the efficiency in reaching the minimum energy state should not be high. We can notice that the previous equations define the scalar field equation of state in a dynamical way: to have accelerated expansion we need $\dot{\varphi}^2 < V(\varphi)$. This is possible if the potential is flat enough, as the scalar field is then expected to slowly roll along the potential: the slow roll approximation is consistent as long as both the potential slope and curvature are small (in Planck units). This is the basis for defining two basic parameters common to the inflationary models, the slow-roll parameters ε and η, which are related to the first and second derivative of the potential. Of course, the potential should also have a minimum in which inflation can end: all these conditions are satisfied for a wide range of potentials, although we are not yet in a position in which there is a fundamental established theory for the inflation field.

Concerning the flatness problem, it can be easily understood defining the critical density $\rho_c = 3H^2/(8\pi G)$, $\Omega = \rho/\rho_c$, and writing the Friedmann equation as the curvature (intended as a deviation from flatness) evolution with time:

$$|\Omega - 1| = C(t) = (\mathrm{d}a/\mathrm{d}t)^{-2}. \tag{3.4}$$

Let us assume that the scale factor increases with a power law, $a/a_i = (t/t_i)^n$, with t_i being some initial time. We can thus write $C = C_i(t/t_i)^{2(1-n)}$. The index n turns out to be 1/2 in the radiation era, and 2/3 in the matter dominated era, as a consequence of the Friedmann equations (these values of the exponents correspond to the usual Friedmanian solutions, as opposite to the inflationary solutions). As the scale factor evolves as the inverse of the temperature, we can write $C = C_i(T/T_i)^{2(n-1)/n}$. The current value for the temperature is $T_0 = 10^{12}$ GeV; we can also assume the Planck scale to set the initial value $T_i = 10^{19}$ GeV. The recent observations of the CMB have confirmed that the Universe is close to flat to the accuracy of few percent [233]: thus, today the density parameter Ω is of the order of unity, $|\Omega - 1| \approx 1$, which implies that the initial curvature $C_i \approx 10^{62(n-1)/n}$. We see that, using, for example, $n = 1/2$ as in the radiation era (but the result would not change appreciably in matter dominated era), at the initial time one should have $\Omega_i = 1 \pm \varepsilon$, with $\varepsilon \approx 10^{-62}$. If, for example, $\varepsilon \approx 10^{-59}$, one would be left today with a density parameter Ω 1,000 times larger (or smaller, depending on the sign of the geometric curvature k) than observed. In practice, a severe fine tuning is required on the initial conditions in order for the Universe not immediately collapse or become empty. It is clear that a value $n > 1$ is needed at least for a certain amount of time. On the other hand, an "inflationary solution" of the type $a \approx t^n$, with $n > 1$ implies an accelerated expansion, and an equation of state $w < -1/3$. This is precisely the solution offered by the inflationary scenario to the flatness problem: assuming such power law with $n > 1$, it can be shown that the situation is now reversed. Even starting with large values of C_i, we would finish today with a Universe very close to flatness. While in the standard scenario, the curvature redshifts away more slowly than matter and radiation, requiring a negligible initial contribution, the inflationary model, with its ability to stretch out any physical scale, naturally induces flatness independently on the initial value of the spatial curvature.

Therefore, inflation is a brilliant framework to bypass the most crucial problems of the Big Bang cosmology. That is not the end, though: since the first inflationary theories were proposed, it was realized that the very mechanism that stretches the physical scales can also account for the origin of primordial perturbations and, ultimately, can explain the current large scale structures despite the astonishing homogeneity of the inflated Universe. In this picture, both density fluctuations and tensor perturbations (gravitational waves) are created by quantum fluctuations, redshifted beyond the horizon during inflation, and finally frozen when the perturbation scale exit the horizon. The scalar perturbations do couple to the matter stress–energy and form the seeds out of which structures form, while gravitational waves do not couple to matter. It is necessary, for each scale of perturbation, to wait to re-enter the horizon, before they can grow and become eventually nonlinear: on sub-horizon scale, it is the gravitational collapse that makes the structure form as we observe them today. It is possible to make robust predictions in the context of inflationary models. First of all, the spectrum of primordial perturbations generated by the inflation is a Gaussian random field. This means that quantum fluctuations during inflation generate irregularities that obey Gaussian statistics. Any (remarkable) deviations from Gaussian statistics in the CMB temperature anisotropies would thus be difficult to reconcile with the inflationary model. Furthermore, the spectra of the density and gravitational wave fluctuations resulting from inflation can be characterized by power law indices n_s and n_t, respectively (for *scalar* and *tensor* modes). The specific model of inflation will lead to slight differences in the predicted values for these spectral indices, but most models of inflation predict nearly scale-invariant fluctuations, corresponding to values $n_s \approx 1$ and $n_t \approx 0$. Indeed, an exactly exponential (de Sitter) expansion during inflation would lead to precisely scale invariant density and gravitational wave spectra with indices $n_s = 1$ and $n_t = 0$, respectively. But, to match on to Friedmann–Lamaître–Robertson–Walker (FLRW) models, the expansion rate must slow down at the end of inflation. This leads to small deviations from a purely scale invariant spectrum in any realistic inflationary model that depend on the precise shape of the potential of a scalar field. (In general, an extreme fine tuning is required to produce inflationary models with spectral indices that lie outside the bounds $0.7 < n_s < 1.2$ and $0.3 < n_t < 0$). The relation between spectral indices and slow roll parameters is given by $n_s \approx 1-4\varepsilon+2\eta$ and $n_t \approx -2\varepsilon$ [164]; a useful way to quantify the relevant importance of density perturbations and gravitational waves is the ratio of their power spectra at horizon crossing, $r = 16\varepsilon = P_t(k)/P_r(k)$ [165]. These quantities characterize the imprint of the inflationary phase on the CMB, and constitute the observable associated with the model.

In summary, inflationary paradigm can explain not only the homogeneity of the Universe, but, most importantly, also its level of inhomogeneity: this is, in my view, the golden goal of the inflationary picture.

Many versions of inflation have been presented. What classes of models are currently disfavored by existing data and what are compatible with observational evidences?

As it often occurs in science, finding answers opens new questions. Despite the stringent observational tests inflation has passed, there are many open questions about the theory. Einstein equations tell us that, to produce accelerated expansion, a fluid with negative pressure is required: this does not correspond to any known form of matter. The current inflationary models easily accommodate a negative pressure through the scalar field implementation. It can be shown that, under certain conditions ("slow rolling" of the scalar field), the dynamics of a scalar field can mimic the behavior of such a fluid. (More precisely, this is realized when the kinetic energy is subdominant with respect to the potential energy of the field). Any specific particle theory contains scalar fields; however, there is no currently known scalar field that can drive inflation, as no fundamental scalar field has been observed. In the simplest inflationary model, originally proposed by Guth, an Higgs field was postulated to activate the process, via a mechanism analogous to the one in which the field gives mass to the elementary particles. The Higgs potential has the form $V(\varphi) = \lambda(\varphi^2 - M^2)^2$ and gives mass when sitting in the minima, while it can provide the potential energy for driving inflation when it is out of equilibrium. It is now known that the Higgs field cannot be the inflation field: unless one imposes a severe fine tuning on the potential of the field, it was not possible to find a reasonable mechanism to exit the inflationary phase, or to make it as long as required for the observed level of large scale homogeneity. Many physicists believe that inflation is included in a fundamental theory like string theory or a supersymmetric Grand Unified Theory [59, 82]. A promising suggestion is brane inflation, where the relevant process is related to generalization of strings as fundamental objects [81]. However, inflationary models have been proposed also without specific reference to underlying particle theories.

In the chaotic inflation by Linde [166], inflation can be induced by whichever scalar field with unbounded potential energy: the polynomial chaotic inflation, for example, can be realized with both the classes of potentials $V(\varphi) = \frac{1}{2}m^2\varphi^2$ and $V(\varphi) = \lambda\varphi^4$. However, in his model, the inflation field necessarily takes values larger than one Planck unit: for this reason, these are often called *large field* models, and the competing new inflation models are called *small field* models. In this situation, the predictions of effective field theory are thought to be invalid, and renormalization should cause large corrections that could prevent inflation: this theoretical problem has not yet been resolved. Some models do not satisfy the condition of a minimum in which inflation ends, thus permitting inflation to continue forever: this is the case for the "natural inflation" model, where $V(\varphi) = V_0[1 + \cos(\varphi/f)]$. The power law inflation, $V(\varphi) = V_0 \exp\left(\sqrt{\frac{16\pi}{p}}\frac{\varphi}{m_{Pl}}\right)$ [169] does not possess a natural mechanism to exit the inflationary expansion. The hybrid inflation model makes use of two fields rolling in a potential [167]: interestingly, a constant term in the potential would give a present day cosmological constant.

Rather than introducing a scalar field to drive inflation, some theories modify the gravitational sector of the Lagrangian of the theory. Models of inflation based on altering gravity are much more constrained than other models and very vulnerable to observations. This is the case, for example, of the extended inflation model [157].

Despite tremendous progress over the last years, the current constraints do not yet distinguish between inflationary models [25, 251], although some indications began to arise from the CMB observations performed by the WMAP satellite. By far the most promising observations appears indeed to be the high resolution CMB observations, as they directly probe the angular imaging of the primordial spectrum. Inflation predicts that the fluctuations should be almost but nonexactly scale invariant, with typical deviations of slope from scale invariance of the order of a few parts in a hundred. The recent analysis of the WMAP data [233] has found consistency with a nearly Harrison–Zel'dovich primordial spectrum ($n_s = 0.96$) with no running and no gravitational waves, within errors. In general, the CMB data seem to favor inflationary models with gentle slope of the potential, extremely slow rolling of the field, and poor or negligible production of gravitational waves: representing the potentials as power laws in the scalar field, $V(\varphi) = \varphi^p$, the model is consistent with the WMAP data for $p = 2$ and ruled out for $p = 4$ [219]. These are quite modest constraints, given the amount of inflationary models proposed in the last 15 years. However, observations have great prospects for improvements, mainly thanks to the incoming data from the *Planck* mission: an experiment with the sensitivity and resolution of *Planck* can measure the scalar spectral index with enough accuracy to accept or reject many of the several models proposed to date. The ability to measure small deviations from a precise scale invariant spectrum will provide tight constraints on the form of the inflationary potential and hence on fundamental physics at energies $>10^{15}$ GeV.

Another prediction of these models is that the rate of change of slope with scale, the "run" of the spectral index $\alpha = \mathrm{d}n_s/\mathrm{dln}k$, should be very small, $\alpha \approx 10^{-3}$. The analysis of the WMAP 5 years data [144] found no evidence of running for the spectral index, thus confirming the prediction of most inflationary models.

Inflation also predicts weak but most important and yet unobserved effects. One is represented by departure from Gaussianity; when treated non-perturbatively or at any order beyond the linear one, each inflationary model predicts a certain amount of non-Gaussianity, primarily affecting the harmonic counterpart of the three point correlation function, that is, the CMB bispectrum. To be able to detect such a distortion, a large scale survey with high sensitivity and angular resolution is needed, such as the *Planck* mission, and yet the community has to be able to control any other possible spurious source of non-Gaussianity from instrumental systematic or residual foreground leaking in the CMB signal, which is extracted from the data.

Finally, the most important prediction that may be able to distinguish between different models is the amount of gravitational waves generated by the inflationary mechanism. Although there are three fundamental parameters in inflation (r, n_s, and n_t), they are not completely independent: as they depend on the two slow roll parameters, we can obtain a relation between these observables, the consistency equation

$$r = -2\pi n_t. \tag{3.5}$$

The two perturbation spectra, scalar and tensor, originate indeed from the same scalar field, and ultimately from the same potential: this is the reason why they are entangled in such a way. Verification of the consistency equation would be a convincing test of the inflationary paradigm; however, it relies on the detection of gravitational waves or of their imprint on the last scattering surface (since tensor perturbations also contribute to the CMB temperature anisotropies). The current data from WMAP put the upper limit $r < 0.4$ (or $r < 0.35$ if combined with data from the Arcminute Cosmology Bolometer Array Receiver (ACBAR)) [144]. In the future, there are good perspectives of performing high precision measurements of the so-called “B-modes” of the polarization of the CMB: they would be evidence of the gravitational radiation produced by inflation, and they would also show whether the energy scale of inflation predicted by the simplest models ($10^{15} \div 10^{16}$ GeV) is correct. These measurements are expected to be performed by the *Planck* satellite, although it is unclear if the signal will be visible, or if contamination from foreground sources will interfere with these measurements. Other experiments for detecting the B-modes of the CMB are being planned with an extraordinary effort in terms of technology and data analysis development (see, e.g., the Legacy Archive for Microwave Background Data Analysis (LAMBDA) web page): these projects are first oriented to sub-orbital missions targeting low foreground areas in the sky, in the perspective of the design of a future B-modes dedicated CMB satellite.

Thank you Francesca. Indeed, large hopes are pinned on the discoveries of primordial B-modes through future CMB polarization anisotropy experiments, a topic that will be widely discussed in Chap. 5. It is remarkable that the primeval background of gravitational waves predicted by inflation could be indirectly (but firmly) detected through the CMB, before it will be observed with much more expensive direct projects targeted on gravitational waves.

In the next interview, Ruth Durrer will widely discuss the theory and the main observational imprints on the CMB of topological defects, often considered as alternative paradigms to the genuine inflationary framework, but whose contribution has also been considered alongside inflation in determining the pattern of CMB anisotropy.

3.3 Topological Defects

Dear Ruth *(Durrer)***, the formation of topological defects is one of the predictions of symmetry breaking phase transitions in the early Universe. Various classes of topological defects have been considered in the literature, often as alternatives to inflation. Could you give a characterization of topological defects and of their main observational peculiarities?**

Topological defects are space–time “regions” of higher energy density, which can appear during symmetry breaking phase transitions[1]. Well known examples are vortex lines in superconductors or “declinations” in liquid crystals. If the manifold of energy minima after the phase transition is sufficiently complex (topologically nontrivial), such defects necessarily form during a phase transition, for example, during the adiabatic expansion and cooling of the Universe. The energy density of such topological defects is intrinsically inhomogeneous and they might therefore represent the “seeds” for the formation of cosmic structures. It has been shown that

[1] For example, when water freezes to ice, the ice crystals assume an arbitrary but fixed direction which breaks rotational symmetry.

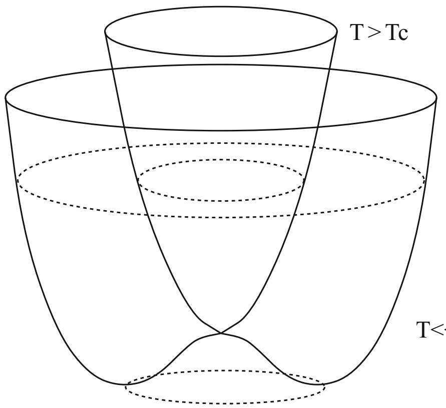

Fig. 3.1 The effective potential of a complex Higgs field for two values of the temperature, $T > T_c$ and $T < T_c$ is shown. The *circle* at the *bottom* is the vacuum manifold $\mathcal{S}$ of the low temperature phase

certain kinds of topological defects always lead to a scale invariant spectrum of fluctuations. In this dialog, we argue that despite this initial success, topological defects are probably irrelevant for structure formation and maybe even for cosmology.

Topological defects can form during symmetry breaking phase transitions. If the vacuum manifold, that is, the manifold of minima of the Higgs field (or order parameter) which is responsible for the symmetry breaking, is topologically nontrivial, regions where the field cannot relax to the minimum generically occur. The simplest example are cosmic strings that form, for example, when a $U(1)$ symmetry is broken. Below a critical temperature T_c, the temperature dependent effective potential $V(\phi, T)$ of the complex Higgs field ϕ changes from a form with a single minimum at $\phi = 0$ to a Mexican hat shape with an entire circle $\mathcal{S}$ of minima, see Fig. 3.1.

During the adiabatic expansion and cooling of the Universe, when the temperature drops below T_c, the field at a given position $\mathbf{x}$ assumes some value in the new vacuum manifold $\mathcal{S}$. The field values at positions that are further apart than the Hubble horizon at T_c are uncorrelated. Therefore, the configuration $\phi(s) = \phi(\mathbf{x}(s))$ along some large closed curve $\mathbf{x}(s)$ in a plane of physical space may well make one (or several) full turns in $\mathcal{S}$. If this happens, to remain continuous, ϕ has to leave the vacuum manifold and assume a value with higher potential energy somewhere in the interior of this curve. Continuing this argument in the third dimension, one obtains a line of higher energy. These lines, which are either closed or infinite, are cosmic strings, see Fig. 3.2. The density of cosmic strings that form in this way during a cosmological phase transition is typically a few per horizon. The mechanism outlined above is called the Kibble mechanism [143, 261] .

As the Universe expands, the Higgs field straightens out. Strings that intersect exchange partners (inter-commutation) and can thereby chop off loops from the network of long strings. In this way the long string network looses energy by shortening the total length of strings. The strings from a broken gauge symmetry interact with other matter components only gravitationally. They shed energy only into a background of gravitational waves that they produce. This process is slow, but sufficiently effective to lead to a mean energy density, ρ_S, in cosmic strings that scales

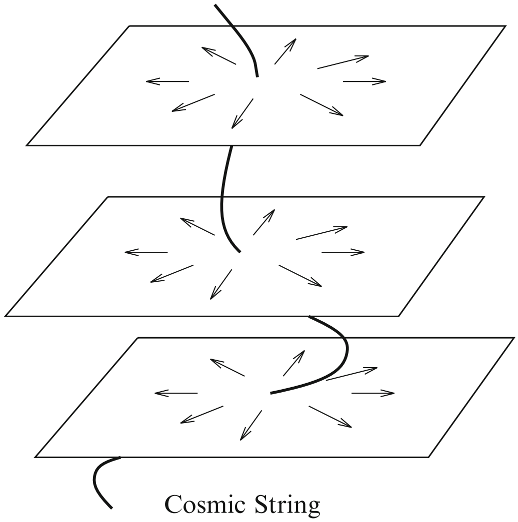

Fig. 3.2 A cosmic string in space is shown with the corresponding configuration of the complex Higgs field, indicated as *arrows*

like the background energy density $\rho_S \propto 1/t^2$, where t is cosmic time. The Hubble parameter as well as the mean density also scale like $H^2(t) \propto \rho(t) \propto 1/t^2$. If $M \simeq T_c$ is the energy scale of the phase transition, we expect $\rho_S \simeq M^2/t^2$, so that

$$\frac{\rho_S}{\rho} \simeq 4\pi G M^2 = 4\pi \left(\frac{M}{M_{Pl}}\right)^2 \equiv \epsilon. \quad (3.6)$$

Here we have introduced the Planck mass[2] $M_{Pl} = 1/\sqrt{G}$. The amplitude of the induced perturbations will be of the order of ϵ. Recalling that the amplitude of CMB anisotropies is roughly 10^{-5}, we infer that the symmetry breaking scale cannot be much smaller than $M \sim 10^{-3} M_{Pl} \sim 10^{16}$ GeV, if such a component is to play a role for CMB anisotropies. Interestingly, this is a Grand Unified Theory (GUT) scale where some drastic changes of physical interactions, as, for example, a phase transition, are expected to occur as the running coupling constants of gauge interactions all intersect at this energy scale. If cosmic strings would be generated at the electroweak transition (which is not the case in the standard model), they would have far too low energy to play a role for structure formation and CMB anisotropies.

Recently, it has been argued that string-like energy distributions are also expected from string theory models with D-branes[3] [60, 134, 211]. These can be either fundamental strings (superstrings) or one-dimensional D-branes. The main difference

[2] In this text, we use unites where $c = \hbar = k_{\text{Boltzmann}} = 1$ so that masses and temperatures are measured in units of energy, for example, electron volts (eV), and lengths and times have the dimension of inverse energy.

[3] D-branes (Dirichlet branes) are sub-manifolds of the higher (10) dimensional space–time of superstring theory on which open strings end. They can have different dimensions, depending on the

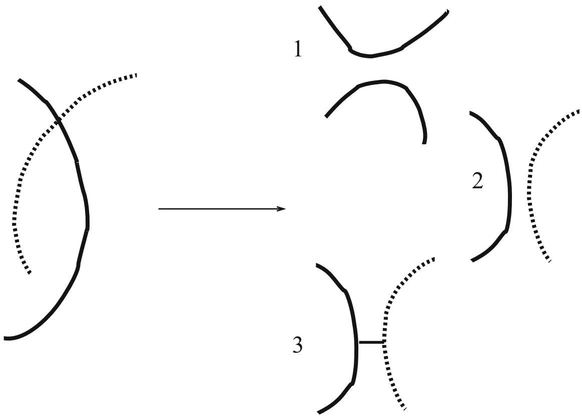

Fig. 3.3 Intersecting cosmic strings can either intercommute (exchange partners, 1), pass through each other (2), or generate a Y-junction (3). Ordinary cosmic strings always do the first. Cosmic superstrings that come in different types do either, depending on the type of string

of such strings from the "old" cosmic strings is that the intercommutations probability is no longer close to one. When they intersect, such strings can intercommute or simply pass through each other, or even form so called Y-junctions by trailing a connecting string, see Fig. 3.3. It is not entirely clear under which conditions such a super string network obeys scaling [61].

Another type of topological defects, called monopoles, occur at symmetry breaking phase transitions if the vacuum manifold of the broken phase, $\mathcal{S}$, has the topology of a sphere. More generically, for the mathematically inclined reader, monopoles form if the second homotopy group, $\pi_2(\mathcal{S})$, is nontrivial. Monopoles are points of higher potential energy. If the broken symmetry is gauged, such massive monopoles cease to interact soon after the phase transition. Their energy density then scales like ordinary matter $\rho \propto a^{-3}$ and soon dominates over the radiation density of the Universe. Therefore, local monopoles are ruled out by observations. Topological defects that form by breaking a gauged (or local) symmetry are strongly localized. Their energy density extends over a scale comparable to the symmetry breaking scale M^{-1}. This corresponds to the thickness of a cosmic string or the size of a gauge monopole.

However, if the symmetry is not gauged, the gradients of the scalar field cannot be compensated by the presence of a gauge field and the energy density decays only slowly. In this case, long range interactions lead to very effective annihilation of monopole–anti-monopole pairs and the remaining energy density has the correct scaling, $\rho_M \propto 1/t^2$.

This is true not only for a symmetry breaking Higgs field. An arbitrary, unordered multicomponent scalar field with a potential minimum at some scale $M \neq 0$ evolves

string theory adopted. The world "branes" probably should remember of "membranes" that are two-dimensional sub-manifolds in 3-space.

in an expanding Universe such that $\rho_S/\rho \simeq GM^2 =$ constant. The field is ordering on the Hubble scale, so that its gradient and kinetic energy are of the order of M^2/t^2. These findings have been confirmed by numerical simulations and they become very accurate for fields with three or more components [80]. For fields with only two components (global strings), this scaling law seems to obtain logarithmic corrections.

One component or real scalar fields do not scale at all. They generically lead to the formation of domain walls that soon come to dominate the energy density of the Universe and are therefore ruled out. (Their vacuum manifold consists of isolated points, so that they have negligible gradient and kinetic energy. Their energy is dominated by potential energy.)

Topological defects from scalar fields with four components are called "texture". Their vacuum manifold is a 3-sphere. A spatial configuration that winds around it has to leave the vacuum manifold in a point in space–time, an event, in order to unwind.

Global scalar fields with more than four components do not lead to topological defects in four-dimensional space–time, but they still scale and therefore are candidates for seeds that can induce structure formation. If the scalar field is "gauged", that is, charged with respect to some gauge group, only cosmic strings scale, gauged monopoles quickly come to dominate the Universe and are therefore ruled out, while gauged textures rapidly thin out and are therefore irrelevant for structure formation.

The distribution of topological defects is intrinsically inhomogeneous and anisotropic. If their mean energy density is scaling, they lead to a scale invariant spectrum of CMB fluctuations. Therefore, they were still a viable candidate of structure formation after the COBE observations, which had uncovered the scale invariant part of the CMB fluctuation spectrum on large scales, see, for example, [76].

3.3.1 Imprints on the CMB

Dear Ruth *(Durrer)*, **in spite of their theoretical relevance, the main properties of the CMB anisotropy recently determined by a large set of experiments, first of all WMAP, and in particular of its angular power spectrum seem consistent with a genuine inflation model with adiabatic fluctuations. One interesting possibility is to consider also a mixture of contributions from inflation and topological defects. Could you discuss how such models can be constrained with present and future CMB observations?**

Let us first discuss the specific signatures imprinted on the CMB by topological defects. The main result is that the presently observed temperature anisotropy spectrum is not compatible with a spectrum from topological defects, but it is compatible with inflation. Nevertheless, it could contain contribution of about 10% on large scales from topological defects.

3.3.1.1 CMB Temperature Anisotropies

Inflation naturally leads to a nearly scale-invariant spectrum of scalar fluctuations (density perturbations, acoustic oscillations) and tensor fluctuations (gravitational waves). These are generated during inflation and evolve freely during the subsequent radiation and matter dominated phase of the Universe. The physics of inflation and its motivation are discussed in detail in the preceding section. However, also scaling topological defects generate a scale invariant spectrum of perturbations.

An important difference in the typical temperature anisotropy spectrum from topological defects as compared to inflation is the absence of the acoustic peak structure that is present for inflationary perturbations, see Figs. 3.4 and 3.5.

This absence of acoustic peaks in the CMB anisotropies from defects has two main reasons, which we now discuss. First of all, topological defects, which are scaled by source gravitational perturbations at the horizon at any moment, typically have important contributions from tensor and especially also from vector perturbations (vorticity, spin-1 perturbations), while inflation does not lead to vector perturbations. Even if the latter were produced during inflation, they simply decay afterwards and cannot leave any imprint on the CMB. For topological defects, the

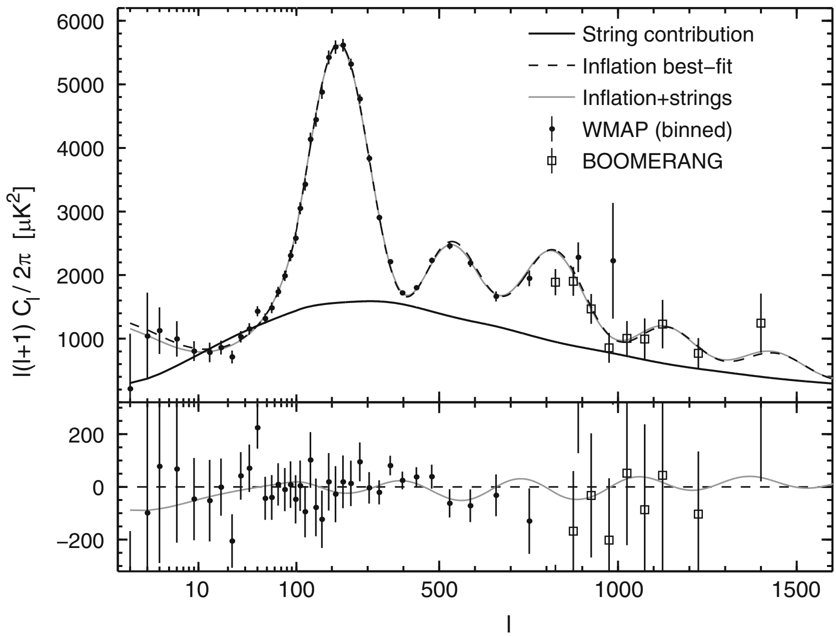

Fig. 3.4 The CMB anisotropies from a model with cosmic strings alone, inflation alone, and a combination of cosmic strings plus inflation compared to the data from WMAP [118] and the balloon observations of millimetric extragalactic radiation and geophysics (BOOMERanG) [135]. The *lower* panel shows the best fit combined model after subtraction of the inflation contribution. Figure from [29]

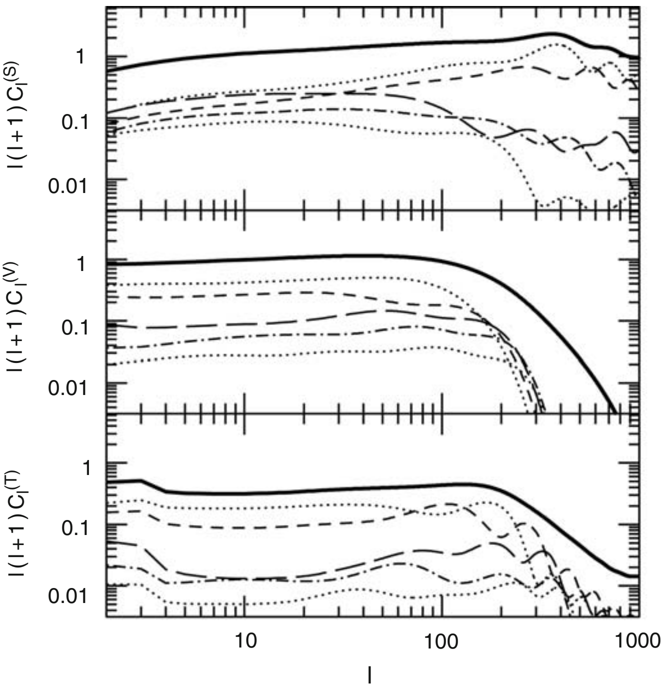

Fig. 3.5 The scalar, vector, and tensor contributions for the texture model of structure formation are shown. The *dashed lines* show the contributions from the first few single eigenfunctions while the *solid line* represents the sum (over 100 eigenfunctions). Note that the single contributions to the scalar and tensor spectrum do show oscillations that are, however, washed out in the sum (vector perturbations do not obey a wave equation and thus do not show oscillations). Data courtesy of Bevis and Kunz (private communication). See also [26]

situation is different. The energy momentum tensor of topological defects does contain significant vector perturbations that induce vector perturbations in the metric at horizon entry. Vector perturbations of topological defects have the same amplitude as scalar perturbations at large scales. Tensor perturbations are somewhat smaller (about 50%), see Fig. 3.5. As both vector and tensor perturbations do not contribute to the acoustic peaks region of the CMB anisotropy spectrum, the acoustic peaks are suppressed in defect models of structure formation.

The second reason is more subtle. One first has to realize that the acoustic peaks from inflation are caused by the fact that all the acoustic oscillations in the radiation/baryon density fluctuations are in phase once they enter the horizon. Only the cosine oscillation, which comes from the growing mode perturbation, survives the long expansion phase after inflation, the sine mode has decayed at horizon entry[4]. Because of this "coherence" of all fluctuations of a given wavelength, they all

[4] This is true for standard adiabatic fluctuations, and for isocurvature fluctuations the situation is opposite, see [75].

oscillate in phase. They are at a maximum, respectively, minimum at the same cosmic time t. This leads to the acoustic peaks in the CMB spectrum, at the harmonics ℓ_n, which correspond to an angle $\theta_n = \lambda_n / D_A(t_{dec})$, where λ_n are the wavelength at which the oscillations are at a maximum (for odd n's) or at a minimum (for even n's) at the moment of recombination, that is, photon decoupling. $D_A(t_{dec})$ is the angular diameter distance to the last scattering surface.

If the fluctuations are generated by topological defects, the situation is different. Topological defects typically induce metric fluctuations of arbitrary phases at the moment of horizon crossing. This decoherence "smears out" the acoustic oscillations, which also reduces their amplitude [6].

Instead of the distinguished acoustic oscillations from inflation, topological defects therefore generate a rather broad "hump" in the acoustic region of the CMB power spectrum, see Fig. 3.5. This is a quite generic result from fluctuations, which have not been initiated at some very early phase of the Universe but only at horizon crossing. I therefore argue that the acoustic peaks observed in the CMB temperature anisotropy spectrum are the most important confirmation of "inflation": they imply that these fluctuations must have been laid down at a time when they were still much larger than the Hubble scale. Hence, the causal horizon must be much larger than the Hubble scale. Therefore, the Universe must have undergone a phase where the causal horizon grew faster than the Hubble scale. If we use such a loose definition of inflation, this would prove that the cosmological fluctuations have to be generated during inflation.

A word of caution is in order here. This "proof" is not watertight, but just generically valid. One can "fine tune" perturbations from causal sources, which are similar to topological defects such that they show acoustic oscillations [78, 79, 255]. But so far nobody could "mimic" inflationary perturbations sufficiently accurately with sources that do not allow for superluminal propagation (Scodeller et al. in preparation).

Even though topological defects are not compatible with CMB data as can be seen in Fig. 3.4, they can still contribute a fraction of about 10% at large angular scales, $\ell \simeq 10$. According to [27], this limits $\mathrm{GM}^2 \lesssim 7 \times 10^{-7}$.

Kaiser Stebbins effect. Because of their discrete nature, cosmic strings lead to an abrupt temperature change or order GM^2 for photons passing on either side of a moving string from the last scattering surface to the observer. This "Kaiser–Stebbins effect" [136] generates small scale structure in the CMB, which is significantly larger than other known small scale effects and will be detected in future CMB observations, which go to $\ell \gtrsim 5{,}000$ if $\mathrm{GM}^2 \gtrsim 2 \times 10^{-7}$, that is, if the defects contribute about 10% to the large scale CMB anisotropies [95]. The difference of this number with the findings of [27, 28], which obtain $\mathrm{GM}^2 \simeq 7 \times 10^{-7}$ for a 10% contribution to the CMB anisotropies, actually lies in the different simulations. I think it is fair to say that the amplitude of the string-induced CMB anisotropies is still uncertain by a factor of about 2. This is not the case for the much simpler global defects, where different simulations are in good agreement.

3.3.1.2 Polarization

Because of the angular dependence of Thomson scattering, CMB fluctuations are slightly polarized at the last scattering surface. This polarization is usually split into so-called E-modes, which represent a gradient field polarization pattern, and B-modes, which represent a curl type polarization pattern on the last scattering surface [75].

The results for temperature anisotropies and polarization comparing cosmic strings and inflationary scalar and tensor modes are shown in Fig. 3.6 for the largest allowed contribution from cosmic strings, about 10% at harmonic $\ell = 10$, and for

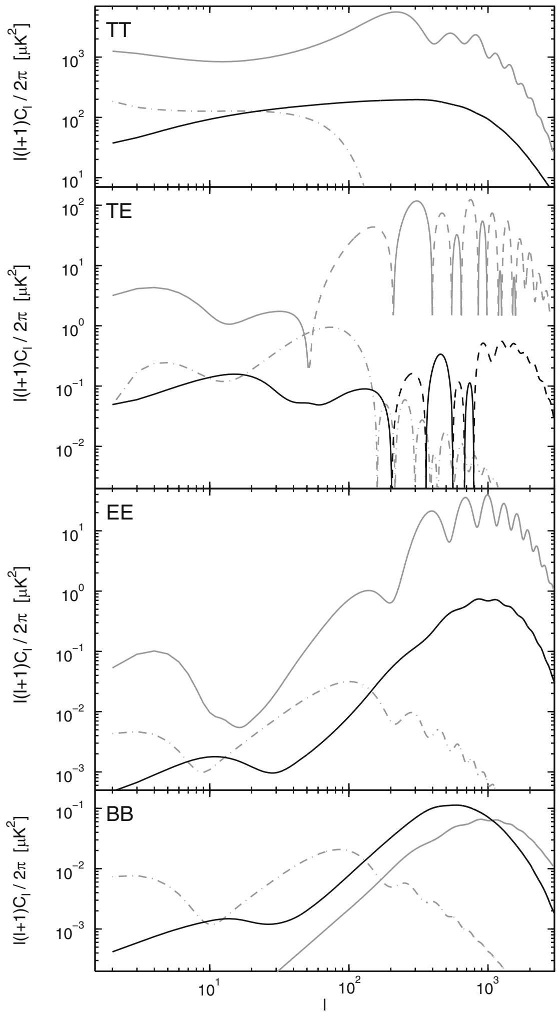

Fig. 3.6 The CMB temperature anisotropy and polarization power spectra from cosmic strings (*black solid*) inflationary scalar modes (*gray, solid*) and inflationary tensor modes (*dashed*). The *dashed* parts in the TE cross correlations for strings and scalar inflationary perturbations denote anticorrelations. The normalization of the string and tensor power spectra correspond to the 95% upper bound allowed from parameter estimation, which corresponds to roughly 11% at $\ell = 10$ for the strings ($GM^2 = 0.7 \times 10^{-6}$) and 15% for the tensors ($r = 0.36$). Figure taken from [28]

the largest allowed tensor contribution, about 15% at harmonic $\ell = 10$ [28], which corresponds to a tensor to scalar ratio $r = 0.35$. The results from global defects are similar, the main difference being that cosmic strings have somewhat more power on smaller scales.

The first interesting remark is that the oscillations of the TE-correlation from cosmic strings is shifted by about half a period from the one of inflationary perturbations. The first zero of the string TE-correlations is at $\ell \sim 200$, where we already find the second zero of the inflationary TE-correlations. This comes from the fact that fluctuations from cosmic strings are closer to the so-called iso-curvature modes (which correspond to sine modes instead of cosine modes in the radiation/baryon density fluctuations).

In the EE power spectrum from cosmic strings, one sees again the effect of "decoherence", fluctuations are smeared out whereas inflationary perturbations show a pronounced peak structure.

The most interesting is the Black Body (BB) power spectrum: despite the fact that the cosmic strings contribute only about 10% to the large scale temperature anisotropies and even less on intermediate scales ($100 \lesssim \ell \lesssim 1000$), they dominate B-polarization in the range $100 \lesssim \ell \lesssim 1000$. The reason for this is that scalar perturbations do not generate intrinsic B-polarization. B-polarization of scalar perturbations comes only from "lensing", that is, the modification of the polarization pattern when the photons pass through the gravitational field of the foreground density fluctuations [75]. This is a second order effect and therefore quite small. Inflationary tensor perturbations do generate B-polarization in first-order perturbation theory but not very much. The most significant contribution to B-polarization comes from vector perturbations that are not at all present in the inflationary case, but which are significant for cosmic strings. It is interesting to note that for topological defects, E- and B-mode polarizations have comparable amplitudes up to $\ell \simeq 100$. For larger ℓs, scalar perturbations, which do not contribute to the B-mode at first order, dominate.

If future B-polarization experiments like Cℓ ObserVER (CℓOVER) [249] only discover the amount of B-polarization expected from lensing of scalar perturbation in the angular range corresponding to harmonics $100 \lesssim \ell \lesssim 1000$, this will allow nearly a factor of 10 tighter limits on a contribution to the CMB fluctuations from topological defects.

3.3.2 Non-Gaussian Anisotropies

Dear Ruth *(Durrer)*, **a remarkable non-Gaussianity level of CMB anisotropy is one of the main predictions of topological defects. On the other hand, several inflation models also predict a certain level of non-Gaussianity. Could you illustrate this point and discuss on the level of precision necessary to disentangle between these two scenarios?**

Topological defects are nonlinear physical objects and generically generate non-Gaussian fluctuations, contrary to inflation. It may therefore be useful to calculate the higher order correlators, like, for example, the bispectrum[5], to distinguish them observationally from inflation [101]. However, such calculations are very time consuming and it is not clear whether they are the most straight forward non-Gaussian signature of topological defects. Already the one-point distribution function becomes significantly non-Gaussian once the amplitude drops below a few percent of the maximum [95].

The most significant non-Gaussianities for cosmic strings will, however, be on small scales of the order of 1 arcmin. Because of the Kaiser Stebbins effect [136], we expect much more pronounced gradients for cosmic strings than in a Gaussian distribution [95]. Fraisse et al. [95] actually conclude that the non-Gaussianity of gradient maps with angular resolution of about 1 arcmin should reveal cosmic strings down to a symmetry breaking scale of $GM^2 \simeq 2 \times 10^{-7}$. However, from their simulation they also find this same limit from the CMB power spectra. Again, the difference of this number with the findings of [27,28] is a result of the difference in the respective cosmic string simulations.

Also this result has been obtained with the WMAP data and it is expected that future experiments with much higher angular resolution will yield significantly tighter limits.

A recent analysis of the remarkable WMAP cold spot seems to support a texture model. Could you comment on the new perspectives opened by this kind of analysis for our comprehension of topological defects?

Recently, it has been proposed that the most prominent cold spot in the CMB sky of about 5° radius might be due to a global texture [63], that is, a topological defect which is formed if a global symmetry is broken and the vacuum manifold has a nontrivial third homotopy group, $\pi_3(\mathcal{S}) \neq 0$. For example, the total breaking of the special unitary group $SU(2)$, which is topologically equivalent to the 3-sphere, leads to textures. In this case, it has been shown that hot and cold spots should appear in the CMB [74,256]. However, numerical simulations indicate that one expects rather of order 3–5 spots of angular scale 5° in the CMB sky from texture unwinding events [77], where the data would show only one.

Of course, there are also other proposed explanations of the observed cold spot, namely that it be a giant void or simply a statistical fluke. Actually, the simulations discussed in [63] show that about 5.8% of all realizations with purely inflationary Gaussian perturbations show the same or even stronger evidence of the presence of a texture than the observed sky. Therefore, such a cold spot is not truly exceptional and to me the interpretation as a texture seems rather far fetched. This opinion is strengthened by the fact that, in contrast to cosmic strings, there is no particle physics motivation for global textures or any exact global symmetry for that matter. Nevertheless, generic self-ordering multicomponent scalar field would probably

[5] The bispectrum is a measure of the 3-point correlations function, that is, the correlation of the temperature fluctuations in three different directions. For Gaussian perturbations, the 3-point correlations function and therefore also the bispectrum vanish.

lead to similar hot and cold spots in the sky. Since the value of the symmetry breaking scale needed to explain the data is $GM^2 \simeq 5 \times 10^{-7}$, we will be able to test it with future polarization experiments like CℓOVER [249].

To Summarize

I have discussed the imprints of cosmological topological defects in the Universe with special emphasis of the candidate whom I consider most interesting, namely cosmic strings. Global defects are not well motivated from particle physics and other local defects do not scale and are therefore either excluded (local monopoles) or cosmologically irrelevant (local texture). The most solid constraints on topological defects come from the CMB anisotropies and polarization[6] to which defects cannot contribute more than about 10% on large scales, this leads to a limit of $GM^2 \lesssim$ a few $\times 10^{-7}$.

Other experimental constraints on cosmic strings and their perspectives will be considered in Sect. 5.10 of Chap. 5.

Thank you Ruth.
The next interview to Martin Bucher will focus on a crucial aspect of cosmology, the nature and classification of the primordial perturbations from which CMB anisotropies, LSS, and, ultimately, all astrophysical objects have formed during cosmic evolution. The existence of perturbations offers in fact a fruitful way to investigate also those properties of the Universe that determine the evolution of its homogeneous "segment" (let us think, e.g., of Λ and H_0) and not only of its inhomogeneities.

3.4 Adiabatic vs. Isocurvature Perturbations

Dear Martin *(Bucher)***, the current data coming from CMB analysis seem to confirm that the cosmological perturbations at the origin of present day structures are essentially adiabatic. However, other kinds of perturbations (e.g., isocurvature) are predicted by many models. Can you briefly describe the problem and specify which observations could potentially solve the present dichotomy?**

The search for isocurvature perturbations constitutes an important test of our understanding of the character of the primordial cosmological perturbations. Before proceeding, it is useful to recall what "adiabatic" means when cosmologists speak of "adiabatic" primordial perturbations, given that the connection to the ordinary

[6] Detailed discussions on present and future CMB experiments can be found in Chap. 2, Sect. 2.8 and in Chap. 5, Sect. 5.3. Theoretical aspects of the CMB are explained in this chapter, Sect. 3.5.

meaning in physics is not entirely self-evident. Adiabatic primordial perturbations are the simplest kind of cosmological perturbations that one can postulate as the "initial conditions" for a Universe that is not completely regular – that is, spatially not completely homogeneous and isotropic. One assumes that at early times the Universe is governed by a common equation of state without spatial variation, so that the total density suffices to determine the densities of the various components (e.g., photons, electrons, neutrinos, baryons, and CDM if we do not set our initial surface too far back in cosmic time) and that in addition these components initially share a common velocity field. To fully appreciate this definition, it is useful to consider some examples of "nonadiabatic" perturbations. One can, for example, as was studied by Peebles and by Efstathiou and Bond, contemplate primordial variations in the baryon-to-photon [191, 192] or the CDM-to-photon number density ratio [34, 193, 194] as well as correlated combinations of the two [40]. To characterize the fluctuations of the equation of state, it is convenient to use photons (or equivalently entropy as the entropy of the Universe resides almost entirely in the photons) as a marker, because the number of photons per unit co-moving volume is conserved at least on superhorizon scales at early times (except for trivial changes by a constant factor occur during the epoch of electron–positron annihilation or during any first-order phase transition). Consequently, if the initial conditions include a spatial variation in the baryon-to-entropy ratio $\eta_{\mathrm{B}} = n_{\mathrm{B}}/n_{\mathrm{S}}$, this spatial variation remains frozen in until late times, when the relevant mode re-enters the horizon and the photons and baryons decouple from each other. We can along the same lines admit initial spatial variations in the analogous quantity $\eta_{\mathrm{CDM}} = n_{\mathrm{CDM}}/n_{\mathrm{S}}$, where n_{CDM} is the Weak Interacting Massive Particles (WIMP) number density. In both these examples, the particles in question at first contribute negligibly to the energy density of the Universe. However, subsequently, as the Universe becomes matter dominated at a redshift of $z \approx O(10^4)$, these perturbations start to manifest themselves gravitationally, and at late times effectively transform into ordinary density perturbations. Consequently, early on, at high z, spatial variations in η_{B} (or in η_{CDM}) contribute negligibly to the total energy density perturbation, which in turn through the Einstein equation is linked to a curvature perturbation. This is why such perturbations in the equation of state are commonly known as "isocurvature" perturbations, and adiabatic perturbations are also known as "curvature" perturbations. In addition to the two above isocurvature modes, the standard cosmology also allows two possible neutrino isocurvature modes, a neutrino density, and a neutrino velocity isocurvature mode [40, 45], and in models with new physics (e.g., with "quintessence") additional modes are possible. While the simplest models of inflation having only a single scalar field necessarily create only adiabatic perturbations, inflationary models with several light scalar fields can also create isocurvature perturbations. Reheating need not occur in the same way at each place because reheating takes places somewhere on a surface rather than always at the same point in field space. On the "reheating surface", the scalar field directions tangential to this surface correspond to the isocurvature modes and the values of these tangential fields can "modulate" the outcome of reheating, for example, by determining how many baryons are produced or how many CDM particles are produced (relative to the entropy density).Because these light fields are disordered during inflation in much the same way as thelongitudinal inflation perturbations, such variations typically have a very red "acausal" spectrum.

While isocurvature perturbations are an intriguing possibility, the present data does not require their presence. The simplest minimal explanation for the present cosmological data (e.g., as described in the 5 year WMAP paper [149]) includes only adiabatic perturbations. However, it remains to be seen what the data will require as observations improve. The precision E-mode polarization data that we expect to obtain from *Planck* promises to greatly improve present constraints on isocurvature modes, or possibly discover something new [41]. While linear combinations of isocurvature modes are possible with very little variation in the CMB temperature anisotropy, this degeneracy is broken by the predicted polarization anisotropy, as indicated in Fig. 3.7.

It is to my mind extremely remarkable how close our present best understanding of the cosmological perturbations is to the earliest speculations on the subject. It should be emphasized that the study of cosmological perturbations substantially predates inflation, and that the first guesses – and guesses they were because at the time the relevant data was not available – essentially coincide with the "predictions" of inflation. The pioneers of the field – such as Peebles and Zel'dovich – started working out the consequences of the simplest hypotheses for the cosmological perturbations well before solid observational data constraining their character was available (see, e.g., the discussion in [190]). The celebrated "scale-invariant" spectrum of Harrison, Peebles, and Zel'dovich [113, 195, 273] predates inflation, a fact that is sometimes forgotten. At that time, mainly motivated by the quest for simplicity, the consequences of small Gaussian perturbations adiabatic in character and parametrized by a power law spectrum were considered. Inflation is certainly very satisfying, in that it offers a plausible theoretical basis for these hypotheses, which when first put forth must have seemed ad hoc and naive. Nevertheless, it is somewhat discomforting to realize that if inflation had not been discovered, cosmological data would probably be analyzed in much the same way as it is now. There are some who believe that inflation has already been proved. However, to my mind, the only truly remarkable new prediction of inflation is the generation of primordial gravity waves, and I hope we shall soon see these gravity waves.

Thank you Martin.
The next wide section collects various interviews devoted to the theoretical implications of the CMB.

In the first interview with John Mather, we come back to a fundamental property of the CMB, its spectrum, or, in a more technical language, the frequency dependence of the intensity of the monopole term (i.e., the average over the whole sky) of the CMB pattern. Although the CMB spectrum is influenced also by physical processes associated to inhomogeneities in the cosmic plasma, it properly refers to the homogeneous "segment" of the properties of the Universe.

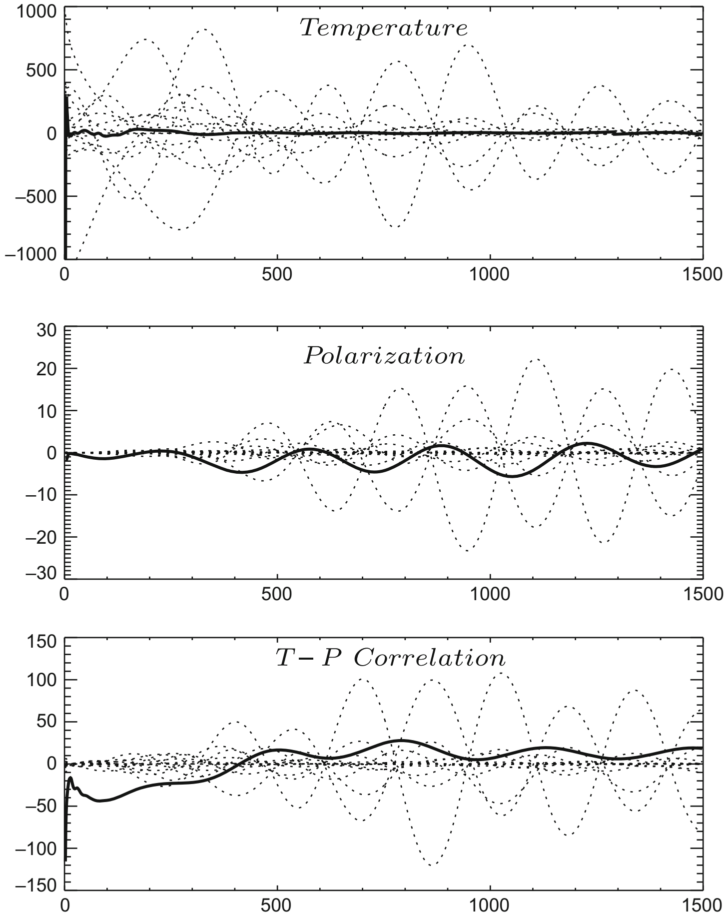

Fig. 3.7 The figure shows the usefulness of polarization data for breaking isocurvature mode degeneracy. A linear combination of isocurvature modes and their correlations is shown for which the temperature anisotropy nearly cancels. Nevertheless, the TE and EE signals allow this degeneracy to be broken, as shown in the two *lower* panels. From [41]

3.5 CMB Theory

3.5.1 Implications of CMB Spectrum Observations

Dear John *(Mather)***, can you summarize the state-of-the-art of CMB spectrum measurements and discuss their main constraints on physical processes at various cosmic times?**

The first scientific results from the COBE are shown in Fig. 3.8. Later analysis of the full data set showed that the best fit temperature is 2.725 ± 0.001 K and that the

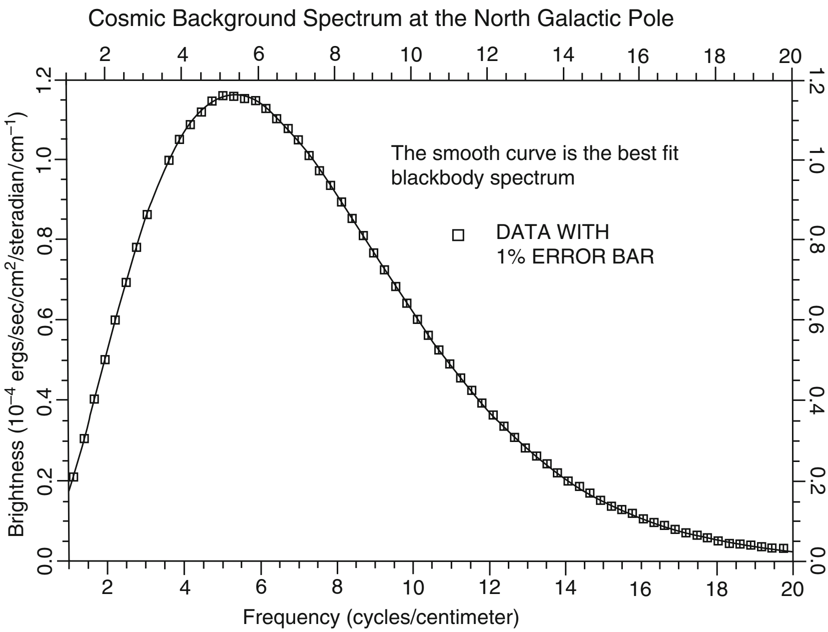

Fig. 3.8 First far infrared absolute spectrophotometer (FIRAS) spectrum, showing 1% error bars and all the observations lying directly on the BB curve. This figure drew a standing ovation at the American Astronomical Society meeting in January 1990

weighted rms error was only 50 parts per million of the peak brightness. In addition, strict limits were set on the two main possible forms of distortion of the spectrum, Bose–Einstein (for high redshifts) and Compton (for lower redshifts). To achieve these results, an extensive least-squares fit was made to all the sky and calibration observations, with terms to represent numerous potential error sources within the instrument, and astrophysical sources such as interplanetary and interstellar dust at a variety of temperatures, and atomic and molecular emissions from the interstellar medium. The error sources to be considered within the instrument included light that passed through the interferometer more than once (after reflection from the detector); thermometer calibration; gain variations due to temperature changes at the detectors; high frequency mechanical oscillation of the mirror mechanism, excited by its servo control system; cosmic ray impacts on the detectors; temperature gradients in the calibrator; radiation leaking past the edge of the calibrator; and radiation (or lack thereof) originating at the detectors or other parts of the optical system, exiting through sky horn, and reflecting back from the calibrator during calibration.

The emission from high-latitude dust in our Galaxy was in some cases difficult to distinguish from broadband distortions of the spectrum. However, it was analyzed by measuring the latitude dependence, and characteristic patterns were seen. After careful subtraction of the latitude-dependent dust emission, it was possible to set strong limits on the broad-band spectral distortions from the early Universe. The Bose–Einstein distribution applies to energy added to the background radiation over the redshift range from 10^5 to 3×10^6. It is characterized by the parameter μ, which is added to the number $x = h\nu/kT$ in the occupation number part of the BB formula. Our measured value was $\mu = -1 \pm 4 \times 10^{-5}$, or a 95% confidence level upper limit of $|\mu| < 9 \times 10^{-5}$. Energy release at a later time produces a mix of BBs at a range of temperatures, and is characterized by the Compton y parameter, an integral over optical depth of the temperature difference between the electrons and the CMB temperature. Our result is that $y = -1 \pm 6 \times 10^{-6}$. In both cases there is a firm upper limit that no more than 1 part in 10^4 of the energy in the background radiation field was added during the specified redshift range.

Of course, the larger implication of the spectrum is the hot Big Bang, the only theory that fits naturally. The Steady State theory has a place for a background radiation field, as the cumulative radiation of stars and other objects throughout time greatly redshifted. However, as we know that the Universe is quite transparent at millimeter wavelengths, there is no good way to explain a spectrum that fits the perfect BB curve, without an equilibration mechanism. Without such a mechanism, it would require a special coincidence to have the sum of radiations from many individual objects at different temperatures to match the radiation of a single object at a single temperature.

Similar arguments apply to the idea of a cold Big Bang, in which the background radiation is produced after the initial explosion and then equilibrated. To produce a large optical depth, one could imagine tiny metallic needles of something like iron. These would have large millimeter wave cross sections, and might be able to convert light from short wavelengths into something resembling a thermal spectrum. However, the extremely close match between the CMB and a BB spectrum again is difficult to explain with these needles.

In the case of the oscillating Universe, there is a possibility that the CMB energy originates in prior cycles. However, in this case, the early parts of this cycle still are extremely hot and dense, and are able to thermalize the heat radiation perfectly. The only quantity we would expect to remain from prior cycles would be the amount of the radiation. So the oscillating Universe cannot be distinguished from the standard hot Big Bang on the basis of the BB spectrum.

There is still the possibility that at wavelengths greater than 1 cm, there will be observable broad-band distortions of the CMB spectrum, due to hot electrons post-recombination. The free-free process could produce a distortion that would not have been detectable with the COBE FIRAS. Experiments are under way to search for or set limits on such distortions (see Sect. 5.3.1 in Chap. 5).

Thank you John.

In the next interviews, we come back to those aspects related to the geometry of the Universe and to its inhomogeneities. When studied through the CMB, they are projected over the celestial sphere and appear as anisotropies (or fluctuations) of the CMB pattern.

3.5.2 CMB Anisotropy

Let us start with the interview with Nicola Vittorio on what is likely the most usual and fruitful way in which, up to now, the properties of CMB anisotropies have been studied.

3.5.2.1 Angular Power Spectrum

Dear Nicola *(Vittorio)*, **the CMB anisotropy pattern is usually analyzed in terms of the so-called angular power spectrum (APS). Can you summarize why and how the main statistical information contained in the CMB maps can be compressed in this estimator? Since the earlier studies devoted to the physics of CMB anisotropy, it was clear that the APS is a very powerful tool keeping memory of the properties and evolution of the Universe. Can you discuss the fundamental steps in the development of the analytical and numerical methods for the computation of CMB anisotropy APS in the context of the current cosmological models?**

One of the most powerful tools available to modern cosmologists is the possibility of reconstructing the status of the Universe when it was just about 3.8×10^5 years old (or about 13.7 billions years ago), by studying the CMB radiation, the fossil relic of the hot and dense phase that followed the big bang. The CMB is a thermal (or BB) radiation – the same kind of radiation emitted by a body at a certain temperature. The CMB radiation filling the present Universe corresponds to a temperature of only about 2.725 K. The CMB looks very isotropic – that is, it appears to have almost the same intensity in every direction of the sky. But there are also slight fluctuations around the average temperature, of the order of tens of a millionth of degree: these were first discovered by the COBE satellite in the early 1990s [231].

The presence of tiny fluctuations in the CMB temperature is a direct consequence of the fact that the early Universe had slightly different densities at different points in space. Those small density perturbations later served as the source for the growth of cosmic structures: as time passed, gravity attracted more and more mass around overdense regions, eventually forming galaxies, clusters of galaxies, etc. But in the hot and dense phase which followed the big bang, matter and radiation were tightly coupled and in thermal equilibrium. Then, any inhomogeneity in the matter density left an imprint in the CMB radiation – roughly speaking, overdense regions resulted in slightly hotter than average CMB temperatures, while underdense regions resulted in slightly colder than average CMB temperatures. The observed pattern of CMB fluctuations, then, literally represents a snapshot of the density distribution in the Universe at a very early epoch of its evolution.

How can we extract useful physical information from the detailed temperature pattern of the CMB? First, we have to keep in mind that theory cannot predict exactly which CMB temperature value we will observe in a particular direction of the sky. According to the conventional scenario, density perturbations were generated by amplifying random quantum fluctuations during an early phase of accelerated cosmic expansion called *inflation* [5]. As a consequence, we expect the CMB temperature to fluctuate randomly on the sky. Our theories have predictive power, however, because we can still calculate the expected statistical properties of the CMB pattern and compare them to the observed values.

By far, the statistical estimator that had the widest application to CMB is the APS. The APS compresses the large amount of information contained in the CMB temperature pattern into a more manageable form: an average quantity that maintains all the information on the temperature distribution of the CMB on the sky. Essentially, the APS tells us what is the average contribution to the CMB anisotropy coming from different angular scales (see Fig. 3.9). To obtain the APS from observations, a discretized image of the CMB (represented by a sky map in which each pixel is the CMB temperature in a certain sky direction) is analyzed using techniques that are very similar to those used in the analysis of sounds (like spectral Fourier decomposition), so that the power from modes of different angular frequency can be singled out. Think of it as decomposing a CMB image into many more images, each one containing only temperature fluctuations that have a precise angular size on the sky,

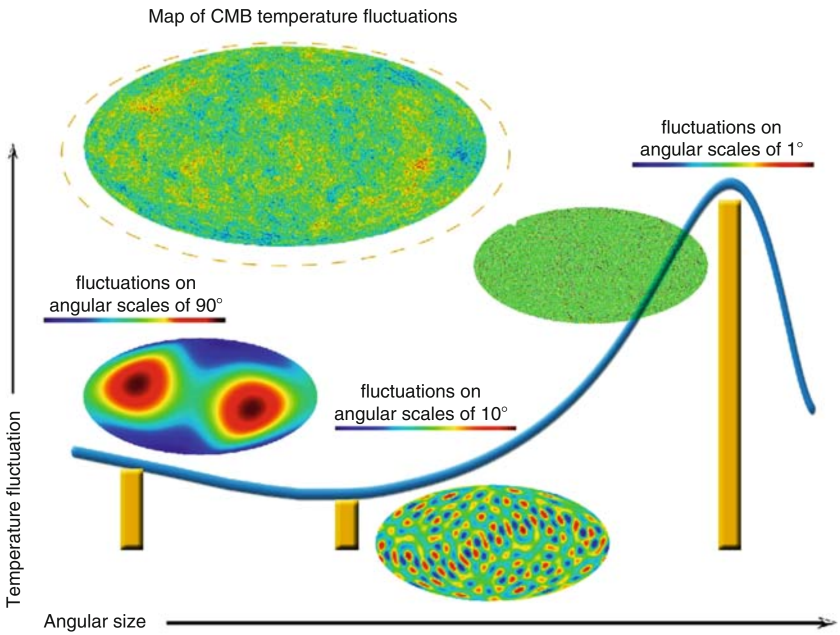

Fig. 3.9 The CMB angular power spectrum is extracted from a CMB map by decomposing the contribution from temperature fluctuations of given angular scale. The height of the spectrum, then, gives a measurement of the level of fluctuations at each given angular scale on the sky, allowing to reconstruct the phase of acoustic oscillation of different density modes at recombination

and assigning a relative weight to each image. In technical terms, if the temperature fluctuation $\Delta T/T$ as a function of the direction θ, ϕ on the sky is expanded into spherical harmonics

$$\frac{\Delta T}{T}(\theta, \phi) = \sum_{lm} a_{lm}, Y_{lm}(\theta, \phi) \tag{3.7}$$

then the angular power spectrum C_l is defined as the average of the quantity $|a_{lm}|^2$ taken over every position in space. Each multipole number is related to an angular scale given approximately by $\theta \approx 180°/l$.

Why is the APS a good way to compress the CMB information? The key is in the physics that generates the CMB temperature anisotropy pattern [121]. Theoretical models tell us that, before recombination, the presence of density perturbations results in the production of sound waves in the matter–radiation fluid. This is because density perturbations oscillate under the competing action of gravity (that tends to cause the collapse of overdense regions) and of internal pressure (that instead forces overdense regions to expand outwards). The frequency of these sound waves is related to the spatial extent of perturbations: larger overdense regions tend to oscillate more slowly (i.e., at lower frequencies) than smaller ones. Sound waves of different frequencies arrive with different phases at recombination. Also, if inflation was the mechanism that generated density perturbations in the early Universe, sound waves with the same frequency arrive with the same phase at recombination. This latter point is crucial, because it allows us to look for a coherent pattern in the APS of the CMB, carrying valuable information on the physical properties of the Universe.

More specifically, the APS extracted from a CMB map will show a series of peaks at certain characteristic frequencies. These frequencies are multiple (or harmonics) of a fundamental frequency. To understand how this pattern is generated, we have to keep in mind that some sound waves were in a phase of maximum compression at recombination. Those were the sound waves associated to density perturbations whose spatial extent corresponded to the distance a sound wave had traveled from the big bang to recombination – the so-called *sound horizon.* The frequency of such sound waves will correspond to the first peak (fundamental) in the APS. Sound waves at twice the frequency (corresponding to density perturbations half the size of sound horizon) oscillated twice before recombination, and were also in an extreme of the oscillations, and so on. These sound waves will produce the higher harmonics in the APS.

Much like studying the overtones of a musical instrument, the detailed pattern of peaks in the APS extracted from a CMB map can be used to obtain information on the specific characteristic of the cosmological model. Of course, not only do we have to measure the APS by performing CMB observations with high sensitivity and high resolution; we also have to produce extremely accurate predictions of the APS peak structure for each particular cosmological model. This task was accomplished over the course of the past decades, starting from pioneering calculations that could provide the average CMB fluctuation expected at a given angular scale after quite lengthy numerical integrations, and finally arriving at modern codes that can produce high-accuracy predictions in a matter of seconds. Analytical study of the theoretical APS were also crucial to disentangle the contribution from various physical processes to the overall peak pattern.

One spectacular application of this kind of investigation is that cosmologists can probe the geometry of the Universe by measuring the angular frequency of the first peak in the APS. The Universe can have one of three possible geometries, defined by a certain constant curvature of space: an open, negatively curved geometry (often visualized using the surface of a saddle as a two-dimensional analogy), a closed, positively curved geometry (a sphere, in two dimensions), and a flat, non-curved geometry (similar to a plan). The latter, Euclidean geometry, is a firm prediction of inflationary models. The way to deduce the geometry of the Universe from the CMB angular power spectrum is as conceptually simple as it is powerful. The sound horizon scale, imprinted in the CMB pattern and in the APS as the fundamental frequency of acoustic waves, acts as a standard ruler: that is, a fixed, known dimension, which we see under a certain angle when we look at it from a certain, unknown distance. Specifically, here the distance corresponds to the path traveled by the CMB photons from recombination to the present (a scale comparable to size of the observable Universe). The geometry of the Universe alters the path of photons along the way, making the trajectories converge or diverge as a consequence of the curvature of space. By comparing the angular size of the sound horizon scale observed in the CMB APS with the theoretical predictions, we can then measure the geometry of the Universe (see Fig. 3.10).

One major milestone in CMB research was the measurement of the first peak frequency performed by the BOOMERanG [65] and Millimeter Anisotropy eXperiment Imaging Array (MAXIMA) [11] experiments in 2000, which allowed the first determination of the geometry of the Universe. The curvature is close to zero, that is, the Universe has a flat, Euclidean geometry, as predicted by inflation. As the curvature of space is related to the overall content of the Universe, this result is also a measurement of the average cosmic density, which is close to the so-called critical value (roughly, $10^{-29}\,\mathrm{g\,cm^{-3}}$). This finding was confirmed to higher accuracy by WMAP in 2003 [232].

Another important quantity that we have estimated from the peak structure of the CMB APS is the average density of atomic matter (or *baryonic* matter) in the Universe. This can be deduced by observing the ratio of even and odds peaks (in particular, of the first to the second) in the spectrum. The reason is that even and odd peaks are associated with sound waves, which were maximally compressed and rarefied at recombination, respectively. Compression of the matter–radiation fluid is assisted by gravity, while rarefaction is driven by the radiation pressure. As a result, if more baryons are in the fluid, relatively to photons, the compression phase will be enhanced, and so will the height of the even peaks with respect to the odd ones. By comparing theoretical predictions to APS measurements, we were able to measure the average density of atomic matter in the Universe to great precision: this is about 4% of the critical density. This result is in remarkable agreement with the one found using a different method, namely the observation of the abundance of light nuclei in the Universe. This concordance is an important consistency check

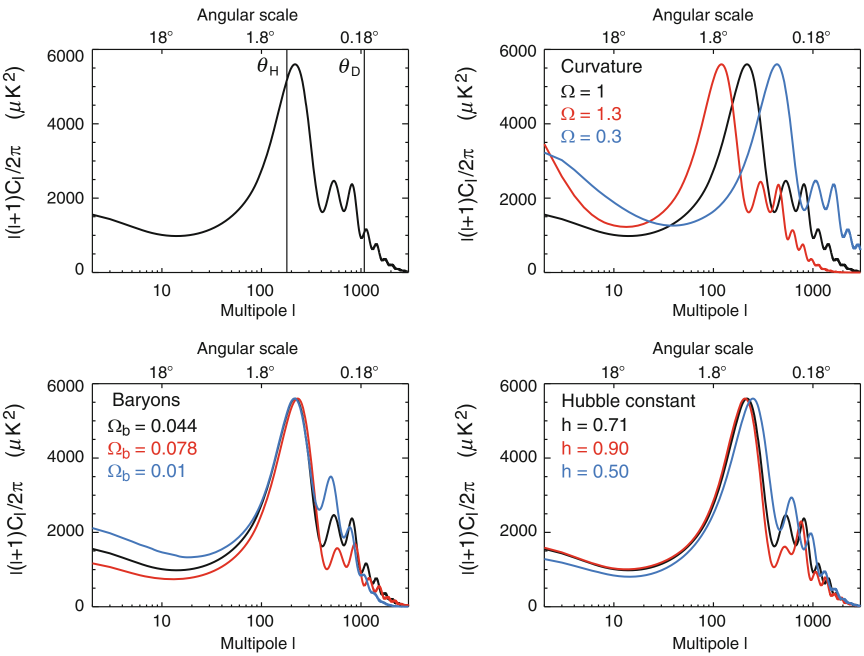

Fig. 3.10 The detailed pattern of peaks and trough in the CMB angular power spectrum is strongly dependent on cosmological parameters. In the *upper left* panel, the theoretical power spectrum corresponding to the inflationary model, which best fits current observations. In this plot, the vertical lines correspond to the approximate angular scale of causal horizon (*left*) and to the damping scale (*right*). In the remaining plots, the effect of varying some cosmological parameters with respect to the fiducial model: *clockwise*, the curvature of the Universe, related to the total density parameter, the reduced Hubble constant h, and the density parameter of baryons

of our cosmological theories. However, it also points to a problem with our cosmic inventory: we currently miss an explanation for more than 95% of the mass needed to obtain the critical density value.

Even if we do not know exactly what is making up the missing mass of the Universe, we can still try to assess how much it contributes to the total density. The result we can obtain using different kinds of cosmological probes is that the Universe contains about 30% of the critical density in form of some unknown CDM (most likely, made up of massive elementary particles with small velocity dispersion), which interacts only gravitationally with the rest of the components of the cosmos. The CMB angular power spectrum encodes information on the Dark Matter (DM) content because acoustic oscillations are sustained only if DM provides the necessary gravitational force (otherwise, radiation pressure rapidly dissipates density perturbations on smaller scales). Then, the height of higher order harmonics is a clear indicator of the quantity of CDM existing in the Universe.

Below a certain physical scale, sound waves are rapidly damped by dissipation processes (the mechanism is similar to the one which prevents very high frequency pitches to propagate through air). This is why we basically find negligible temperature fluctuations at very small angular scales. Measuring the shape of the damping tail of the CMB angular power spectrum, however, provides us with an interesting test of the physics of the matter–radiation coupling before recombination.

On large angular scales, the CMB APS is basically featureless. This is because those angular scales encompass regions of the Universe, which had not enough time to exchange physical information before recombination (i.e., they were larger than the causal horizon). This does not mean that large angular scales are uninteresting: quite the contrary. The information on the status of the density perturbations in the very early Universe is preserved in the shape of the APS at those scales (since no physical process could have altered that), and we can use this fact as a test of the post-inflation, primordial perturbation spectrum. Moreover, large angular scales are also relevant because they are subtly modified by post-recombination processes, for example, by the passage of the CMB photons through ionized regions of the intergalactic medium, or through zones of structure formation. The latter process (called Integrated Sachs–Wolfe (ISW) effect) is very sensitive to the presence of a cosmological constant or Dark Energy (DE) in the Universe. Understanding the nature of this unknown component (currently thought to make up almost 70% of the critical density) is one of the greatest challenges posed to modern cosmology.

The APS is a rich source of cosmological information, which cosmology has just started to exploit. The next few years will see increasingly more accurate observations of its structure that, when combined with the detailed modeling of the physical mechanism that generated it, will further our knowledge of the structure and evolution of the Universe.

Thank you very much Nicola.
As discussed in the earlier interview, under certain hypotheses, the APS fully describes the physical information contained in CMB anisotropies. On the other hand, this is no longer true under more general conditions and, in general, it is crucial for cosmology to set stringent verifications to the validity of assumptions like homogeneity, isotropy, and Gaussianity. Peter Coles will address these aspects in the next interview.

3.5.2.2 Analysis of Phases

Dear Peter *(Coles)*, **together with your collaborators you have recently proposed to analyze the phases of the CMB anisotropy pattern to complement the information contained in the angular power spectrum. Could you describe how these two different kinds of analysis complement each other?**

In the simplest theories of cosmological structure formation, the primordial density fluctuations were generated by an episode of inflation and these fluctuations comprise a Gaussian random field. To test for departures from this assumption, and to tap the rich source of information provided by present and future CMB maps, it is important to devise as many independent statistical methods as possible to detect, isolate, and diagnose the various possible forms of such departures. One particularly fruitful approach is to look at the behavior of the complex coefficients that arise in a spherical harmonic analysis of CMB maps. Chiang et al. [47, 48] and Coles et al. [57] have focused on the phases of these coefficients on the grounds that the definitive property of a statistically homogeneous and isotropic Gaussian random field is that these phases are random.

It is normal practice to describe the distribution of temperature fluctuations in the microwave background over the celestial sphere using a sum over a set of spherical harmonics:

$$\Delta(\theta,\phi) = \frac{T(\theta,\phi)-\bar{T}}{\bar{T}} = \sum_{l=1}^{\infty}\sum_{m=-l}^{m=+l} a_{l,m} Y_{lm}(\theta,\phi). \tag{3.8}$$

Here $\Delta(\theta,\phi)$ is the departure of the temperature from the average at angular position (θ,ϕ) on the celestial sphere in some coordinate system, usually Galactic. The $Y_{lm}(\theta,\phi)$ are spherical harmonic functions that we define in terms of the Legendre polynomials P_{lm} using

$$Y_{lm}(\theta,\phi) = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}\, P_{lm}(\cos\theta)\, \mathrm{e}^{im\phi}; \tag{3.9}$$

this is the Condon–Shortley phase convention. In this equation, the $a_{l,m}$ are complex coefficients, which can be written

$$a_{l,m} = x_{l,m} + iy_{l,m} = |a_{l,m}| \exp[i\phi_{l,m}]. \tag{3.10}$$

Note that, as Δ is real, this definition requires that the phases modes at m and $-m$ be equal and opposite. From this it is clear that the $m = 0$ mode always has zero phase. There are consequently only l independent phase angles describing the harmonic modes at a given l. Without loss of information, we can therefore restrict our analysis to $m \geq 0$.

If the primordial density fluctuations form a Gaussian random field in space, the temperature variations induced across the sky form a Gaussian random field over the celestial sphere. This means that

$$\langle a_{l,m} a^{*}_{l',m'} \rangle = C_l \delta_{ll'} \delta_{mm'}, \tag{3.11}$$

where C_l is the angular power spectrum, the subject of much scrutiny in the context of the CMB (e.g., [117]), and $\delta_{xx'}$ is the Kronecker delta function. As the phases are random, the stochastic properties of a statistically homogeneous and isotropic Gaussian random field are fully specified by the C_l, which determines the variance of the real and imaginary parts of $a_{l,m}$, both of which are Gaussian:

$$\sigma^2(x_{l,m}) = \sigma^2(y_{l,m}) = \sigma_l^2 = \frac{1}{2}C_l. \tag{3.12}$$

Notice that C_l is constructed in such a way as to remove all phase information. Using phases will provide information that is essentially complementary to the power spectrum.

Note also that the bispectrum,

$$B(l_1, m_1, l_2, m_2, l_3, m_3) \equiv \langle a_{l_1,m_1} a_{l_2,m_2} a_{l_3,m_3} \rangle, \tag{3.13}$$

which involves a generalization of the power spectrum to products of three modes taken in such a way that B is rotationally invariant, can only be nonzero if there are correlations in the phases of the relevant modes; see, for example, [265]. The bispectrum is therefore in some sense a measure of phase correlations (and so are its higher-order extensions, the polyspectra).

A graphic demonstration of the importance of phases in determining the morphology of a pattern generally is given in Fig. 3.11. This figure shows a two-dimensional simulation of galaxy clustering, which is expanded in Fourier modes rather than spherical harmonics, but the relevance is clear.

As the amplitude of each Fourier mode is unchanged in the phase reshuffling operation, these two pictures have exactly the same power-spectrum, $P(k) \propto |\tilde{\delta}(\mathbf{k})|^2$. In fact, they have more than that: they have exactly the same amplitudes for all $\mathbf{k}$. They also have totally different morphology. Sharply localized features are associated with correlated phases, rather than with features in the power spectrum; see [56].

While the phases clearly contain information, extraction of useful insights from them is not so straightforward. One simple approach is to start with a set of phases $\phi_{l,m}$ corresponding to a set of spherical harmonic coefficients $a_{l,m}$ obtained from a data set, either real or simulated. We can also form phase differences according to

$$D_m(l) = \phi_{l,m+1} - \phi_{l,m}. \tag{3.14}$$

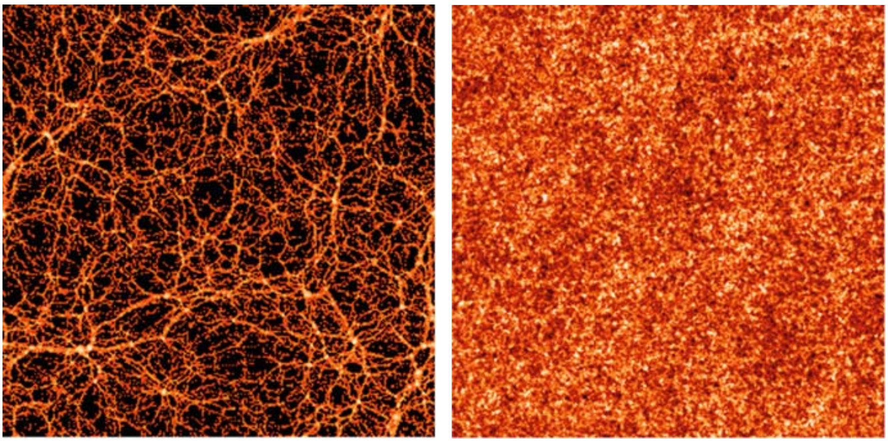

Fig. 3.11 Numerical simulation of galaxy clustering (*left*) together with a version generated by randomly reshuffling the phases between Fourier modes of the original picture (*right*)

If the orthodox cosmological interpretation of temperature fluctuations is correct, the phases of the $a_{l,m}$ should be random and so should phase differences of the form $\phi_{l,m+1} - \phi_{l,m}$ and $\phi_{l+1,m} - \phi_{l,m}$. Let us assume, therefore, that we have n generic angles, $\theta_1 \ldots \theta_n$. Under the standard statistical assumption, these should be random, apart from the constraints described in the previous section. The first thing we need is a way of testing whether a given set of phase angles is consistent with being drawn from uniform distribution on the unit circle. This is not quite as simple as it seems, particularly if one does not want to assume any particular form for actual distribution of angles, such as a bias in a particular direction; see [94]. Fortunately, however, there is a fully nonparametric method available, based on the theory of order statistics, and known as Kuiper's V statistic [153]. In this context, it is convenient to determine significance levels using Monte Carlo simulations of the "null" hypothesis of random phases.

The first point to mention is that a given set of phases, say belonging to the modes at fixed l, is not strictly speaking random anyway, because of the constraints noted in the previous section. One could deal with this by discarding the conjugate phases, thus reducing the number of data points, but there is no need to do this when one can instead build the required symmetries into the Monte Carlo generator. In addition, suppose the phases of the temperature field over the celestial sphere were indeed random, but observations were available only over apart of the sky, such as when a Galactic cut is applied to remove parts of the map contaminated by foregrounds. In this case, the mask may introduce phase correlations into the observations and so the correct null hypothesis would be more complicated than simple uniform randomness. As long as any such selection effect were known, it could be built into the Monte Carlo simulation. One would then need to determine whether V from an observed sky is consistent with having been drawn from the set of values of V generated over the Monte Carlo ensemble.

There is also a more fundamental problem in applying this test to spherical harmonic phases. This is that a given set of $a_{l,m}$ depends on the choice of a particular coordinate axis. A given sky could actually generate an infinite number of different sets of $\phi_{l,m}$, because the phase angles are not rotationally invariant. One has to be sure to take different choices of z-axis into consideration when assessing significance levels, as a random phase distribution has no preferred axis while systematic artifacts may. A positive detection of nonrandomness may result from a chance alignment of features with a particular coordinate axis in the real sky unless this is factored into the Monte Carlo simulations too. For both the real sky and the Monte Carlo skies, we therefore need not a single value of V but a distribution of V-values obtained by rotating the sky over all possible angles. Details of how to implement this are given in [57].

To apply these ideas to make a test of CMB fluctuations, we first need a temperature map from which we can obtain a measured set of $a_{l,m}$. Employing the above transformations with some choice of Euler angles yields a rotated set of the $a_{l,m}$. It is straightforward to choose a set of angles such that random orientations of the coordinate axis can be generated. Once a rotated set has been obtained, Kuiper's

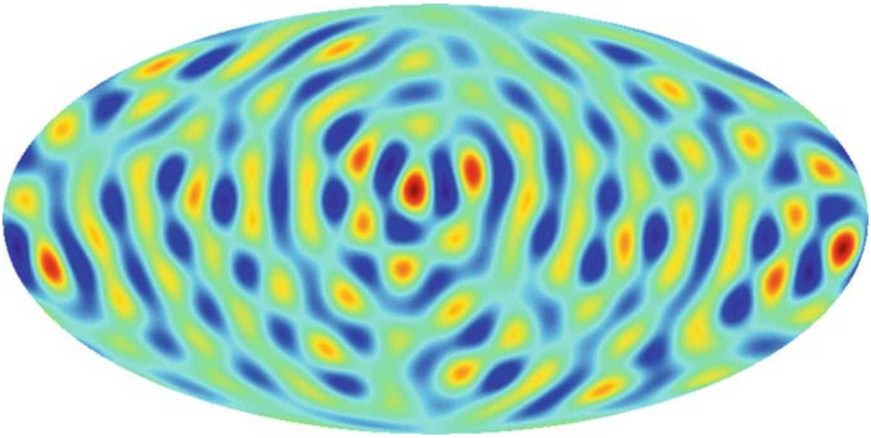

Fig. 3.12 A reconstruction of the WMAP internal linear combination (ILC) made using the spherical harmonic mode amplitudes $a_{l,m}$ for $l = 16$ only. Our analysis method [57] shows that these modes at different m have correlated phases in harmonic space, and the reconstructed sky shows this is aligned with the Galactic Plane

statistic is calculated from the relevant transformed set of phases. For example, [57] generated 3,000 rotated sets of each CMB map using this kind of resampling of the original data, producing 3,000 values of V_{cmb}. The values of the statistic were then binned to form a measured (re-sampled) distribution of V_{cmb}. The same procedure is applied to the 1,000 Monte Carlo sets of $a_{l,m}$ drawn from a uniformly random distribution, that is, each set was rotated 3,000 times and a distribution of V_{MC} under the null hypothesis is produced. These realizations were then binned to create an overall global average distribution under the null hypothesis. Application of this relatively straightforward method to the WMAP first-year data shows the existence of phase correlations, as demonstrated in Fig. 3.12.

In this particular case, we have found a departure from statistical homogeneity rather than Gaussianity per se. The effect we have found is a residue of Galactic foreground subtraction and is not unexpected in the ILC map. It does, however, demonstrate the potential of using phases to locate unexpected properties of maps like this.

Thank you Peter.
As seen earlier, CMB takes precise information on the cosmological parameters which, ultimately, can be seen as a set of "numbers" that in a given cosmological model determine the global evolution of the Universe and the statistical properties of its inhomogeneities via their growth from small primordial perturbations to current structures. On the other hand, the definition of the set of cosmological parameters is far from trivial. We will discuss the aspects related to this issue in the next interview to David Spergel, mainly in the light of WMAP data.

3.5.3 Cosmological Parameters from WMAP

Dear David *(Spergel)*, **the rapid advances of observational cosmology have led to the establishment of the Concordance Model, characterized by a certain set of the so-called "cosmological parameters". This term, however, may be misleading. Would you like to precise what are these parameters and how much we can trust in their current values?**

There are two parts to this question: is the ΛCDM model correct? Do we trust the observations that are used to determine the best-fit cosmological parameters? In the case of the CMB observations from the Wilkinson Microwave Anisotropy Probe (WMAP), I think that the answer to the second question is "yes". There are multiple checks of experimental systematics and there are relatively small corrections for foregrounds. The answer to the first question, the more interesting question, is less clear.

Our current best fit model assumes that the Universe is flat and is composed of atoms, matter, and DE. Some process in the early Universe, usually assumed to be inflation, generates Gaussian random-phase nearly scale-invariant fluctuations. In many ways, this is a very simple model. The fluctuations are characterized by only two numbers, an overall amplitude and a slope. These two numbers determine all of the statistical properties of the fluctuations. Once we have specified these two numbers, the expansion rate of the Universe (the Hubble constant) and the density in these components, we have completely determined the large-scale properties of the Universe.

While the model is simple, is it correct? Given that we do not understand either the physics of the DM or the physics of DE, it seems quite plausible that the current Concordance Model is a phenomenological model that fits currents data but does not fully capture the underlying physics.

If we assume that the ΛCDM model is a good description of the Universe, then in the current WMAP data a Universe is best fit by a Hubble constant of $73\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, a matter density of $\Omega_\mathrm{m}h^2 = 0.1277$, an atomic density of $\Omega_\mathrm{b}h^2 = 0.02223$, a slope of 0.958, and an optical depth of 0.089. The Hubble constant is determined to about 5%, the matter density to about 10%, the atomic density to 5%, the slope to 2%, and the optical depth has a 30% uncertainty. These values, determined only from the WMAP data, agree remarkably well with measurements done by other techniques: determination of the atomic density based on the abundance of Deuterium and Helium, observations of large-scale structure, observations of gravitational lensing, X-ray cluster measurements, and determination of the Hubble constant from HTS. This concordance suggests that the model is a good description of our current observations.

If, on the other hand, we assume that there is no DE, the best fit parameters are not consistent with a host of astronomical measurements. For example, the best fit value for the Hubble constant is 30. The large-scale structure, lensing, and X-ray measurements are not consistent with the parameters in this model.

For the WMAP data, there are two possible types of systematic experimental error: problems with the sky maps and problems with our analysis of the sky maps.

As WMAP measures the sky with ten independent detectors at five different frequencies, we have many independent tests of the data. We also remeasure the sky every year. By looking at the statistical properties of the differences between the various versions of our maps (different years and detectors), we can test our understanding of the statistical properties of the maps. On the large scales, we can also compare our maps to measurements by COBE, WMAP predecessor. On smaller scales, we can compare our maps to measurements done by various balloon-based and ground-based experiments. At present, all of the experiments are in agreement.

The WMAP team has made its sky maps (and its raw data) publically available at LAMBDA (see web page list). Many different groups have downloaded our data and reanalyzed our maps with different foreground removal techniques and with different statistical tools. Generally, they have reached the same conclusions.

What is the impact of the choice of a given set of cosmological parameters in the interpretation of experimental data? How the use of more kinds of astronomical and cosmological data allows one to test more complex cosmological models possibly defined by a larger number of parameters?

Our simplest model assumes that the DE is a cosmological constant. If we assume a more general model and allow the DE properties to evolve with time, then there are a degenerate set of parameters consistent with the CMB data alone. The CMB data is mostly probing physical conditions at the surface of last scatter (a redshift of 1,100) and is not very sensitive to physics at lower redshift. If, however, we simultaneously fit the CMB data and other data sets, then we can often eliminate this degeneracy and determine cosmological parameters in a more complicated model with larger parameters.

There are many possible generalization of the ΛCDM model: The primordial fluctuations could be an admixture of adiabatic and isocurvature fluctuations; the geometry of the Universe could be curved; the neutrinos could make a significant contribution to the energy density of the Universe; the primordial power spectrum could be complicated; or the fine structure constant could be evolving with time. By adding other astronomical data, we can test these models.

A particularly promising combination involves using measurements of the clustering of galaxies. We can measure the same features (baryon oscillations) in both the CMB and the galaxy clustering. As we can measure these features at different redshifts, they serve as a cosmic ruler. Measurements of cosmic distance determine the evolution of the density of the Universe and can constrain many of these more general models.

We are in the midst of a "golden age" in cosmology. While the WMAP data was a significant improvement over previous measurements, upcoming experiments will measure microwave background fluctuations with significantly higher resolution. Measurements of galaxy clustering and gravitational lensing are also rapidly improving. The combination of these observations will enable improved and more precise tests of the ΛCDM model.

Recently, a certain attention has been devoted in the cosmological community to models with a running spectral index or a large scale cut-off in the power spectrum. There are controversial claims about their relevance for the accurate interpretation of current cosmological data. Can you provide your opinion on their ability to explain such data? How these models are linked with inflation?

Both WMAP and COBE observed a very intriguing feature in the microwave background sky: the amplitude of fluctuations on the largest angular scale (the microwave background quadrupole) is low. Fewer than 5% of all observers in the Universe are expected to see such a low value for the quadrupole. Some cosmologists interpret this as a statistical fluke, significance of which is not worthy of detailed exploration. Others view it as evidence for new physics.

Several different explanations have been put forward for the low quadrupole. One explanation is to assume that there is a special feature in the power spectrum that cuts off the amplitude of fluctuations on large angular sky. While the addition of this parameter slightly improves the fit, it makes no other easily testable prediction and is finely tuned.

Another possibility is represented by a running spectral index. This implies that the slope of the power spectrum changes with the scale. This model has the advantage of improving the fit on both large scales and small scales. While the running spectral index model is a slightly better fit to the data, the statistical significance of the improved fit is not large enough to strongly favor this model over the simpler (and better motivated) constant spectral index model.

Future data will soon test the running spectral index model. In the next few years, the *Planck* satellite (scheduled to be launched in 2009) and a series of ground-based experiments will measure the amplitude of CMB fluctuations on small angular scales. If the running spectral index model is correct, these experiments will find a relatively low amplitude of fluctuations on small scales.

Thanks a lot David.
Likely, the bulk of the results achieved on cosmological parameters is not so strongly dependent on the global geometrical properties of the Universe, the Concordance Model being supported by a wide set of observations, as presented in Chap. 2. On the other hand, the high precision of the CMB anisotropy measures achieved with WMAP allows us to investigate more deeply the geometry of the Universe. This is crucial to test the validity of the FLRW models and the plausibility of alternatives or generalizations of them. Peter Coles will discuss these aspects in the next interview.

3.5.4 Geometry of the Universe

Dear Peter *(Coles)*, **the recent WMAP data seem to indicate large scale anomalies that could support modifications of our simple view of the geometry of the Universe. Can you describe the most significant theoretical models (Bianchi**

models, coherent cosmic magnetic field, etc.) alternative to the standard FLRW model accounting for different topology and/or geometry of the Universe?

Recent observational advances have led to the establishment of a standard Concordance Model of cosmology, which seems to fit virtually all available experimental data. This model is based on two key assumptions: that the background Universe is described by one of the FLRW models, and that the primordial fluctuations superimposed on it comprised a statistically homogeneous Gaussian random field, possibly generated during inflation. A key goal for modern research is to exploit the success of this scenario to refine estimates of the parameters of the standard model using relatively straightforward statistical tools, such as the power spectrum. At the same time, however, it is also essential to probe for possible evidence of physics beyond the standard theory. Some aspects of this alternative approach to cosmology are described here. Since the release of the first (preliminary) WMAP data set, it has been subjected to a number of detailed independent analyses that have revealed some surprising features. For instance, Eriksen et al. [89] have pointed out the existence of a North–South asymmetry, suggesting that the CMB revealed by WMAP data is not statistically homogeneous over the celestial sphere. This is consistent with the results of Coles et al. [57] who found evidence for phase correlations in the WMAP data that should not exist if the sky pattern were generated by Gaussian random perturbations on a standard FLRW background. The low-order multipoles of the CMB also display some peculiarities, such as unexpectedly low variance and curious alignments [67, 83, 158, 187]. Vielva et al. [260] found significant non-Gaussian behavior in the form of a localized cold spot, using a wavelet analysis of the same data. Evidence for these various quirks and anomalies has persisted through further data releases from WMAP. On the other hand, various other analyses of the statistical properties of the WMAP have yielded results consistent with the standard form of fluctuation statistics [58, 148]. The unusual properties found in some studies may in principle be generated by residual foreground contamination [12, 68, 182] or other systematic effects, but may also provide the first hints of physics beyond the standard cosmological model.

A key element of the standard cosmological model is the Cosmological Principle, that is, that the Universe on large scales is homogeneous and isotropic. This principle is expressed mathematically by the Robertson–Walker metric,

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - a(t)^2\left[\frac{\mathrm{d}r^2}{1-Kr^2} + \mathrm{d}\Omega^2\right], \tag{3.15}$$

which imposes global symmetries that enable a special family of solutions to the Einstein equations to be obtained. The metric here is given in polar coordinates with the angular components included in the solid angle element $\mathrm{d}\Omega$. The Cosmological Principle imposes a preferred time coordinate (cosmological proper time, t) enabling the slicing of four-dimensional space–time into hypersurfaces on which the density is uniform. The geometry of these spatial sections is controlled by the constant K: if $K > 0$ they have positive curvature, like a sphere; if $K < 0$ they have negative curvature, like a hyperboloid; and if $K = 0$ the spatial surfaces are flat (Euclidean). The solutions derived from this geometrical assumption are called the Friedmann–Lemaître–Robertson–Walker (FLRW) solutions and they provide the analytical backbone of the Big Bang model.

The FLRW models describe featureless Universes, and so some mechanism must be added that can generate the galaxies and clusters that characterize the observed large-scale distribution of matter. The structure that we observe today in the Universe develops by gravitational instability, "seeded" by primordial density fluctuations imprinted on the homogeneous background by quantum effects in the early Universe. These are generally expected to be of the form of a statistically homogeneous and isotropic Gaussian random field, which means that while the density varies from place to place, their statistical properties do not. These density variations are treated as small perturbations of a FLRW Universe, at least until they become large comparable to the mean density at which point other methods must be used.

This general approach has been immensely successful, but it is part of the scientific method that hypotheses should be questioned, and particular so in the case of cosmology when there are suggestions of departures from global symmetry. There is of course a possibility that the primordial density fluctuations may not be statistically isotropic or Gaussian. Indeed, some level of departure from Gaussianity is expected in inflationary models. However, a more radical idea is to imagine that the background Universe may deviate from the standard assumption.

The first possibility worth examining is that the background Universe might not be of FLRW form. Following this line of thought poses some fundamental problems because there are few exact solutions for cosmologies in which the Cosmological Principle does not apply. However, there is a well-defined class of models that are homogeneous but not isotropic. Recall that in the Friedmann models, the spatial hypersurfaces are those on which the matter density is constant and these are surfaces of constant time. We can give a more general definition of homogeneity by requiring that all observers moving with the cosmological expansion see essentially the same version of cosmic history. This means mathematically that there must be a symmetry that relates what is seen in a coordinate system centered on Observer A and that seen in a similar system centered on Observer B.

The possible symmetries can be classified into the *Bianchi types*, which are constructed as follows. On a homogeneous spatial hypersurface, it is possible to define at least three independent vector fields ξ_α that satisfy the constraint

$$\xi_{i;j} + \xi_{j;i} = 0 \tag{3.16}$$

(the indices i and j run from 0 to 3 and the semicolons denote covariant derivatives). This is called Killing's equation and the vectors that satisfy it are called Killing vectors. From these it is possible to construct structure constants

$$C^{\delta}_{\alpha\beta} = \xi_\alpha \xi_\beta - \xi_\beta \xi_\alpha. \tag{3.17}$$

The properties of these structure constants relate to the kind of group that describes the required symmetries. The Bianchi types correspond to the size and subgroup structure of this group. In general, if there are n Killing vectors, it will be described by an n-dimensional group. The bigger the group, the more "free parameters" are needed to describe it and, in some sense, the more general is the form of asymmetry described by the relevant Bianchi type.

The simplest such model is Bianchi I, illustrated by the Kasner solution [139]

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - \sum_{i=1}^{3} X_i^2\,\mathrm{d}x_i^2, \tag{3.18}$$

with spatial coordinates X_i, has a different expansion rate in each spatial direction but is otherwise like the FLRW solution. More complicated Bianchi models can involve global rotation, and/or shear as well as curvature, and thus require more parameters to describe them than the Friedmann models. The Friedmann models form special (degenerate) cases of the Bianchi types: the flat Friedman model is a special case of either Bianchi I or Bianchi VII_0. The open Friedman model belongs to type V or VII_h and the closed version belongs to Bianchi type IX.

The Bianchi types have long been regarded by most cosmologists as mathematical oddities rather than realistic models. However, with the emergence of apparently significant departures from global symmetry in the WMAP temperature pattern, interest in them has to some extent been revived. A FLRW Universe produces a completely smooth last scattering surface and hence a totally featureless CMB sky. Any hot or cold spots must be generated by random perturbations. However, a localized cold spot like that seen in WMAP can be produced naturally in Bianchi Universes without having to add random perturbations at all. For example, in Bianchi I, each electron on the last scattering surface sees an anisotropic radiation bath. This means that there will be global asymmetry in the microwave background temperature pattern, which in the simplest models is of pure quadrupole form. Of more interest from the point of view of the observed properties of the CMB is the Bianchi VII_h model, which possesses both rotation and negative spatial curvature. What happens in this model is that the initial anisotropy is twisted by the rotating space-time into a spiral pattern. The effect of spatial curvature is then to focus this pattern into part of the celestial sphere [17], producing the localized cold spot.

Detailed modeling of the CMB sky (e.g., [132]) shows that a good fit to the observed pattern can be achieved in such a Bianchi model, but there is no fundamental theory that motivates this particular explanation. Moreover, there is a relatively simple test that already appears to exclude it. Global anisotropy produces not only temperature fluctuations, but also variations in the polarization of the CMB sky. In Bianchi VII_h, the presence of rotation will generate curl-type B-mode polarization at a level that is already above the upper limits imposed by WMAP [200] (Fig. 3.13).

Another way in which global anisotropy could influence the CMB pattern is through the possible existence of large-scale magnetic fields, which will impose a preferred direction. If these fields are uniform on the scale of the observed Universe, they can be described by Bianchi models. Smaller-scale fluctuations in the CMB could also be generated by tangled magnetic fields present at the recombination epoch [46, 183].

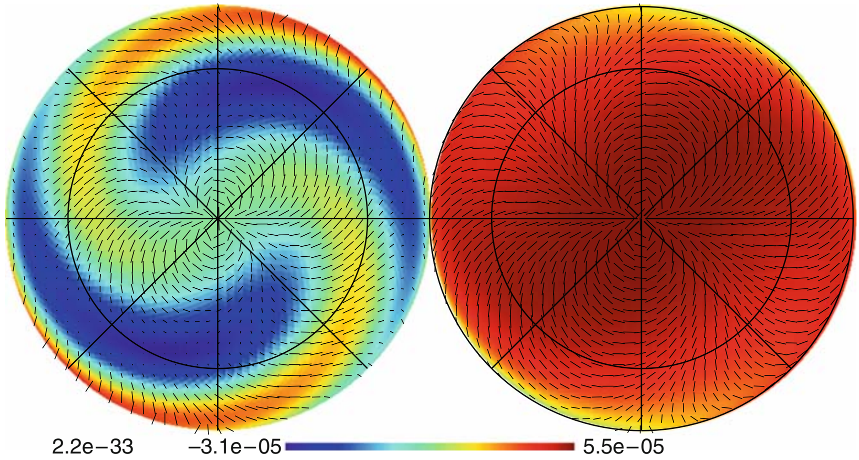

Fig. 3.13 Calculation of the CMB temperature and polarization pattern in an (open) Bianchi VII_h model. In the absence of curvature, the spiral pattern would extend over the entire celestial sphere, but in this example it is focused into a much smaller angle

Even if we accept the applicability of the restricted geometry implied by the Robertson–Walker metric in (3.15), there remains another way to construct a non-standard cosmology. Einstein's equations are local, in the sense that they relate the space–time geometry at a point to the energy–momentum tensor at the same four-dimensional position. What the equations do not specify is how the space–time is constructed globally. In other words, the equations specify geometry but not topology. It is possible to construct spaces with the same local geometry but different global structure. In standard cosmologies, each choice of curvature K is taken to imply the simplest topology consistent with it. For example, a flat Universe is generally assumed to be spatially infinite. However, it is mathematically possible to conceive of flat spaces that are finite, such as the space obtained by identifying opposite faces of a cube either directly, which gives the equivalent of a torus, or with a twist, which gives a Klein bottle. Alternatively, the space might be infinite in one dimension but not in others. There is also no particular reason why the basic unit need be cubic: any fundamental domain that tessellates the space, such as a hexagonal prism, would do as well. A Universe with negative spatial curvature need not be infinite either, although the presence of spatial curvature imposes restrictions on the form of topology that is allowed. A particularly interesting example is a closed Universe in which the fundamental domains are dodecahedral (e.g., [170]).

If the scale of a topological modification is larger than that of the observable Universe, then there is no possibility to detect it. If we live in a world with a very small topological scale, we would see large numbers of copies of local structures in deep observations, which we do not. The most plausible form of nonstandard cosmology is therefore one in which the topological scale is similar to that of the horizon. This leads us more-or-less directly to the CMB pattern as the most viable probe of topological oddities.

There are basically three ways in which the CMB may reveal the presence of a topological scale. First, if the Universe is very small, there is no space in it for long-wavelength fluctuations. This leads to a suppression of large-angle anisotropy. This is interesting because there is indeed marginal evidence that the quadrupole and other low-order modes of the CMB are suppressed relative to the concordance spectrum. Second, the presence of special directions (in which topological identifications are made) leads to a breakdown of statistical homogeneity. This can be tested using a number of methods, including phase correlations [69], but there is yet no compelling evidence for their presence in actual data. Finally, if the last scattering surface is larger than the fundamental topological domain, then it contains copies of itself along circles corresponding to its intersection with these domains.

Interesting though these alternative models undoubtedly are, the evidence for any of them is at best questionable. However, it is definitely important to continue questioning the fundamental assumptions of the standard cosmological model. It is important for cosmologists to continue being radical in generating new ideas, but also to be conservative in their treatment of evidence.

Thank you Peter.
As discussed in Sect. 3.5.2.1, the CMB anisotropy APS depends significantly on the physical processes that occurred during recombination era. The renaissance of attention to this cosmic epoch in the recent literature is remarkable, driven by the necessity to precisely characterize the CMB anisotropy APS for an accurate comparison with data from WMAP and, even more, with those expected in the near future by Planck. Also, as pointed out in Sect. 3.5.3, recent data indicates a nontrivial Thomson optical depth. Although not so "intriguing" as that claimed on the basis of the first release of WMAP data, it calls for a deep understanding of (at least) the astrophysical processes, which occurred during the dark and dawn ages, when first structures and objects began to form and the Universe reionized. We will discuss these themes with Pavel Naselsky in the next interview.

3.6 The Ionization History

Dear Pavel *(Naselsky)*, **the recombination epoch marks the transition between the ages of the Universe opaque and transparent to radiation, following the transition from the radiation dominated era to the matter dominated era. Can you briefly summarize the cosmological events that characterize this epoch of the Universe? Which are the consequences of the finite thickness of the last scattering surface for CMB temperature fluctuations and spectrum? Do exist alternative mechanisms that can bring to the same observational evidences?**

The ionization history of the cosmic plasma is one of the most important part of the Big Bang model. Based on well known atomic physics principles, this theory provide remarkable information about the most fundamental properties of the matter in the Universe through the anisotropy and polarization of the CMB. Here I discuss the theoretical basis and observational consequences of the standard model of recombination and different modification of this model, the possible explanation of the re-ionization of the cosmic plasma at redshifts $z \sim 10 - 14$ by the first quasars, and restrictions on the long-lived particles, which can leave some imprints on the power spectrum of the CMB anisotropy and polarization through distortion of the history of ionization.

3.6.1 Recombination

The Inevitability of Hydrogen Recombination

The modern cosmology and the physics of CMB are tightly related to the physics of interactions between quanta and electrons. The electrons are the lightest charged particles in the Universe that annihilate with positrons in the process of cooling of the cosmic plasma during the cosmological expansion at the time when the temperature of the plasma was falling down to 10^9 K. Remarkably, that tiny fraction of electrons survive during the process of annihilation and after that they play a crucial role in the ionization history of the Universe. It is a good question why the number density of electrons is not exactly equivalent to the number density of the positrons before the epoch of e^+e^- annihilation. This fundamental question is closely related to the question, why the Universe is preferably baryonic (protons, neutrons, etc., but not antiprotons, antineutrons!) without more or less reasonable big amount of antimatter? Shortly speaking, asymmetry of number densities of electrons and positrons just reflect asymmetry of the baryonic charge and fundamental observational fact that the matter in the Universe is neutral in respect to electrical charge.

After the epoch of e^+e^- annihilation, the remaining electrons, by now nonrelativistic, constitute the most important factor for the kinetic of the CMB photons due to the Compton scattering that becomes the dominant mechanism of formation of the CMB anisotropy and polarization.

In quantum physics, the interaction between different particles is characterized by the cross-section, which for Compton scattering is the Thomson cross-section $\sigma_{\rm T} = 6.65 \times 10^{-25}\,{\rm cm}^2$. However, to describe the kinetic of the photon–electron plasma in the Universe, we need to use an additional parameter, the optical depth $\tau = \int \sigma_{\rm T} n_{\rm e} c\, {\rm d}t$ changes in time because of two most important factors: the expansion of the Universe and the dynamics of evolution of electron concentration. Here $n_{\rm e}$ is the number density of electrons, c is the speed of light. As a rule, this process is described in terms of the fraction of ionization of the plasma, concentration of baryons in the plasma. If the plasma temperature definitely exceeds 10^5 K, electrons must be free (not bound to protons) as otherwise a gigantic number of ionizing quanta would immediately destroy hydrogen atoms. In other words, the efficiency of the reaction $H + \gamma \rightarrow p + e$ is so high that beyond any doubt the amount of neutral hydrogen in cosmological matter is quite inconsequential. In this case, therefore, the fraction of ionization $x_{\rm e}$ equals 1 with high accuracy and changes in the optical depth of the plasma relative to Thomson scattering are caused only by the expansion of the Universe. Note that even when the fraction of plasma ionization does not change with time ($x_e = 1$ =const), the anticipated plasma depth continues to diminish anyway as a result of the cosmological expansion [120]

$$\tau = \int_t^{t_{now}} \sigma_T x_e n_b c \, dt \simeq 4.1 \times 10^{-2} \frac{\Omega_b}{\Omega_m} h \left\{ \left[\Omega_\Lambda + \Omega_m (1+z)^3 \right]^{1/2} - 1 \right\}, \quad (3.19)$$

where $\Omega_\Lambda + \Omega_m = 1$. Equation (3.19) clearly shows that if $z > z_{cr}$, where z_{cr} is found from the condition $\Omega_\Lambda \simeq \Omega_m (1 + z_{cr})^3$, the behavior of the optical depth follows the relation $\tau \propto (1+z)^{3/2}$ while for $z \to 0$ we have $\tau(z) \propto \frac{3}{2}\Omega_m \cdot z \to 0$.

In fact, the first important conclusion that follows from this analysis of the extremal asymptote $x_e = 1$, regardless of the value of the redshift, is that relative to the Thomson scattering today's Universe must be optically thin – to within $\leq 1\%$. As we see from (3.19), the behavior of the optical depth for $z \gg 1$ is independent of Ω_Λ,

$$\tau(z) \approx 4.1 \times 10^{-2} \Omega_b \Omega_m^{-1/2} h (1+z)^{3/2} \quad (3.20)$$

and formally the zone of the "last scattering" of quanta by electrons ($\tau(z) = 1$) corresponds to the redshift $z_* \simeq 8.4 \Omega_b^{-2/3} \Omega_m^{1/3} h^{-2/3}$. To be specific, we assume $\Omega_b h^2 \simeq 0.02$, $\Omega_m \simeq 0.3$, and $h \simeq 0.7$ and finally obtain $z \simeq 60$. Therefore, with cosmic plasma completely ionized, the maximum redshift after which the cosmic background radiation propagates freely is a relatively low $z \simeq 60$. There is still a question of whether the total hydrogen ionization can be self-maintained down to this redshift. This question can be answered using the following qualitative reasoning.

To maintain the fraction of ionization at the level $x_e = 1$ it is necessary for the fraction of quanta having energy above the hydrogen ionization potential $I \simeq 13.6$ eV to reach approximately one quantum per baryon. As the Universe is "hot", the number density of the CMB photons is about 10 orders of magnitude greater than the number density of the protons ($\xi = \frac{n_{cmb}}{n_{bar}} \simeq 5 \times 10^{10}$). Thus, the assumption $x_e = 1$ needs constant temperature of the plasma in order to have $T(z) \sim T_i \ln^{-1}(\xi^{-1}) \approx 3.8 \times 10^3 \, K$, where $T_i = \frac{I}{k} \simeq 1.5 \times 10^5$ K, is the temperature corresponding to the potential of ionization. In fact, the temperature of the CMB evolve as $T(z) = T_0(1+z)$, where $T_0 = 2.736$, as it follows from the COBE data. Thus, we can see that the ionizing (Wien's) part of the CMB spectrum cannot sustain the fraction of ionization at the level $x_e \simeq 1$ at redshifts $z < 1{,}400$. It is therefore inevitable that the cosmological hydrogen must undergo recombination.

The estimates given above yield an obvious conclusion that the ionization history of the cosmic plasma is one of the most important probes for studying the properties of cosmic matter in the epoch of redshift $z \leq 1{,}400$. Any information on fraction of ionization in this period is inevitably tied to testing the processes of energy release, and therefore, to an identification of the possible sources of this energy release. In fact, the situation becomes even more dramatic if we take into account the observation of the hydrogen line $\lambda = 21$ cm and the Lyα absorption in the spectra of remote quasars: these show that the cosmological hydrogen must be ionized up to $x_e \simeq 1$ already at redshifts $z \sim 5 \div 6$ (see next subsection). Therefore, the very idea of nonequilibrium sources of energy release is directly confirmed, but for small redshifts only. What is the situation in the range of redshifts $60 < z < 1{,}400$? What can we say about the presence or absence of sources of nonequilibrium ionization (nonequilibrium relative to the "primordial" CMB)? Answering to that question lets briefly discuss the standard model of recombination.

Standard Model of Hydrogen Recombination

Here we turn our attention to the standard model of hydrogen recombination. Its fundamentals were formulated at the end of the 1960s in pioneer articles [189, 275]. It is necessary to point out that at that time the role played by DM in the kinematics and dynamics of the evolution of the Universe was underestimated, unlike the present day. Thus, all the results of the hydrogen recombination theory in the baryonic Universe needed certain corrections that would take into account the simple fact that the density of the DM exceeds that of baryonic matter. Consequently, the rate of expansion of the Universe should follow the law $a \propto t^{2/3}$ starting from the redshifts $z_{eq} \simeq 1.2 \times 10^4 \Omega_m h^2$ when its density becomes equal to that of the CMB. At the same time, the redshift z_{eq} at low density of baryonic matter $\Omega_b h^2 \simeq 0.02$ would correspond to $z_{eq} \sim 240 \left(\frac{\Omega_b h^2}{0.02}\right)$ (with the DM background neglected) and hydrogen recombination would be completed already at the radiation-dominated phase. The disbalance of hydrogen recombination reaction rates (affected by Ω_b) and the rate of cosmological expansion (dictated by Ω_m) is the principal distinctive feature of the "standard" models when the DM is taken into account. This factor was pointed to first by [272] (see also [133, 152, 171]). Simultaneously, with including the DM factor, [19, 152, 210] improved the model of transfer of resonance quanta in the expanding Universe, and [161] calculated the effect of the ionization regime on the molecular synthesis at later stages of hydrogen recombination ($z \ll 400$). A new element due to the unique accuracy of the future CMB experiments was that the effect of He^4 on the kinetics of the cosmological hydrogen recombination and on its residual ionization was to be taken into account. Seager et al. [217] analyzed a multilevel model of the hydrogen atom (~300 levels) and gave a systematic summary of the main achievements of the theory (see also [248] and references therein). The effort was completed with creating a specialized programs package RECFAST, which at the moment is the most successful tool for calculating the dynamics of cosmological hydrogen recombination. This package allow us to estimate the dependence of ionization fraction on the most fundamental cosmological parameters (Ω_b, Ω_m, h etc.) and to analyze the anisotropy and polarization of the CMB with unprecedented accuracy.

The Three-Level Approximation for the Hydrogen Atom

Lets consider the three-level model of the hydrogen atom comprising the ground state HI and the $2P$ and $2S$ states as the initial approximation for describing its

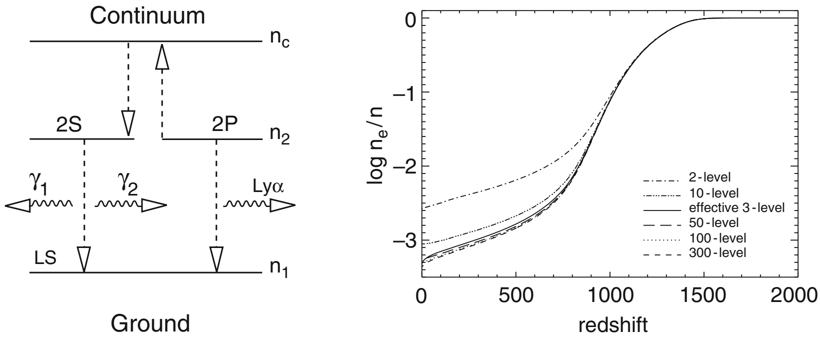

Fig. 3.14 (*Left*) The three-level model of hydrogen atom. (*Right*) The comparison of ionization fractions for multilevel models of recombination (see the end of this section). Adapted from [217]

recombination kinetics. We will neglect the contribution of He^4 atoms to this process, suggesting to take it into account as small correction of the standard model. The model is schematically presented in Fig. 3.14 and indicates the possible directions of electron transitions in the hydrogen atom.

Lets take into account the following specifics of the level structures of the hydrogen atom and also the properties of the cosmological plasma at the redshifts $z \sim 10^3$. First of all, the temperature of the plasma is sufficiently low ($<10^4$ K) for the collisional ionization to be negligible in comparison with radiative processes. Second, the time of transition of an electron in hydrogen atom from the $2P$ level to the ground state is much shorter than the cosmological time. As for the BB CMB, the fraction of the "soft photons" (with energy below the potential of ionization I) is significantly high in comparison with the "hard" photons, the population of higher levels obeys thermal equilibrium distribution. Since the pioneering papers by [189, 275], two additional remarks needed to be taken into account: the population of the level $2S$ satisfies the condition $n_{2S} \ll n_1$, where n_1 is the hydrogen population at the ground level, and each act of recombination to the $2P$ level creates a resonant Lyα quantum, and each decay of the $2S$ level produces two low-energy photons. This means that the reaction $e + p \leftrightarrow H + \gamma$ occurs in both directions, in such a way that the equilibrium concentrations of electrons, protons, and neutral hydrogen atoms obey the Saha formula

$$\frac{n_e n_p}{n_H} = \frac{g_e g_p}{g_H} \frac{(2\pi m_e kT)^{3/2}}{h^3} e^{-I/kT}, \qquad (3.21)$$

where g_i are the statistical weights of each component. We now introduce the fraction of ionization $x_0 \equiv \frac{n_e}{n_p + n_H}$. Then from (3.21) we gets

$$\frac{x_0^2}{1 - x_0} \simeq 4.4 \times 10^{22} \left(\frac{\Omega_b h^2}{0.02}\right)^{-2} (1+z)^{-3/2} \exp\left[-\frac{5.77 \times 10^4}{1+z}\right]. \qquad (3.22)$$

Note, however, that the Saha formula (3.22) does not provide the entire detailed picture of the formation of neutral hydrogen atoms. The point is that each act of recombination $p + e \rightarrow H + \gamma$ is accompanied by the emission of one Lyα quantum with energy $h\nu_\alpha = \frac{3}{4}I = 10.2\,\text{eV}$, which immediately excites a hydrogen atom, while quanta with $h\nu_\alpha = \frac{1}{4}I = 3.4\,\text{eV}$ that are abundant in the BB radiation immediately ionize this atom. The corresponding optical depth $\tau_\alpha \simeq \sigma_I n_\text{b} ct$ is found to be $\geq 4 \times 10^8$ [189] and therefore, Lyα quanta generated in each recombination event should be immediately absorbed by the generated hydrogen atoms.

Relatively low-energy quanta are needed to ionize hydrogen atoms from the $2P$ level: $E = I - h\nu_\alpha = \frac{1}{4}I \simeq 3.4\,\text{eV}$. Owing to the Wien character of the spectrum of the CMB photons, the number of such "soft" quanta is found to exceed that of "hard" quanta (with energy $E \simeq I$) by a factor of approximately $\text{e}^{-I/4kT}/\text{e}^{-I/kT} \simeq \text{e}^{\frac{3}{4}\frac{I}{kT}} \gg 1$. Therefore, hydrogen atoms in the $2P$ state are immediately ionized by quanta from the "soft" part of the CMB spectrum. The right-hand side branch in Fig. 3.14 that describes the dynamics of population of the hydrogen atom $2P$ state is at the same time a sort of "engine" for producing and accumulating Lyα quanta in the process of hydrogen recombination over their equilibrium concentration in the Wien range of the spectrum. The reaction channel on the left in Fig. 3.14 is directly responsible for the formation of neutral hydrogen – via the metastable $2S$ level, as was first shown in [189, 275].

Let us consider the kinetics of this process in more detail. Introducing now the fraction of ionization $x_\text{e} = n_\text{e}/(n_\text{p}+n_{1\text{S}})$, [189,275] came to the following equation:

$$-\frac{\text{d}x_\text{e}}{\text{d}t} = D\left[\alpha_\text{c} n x_\text{e}^2 - \beta_\text{c}(1 - x_\text{e})\,\text{e}^{-\frac{B_1 - B_2}{kT}}\right], \tag{3.23}$$

where

$$D = \frac{1 + K\Lambda_{2\text{S},1\text{S}} n(1 - x_\text{e})}{1 + K(\Lambda_{2\text{S},1\text{S}} + \beta_\text{c}) n(1 - x_\text{e})}, \qquad \beta_\text{c} = \alpha_\text{c}\frac{2\pi m_\text{e} kT}{h^3}\,\text{e}^{-\frac{B_2}{kT}}, \tag{3.24}$$

and $K = \frac{\lambda_\alpha^3}{6\pi H(t)}$, $B_1 - B_2 = \frac{I}{4}$, $B_2 = 3I/4$, and $\Lambda_{2\text{S},1\text{S}} = 8.227\,\text{s}^{-1}$ is the decay rate from $2S$ state to $1S$ for the hydrogen atom. For the recombination coefficient, we use an improved value [123, 197, 259].

$$\alpha_\text{c} = 10^{-13}\frac{at^b}{1 + ct^d}\,\text{cm}^3\text{s}^{-1}, \tag{3.25}$$

where $a = 4.309$, $b = -0.6166$, $c = 0.6703$, $d = 0.5300$, and $t = T_\text{M}/10^4\,\text{K}$. Here T_M is the plasma temperature, which is assumed to be equal to the CMB temperature in the three-level recombination model chosen here.

However, one can ask how accurate is the standard model of recombination? One of the main reasons for that question hides in the need to develop a modern theory of formation of the CMB anisotropy. As we already know, the fluctuations $\Delta T/T$ on a scale of several minutes of degree are formed during the epoch of recombination at $z \sim 1000$.

Another important reason follows from the high accuracy of measurements of the CMB anisotropy and polarization in the WMAP experiment and in the *Planck* mission planned for 2009. If we chose the relative error of determining ΔT as $\leq 3 \div 10\%$ – as a very conservative estimate of determining the characteristics of CMB anisotropy – we need to be absolutely sure that these 3–10% are not "accumulated" only as a result of inaccuracies of theoretical predictions of the dynamics of thinning of cosmic plasma for primordial radiation. Therefore, a detailed theory of recombination must be capable of predicting the behavior of $x_e(z)$ with an error $\leq 10\%$, and possibly even better.

Finally, there is the third cause, mostly of predictive nature. Hydrogen recombination results in considerable distortions of the background radiation spectrum in the Lyα frequency range at the expansion stage $z \sim 10^3$. As a result of redshifting, these distortions must be represented today in the long-wavelength range $\lambda \simeq cz_{\rm rec}/\nu_\alpha \sim 11.3 \times 10^{-2}$ cm, that is, in the near infrared range of cosmic radiation spectrum. An experimental detection of a specific electromagnetic "echo" of the recombination epoch would be fantastically important for testing the properties of the cosmic plasma at redshifts $z \sim 10^3$. Unfortunately, the high level of infrared background in this range does not offer us any hope of rapid experimental solution of the problem. However, a detailed prediction of the shape of these distortions would be of extreme interest for understanding the mechanisms of formation of radiation in the near infrared band.

It must be specially emphasized that the ionization history of the cosmic plasma at the stage $z \ll 10^3$ is extremely important for understanding the processes of formation of primary molecules of cosmological origin – those that could act as an efficient "coolant" in dense clouds, facilitating the formation of first-generation stars. We must again emphasize that the development of a detailed theory of hydrogen recombination does not mean in any way a diversion from the simplified three-level recombination model whose fundamentals were created more than 30 years ago. Moreover, this simplified model reproduces qualitatively and very often quantitatively the main physical principles and processes that resulted in the transformation of ionized hydrogen to neutral state. Assuming this model as a basis, I could now list the main features of the detailed theory of cosmological recombination [218].

- *Dynamics of excited states of hydrogen and helium.* The main difference lies in the increase in the number of hydrogen levels to $N = 300$ and in adding higher levels of He^4 to the analysis (see Fig. 3.14 the right plot).
- *Radiation kinetics.* As we saw in the three-level model of the hydrogen atom, the kinetics of cosmological hydrogen recombination is very sensitive to the behavior of hydrogen resonance lines. Resonance quanta take part in absorption, scattering and emission by hydrogen and helium atoms, undergoing redshift because of the expansion of the Universe for $z < 10^3$.
- *Thermal history of matter.* Details of the thermal history of matter are extremely important when analyzing the asymptotic of recombination ($z \ll 10^3$), when the Compton scattering of the CMB quanta fails to sustain the thermal contact $T_M = T_R$. When constructing the thermal history of cooling of the matter we

need to take into account, along with the Compton process, the free-free transitions and the cooling via photorecombination radiation in the plasma. A detailed balance of these processes in a hydrogen–helium plasma was described in [218].

What we learn about the Universe from the epoch of recombination?

We will now consider various modes of recombination of hydrogen and helium as functions of parameters of the cosmological plasma in the framework of the mathematical formulation of the problem discussed earlier.

On the basis of the available numerical data, [218] proposed a generalized model of the three-level recombination that takes into account the contribution of helium recombination. Using this modification, we consider the dependence of the fraction of ionization $x(z)$ on parameters of the cosmological model:

(a) The function $x_e(\Omega_{DM})$. As our base model, we use the frequently cited ΛCDM cosmological model with the following set of parameters: $\Omega_m = 0.3$, $\Omega_b h^2 = 0.02$, $h = 0.65$, and $1 = \Omega_m + \Omega_b + \Omega_\Lambda$. To analyze the function $x_e(\Omega_m)$, we fix the parameters $\Omega_b h^2$ and h of this model but at the same time vary Ω_m and Ω_Λ, so that the condition Ω_{tot} continues to hold. Using RECFAST, we calculate the function $x_e(z)$, taking into account that the helium mass concentration is independent of the value Ω_m. The results of these calculations are plotted in Fig. 3.15. This plot shows that as the density of DM Ω_m decreases (while $\Omega_m + \Omega_\Lambda \approx 1$), the degree of ionization drops systematically in the whole redshift range $0 \leq z \leq 1,500$. In Fig. 3.15, we illustrate the corresponding ratio for ionization degrees $x_e(\Omega_m)/x_l(\Omega_m = 0.3) = \Gamma(\Omega_m)$. Figure 3.15 demonstrates that differences in ionization degrees during the $z \sim 10^3$ epoch are at the 10–25% level, while the residual fraction of ionization (for $z = 0$) for $\Omega_m = 0.1$ is lower by a factor of approximately 1.5 than in the model with $\Omega_m = 0.3$. Qualitative arguments support this result. Namely, a decrease in Ω_m is accompanied by a rise in Ω_Λ and, as a consequence, by an increase in the age of the Universe. In its turn, the rate of expansion for $z > 1$ mostly depends on Ω_m and is practically independent of Ω_Λ: $t_{exp} \sim a/\dot{a} \sim H_0^{-1}\Omega_m^{-1/2}z^{-3/2}$. It is clear now that in models with lower values of

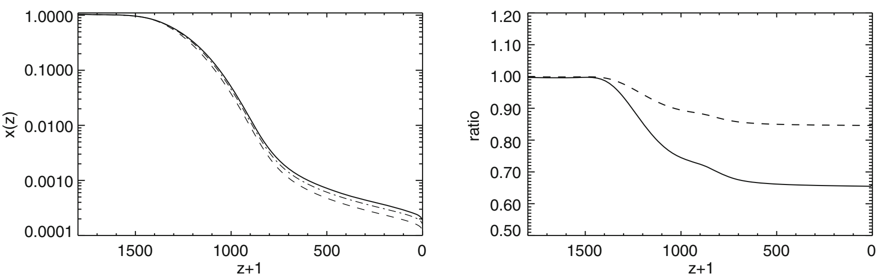

Fig. 3.15 (*Left*) Degree of ionization $x(z)$ as a function of redshift z for a number of values of the parameter Ω_m. *Solid curve* corresponds to $\Omega_m = 0.3$, *dash-dot curve* – $\Omega_m = 0.2$ and *dashed curve* – $\Omega_m = 0.1$. (*Right*) Parameter $\Gamma(\Omega_m)$ (ratio) as a function of redshift in a cosmological model ΛCDM. *Solid curve* corresponds to $\Omega_m = 0.1$, *dashed curve* – $\Omega_m = 0.2$

Ω_m, the age of the Universe for the same z is higher, and therefore, a larger fraction of hydrogen atoms have sufficient time to recombine.

(b) The function $x_e(\Omega_b h^2)$. The quantity $x_e(z)$ is plotted in Fig. 3.16 as a function of density of the baryonic fraction of matter in the base ΛCDM model. The main conclusion is that as $\Omega_b h^2$ increases, the recombination rate α_H grows, resulting in reduced ionization fraction of the plasma. In Fig. 3.16, we plot the ratio of the corresponding ionization fractions $\Gamma(\Omega_b) = x_l(\Omega_b h^2)/x_l(\Omega_b h^2 = 0.02)$ for various values of the parameters $\Omega_b h^2 = 0.01$ and 0.03 in the entire range of variation of z. As we see from these curves, the residual fraction of ionization varies roughly by a factor of 2 in comparison with the $\Omega_b h^2 = 0.02$ model.

(c) The function $x_e(h)$. Figure 3.17 gives the results of numerical calculations of the ionization degree for three values of the Hubble constant: $H_0 = 50$, 65, and 100. The effect of this parameter is not as trivial as that of Ω_m because, on one hand, it dictates the rate of expansion of the Universe, and on the other hand, it determines the recombination rate via the parameter $\Omega_b h^2$. Figure 3.17 plots the ratio of ionization degrees $x(h = 0.5)$ and $x(h = 1)$ to $x(h = 0.65)$ as a function of redshift z. As we see from this figure, the effect of this parameter is comparable to that of Ω_b.

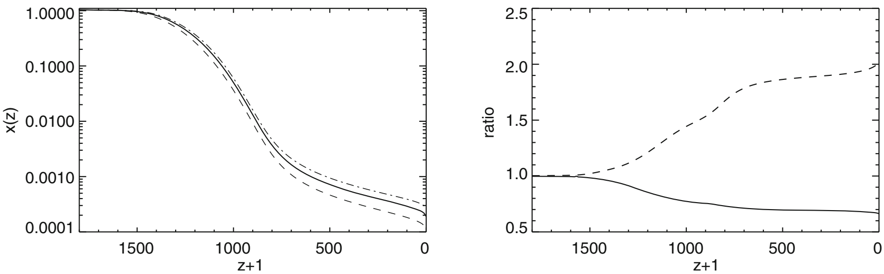

Fig. 3.16 *Left.* Degree of ionization as a function of redshift in the ΛCDM model for various densities of the baryonic fraction. *Solid curve* – $\Omega_b h^2 = 0.02$, *dashed curve* – $\Omega_b h^2 = 0.03$ and *dash-dot curve* – $\Omega_b h^2 = 0.01$, $h = 0.65$. *Right.* Parameter $\Gamma(\Omega_m)$ as a function of redshift. *Solid curve* corresponds to $\Omega_b h^2 = 0.03$, *dashed curve* – $\Omega_b h^2 = 0.01$

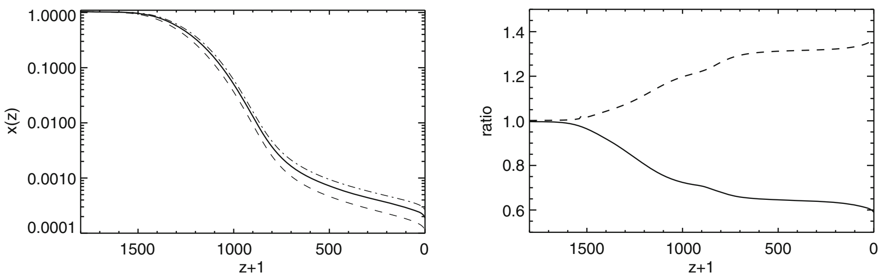

Fig. 3.17 *Left.* Degree of ionization as a function of z for various values of the Hubble constant. *Solid curve* – $h = 0.65$, *dashed curve* – $h = 1$, *dash-dot curve* – $h = 0.5$. The base model is ΛCDM with $\Omega_{tot} = 1$, and Ω_b is identical for all models. *Right.* Ratio of ionization degrees $\Gamma(h)$ as a function of z for $h = 0.5$ (*solid curve*) and $h = 2$ (*dashed curve*)

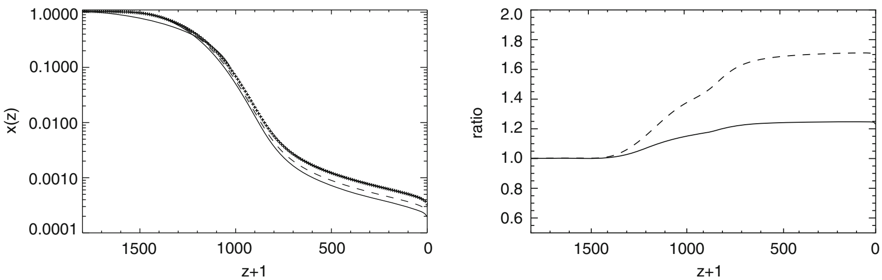

Fig. 3.18 *Left*. Degree of ionization $x(z)$ in an "open" model: $\Omega_\Lambda = 0$, $\Omega_b h^2 = 0.02$, $h = 0.65$. *Solid curve* – $\Omega_m = 0.3$, *dashed curve* – $\Omega_m = 0.5$, *marked curve* $\Omega_m = 1$. *Right*. Ratio of ionization degrees $\Gamma(\Omega_m) = x(\Omega_m)/x(\Omega_m = 0.3)$ in "open" models. *Solid curve* – $\Omega_m = 0.5$; *dashed curve* – $\Omega_m = 1$

(d) The function $x(\Omega_m)$ in "open" models. In this class of models, we drop the condition $\Omega_m + \Omega_b + \Omega_\Lambda = 1$, analyzing differences in ionization modes in the so-called *open* models of the Universe. In all models with $\Omega_{tot} \le 1$, we fix the parameters $\Omega_b h^2 = 0.02$ and $h = 0.65$ and vary the parameter Ω_m from $\Omega_m = 0.3$ up to $\Omega_m = 1$. The results of calculations of ionization degrees and their ratios are given in Fig. 3.18. On the basis of this numerical calculations, we can suggest the following approximation for $x(\Omega_b, \Omega, h)$ [35]:

$$x \sim \Omega_m^{1/2} \Omega_b^{-1} h^{-1}. \tag{3.26}$$

This approximation takes into account both quantitatively and qualitatively all specific features of the function $x(\Omega_m, \Omega_b, h)$ in the framework of the standard model of recombination of cosmological hydrogen.

3.6.2 Reionization

Dear Pavel *(Naselsky)*, **the processes that resulted in the Universe reionization along his history are of great importance for cosmology. Can you discuss the standard scenarios related to primeval galaxy and star formation and alternatives physical mechanisms in relation to the imprints of reionization on the CMB?**

The Inevitability of Hydrogen Reionization

The standard model of hydrogen recombination presented before is based on one very drastic assumption whose legitimacy is not necessarily obvious. We mean the hypothesis that beginning with redshifts $z \sim 3\times10^3$ and ending with $z = 0$, the Universe never contained any other sources of ionization of cosmic plasma in addition to the microwave background radiation. This assumption is certainly wrong for the epoch of $z < 10$, because the very fact of existence of galaxy clusters and especially quasars with high redshifts $z_q \approx 3 \div 6$. These gravitationally bound structures were forming in the neutral gas, constituting potential sources of gas reionization. The study of this epoch of reionization of cosmological hydrogen is a separate important chapter of modern cosmology, which continues to be actively researched now.

The observation of Lyα lines in the spectrum of remote quasars provides an experimental foundation for the conclusion on the inevitability of the hydrogen reionization epoch. Martin Schmidt [215] was the first to conduct the observation of the Lyα line in the spectrum of the 3C9 quasar, which stimulated the famous work of Gunn and Peterson [108]. In this paper, the authors formulated for the first time the conclusion that the fact of observation of Lyα lines in quasars with redshift $z \geq 2$ signifies that at this z hydrogen is practically completely ionized. This conclusion is in agreement with the latest analysis of Lyα absorption in the spectra of the 19 highest redshift Sloan Digital Sky Survey (SDSS) quasars, which reveal a strong evolution of the Gunn–Peterson Lyα opacity at $z \sim 6$ [92, 100]. According to [108], the optical depth of neutral hydrogen is calculated on the basis of Lyα line absorption using the following expression (see also [15]):

$$\tau_{\rm GP} = \frac{\pi e^2 f_\alpha \lambda_\alpha n_{\rm HI}(z_{\rm S})}{m_{\rm e} c H(z_{\rm S})} \approx 4.3 \times 10^5 x_{\rm HI} \left(\frac{\Omega_{\rm b} h^2}{0.02}\right) \times \left(\frac{\Omega_{\rm m}}{0.3}\right)^{-1/2} \left(\frac{1+z_{\rm S}}{10}\right)^{3/2} . \tag{3.27}$$

Here $f_\alpha = 0.4162$ is the oscillator strength for the line $\lambda_\alpha = 1216\mathring{A}$, $H(z_{\rm S})$ is the value of the Hubble parameter for the redshift $z_{\rm S}$ of the source, $x_{\rm HI}(z_{\rm S})$ is the neutral hydrogen concentration for $z = z_{\rm S}$, and $n_{\rm HI}$ is the neutral hydrogen density at the source redshift $z_{\rm S}$.

Treating the recent data of [92, 100] on identifying the lines of SDSS quasars in the spirit of the paper [108] one can evaluate the expected fraction of hydrogen ionization. Assuming in (3.27) that $z_s = 5.8$ and $\tau_{\rm GP} \leq 0.5$, we find that the fraction of neutral hydrogen $x_{\rm HI}$ must be infinitesimally small: $x_{\rm HI} \leq 10^{-6}$. Therefore, we can be absolutely certain that already at $z \sim 6$ the epoch of neutral hydrogen was replaced by the epoch of its complete ionization. We need to mention that regardless of the specific mechanisms that produce the reionization of cosmological hydrogen at such a high redshift as $z \sim 6$, the current data on small-scale anisotropy of CMB show that the optical depth relative to the Thomson scattering is unlikely to exceed $\tau_{\rm T} \simeq 0.1 \div 0.15$. A simple estimate of the maximum redshift at which secondary ionization of hydrogen could take place for this constraint follows from the definition of $\tau_{\rm T}(z)$ in (3.20):

$$z_{\max} \simeq 20 \left(\frac{\tau_{\rm T}}{0.2}\right)^{2/3} \left(\frac{\Omega_{\rm b} h^2}{0.02}\right)^{2/3} \left(\frac{\Omega_{\rm m} h^2}{0.126}\right)^{-1/3} . \tag{3.28}$$

According to the 3 year WMAP data, the electron scattering optical depth is close to $\tau = 0.09 \pm 0.03$. This estimate is consistent with "minimal reionization models", which do not require the presence of very massive ($M > 100\,{\rm M}_\odot$) Pop III stars [49, 105] and restrict the maximum redshift $z_{\max} \simeq 9 \div 14$ for most probable values of the cosmological parameters in (3.28). Thus, the whole history of changes in the hydrogen reionization plays out within a relatively narrow range of redshift $1 \div 2 \leq z \leq z_{\max}$ at which the formation of the very first objects in the Universe occurs. We have to emphasize that the specifics of secondary ionization of hydrogen, including that at the maximum redshifts $z_{\max}$ at which reionization actually begins, strongly depend on reionization DM [15] and models of the structure formation (see Burigana et al. [43] and references therein).

Unfortunately, there is no standard model of reionization that provides clear explanation of the temperature balance and radiation kinetics in multicomponent gas during the structure formation at low redshifts. The reason for that is the complication of the effect of back reaction of ionizing quanta on the process of structure formation, including first quasars and first stars. In the modern literature, this effect is called the "radiative feedback", which is the subject of modeling in the vast collection of papers.

Recently, [216] (but see also [43, 49]) propose two alternative prescriptions for the radiative feedback, based on different assumption on the stars formation in halos. The first one assumes that in photoionized regions halos can form stars under conditions that their circular velocities exceeds the critical value $v_{\rm crit} = 2kT/\mu m_{\rm p}$, where k is the Boltzmann constant, T is the average temperature of ionized regions, μ is the mean molecular weight. The second model is based on the assumption that the average baryonic mass $M_{\rm b}$ within halos in photoionized zones is proportional to the ratio $\Omega_{\rm b}/\Omega_{\rm m}$ [105]. Recent analysis of these models made by [43] shows that only 16% of the volume is reionized at $z = 10$ for the first model, while for the second one we may expect to get about 85% of reionized volume at $z = 10$. Thus, one can see the importance of the CMB data in respect to determination of the history of reionization. If the future *Planck* experiment confirm the WMAP estimate $\tau \simeq 0.1$, it would be significant evidence in favor of the second model. Moreover, [43] have pointed out that using the *Planck* polarization data, it will be possible to observe direct manifestation of the process of reionization through the analysis of the EE power spectrum at the multipole range $l \leq 100$.

3.6.3 Alternative Ionization Histories

Dear Pavel *(Naselsky)*, **could it be other history of recombination?**

The standard model of hydrogen recombination predicts rapidly decreasing concentration of free electrons already at redshifts $z \leq 1,400$. In a realistic model with $\Omega_{\rm tot} = \Omega_{\rm DM} + \Omega_{\rm b} + \Omega_{\Lambda} = 1$, and $\Omega_m = 0.3, = \Omega_b h^2 \simeq 0.02$, $h = 0.7$, the fraction of ionization is found to be close to $x_{\rm e} \simeq 0.1$ already at $z \simeq 10^3$ and reduces to $x_{\rm e} \simeq 10^{-2}$ for $z \simeq 800$. The optical depth of the plasma with respect to the Thomson scattering then becomes quite low ($\tau \ll 1$) and at $z < 800$ the CMB quanta propagate freely, not being scattered on free electrons. This picture, standard for each cosmological model operating with its own set of parameters $\Omega_{\rm tot}$,

$\Omega_{\rm DM}$, $\Omega_{\rm b}$, Ω_{Λ}, and h, is based on the assumption that it is precisely in the epoch of redshift $z \leq 1,400$ that the cosmic plasma contains no sources of nonequilibrium ionization of hydrogen that would supply plasma with additional ionizing quanta not connected with the kinetics of the Lyα part of the CMB spectrum.

It is clear that if the energy density of the nonequilibrium Lyα exceeds the CMB, then the kinetics of hydrogen recombination should evolve according to a scenario that differs in principle from the standard model. Therefore, the characteristics of CMB anisotropy, shaped during the period of cosmological recombination, should differ from those of fluctuations that are formed in the "standard" model of recombination.

A reservation is necessary here: the epoch with redshifts $z \sim 10^3$ is definitely "peculiar" for any models explaining the origin of structures in the Universe. The formation of gravitationally bound structures with masses $M \sim M_{\rm G} \sim 10^{12}\,{\rm M}_{\odot}$ proceeds in the framework of CDM models mostly at relatively low redshifts $z \leq 2 \div 3$. The low-mass part of the spectrum ($M \sim 10^5 \div 10^6\,{\rm M}_{\odot}$) is responsible for the formation of objects at $z \leq 25 \div 30$ close to the Jeans mass $M_{J(b)}$ in the baryonic fraction of matter precisely at the moment when the plasma becomes transparent for the CMB radiation at ($z \sim 10^3$). This means that formally the spectrum of density fluctuations in the DM gas contains perturbations with $M \ll M_{J(b)}$ that could reach a nonlinear mode at $z \sim 10^3$. However, the emerging low-mass nonlinear structures would not in fact involve the baryonic matter. We can add that at $z \sim 10^3$ the age of the Universe was only

$$t_{\rm rec} \simeq \frac{2}{3H_0 z_{\rm rec}^{3/2}\sqrt{\Omega_{\rm m}}} \approx 10^6 \left(\frac{\Omega_{\rm DM} h^2}{0.15}\right)^{-1/2} \quad \text{years},$$

which is insufficient for transforming the rest energy of baryons into ionizing photons even if the primary stars were supermassive ($M \geq M_{j(b)}$) [250].

However, in addition to stellar energy sources, there exists another mechanism that transforms the rest energy of matter into radiation. We speak here of the electromagnetic decay of massive particles $X \rightarrow X' + \gamma$ or $X \rightarrow X' + e^+ + e$, in which the initial particle X is transformed into a new particle X' and a γ quantum or an electron–positron pair is emitted. Moreover, it is not at all necessary for the electromagnetic channel to dominate the X-particle decays. It would be sufficient for the supermassive ($m \gg 10^3$) X particle to generate a quark–antiquark jet ($X \rightarrow q + \overline{q}$) and then the annihilation of quarks would result in a rapid ionization of decay products and in the formation of the electromagnetic component. Note that this mechanism is considered nowadays as one of the sources of generation of ultra high-energy cosmic rays ($E \geq 10^{20}$ eV) in the so-called Top-Down scenario (for details see [30]). This scenario practically coincides with the model of evaporation of primordial Black Holes (BHs); the possibility of primordial BH formation in the early Universe was first considered by [116, 274]. These objects are quite special in that their formation requires only a relatively high – in comparison with galactic scales – amplitude of adiabatic inhomogeneity $\frac{\delta\rho}{\rho} \simeq (3 \div 10) \times 10^{-2}$ at the moment $t \simeq \frac{2GM_{\rm BH}}{c^3}$ (here $M_{\rm BH}$ is the mass of matter collapsing onto a BH on the scale of the cosmological horizon).

An important feature of this potential remnant of the very early Universe is the effect of quantum decay of primordial BHs into particles [116]. The characteristic energy of particles created in this decay is related to the mass of primordial BH as $E_{\rm BH} \simeq \frac{hc}{\lambda} = \frac{hc}{r_g({\rm M})} \propto \frac{hc^3}{GM_{\rm BH}}$, and the life-time is $\tau_{\rm BH} \simeq t_{\rm u}\left(M_{\rm BH}/10^{14.5}{\rm g}\right)^3$, where $t_{\rm u}$ is here the present age of the Universe (and g means gram). A comparison of $\tau_{\rm BH}$ and the characteristic time of the onset of the hydrogen recombination epoch $t_{\rm rec} \simeq 10^6(\Omega_{\rm m}h^2)^{-1/2}$ shows that $\tau_{\rm B} \simeq t_{\rm rec}$ for BHs with mass $M_{\rm BH} \simeq 10^{14.5} \times z_{\rm rec}^{-1/2} \simeq 10^{13}$g. The characteristic energy of electron–positron pairs, γ-quanta and neutrinos equals then $\overline{E}_{\rm BH} \simeq 1.5\,$GeV, which is close to the proton rest energy.

As a last step, following [159], I would like to mention another potential channel for distorting of hydrogen recombination kinetics, not related directly to injection of additional photons. I mean here time-dependent fundamental physical constants; this would inevitably result in time-dependent atomic constants, which in principle may not be equal to their current values [129,258]. As we saw earlier, the cosmological plasma becoming transparent for the CMB radiation depending on the kinetics of Lyα photons via the rate of the two-quanta decay of metastable $2S$ state of the hydrogen atom; hence a weak variation of the fundamental constants can be accompanied with strong changes in the kinetics of recombination.

In fact, I am talking here about investigation of the stability of recombination kinetics, especially at its initial stages at $z \sim 10^3$, that is, the actual time of formation of anisotropy of the CMB. Clearly, regardless of the specifics of mechanisms and sources of energy release, this aspect is of independent interest.

Following Peebles, Seager, and Hu [196], we can offer a sufficiently general phenomenological description of a "nonequilibrium" hydrogen recombination by formalizing the effects of various mechanisms of "pumping" ionizing quanta into the plasma. Namely, we introduce the number density of ionizing and Lyα quanta injected into the plasma,

$$\frac{{\rm d}n_i}{{\rm d}t} = \varepsilon_i(t)n_{\rm H}H(t), \qquad \frac{{\rm d}n_{Ly_\alpha}}{{\rm d}t} = \varepsilon_\alpha(t)n_{\rm H}H(t) \tag{3.29}$$

where $n_{\rm H}$ is the concentration of neutral hydrogen atoms, $H(t) = \dot{a}/a$ is the Hubble parameter, and $\varepsilon_i(t), \varepsilon_\alpha(t)$ are the efficiencies of transformation of the spectrum of injected high-energy particles into ionizing and Lyα photons (see for details [72]). In an analysis of hydrogen recombination kinetics in the presence of an ionizer, as in (3.29), two characteristic time scales can be pointed out, differing in the role played by Lyα quanta in the formation of the ionization equilibrium. The first of them corresponds to redshifts $1{,}000 \leq z < 1{,}400$ when Lyα quanta of CMB play a decisive role in the formation of neutral hydrogen, and the second – to redshifts $z \leq 800$ when the role of Lyα quanta becomes insignificant and recombination processes dominate over ionization processes in the absence of additional ionization sources.

Figure 3.19 plots various ionization curves in the model of "delayed" recombination for $\Omega_{\rm b}h^2 = 0.02$, $\Omega_{\rm tot} = 1$, $\Omega_{\rm DM} = 0.3$, $h = 0.7$, and $\Omega_\Lambda = 0.7$. We see from this figure that as the power of the ionizer (ε_α) increases, the curve $x_l(\varepsilon_\alpha, z)$ is

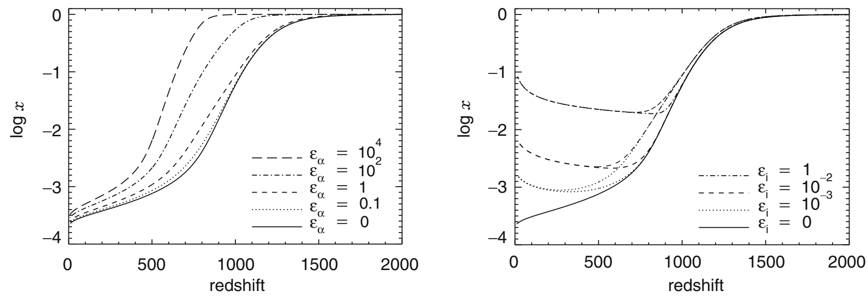

Fig. 3.19 Ionization modes in models with "delayed" recombination. The figure on the *left* corresponds to the resonance photons, that on the *right* – to an ionizer, active at small redshifts ($z < 10^3$). Adapted from [196]

"flattened" in the range $700 \leq z \leq 1{,}400$. Hydrogen recombination becomes more and more delayed even though the change in the residual fraction of ionization (at $z = 0$) is not so large in comparison with a drop of $2 \div 3$ orders of magnitude in the function $x_l(z)$ for $z \simeq 800 \div 10^3$. Obviously, low values of $\varepsilon_\alpha \simeq 0.1 \div 1$ result in insignificant distortions of the ionization mode at $z \simeq 10^3$, which is the most important range for the formation of temperature fluctuations of the CMB.

In the limit $z \ll 1{,}000$, when the effect of excess Lyα quanta on the recombination kinetics becomes negligible, the main mechanism of distortions is the ionization of the $1S$ state of the hydrogen atom. Figure 3.19 gives the results of calculations of the fraction of the ionization $x_e(z)$ in this model for various values of ionizer power $\varepsilon_i = \text{const}$ ([196]; see also [72]). As we see from this figure, nonequilibrium hydrogen ionization leads to considerable distortions of the function $x_e(z)$ at $z < 10^3$ even at relatively low values of the parameters $\varepsilon_i = 10^{-3} \div 10^{-2}$.

Distortion of the CMB Anisotropy and Polarization Power Spectrum for Nonstandard History of Ionization

As was mentioned in [196] (see also [21, 72]), the distortions of recombination change the redshift of the last scattering surface, z_r, and its thickness, Δ_z. In turn, these variations shift positions of the peaks in the CMB anisotropy and polarization power spectrum [122] and change their amplitudes. Here we give the rough analytical estimates of these influences. More accurate numerical results are presented below.

Both parameters of the last scattering surface, z_r and Δ_z, are roughly expressed through the Thomson optical depth, τ_T, and its derivations, $\tau_T^n = d^n\tau_T/dz^n$.

The redshift of the last scattering surface, z_r, is defined by the position of the maximum of the so-called visibility function, $g(\tau_T) = \tau_T^{'} \exp(-\tau_T)$, which gives us the equation

$$\frac{dg(\tau_T)}{dz}|_{z=z_r} = 0 \Longrightarrow \quad \tau_T^{''}(z_r) = \left(\tau_T^{'}(z_r)\right)^2, \tag{3.30}$$

In the vicinity of the maxima, we can use the Taylor series representation of the visibility function and gets the width of the last scattering surface as

$$\Delta_z \simeq \left(2\left(\tau_T^{'}(z_r)\right)^2 - \frac{\tau_T^{'''}(z_r)}{\tau_T^{'}(z_r)}\right)^{-1/2} \quad (3.31)$$

(see e.g., [181]). Taking into account that recombination is a fast process and $\Delta_z \simeq 0.1z_r$, we can further consider only redshift variations of the ionization fraction, x_e. Hereafter z_r and z_* are the redshifts of the last scattering surface in the standard and the distorted models of recombination.

For small distortions of ionization fraction we have [72]

$$z_* \simeq z_r(1 - \varepsilon_\alpha(z_r)\nu), \qquad \Delta_z(z_*) \sim \Delta_{z,st}[1 + 0.5\varepsilon_\alpha(z_r)\Phi^{-1}(z_r)x_{e,st}^{-2}]. \quad (3.32)$$

Here $x_{e,st}$ and $\Delta_{z,st}$ are the ionization fraction and the thickness of the last scattering surface for the standard model ($\varepsilon_\alpha \equiv 0$), $\nu = 0.5\Phi^{-2}(z_r)x_{e,st}^{-3}$, and $\Phi = c\sigma_T n_b(z_r)/H(z_r)$.

As one can see from (3.32) the delay of recombination shift the redshift of the last scattering and increase the width $\propto \varepsilon_\alpha$. This means that for the models with nonstandard history of ionization, the power spectrum of anisotropy and polarization of the CMB should have some features related to the shifting of the last scattering surface and its width. For the anisotropy of the CMB, the leading order dependence of the first acoustic peak location is [196]

$$l_1(\varepsilon_\alpha) \simeq l_{1,st}\sqrt{z_*/z_r} \simeq l_{1,st}[1 - 0.5\nu\varepsilon_\alpha(z_r)]. \quad (3.33)$$

For the second peak of the CMB power spectrum, the shift of the position is practically the same as for the first one, but the ratio between their amplitudes is a linear function of ratio z_*/z_r. For the next few peaks, growth of the thickness of the last scattering surface, Δ_z, increases the damping and decreases the amplitudes of the peaks. In contrast, for models with accelerated recombination, the parameter ε_α is negative [72, 180], and peaks are shifted to larger l and the damping of perturbations decreases together with Δ_z as compared with the standard model.

To get some restrictions on the parameters of resonance and ionizing photons ε_α and ε_i, we can take under consideration the WMAP and Cosmic Background Imager (CBI) data for the CMB anisotropy, covered the range of multipoles $l \leq 1,500$. The simplest approach is to fix the most probable value of the cosmological parameters obtained by the WMAP team and then to analyze the perturbations of the likelihood function coursed by nonstandard ionization history of the plasma. The results of these investigation are presented in Fig. 3.20.

These maps give us the best fit for the $\varepsilon_\alpha \leq 0.1$ and $\varepsilon_i \leq 3 \times 10^{-3}$ parameters from the recently available CMB data. The values of these parameters are in agreement with the results obtained by analysis of the WMAP data and SDSS survey [21]: $\varepsilon_\alpha \leq 0.3$, $\varepsilon_i \leq 4 \times 10^{-3}$. However, as it was pointed out by [21], both

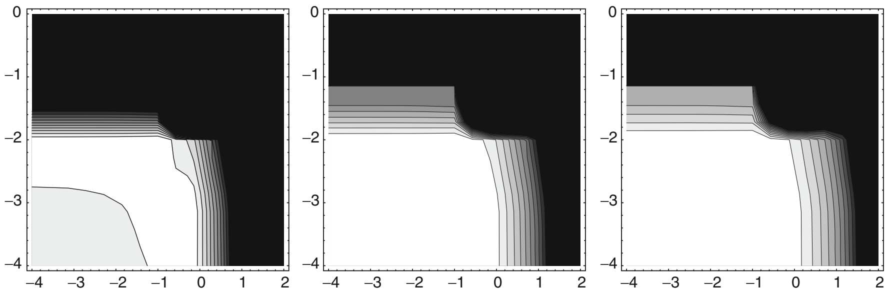

Fig. 3.20 The $\varepsilon_\alpha - \varepsilon_i$ diagram for significance level, obtained by combination of the WMAP and CBI data sets. *Vertical* axis corresponds to the ε_i-parameter, *horizontal* one is for the ε_α parameter (both in logarithmic scale). The *left* plot corresponds to the 68% confidential level, the *middle* plot is for 95%, the *right hand side* plot corresponds to 99.9%

the parameters can produce significant renormalization of the most important cosmological parameters, such as spectral index of adiabatic perturbations, the density of the DE, the tensor/scalar ratio, the neutrino rest masses, etc. As it was mentioned by [72], the most important information about the epoch of recombination can be obtained by using the high multipole range of the CMB polarization. This is why the future *Planck* mission data seem to be extremely valuable for reconstruction of the ionization history of the Universe.

To Summarize

I have made a systematic discussion of the ionization history of the cosmic plasma and its relation to the physics of the CMB. This field of the modern cosmology involves an enormous diversity of processes that play their role in the Universe. The CMB proved to be a true "goldmine" for extracting scientific information on these processes; it has in fact grown into the central branch of modern cosmology. Comparing this highly perfected theory with observational data makes it possible to obtain essential information on the early Universe and on the physical parameters of the Universe as a whole. Results from BOOMERanG, MAXIMA-I, *Archeops*, CBI, Degree Angular Scale Interferometer (DASI), and WMAP were so important that they have taken the field into the era of "Precision Cosmology", producing impressive constraints on many fundamental cosmic parameters and have led to a very definite picture of the structure and evolution of the Universe. As a result, one could become overexcited and declare that almost everything in the ionization history of the cosmic plasma is known. However, we want to emphasize that even after those remarkable projects, the study of CMB physics is not coming to an end. There are still many unsolved problems in cosmology and another generation of satellite experiments, as well as ground-based and balloon-borne experiments, is needed. The deeper science penetrates and the more mysteries it solves, the more problems it discovers, each more daunting and less predictable than the last. Here we want to emphasize that the study of the re-ionization process is a crucial test of the correctness of our knowledge of the processes of the formation of structure in the Universe. It also tests our knowledge of the possible nature of the hypothetical unstable particles, the decays of which influenced the kinetics of hydrogen recombination.

After the beginning of the era of "Precision Cosmology", the number of questions affecting the basic fundamentals of cosmology increased significantly. I believe there is plenty of room for surprises. The cosmological and physical community is now working on future projects such as the *Planck* mission, the Atacama Large Millimeter/submillimeter Array (ALMA), the Laser Interferometer Space Antenna (LISA), etc., which will uncover yet more cosmological surprises.

Thank you very much Pavel. Great expectations are indeed posed in new challenges, some of them described in Chap. 5, aimed at mapping the structure formation as its infancy, observed in CMB anisotropies, to the subsequent evolution. We hope to observe in the future the signals from astrophysical sources responsible for (at least the later stages of) the reionization of the Universe. This would allow to more accurately model the ionization history associated with structures and then, indirectly, to better understand or constrain, decay models.

The previous section has shown how deeply CMB studies are linked to the understanding of cosmic structures. This aspect will be highlighted further in the next interview to Amedeo Balbi who will discuss the physical link between CMB and LSS and the most important effects and cosmological results based on it.

3.7 Large Scale Structure

Dear Amedeo *(Balbi)*, **CMB anisotropy data are often analyzed jointly to other kinds of cosmological and astrophysical data. In particular, both the CMB anisotropy and the large scale structure of the Universe mapped by galaxy surveys are related to the genesis and evolution of primordial perturbations and to the cosmological model and parameters. Could you comment on the synergy between CMB anisotropy and large scale structure information for the current comprehension of the Universe?**

The present Universe is very complex. The estimated number of galaxies in the observable Universe is of the order of one hundred billion; a typical galaxy, in turn, contains hundreds of billions of stars. It is not just the number of structures found in the Cosmos to be mind-boggling. Galaxies are arranged into a magnificent architecture: groups of tens, hundreds, even thousands of them are bound together into clusters; clusters clump with other clusters to form superclusters. Matter is distributed in a hierarchy of larger and larger structures – huge voids are interspersed with high-density knots, filaments, walls, and so on, forming a foam-like pattern of stunning complexity.

On the other hand, the early Universe was much simpler. Matter was spread almost uniformly in space, except for the presence of tiny random inhomogeneities, most likely generated during a primordial phase of accelerated expansion called inflation [5]. Those slight perturbations were the seeds that, over the past 13.7 billion years, gravity used to build the complex cosmic structure we can now observe [235]. When matter decoupled from radiation – at an epoch called recombination, situated about 3.8×10^5 years after the big bang – those seeds left an observable imprint as temperature fluctuations in the CMB [121].

It is natural, then, to investigate the relation between the pattern of CMB temperature fluctuations and the present matter distribution. The CMB pattern is basically the blueprint for every cosmic structure we observe today. Putting together the information on how the matter was distributed a few hundred thousand years after the big bang (derived from CMB observations) with the present large-scale structure (reconstructed from galaxy surveys) may then bridge a gap of billions of years in reconstructing the cosmic evolution.

We are just starting to have the high-quality cosmological data needed in order to perform this kind of study. On the one hand, CMB observations have become increasingly accurate over the past decade. We now have full-sky high-resolution microwave maps produced by WMAP, reconstructing subtle temperature variations in the CMB of the order of tens of a millionth of a degree [119]. On the other hand, the gigantic task of mapping the three-dimensional spatial distribution of galaxies on large volumes of the observable Universe has been successfully tackled by such international collaborations as the SDSS [3] and the 2-Degree Field Galaxy Redshift Survey (2dFGRS) [198]. We then have, for the first time, the possibility of comparing two snapshots of the Universe taken at very different epochs, which we can use to infer the physical processes that governed its evolution – a bit like paleologists can reconstruct the history of species by combining fossil records.

One rather straightforward way of exploiting the combination of CMB and large-scale structure data is to look into the way the distribution of perturbations was altered over the course of cosmic evolution – for example, how the primordial distribution was shaped by the physical mechanisms that governed the collapse of structures at much later times. The CMB is a powerful probe to investigate the statistical properties of the density inhomogeneities at very early times. From the sky pattern of CMB fluctuations at large angular scales, we can infer the initial state of the inhomogeneities, as such scales encompass regions of the Universe that were not in casual contact at the time of recombination. In other words, whatever seeds were left in the Universe after inflation, they should be found "frozen" in the large-scale angular distribution of CMB fluctuations. Basic models of inflation predict that the random density perturbations were spread in the early Universe with the same amplitude independently of their spatial extent. This property of the perturbations, called "scale invariance", was first postulated at the beginning of the 1970s by Edward Harrison in the United States and Yakov Zel'dovich in the Soviet Union, and only later found to be a natural outcome of the inflationary phase. Scale invariance is a sort of "democratic principle", preventing the preferential formation of structures on some particular scale – for example the rapid growth of small, compact objects, or of very large aggregations of matter. More formally, inflationary models predict that the power spectrum of primordial density perturbations should take the form

$$P(k) = Ak^n, \tag{3.34}$$

where k is the wave-number of each plane wave contribution to the total density perturbation (roughly speaking, a spherical overdensity or underdensity with radius $\lambda = 2\pi/k$), A is the amplitude of the power spectrum at some given reference pivot scale k_0, and n is a spectral index, which is equal to 1 in the case of perfect scale invariance. Inflation predicts a spectral index very close to 1.

The prediction of scale invariance has been confirmed using CMB data ever since COBE's first results, and to much higher accuracy by WMAP. Slight deviations from scale invariance are possible, however, and actually can be used to select specific models of inflation. When CMB data are combined with large-scale structure data (see Fig. 3.21), we cannot only refine our measurements of deviations from scale

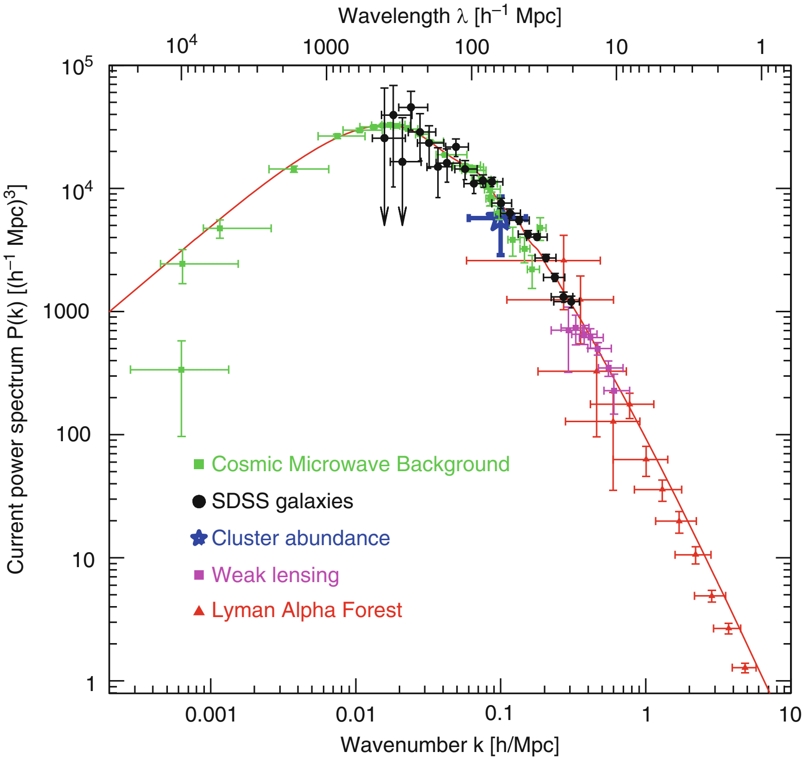

Fig. 3.21 The power spectrum of matter density fluctuations as measured by a number of different cosmological probes: the CMB, the distribution of galaxies in the SDSS, the abundance of galaxy clusters, the weak lensing in galaxy catalogues and the Lyman alpha forest. The continuous curve is a fiducial power spectrum for scale-invariant initial condition and a flat cosmology From [252]

invariance, but we can also try to measure any variation in the spectral index of density perturbations with respect to their scale (i.e., a "running" spectral index $n = n(k)$). We can also use the CMB data as a large-scale "pivot" point to interpret the matter distribution in today's Universe on much smaller scales.

The combination of CMB and LSS data was crucial to establish, in the 1990s, the current standard model of structure formation based on the so-called CDM paradigm, as opposed to the rival Hot Dark Matter (HDM) one. The two differed by the nature of the (so far unknown) particle species making up the bulk of matter in the Universe. CDM has small velocity dispersion, so that gravitational clumping is more effective on smaller scales, leading to the formation of galaxies and then, hierarchically, of larger structures. HDM, on the contrary, has large velocity dispersion (the typical example being a massive neutrino species), so that structure formation has to proceed from the top-down: large structures form first, only later fragmenting into galaxy-size pieces. If HDM were the dominant form of matter in the Universe, the observed amplitude of fluctuations in the CMB would be too small for structures to form, given the time available from recombination to the present. The CDM paradigm for structure formation, instead, can fit both the CMB and the large-scale structure data, and is now currently preferred, even if no known particle species is suitable to act as CDM.

Actually, the current standard cosmological model includes an even greater unknown than CDM. Most of the total energy density of the Universe (roughly, 70%) is now thought to be in the form of a so-called *DE*, whose simplest candidate is the cosmological constant Λ first postulated by Albert Einstein in 1917. (This is why the current standard cosmological model is often referred to as the ΛCDM model.) Assessing the nature of DE is one of the greatest challenges of modern cosmology. In fact, it is a huge problem for fundamental physics at large, as any attempt to theoretically estimate the magnitude of the cosmological constant (which modern quantum field theory interprets as the energy density of the empty spacetime, or vacuum) is in strong conflict with cosmological observations, by many orders of magnitude (40–120, according to most calculations) [266]. DE has bizarre physical properties, acting as a source of repulsive gravity on large scales; this drives an acceleration of the expansion of the Universe. The existence of such acceleration was first detected in 1998 by two separate teams (led by Saul Perlmutter of Lawrence Berkeley Lab [199] and Adam Riess of the Hubble Space Telescope Science Institute (HST-ScI) [205]) analyzing the redshifted spectra of distant type Ia Supernovae (SNe).

The combination of CMB and LSS data promises to be one of the most powerful tools to investigate the nature of DE. CMB alone cannot provide much information about DE, as at the time of recombination its contribution to the total energy density of the Universe was negligible. Only in relatively recent times did the DE start to become a relevant portion of the cosmic budget. This is because the DE density remains nearly constant during cosmic evolution, while the matter density decreases as the volume of the Universe expands. Why did we live precisely in the epoch of cosmic evolution when the densities of matter and DE happen to be of the same order of magnitude is yet another mystery to be solved. Anyway, when DE starts dominating the cosmic expansion, structure formation is already under way. The onset of cosmic acceleration alters in a peculiar way the curvature of spacetime associated with matter aggregations. CMB photons passing through these changing gravitational "wells" have their energy changed by a certain amount, resulting in a CMB temperature variation with respect to average given by

$$\frac{\Delta T}{T} = 2 \int \mathrm{d}t \, \dot{\Phi} \tag{3.35}$$

where $\dot{\Phi}$ is the time derivative of the gravitational potential and the integral is done over the path traveled by the photons. To understand this, we have to keep in mind that photons gain energy (i.e., they are "blueshifted") when they approach a region of space where gravity is stronger, and lose energy (i.e., they are "redshifted") when they abandon it. If the gravitational well remains unaltered during the passage of photons, the net energy difference is zero (because the energy lost abandoning the well compensates the energy gained entering the well). However, the forming structures encompass very large regions of space, so that it takes a considerable amount of time for the CMB photons to travel across them. The time variation of gravitational wells during the passage of CMB photons then results in a nonzero net energy difference for the photons. This means that regions of structure formation imprint an additional signature on the CMB – a signature that was not present at the time of recombination but is generated at much later times. This signature carries valuable information on the DE, as the precise variation of the gravitational well depends on the specific DE model.

The energy change of photons from recombination to the present, generated by forming structures encountered along the photons' path, is called ISW effect (see Fig. 3.22). To extract the contribution of the ISW effect (which is crucial to investigate the nature of DE) from the observed CMB pattern, we have to cross-correlate CMB maps with tracers of the present matter distribution (obtained, e.g., from galaxy surveys). The effect is very subtle and requires both sophisticated statistical techniques and accurate data to be discerned. Early attempts to detect the ISW effect were made at the end of the 1990s by combining COBE's microwave sky maps with X-ray maps, although only in 2004 a positive detection was announced (by Stephen Boughn and Robert Crittenden [36]) from the combination of WMAP first-year data with X-ray maps and radio catalogues. Later, many different teams have confirmed the presence of a ISW signature at a high confidence level, and found that it is indeed compatible with the presence of a cosmological constant accelerating the cosmic expansion.

One well known feature of the CMB anisotropy pattern is its ability to act as a standard ruler – that is, to provide a known reference physical scale that can be used to probe the large-scale geometry of the Universe. Density perturbations in the early Universe evolve under the competing action of gravity and pressure, resulting in the production of sound waves. The length traveled by such sound waves before recombination (the so-called *sound horizon*) remains impressed in the CMB and can be compared to theoretical predictions. This method was successfully used by the BOOMERanG [65] and MAXIMA [11] experiments in 2000, and then by

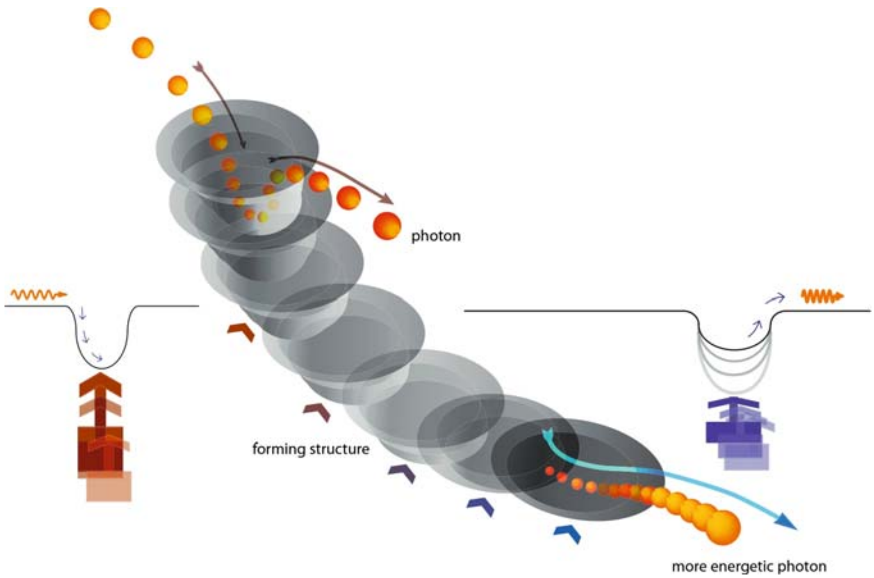

Fig. 3.22 A depiction of the ISW mechanism: when CMB photons pass through an overdense region, they gain and lose the same amount of energy if the gravitational potential does not vary (*left*). However, when the gravitational potential changes during the passage of photons, they can gain energy, resulting in an additional anisotropy in the CMB pattern

the WMAP satellite in 2003 [232], to prove that the Universe has a flat (that is, Euclidean) geometry. The very same sound waves should also leave an imprint in the large-scale distribution of matter in the Universe, at a much later time in cosmic evolution. This prediction was actually confirmed in 2005 by an analysis of the SDSS galaxy survey [85]. The presence of such baryon acoustic oscillation feature in both the CMB and the local distribution of matter allows us to compare how the same physical length (that is the sound horizon) is projected when is seen at very different distances. This can then be used to determine the properties of DE, the geometry of the Universe, and its expansion rate.

There is at least another obvious example of the interplay between the galaxy distribution and the CMB pattern. The space–time curvature caused by mass concentrations bends the CMB photons' path from recombination to the present – a manifestation of the gravitational lensing phenomenon first theorized by Albert Einstein. This is an even subtler effect than the ISW, and does not alter the energy of CMB photons, but only their trajectories. The overall result of many such deviations is a slight "blurring" of the CMB pattern, which can be confused with a genuine reduction of anisotropy at the time of recombination. The correlation of CMB with LSS observations can be used to separate the effect of gravitational lensing, leading to improved constraints on cosmological parameters (e.g., on the mass of neutrinos). A detection of gravitational lensing of the CMB by large-scale structure was recently announced by Kendrick Smith and collaborators [230] and by Christopher Hirata and collaborators [248] using the WMAP data and different tracers of the large-scale distribution.

Of course, much more accurate observations of the CMB and of the LSS of the Universe will be performed in the near future. The prospect for increasing our knowledge of the mechanisms governing the evolution of the Cosmos by combining information from different cosmic epochs looks now more promising than ever.

Thanks a lot Amedeo.
Another important analogy between CMB and LSS, originated from the matter-radiation coupling in the plasma, is discussed by Charles Bennett in the next interview.

3.7.1 Baryon Acoustic Oscillations

Dear Charles *(Bennett)*, **the statistical analysis of galaxy distribution in terms of power spectrum can be used to understand the coupling between matter and radiation. Can you give the fundamental concepts for the comprehension of the baryon acoustic oscillations and their cosmological relevance? Existing surveys are already able to identify this observable?**

The effect of the sound waves, discussed in the sections dedicated to CMB anisotropies, is not limited only to the photons that we observe as the CMB. The sound waves also perturbed the cosmic gas, which constitutes about 1/6 of the total matter in the Universe. Hence, the length scale that was imprinted by the sound waves persists to the present time in the clustering of matter on large scales in the Universe.

This preferred galaxy separation scale was measured from the galaxies seen in both the SDSS [85] and the 2dFGRS [55]. These data show an excess galaxy correlation at the expected characteristic scale based on the WMAP measurements and cosmological model fit. As we understand the bump in the power spectrum based on a ruler whose physics is fully understood, the position of the bump can be a powerful probe of the cosmological model as the angular separation of the galaxies also depends on the expansion history of the Universe.

The primordial Universe was peppered with slightly over-dense, over-pressure regions, initiating the propagation of sound waves of all wavelengths. Consider a primordial perturbation where a small patch of space is a little more dense than the surrounding space. In the standard cosmological theory, the initial perturbations are adiabatic, so all of the species (neutrinos, baryons, CDM, photons, etc.) are perturbed the same fractional amount.

Neutrinos are only very weakly interacting, so they stream away from the initial perturbation. CDM, with no effective internal motion, moves sluggishly, only in response to gravity. As the perturbation is overdense, the surrounding matter is attracted, causing more CDM to fall towards the center. In the ionized plasma of baryons and electrons, the mean free path of the photons is short due to Thomson electron scattering. The baryonic gas and photons are locked together into a single fluid.

The combined high pressure fluid generates an expanding spherical sound wave, with a sound speed of 57% of the speed of light. The CDM collects in the overall density perturbation so the CDM peak remains centrally concentrated, but starts to spread. This is what causes the CDM power spectrum to turn over.

The density of the expanding spherical wave drops as its energy is spread over an increasingly large area. As the expanding Universe cools, the electrons and nuclei eventually combine to form neutral atoms. The sound speed drops suddenly due to the reduced photon–baryon coupling, stopping the expansion of the pressure wave. The photons then free-stream away.

The baryons and the CDM gravitationally attract, locking together. This results in the acoustic peak. A CDM perturbation remains at the original center along with a small perturbation in a spherical shell with a radius of $\sim$150 Mpc.

At late times, galaxies form in the overdense regions at the locations of the initial overdensities and the 1% enhancements in the 150 Mpc surrounding shell. We see this acoustic peak in the correlation function of galaxies (see Fig. 3.23), or equivalently, as a series of acoustic oscillations in the power spectrum. The Universe actually had innumerable underdense and overdense regions, and their small fluctuations simply sum linearly.

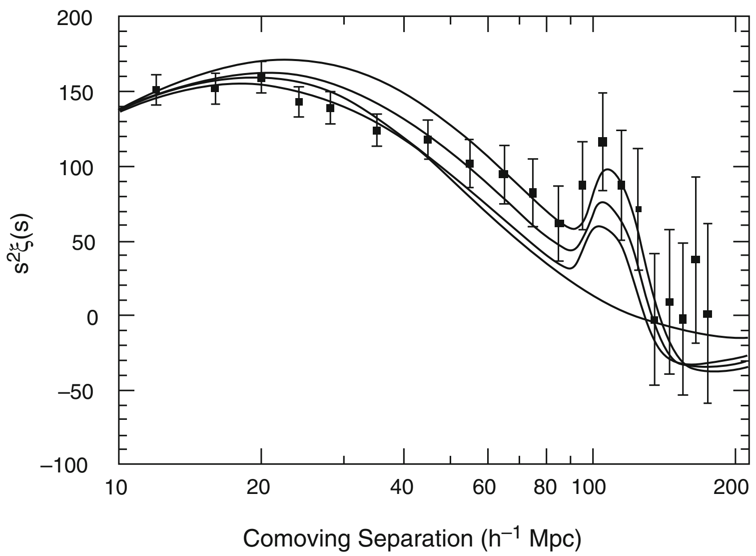

Fig. 3.23 Large-scale redshift-space correlation function of the SDSS LRG sample multiplied by s^2. The models are: $\Omega_m h^2 = 0.12$ (*top line*), 0.13 (*second line*), and 0.14 (*third line*) all with $\Omega_b h^2 = 0.024$ and $n = 0.98$ and with a mild nonlinear folded in. The *bottom line* shows a pure CDM model ($\Omega_m h^2 = 0.105$), which lacks the acoustic peak but that is close to the best fit due to the data points on intermediate scales. The bump 1 Mpc scale is statistically significant. From [85], where further details can be found

Thank you Charles.
In the previous sections, we entered in cosmic epochs when the complexity of cosmic structures and of their feedback played crucial roles. The next discussion with Peter Coles describes important aspects related to the modeling of structure formation.

3.7.2 Large Scale Structure Through Simulations

Dear Peter *(Coles)*, **the dynamics of structure formation involve most of the present day knowledge of the astrophysical processes occurring between matter and radiation. In particular, the nonlinear phase of structure formation can be tested only by semi-analytic methods and N-body techniques. Would you like to discuss pros and cons of these two different approaches?**

In modern theories of cosmic structure formation, small-amplitude initial fluctuations in the density grow by a process of gravitational instability as the Universe expands. In the early stages, this growth is linear and can be handled analytically using standard perturbation theory. When the density contrast, defined by

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}, \tag{3.36}$$

is of order unity, so that variations in density ρ are of order the mean density $\bar{\rho}$, these methods break down. In a nutshell, what happens is that bound structures form by separating from the cosmic expansion and in models dominated by CDM this happens in hierarchical fashion, with small structures collapsing first and then becoming incorporated in objects of increasing mass as time goes on. Simple theoretical models, such as the Press-Schecther theory [202], have proved successful at describing some aspects of this process, but they do not represent a complete understanding of the evolution of the density field in the nonlinear regime. Moreover, pure gravitational instability calculations can describe only the CDM component, because this neither exerts nor feels any nongravitational forces. The baryonic component of the Universe has a much larger repertoire of possible interactions, including complex hydrodynamic and radiative processes, and the scope for modeling these using analytical methods is extremely limited. The latter stages of structure formation, involving the collapse of baryons into DM haloes and subsequent Star Formation (SF) and feedback, therefore pose stiff theoretical challenges.

A "brute-force" approach would be to deploy fully numerical methods for the evolution of both the DM and baryonic components, using either smoothed-particle (Lagrangian) or grid-based (Eulerian) techniques. The first steps towards the creation of the cosmological simulation industry involved DM-only simulations on relatively small computers (by modern-day standards). These generally used a particle-mesh (PM) technique, which allows a faster calculation of long-range forces than is possible by direct summation of the Newtonian gravitational forces between particles used to represent the overall matter distribution. To prevent numerical artifacts caused by the discrete nature of the simulation particles, a softening length is usually introduced so that the force law is usually of the form

$$\mathbf{F}_{ij} = \frac{G M_i M_j (\mathbf{x}_j - \mathbf{x}_i)}{\left(\epsilon^2 + |\mathbf{x}_i - \mathbf{x}_j|^2\right)^{3/2}}. \tag{3.37}$$

Basically, one takes a set of particles and then uses them to construct a density field by smoothing onto a discrete mesh. The gravitational potential corresponding to this density field is then computed on the grid using Green's function methods; forces at actual particle positions are then found by interpolation. The use of the grid makes it possible to use Fast Fourier Transform methods, which mean that this problem scales with particle number N as $N \log N$ rather than $N(N-1)$, which would be the case of direct summation of the inter-particle forces. Higher resolution of the short-range forces can be achieved by computing some of the interactions directly, such as in the particle–particle–particle–mesh P^3M method, which was the mainstay of early CDM calculations [64].

To include hydrodynamics as well as DM, one has to extend the code to describe gas pressure forces and other physics. There are two principal algorithms for doing this, based either on particles or a grid. In smoothed particle hydrodynamics (SPH), for example, two sets of particles are used. One traces the DM distribution as before and the other obeys more complex rules intended to follow the gas physics. In particular, the baryonic component feels pressure forces and can also lose energy through cooling. Thermodynamic properties of the gas are computed using a kind of average over neighboring particles. This kind of approach has the advantage that particle move into regions where the density is high, thus automatically increasing the resolution there. For many years, SPH out-performed the competing grid-based codes for that reason (e.g., [234]). However, more recently, adaptive mesh codes have been developed that can exploit a finer mesh were extra resolution is required.

Generally, the approach used involves two stages. First, a large volume is simulated, which is intended to be representative of the Universe as a whole. This can establish broad properties of the structures formed and generate meaningful distributions of halo masses, etc. In the second stage, individual haloes are extracted from the big box and resimulated with much higher resolution, that is, with lower particle mass in order to test more detailed aspects of individual galaxies and their parent haloes.

However, galaxy formation entails such a huge range of physical scales that resolving everything simultaneously poses extreme difficulties even for the largest supercomputers. An alternative approach is therefore to encode the non-gravitational physics into a series of simplified rules for the "sub-grid" physics to be incorporated in a code, which evolves the larger-scale matter distribution. This is the so-called "semi-analytic" approach (e.g., [20]). Ingredients involved in galaxy formation recipes of this type include prescriptions for cooling and disk formation within DM haloes, rules for the merging of these disks to form spheroids, SF and astrophysical feedback (including starbursts and AGN), stellar population models, rules for the generation of dust and chemical evolution generally. The outputs of the calculation are "observable" galaxy properties, such as disk sizes and angular momenta, synthetic spectra, and so on.

Current semi-analytic models owe a great deal to earlier work by White and Rees [267], who proposed that galaxy formation basically happens in two stages, with haloes forming by the collapse of DM haloes (described by the Press–Schechter theory) and galaxies subsequently forming in these haloes following radiative cooling of the baryonic component. The loss of energy by the baryons allows them to collapse further than the DM, forming tighter bound objects that are stabilized against the subsequent merger of haloes during the continued hierarchical growth of clustering. One of the later extensions of this approach was to take greater account of the hierarchical nature of the process by introducing merger trees that represent the assembly history of a halo. Two haloes of the same mass but with different merger histories will probably end up housing galaxies with different observed properties. The first full semi-analytic models [54, 141] were based on this approach, and although there are now many more sophisticated versions, this argument is still central to the general idea.

Critics of the semi-analytic approach point to the apparently arbitrary nature of some of the rules and the plethora of adjustable parameters such models involve. On the other hand, the alternative ab initio approach of computing sub-grid properties fully is simply not possible from a practical point of view. Moreover, one should always remember that even the basic DM calculations involve some approximations.

It is also pertinent to point out that the difference between semi-analytic and numerical approaches is perhaps not as great as many people seem to think. The use of resimulation methods is not conceptually different from embedding theoretical merger trees in a larger-scale simulation. Both approaches seem to have the same problems too. For example, they both have a problem with excessive gas cooling unless some feedback mechanism, either thermal or kinetic, is introduced to regulate the process. It remains likely that there will always be physical properties that evade the resolving power of any computer, and semi-analytic approaches are a reasonable way to patch up the gaps in the overall picture at least until someone thinks of something better!

Thank you very much Peter.
As discussed in many sections of Chap. 2 and of this chapter, neutrinos play a non-negligible role for many aspects of cosmology and astrophysics, from Big Bang Nucleosynthesis (BBN) to DM, from primordial perturbations to CMB and LSS, from ionization history to energy losses in stars, in spite of the fact that the neutrino energy density is likely always subdominant in the various cosmic epochs. In the next interview, Lucia Popa will give a unified overview on the role of neutrinos in cosmology, after a description of the fundamental results coming from dedicated experiments.

3.8 Neutrino Physics and Its Cosmological Implications

Dear Lucia *(Popa)*, the comprehension of the physical properties of neutrinos is one of the most interesting topic where a substantial progress has been recently achieved by combining information from particle physics experiments, astrophysical observations, and astroparticle projects. Could you summarize the overall picture and the questions still open in neutrino physics?

Neutrino cosmology is a fascinating example of the fecund interaction between the particle physics and astrophysics, the advance in this field being the result of the combined efforts in the two areas.

Nowadays we know from the Z^0 boson decay width that the number of active neutrino flavors is $N_\nu = 2.944 \pm 0.012$ [84]. The solar and atmospheric neutrino oscillation experiments indicate the existence of nonzero neutrino masses in eV range [7, 97]. The oscillation experiments can only measure the difference between the squared masses of neutrinos, being insensitive to the absolute neutrino mass values, leading to two possible mass schemes that leave one neutrino mass unconstrained. These two schemes are known as normal and inverted hierarchies, characterized by the sign of Δm^2_{31} being positive and negative, respectively. Neutrino oscillations impose a lower limit on the heaviest neutrino mass of about 0.05 eV. There are also indications from neutrino oscillations with larger mass-squared difference coming from short baseline oscillation experiments [4, 10] that can be explained by adding one or two sterile neutrinos with eV mass to the standard scheme with three active neutrino flavors (see [173] for a recent analysis).

Figure 3.24 presents the relation between the individual masses and the total neutrino mass [162] obtained by taking as reference the 3σ ranges of the differences of the squared neutrino masses obtained by atmospheric and solar neutrino oscillation experiments [173].

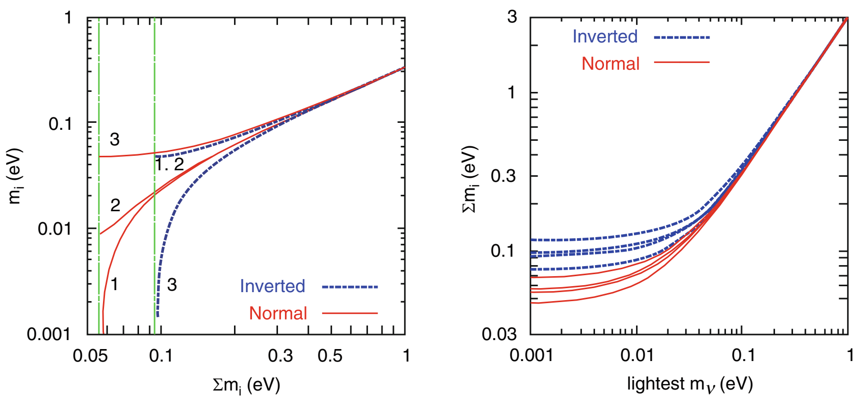

Fig. 3.24 The expected values of neutrino masses according to the values of Δm^2 relevant for solar and atmospheric neutrino oscillations for normal and inverted hierarchies [162]. (*Left*) Individual neutrino masses as function of total neutrino mass. (*Right*) Ranges of the total neutrino mass as function of lightest neutrino mass sate

The precise measurement of the electron spectrum in β-decays is the only laboratory technique for the direct measurement of a small neutrino mass, without additional assumptions on the character of neutrinos being Dirac or Majorana particle (see e.g., [86] for a recent review). The neutrino mass, or an upper limit to it, is inferred from the shape of the energy spectrum near its kinematical end point, usually defined as Q-value of the decay. In practice, the most sensitive experiments use the β-decay of tritium. Two groups, one from Mainz experiment [168] and the other from Troitzk experiment [150], ended up with a similar upper limit on the effective neutrino mass, $m_\beta < 2.3\,\text{eV}$ (2σ). A new experiment, Karlsruhe Tritium Neutrino (KATRIN) [140], currently being in construction, aims for a tenfold improvement in sensitivity, down to $m_\beta <0.2\,\text{eV}$.

The neutrinoless double beta decay ($0\nu2\beta$) is a rare nuclear process where the lepton number is violated and whose observation would mean that neutrinos are Majorana particles. The results from $0\nu2\beta$ measurements are converted into upper limits of effective neutrino mass $m_{\beta\beta}$. The present 2σ upper limit is $m_{\beta\beta} < (0.44 \div 0.62)$ [237]. The above range corresponds to the results of $0\nu2\beta$ experiments such as Heidelberg–Moscow [145], International Germanium Experiment (IGEX) [1], and Cuoricino[7] [8]. It is expected that the future $0\nu2\beta$ experiments will improve the current sensitivities down to values of order $m_{\beta\beta} \sim 0.01 \div 0.05\,\text{eV}$ [88].

Neutrino left over from the early epochs of the evolution of the Universe must have at present a number density of about $n_\nu = 339$ neutrinos and antineutrinos per cm^3, an equilibrium Fermi–Dirac spectrum, and the present temperature $T_{\nu_0} = (4/11)^{1/3}\, T_\gamma = 1.945\,\text{K}$, where $T_\gamma = 2.73\,\text{K}$ is the present temperature of the CMB photons. As the Universe cools, the averaged neutrino weak interaction rate falls below the expansion rate given by the Hubble parameter, and neutrinos decouple from the rest of the plasma. An estimate of decoupling temperature, $T_{\text{dec}} \simeq 1$ MeV, can be found by equating these two quantities and assuming the standard picture of the instantaneous neutrino decoupling. Neutrinos with eV mass are ultra-relativistic at the decoupling time. Shortly after the neutrino decoupling, the photon temperature drops below the electron mass, favoring e^+e^- annihilations that heat the photons. However, some relic interactions between e^+, e^-, and neutrinos exist, leading to nonthermal distortions of the neutrino spectra and to a slightly smaller increases of the comoving photon temperature (see, e.g., [137] and references therein).

A proper calculation of the noninstantaneous neutrino decoupling involves the computation of the momentum-dependent neutrino spectra [90], the inclusion of the finite temperature Quantum Electro Dynamics (QED) corrections to the electromagnetic plasma [174], and the effects of neutrino flavor oscillations [175]. In practice, these effects have only small impact on the evolution of the cosmological perturbations and for many purposes can be safely neglected. Neutrinos fix the expansion rate during the cosmological era when the Universe is dominated by radiation.

[7] A smaller scale experiment to test the technical feasibility of the Cryogenic Underground Observatory for Rare Events (CUORE) project.

It is common to write the energy density of neutrinos and anti-neutrinos, when still relativistic, in terms of *the effective number of neutrino species*, $N_{\rm eff}$, as

$$\rho_\nu = N_{\rm eff}\frac{7\pi^2}{120}T_\nu^4,$$

where T_ν is the neutrino temperature. Taking into account the corrections due to the noninstantaneous neutrino decoupling, the three active neutrino flavors contribute as $N_{\rm eff} = 3.046$ [174]. In particular, the primordial light element abundance predictions in the standard theory of BBN [44, 185, 264] depend on the baryon-to-photon ratio, $\eta_{\rm B}$, and on the radiation energy density at the BBN epoch parametrized by $N_{\rm eff}$. Any departure from the standard value $N_{\rm eff} = 3.046$ would be due to nonstandard neutrino features or due to the possible existence of new particles as axions and gravitons, the time variation of the physical constants and other nonstandard scenarios (see, e.g., [212] and references therein) at the BBN epoch.

At the same time, more phenomenological extensions to the standard neutrino sector have been studied, the most natural being consideration of the leptonic asymmetry [96, 138, 207, 208], parametrized by the neutrino degeneracy parameter $\xi_\nu = \mu_\nu / T_{\nu 0}$, where μ_ν is the neutrino chemical potential and T_{ν_0} is the present temperature of the neutrino background (see also the early works [22, 23, 270]). Although the standard model predicts the leptonic asymmetry of the same order as the baryonic asymmetry, $B_{\rm as} \sim 10^{-10}$, there are many particle physics scenario in which a leptonic asymmetry much larger can be generated [51, 229]. One of the cosmological implications of a larger leptonic asymmetry is the possibility to generate small baryonic asymmetry of the Universe through the non-perturbative (sphaleron) processes [42, 91, 156]. As the measured neutrino mixing parameters implies that neutrinos reach the chemical equilibrium before BBN [2, 70, 269], all neutrino flavors can be characterized by the same degeneracy parameter, ξ_ν, at this epoch. The most important impact of the leptonic asymmetry on BBN [14, 147, 236] is the shift of the beta equilibrium between protons and neutrons and the increase of the radiation energy density. The BBN constraints on $\rm N_{eff}$ have been recently reanalyzed from the comparison of the theoretical predictions and experimental data on the primordial abundances of light elements, by using the baryon abundance derived from the WMAP CMB anisotropy measurements [118, 186, 233]: $\eta_{\rm B} = 6.14 \times 10^{-10}(1.00 \pm 0.04)$. In particular, the ^{4}He abundance, $Y_{\rm p}$, is quite sensitive to the value of $N_{\rm eff}$. The conservative error analysis of helium abundance, $Y_{\rm P} = 0.249 \pm 0.009$ [184], yielded to $N_{\rm eff} = 3.1^{+1.4}_{-1.2}$ (2σ) in good agreement with the standard value [176], but still leaving some room for nonstandard values, while more stringent error bars of helium abundance, $Y_{\rm p} = 0.2516 \pm 0.0011$ [131], leaded to $N_{\rm eff} = 3.32^{+0.23}_{-0.24}$ (2σ) [126] (note that the quoted errors depends on the type of statistical analysis, e.g., maximum likelihood, minimization or marginalization).

The stronger constraints on the degeneracy parameter obtained from BBN [222] gives $-0.04 < \xi < 0.07$ (1σ), adopting the conservative error analysis of $Y_{\rm p}$ by [184] and $\xi = 0.024 \pm 0.0092$ (1σ), adopting the more stringent error bars of $Y_{\rm p}$ by [130].

The CMB anisotropies and the LSS matter density fluctuations power spectra carry the signature of the energy density of the Universe at the time of matter–radiation equality, making possible to measure $N_{\rm eff}$ through its effects on the growth of cosmological perturbations. More effective number of relativistic neutrino species enhances the ISW effect on the CMB APS, leading to a higher first acoustic Doppler peak amplitude. Also, the delay of the epoch of matter–radiation equality shifts the LSS matter power spectrum turnover position toward larger angular scales, suppressing the power at small scales. In particular, for the leptonic asymmetric models, the neutrino mass is lighter than in the symmetric case. This leads to changes in neutrino free-streaming length and neutrino Jeans mass due to the increase in the neutrino velocity dispersion [127, 160].

Our ability to measure the matter energy density, $\Omega_{\rm m}h^2$, through the redshift of the matter–radiation equality epoch, $z_{\rm eq}$, depends on how exactly $N_{\rm eff}$ is known, as $N_{\rm eff}$ and $\Omega_{\rm m}h^2$ are linearly correlated, the width of the degeneracy line being given by the uncertainty in the determination of $z_{\rm eq}$ [149]:

$$1 + z_{\rm eq} = \frac{\Omega_{\rm m}h^2}{\Omega_\gamma h^2}\frac{1}{1 + 0.227\, N_{\rm eff}}.$$

In the above equation, $\Omega_\gamma h^2 = 2.469 \times 10^{-5}$ is the present time photon energy density.

After WMAP3 data release, there are many works aiming to constrain $N_{\rm eff}$ from cosmological observations [52, 111, 125, 176, 220, 233]. Their results suggest large values for $N_{\rm eff}$ within 2σ interval, some of them not including the standard value $N_{\rm eff} = 3.046$ [176, 220, 233].

Discrepancies between BBN and cosmological data results on $N_{\rm eff}$ have been interpreted as evidence of the fact that further relativistic species are produced by particle decays between the BBN epoch and the structure formation epoch [52, 125]. Other theoretical interpretations include the violation of the spin-statistics in the neutrino sector [71], the possibility of an extra interaction between the DE and radiation or DM, the existence of a Brans–Dicke field, which could mimic the effect of adding extra relativistic energy density between BBN and structure formation [66]. The recent WMAP 5 year CMB measurements [149] constrain the effective number of neutrino species to a value, $N_{\rm eff} = 4.4 \pm 1.5$ (1σ), consistent with the standard value $N_{\rm eff} = 3.046$ (see also [73] for another analysis of the WMAP team and [201, 228] for combined analyses of WMAP with LSS data; note again that differences in the error bars are related to the type of statistical analysis used).

Massive neutrinos with eV mass scale are excellent candidates for contributing to the DM energy density in the Universe. DM particles with large velocity dispersion such that of neutrinos with eV masses are HDM candidates. *A single cosmological bound on the neutrino mass does not exist.* By using the combination of data from WMAP 3 year CMB measurements [118, 186, 233], the 2dFGRS [55] and SDSS [251, 252] power spectra together with the type Ia SNe data [9, 206], an upper limit of sum of the neutrino masses of $m_\nu < 0.66$ eV has been found [233]. An important improvement on this value, $m_\nu < 0.17$ eV, was obtained [220] when the same datasets was supplemented with the BAO data [85] and Lyα forest [178] constraints. There is presently a valid concern about the dependence of the bound on the neutrino mass on the cosmological data sets and physical assumptions used by different analysis [87, 110, 276]. In a recent analysis [149], the WMAP team constraints the sum of neutrino masses by combining the WMAP 5 year CMB measurements with the cosmological distance information obtained from the analysis of BAO and SN data. They find an upper limit for the sum of neutrino masses of $m_\nu < 0.61\,\mathrm{eV}$ (2σ), by considering a equation of state for the DE $w = -1$ and $m_\nu < 0.66\,\mathrm{eV}$ (2σ) for $w \neq -1$.

From WMAP 5 year measurements alone, the WMAP team find $m_\nu < 1.3\,\mathrm{eV}$ (2σ) for $w = -1$ and $m_\nu < 1.5$ eV (2σ) for $w \neq -1$. These bounds on the neutrino masses are free from the uncertainty in the normalization of the large scale structure power spectra.

Since the first studies in which the structure formation process has been investigated considering also massive neutrinos as a DM candidate responsible for gravitational instability, a CDM scenario seems now to explain the main observational aspects of structure formation. In spite of this, neutrinos could leave their imprints on various cosmological observables, such as the large scale structure and the cosmic microwave background. Could you physically discuss the dependence of such features on the neutrino properties?

The standard flat ΛCDM model including three neutrino flavors with degenerated masses can be described with the following cosmological parameters: the energy density parameter associated with the cosmological constant, Ω_Λ, the total nonrelativistic energy density of the CDM particles, Ω_{cdm}, the baryon energy density, Ω_{b}, the neutrino energy density, Ω_ν, the relativistic energy density, Ω_{r}, the primordial spectrum amplitude, A_{s}, its spectral index, n_{s}, and the optical depth to reionization, τ_{es}. The total matter energy density is then given by $\Omega_{\mathrm{m}} = \Omega_{\mathrm{b}} + \Omega_{\mathrm{cdm}} + \Omega_\nu$. Neutrinos became nonrelativistic after at a redshift $z_{\mathrm{rel}} = m_\nu c^2/3k_{\mathrm{B}}T_{\nu,0}$, making the transition from radiation to matter. At early times, when neutrinos are relativistic, they contribute together with the CMB photons to the total radiation energy density, Ω_{r}, with an energy density $\Omega_\nu = 0.68\Omega_\gamma$, where $\Omega_\gamma = 2.3812 \times 10^{-5}\mathrm{h}^{-2}\Theta_{2.7}^4$ ($\Theta_{2.7} = T_{\mathrm{cmb}}/2.7$). At latter times, when neutrinos are nonrelativistic, they act as an extra sub-dominant DM component. Their present energy density contribution expressed in units of the critical energy density is given by $\Omega_\nu h^2 = \sum m_\nu/94\,\mathrm{eV}$. The various physical effects induced by neutrinos properties affects the shape and the amplitude of CMB and LSS power spectra. A later *radiation/matter equality* due to the presence of relativistic neutrinos leads to the decrease of the total matter energy density, Ω_{m}. The net effect of postponing the radiation/matter equality is the suppression of the LSS power spectrum at small scales, reducing the overall matter power spectrum normalization. For the CMB anisotropy power spectrum, the effect is opposite: small scale perturbation are boosted, inducing higher amplitude of the CMB acoustic Doppler peaks. On the other hand, the decrease of Ω_{m} makes the *time of matter/DE equality* to take place earlier. As a consequence, the gravitational potential start to decay earlier, leading to a slightly suppression of the normalization of the matter power spectrum. At the same time, the time variation of the metric perturbations leads to a net reshifting of the CMB photons traveling across the gravitational potential (early ISW effect), shifting the first acoustic Doppler peak position toward lower multipoles. This shift can be compensated by the reduction of the value of the Hubble expansion rate, H_0 [124].

As neutrinos are collisionless particles, they can significantly interact with photons, baryons, and CDM particles only via gravity. The neutrino phase space density is constrained by the Tremaine–Gunn criterion [253] that put limits on the neutrino energy density inside the gravitationally bounded objects. Neutrinos can cluster via gravitational instabilities only on distances below a characteristic distance, the free streaming distance, analogous to the Jeans scale of self-graviting systems. Because the formation of galaxies and clusters is a dynamical time process, the differences introduced in the gravitational potential due to neutrino gravitational clustering generate metric perturbations that affect the evolution of the density fluctuations of all the components of the expanding Universe, leading to a *scale-dependence of the growth rate of perturbations*. Neutrinos with eV masses can cluster gravitationally on small scales at latter times, damping the amplitude of the density perturbations at those scales.

Thank you Lucia.
The process of structure formation is mainly driven by gravitation which, in spite of its intrinsic weakness, has an additive nature responsible for its manifestation at cosmic distances.

On the other hand, it is well known that magnetic forces play a relevant role in many astrophysical contexts, from compact objects to galaxies and galaxy clusters, from the interstellar to the intergalactic medium. Could magnetism be also relevant for cosmology and, if yes, when and how cosmic magnetic fields form? We conclude this chapter with the interview to Kandaswamy Subramanian who highlights this intriguing topic.

3.9 Cosmic Magnetism

Dear Kandaswamy *(Subramanian)***, the origin of magnetic fields found in galaxies and in cluster of galaxies is poorly known. It has been claimed that magnetic fields on cosmological scales have been produced well before the epoch of recombination during the production of first density fluctuations. These fields may also been active during the primordial structure and star formation, affecting it. Do you like to comment on the genesis and evolution of cosmic magnetic fields? How do they manifest themselves?**

The origin of cosmic magnetism is an issue of fundamental importance in astrophysics. We consider some of the ideas of how large scale magnetic fields in the Universe, particularly in galaxies and galaxy clusters could arise. The popular paradigm involves the generation of a seed magnetic field followed by turbulent dynamo amplification of the seed field. The dynamo hypothesis is not without difficulties. It is particularly important to understand the nonlinear saturation of dynamos, and whether the fields produced are coherent enough on large-scales to explain the observed fields in galaxies and clusters. At the same time, the alternative possibility of a primordial field produced in the early Universe lacks firm theoretical support but can have very interesting observational consequences.

3.9.1 The Magnetic Universe

Magnetic fields are crucial for understanding a number of physical processes in the Universe. They are observed in a wide variety of astronomical objects from planets and stars to galaxies and clusters of galaxies. The origin of the largest scale magnetic fields are of great interest to cosmologists. Observations of nearby spiral galaxies show that they host magnetic fields with average total field strengths $B \sim 10\,\mu\text{G}$ with a large-scale component about half this value, coherent on scales of several kpc to tens of kpc, and often correlated with the optical spiral arms [24]. Galaxy clusters also seem to host magnetic fields of several μG, and correlated on several kpc scales [53, 263]. How do such ordered large-scale fields arise in galaxies and clusters? Do they require invoking exotic processes occurring in the early Universe, well before recombination? Or could they arise by purely tapping energy from fluid motions to amplify a small seed field by dynamo action, in more recent epochs after galaxies and clusters form? We consider some of the ideas that have been put forth on how the Universe got magnetized [38, 241].

In galaxies, clusters, and indeed in many other astrophysical settings, the gas is partially or fully ionized and can carry electric currents that, in turn, produce magnetic fields. The evolution of the magnetic field and the plasma treated as a fluid is usually studied in the framework of Magneto-Hydro-Dynamics (MHD), where Maxwell's equations of electrodynamics are combined with the fluid equations, including also the Lorentz forces due to electromagnetic fields. Further, the fluid velocities in galactic and cluster plasma are in general nonrelativistic. In this case, the evolution of the magnetic field $\mathbf{B}$ is governed by the induction equation

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B} - \eta \nabla \times \mathbf{B}) \tag{3.38}$$

got from combining Maxwell's equations (neglecting the displacement current) and a simple form of Ohm's law. Here $\mathbf{U}$ is the fluid velocity and η the microscopic resistivity. The field back reacts on $\mathbf{U}$ via the Lorentz force $\mathbf{J} \times \mathbf{B}$, where $\mathbf{J}$ is the current density. Resistivity would cause the field to decay unless compensated by the inductive $\mathbf{U} \times \mathbf{B}$ term. In astrophysical plasma, the ratio of the inductive and resistive terms measured by the magnetic Reynolds number $R_{\rm m} = (UB)/(\eta B/L) = UL/\eta$ is much larger than unity (here L is the typical length scale over which the field varies). Then the magnetic flux $\Phi = \int_S \mathbf{B} \times \mathrm{d}\mathbf{S}$ through any surface moving with the fluid is almost "frozen" in the sense that $\mathrm{d}\Phi/\mathrm{d}t \to 0$, as $\eta \to 0$. In galaxies, we expect naively $R_\mathrm{m} \sim 10^{18}$ and in clusters $R_\mathrm{m} \sim 10^{29}$ based on the Spitzer resistivity [38]. Plasma instabilities could produce fluctuating magnetic fields, which scatter charge particles and affect various transport processes (cf. [214]). But to affect resistivity and hence R_m, by scattering electrons, these fluctuations need to have significant power down to the electron gyro radius.

At this stage, it is important to clarify the following: it is often mistakenly assumed that if one has a preexisting magnetic field in a highly conducting medium, and the resistive decay time is longer than say a Hubble time, then one does not need any mechanism for maintaining such a field. This is not true in general, because given a tangled field the Lorentz forces would drive motions in the fluid. These would either dissipate due to viscous forces if such forces were important (at small values of the fluid Reynolds number $\mathrm{Re} = UL/\nu$, where ν is the fluid viscosity) or if viscosity is small, drive decaying MHD turbulence, with a cascade of energy to smaller and smaller scales and eventual dissipation on the dynamical timescales associated with the motions. This timescale can be much smaller than the age of the system. For example, if clusters host a few μG magnetic fields tangled on say $l \sim 10\,\mathrm{kpc}$ scales, such fields could typically decay on the Alfvén crossing timescale $l/V_\mathrm{A} \sim 10^8$ year, where we have taken a typical Alfvén velocity $V_\mathrm{A} \sim 100\,\mathrm{km\,s^{-1}}$. Although the energy density in MHD turbulence decays with time as a power law, this time scale is still much shorter than the typical age of a cluster, which is thought to be several billion years. Similarly, if the fluid is already turbulent, the associated larger turbulent resistivity will lead to the decay of large-scale fields. Therefore, one has to provide explicit explanation of the origin and persistence of large-scale cosmic magnetic fields; reference to the low Ohmic resistivity of the plasma is not sufficient if the gas is turbulent or the magnetic field is tangled.

The Universe probably did not start magnetized. Thus the origin of the observed fields can be split into two parts. First, the generation of the first fields from a zero initial field, and then their further amplification and maintenance by fluid motions. The first process is known as a battery effect and the second as a dynamo. There are several astrophysical battery effects that can produce coherent seed magnetic fields. It may also be possible to produce coherent seed magnetic fields in the early Universe. But such fields generally turn out to be much smaller than observed fields in galaxies and clusters and so one needs the motions ($\mathbf{U}$) to act as a dynamo and amplify the field further. It turns out to be then crucial to understand how dynamos work and saturate. Further, as we see later, dynamo theories have several potential difficulties to be overcome. The dynamo itself may be helped if one can kick-start the process with a fairly strong coherent field in the first place. These fields may themselves arise from earlier epochs of dynamo action. It is also of considerable interest to look for their origin in the early Universe; that is to search for early Universe mechanisms, which produce not just seed magnetic fields, but fairly strong fields of order a nano Gauss (nG) as redshifted to the present epoch. Such a field will have other effects on cosmology and structure formation, which could impact significantly on the evolution of the Universe. We discuss several of these issues below.

Astrophysical Batteries

The basic problem any battery has to address is how to produce finite currents from zero currents? Note that $\mathbf{B} = 0$ is perfectly valid solution to the induction equation! So magnetic fields can arise only from a zero field initial condition, if the usual form of Ohm's law is violated by a "battery term". Most astrophysical mechanisms use the fact that positively and negatively charged particles in a charge-neutral Universe do not have identical properties. For example, if one considered a gas of ionized hydrogen, then the electrons have a much smaller mass compared to protons. Thus, for a given pressure gradient of the gas, the electrons tend to be accelerated much more than the ions. This leads in general to an electric field, which couples back positive and negative charges. If such a thermally generated electric field has a curl, then from Faraday's law of induction a magnetic field can grow. The resulting battery effect, known as the Biermann battery, was first proposed as a mechanism for the thermal generation of stellar magnetic fields [31].

The thermally generated electric field is given by $\mathbf{E}_{bier} = -\nabla p_{\rm e}/en_{\rm e}$ got by balancing the forces on the electrons, due to pressure gradient and the electric field and assuming the protons are much more massive than the electrons. The curl of this term leads to an extra source term in the induction equation, which if we adopt $p_{\rm e} = n_{\rm e}k_{\rm B}T$, where $k_{\rm B}$ is the Boltzmann constant, is given by $-(ck_{\rm B}/e)[(\nabla n_{\rm e}/n_{\rm e}) \times (\nabla T)]$. This extra *source term* in the induction equation is nonzero if and only if the density and temperature gradients, $\nabla n_{\rm e}$ and ∇T, are not parallel to each other. In the cosmological context, such nonparallel density and temperature (or pressure) gradients can arise in a number of ways. For example, in cosmic ionization fronts produced when the first ultraviolet photon sources, like starbursting galaxies and quasars, turn on to ionize the Inter Galactic Medium (IGM), the temperature gradient is normal to the front. However, a component to the density gradient can arise in a different direction, as density fluctuations, which will later collapse to form galaxies and clusters, will in general have no correlation to the source of the ionizing photons. The resulting thermally generated electric field has a curl, and magnetic fields correlated on galactic scales can grow. After compression during galaxy formation, they turn out to have a strength $B \sim 3\times10^{-20}$ G [244]. This scenario has in fact been confirmed in detailed numerical simulations of IGM reionization [106]. The Biermann battery can also generate both vorticity and magnetic fields in oblique cosmological shocks, which arise during cosmological structure formation [155].

The asymmetry in the mass of the positive and negative charges can also lead to battery effects during the interaction of radiation with ionized plasma. Note that the Thomson cross section for the scattering of photons with charged particles depend inversely on the mass of the particle. So the electron component of an ionized plasma is more strongly coupled with radiation than the proton component. Suppose one has a rotating fluid element in the presence of a radiation bath. The interaction with photons will brake the velocity of the electron component faster than the proton component and set up a relative drift and hence lead to magnetic field generation [112]. In the modern context, second order effects during recombination could perhaps also lead to both vorticity and magnetic field generation due to the

$\gamma - e/p$ scattering asymmetry. The predicted magnetic fields are again very small, $B \sim 10^{-30}$ G on Mpc scales up to $B \sim 10^{-21}$ G at parsec scales [107, 128, 177].

As can be gleaned from the above examples, battery mechanisms give only very small fields, and one therefore needs some form of dynamo action to explain the observed galactic and cluster fields.

Dynamo Origin of Magnetic Fields

Magnetic fields in a conducting medium can be amplified by the inductive effects associated with the motions of the medium. In this process, generally referred to as a dynamo, the kinetic energy associated with the motions is tapped to amplify magnetic energy. The plasmas in galaxies and clusters are most often turbulent. The dynamo in this context is referred to as a "turbulent dynamo", and its analysis must rely on statistical methods or direct numerical simulations. Turbulent dynamos are conveniently divided into fluctuation (or small-scale) and mean-field (or large-scale) dynamos. The fluctuation dynamo produces magnetic fields that are correlated only on scales of the order of or smaller than the energy-carrying scale of the random motions. On the other hand, large-scale dynamos correspond to those where magnetic fields correlated on scales larger than the forcing scale can grow.

Fluctuation Dynamos in Galaxies and Clusters

The importance of fluctuation dynamos in cosmic objects obtains because they are generic in any random flow where R_m exceeds a modest critical value $R_\mathrm{m,cr} \sim 100$. Fluid particles in such a flow randomly walk away from each other. A magnetic field line frozen into such a fluid will then be extended by the random stretching (if R_m is large enough). Consider a small segment of a thin flux tube of length l and cross-section A, and magnetic field strength B, in a highly conducting fluid. Then, as the fluid moves about, conservation of flux implies BA is constant, and conservation of mass implies $\rho A l$ is constant, where ρ is the local density. So $B \propto \rho l$. For a nearly incompressible fluid, or a flow with small changes in ρ, one will obtain $B \propto l$. Any random shearing motion which increases l will also amplify B; an increase in l leading to a decrease in A (because of incompressibility) and hence an increase in B (due to flux freezing).

Of course, as the scale of individual field structures decreases, (that is since $A \sim 1/B$), as the field strength increases, the rate of Ohmic dissipation increases until it compensates the effect of random stretching. For a single scale random flow, this happens when $v_0/l_0 \sim \eta/l_B^2$, where v_0 is the typical velocity variation on scale l_0 and l_B is the scale of the magnetic field. This gives $l_B = l_\eta \sim l_0/{R_\mathrm{m}}^{1/2}$, the resistive scale l_η for the flow, where $R_\mathrm{m} = v_0 l_0/\eta$. What happens after this can only be addressed by a quantitative calculation. For a random flow, which is delta-correlated in time, it was shown by Kazantsev [142] that magnetic field can grow provided $R_\mathrm{m} > R_\mathrm{m,cr} \sim 30-100$, depending on the form of the velocity correlation function. In the kinematic regime, the field grows exponentially roughly on the eddy turnover time l_0/v_0 and is also predicted to be intermittent, that is, concentrated into structures whose size, in at least one dimension, is as small as the resistive scale l_η (e.g., [38, 209]). In Kolmogorov turbulence, where the velocity variations on a scale l is $v_l \propto l^{1/3}$, the e-folding time is shorter at smaller scales, $l/v_l \propto l^{2/3}$, and so smaller eddies amplify the field faster. If $P_\mathrm{m} = R_\mathrm{m}/\mathrm{Re} \gg 1$, as relevant for galactic and cluster plasma, even the viscous scale eddies can exponentially amplify the field.

The fluctuation dynamo has since been convincingly confirmed beyond the limits of Kazantsev's theory by numerical simulations of forced flows [114, 115, 213], especially when $R_\mathrm{m} \geq \mathrm{Re}$. Such simulations are also able to follow the fluctuation dynamo into the nonlinear regime where the Lorentz forces becomes strong enough to affect the flow as to saturate the growth of magnetic field.

In the context of galaxy clusters, turbulence would mainly be driven by the continuous and ongoing merging of subclusters to form larger and larger mass objects. One typically expects the largest turbulent scales of $l_0 \sim 100\,\mathrm{kpc}$ and turbulent velocity $v_0 \sim 300\,\mathrm{km\,s^{-1}}$, leading to a growth time $\tau_0 \sim l_0/v_0 \sim 3\times10^8$ year; thus for a cluster lifetime of a few Gyr, the fluctuation dynamo would significantly amplify the field (cf. [246]). The seed field for this dynamo probably arises from magnetized outflows from active and starbursting galaxies that the cluster hosts.

And in the case of galactic interstellar turbulence driven by SNe, if we adopt values corresponding to the local Inter Stellar Medium (ISM) of $l_0 \sim 100\,\mathrm{pc}$, $v_0 \sim 10\,\mathrm{km\,s^{-1}}$ one gets $\tau_0 \sim 10^7$ year. Again the fluctuation dynamo would rapidly grow the magnetic field even for very young high redshift protogalaxies.

The major uncertainty, in case we want to use the fluctuation dynamo to explain observed Faraday rotation measures in galaxy clusters or that inferred in some high redshift protogalaxies, is how intermittent the field remains when the dynamo saturates. A simple model exploring an ambipolar drift type nonlinearity [238, 239] suggests that the smallest scale of the magnetic structures will be renormalized in the saturated state to become $l_B \simeq l_0 {R_\mathrm{m,cr}}^{-1/2}$ instead of the resistive scale l_η. In this case, one could indeed obtain significant Faraday rotation measure through such a fluctuation dynamo generated field, as indeed found by directly measuring the rotation measure in numerical simulations [246]. On the other hand, it has been argued that the fluctuation dynamo generated field remains highly intermittent even at saturation, with field reversals typically occurring on the resistive scale l_η [213]. As the cluster $R_\mathrm{m} \sim 10^{29}$, if one uses naively the collisional Spitzer resistivity, one hardly expects to see any Faraday rotation from such a field. Plasma effects would then be important to renormalize the effective R_m, much below this ridiculously large value [214]. One needs a better theoretical understanding of the nonlinear saturation of fluctuation dynamos to make further progress.

Mean Field Dynamos and Galactic Magnetism

A remarkable change in the turbulent dynamo action occurs if the turbulence is helical. This can be clearly seen, for example, in the simulations by Brandenburg [37], where a large scale field on the scale of the box develops when a helical forcing is employed, even though the forcing itself is on a scale about 1/5th the size of the box. The large scale field, however, in these closed box simulations develops only on the long resistive timescales. It is important to understand how such a field develops and how one can generate a large scale field on a faster timescale. The possible importance of helical turbulence for large-scale field generation was proposed by Parker [188], and is in fact discussed in text books [151, 179]. We summarize briefly below the theory of the Mean Field Dynamo (MFD) as applied to magnetic field generation in disk galaxies, turn to several potential problems that have been recently highlighted and their possible resolution.

Suppose the velocity field is split into the sum of a mean, large-scale velocity $\overline{\mathbf{U}}$ and a turbulent, stochastic velocity $\mathbf{u}$. The induction equation becomes a stochastic partial differential equation. Let us split the magnetic field also as $\mathbf{B} = \overline{\mathbf{B}} + \mathbf{b}$, into a mean field $\overline{\mathbf{B}} = \langle \mathbf{B} \rangle$ and a fluctuating component $\mathbf{b}$. Here the average $<>$, is defined either as a spatial average over scales larger than the turbulent eddy scales (but smaller than the system size) or as an ensemble average. Taking the average of the induction (3.38), one gets the MFD equation for $\overline{\mathbf{B}}$,

$$\frac{\partial \overline{\mathbf{B}}}{\partial t} = \nabla \times \left(\overline{\mathbf{U}} \times \overline{\mathbf{B}} + \mathcal{E} - \eta \nabla \times \overline{\mathbf{B}} \right). \tag{3.39}$$

This averaged equation now has a new term, the mean electromotive force $\mathcal{E} = \overline{\mathbf{u} \times \mathbf{b}}$, which crucially depends on the statistical properties of the small-scale velocity and magnetic fields. A central closure problem in MFD theories is to compute the mean electromotive force $\mathcal{E}$ and express it in terms of the mean field itself. In the two-scale approach, one assumes that the mean field is spatially smooth over scales bigger than the turbulence coherence scale l, and expresses the mean electromotive force $\mathcal{E}$ in terms of the mean magnetic field and its first derivative [151, 179]. For isotropic, homogeneous, helical "turbulence" in the approximation that the correlation time τ is short (ideally $u\tau/l \ll 1$, where u is the typical turbulent velocity), one employs what is known as the First Order Smoothing Approximation (FOSA) to write

$$\mathcal{E} = \overline{\mathbf{u} \times \mathbf{b}} = \alpha_K \overline{\mathbf{B}} - \eta_t \nabla \times \overline{\mathbf{B}}. \tag{3.40}$$

Here $\alpha_K = -(\tau/3)\overline{\boldsymbol{\omega} \times \mathbf{u}}$ is the dynamo α-effect, proportional to the kinetic helicity and $\eta_t = (\tau/3)\overline{\mathbf{u}^2}/3$ is the turbulent magnetic diffusivity proportional to the kinetic energy of the turbulence [151, 179].

In the context of disk galaxies, the mean velocity $\overline{\mathbf{U}}$ is that of differential rotation. This leads to the Ω-effect, that of shearing radial fields to produce toroidal fields, while the α-effect is crucial for regeneration of poloidal from toroidal fields. A physical picture of the α-effect is as follows: the ISM is assumed to become turbulent, due to, for example, the effect of SNe randomly going off in different regions. In a rotating, stratified (in density and pressure) medium like a disk galaxy, such turbulence becomes helical. Helical turbulent motions of the gas perpendicular to the disk draws out the toroidal field into a loop, which looks like a *twisted* Ω. Such a twisted loop is connected to a current, which has a component parallel to the original toroidal field. If the motions have a nonzero net helicity, this parallel component of the current adds up coherently. A toroidal current then results from the toroidal field. Hence, poloidal fields can be generated from toroidal ones. (Of course microscopic diffusion is essential to make permanent changes in the field.) This closes the toroidal–poloidal cycle and leads to exponential growth of the mean field. The turbulent diffusion turns out to be also essential for allowing changes in the mean field flux. The kinematic MFD (3.39) gives a mathematical foundation to the above picture. One finds exponentially growing solutions of the MFD equations, provided a dimensionless dynamo number has magnitude $D = |\alpha_0 G h^3 \eta_t^{-2}| > D_{crit} \sim 6$ [209, 226], a condition that can be satisfied in disk galaxies. (Here h is the disk scale height and G the galactic shear, α_0 typical value of α, and we have defined D to be positive.) The mean field grows typically on time-scales a few times the rotation time scales, of order $1\text{–}10 \times 10^8$ year.

This picture of the galactic dynamo faces several potential problems. First, while the mean field dynamo operates to generate the large-scale field, the fluctuation dynamo is producing small-scale fields at a much faster rate. Also the correlation time of the turbulence measured by $u\tau/l$ is not likely to be small in turbulent flows. So the validity of FOSA is questionable. Indeed, based on specific imposed (kinematic) flow patterns, it has been suggested that there is no simple relation between α and helicity of the flow; see [62]. We have recently measured $\mathcal{E}$ directly in numerical simulations of isotropic, homogeneous, helical turbulence [247]. These simulations reach up to a modest $R_\mathrm{m} \sim 220$. We find, somewhat surprisingly that for isotropic homogeneous turbulence the high conductivity results obtained under FOSA and other closures are reasonably accurate up to the moderate values of R_m that we have tested. Interestingly, this agreement is obtained even in the presence of a small-scale dynamo, where $\mathbf{b}$ is growing exponentially to a value much larger than $\overline{\mathbf{B}}$. This suggests that the exponentially growing part of the small-scale field does not make a contribution to the mean electromotive force $\mathcal{E}$, correlated with the mean field $\overline{\mathbf{B}}$. However, by the end of the simulation, the fluctuations in α and η_t grow significantly. This perhaps reflects the fact that when $\mathbf{b}$ and $\mathbf{u}$ are large compared to $\overline{\mathbf{B}}$, fluctuations in the mean electromotive force $\mathcal{E}$ can dominate over the steady contributions. It is essential to extend these results to even higher values of R_m, and also explore more fully mean field dynamo models taking into account such fluctuations. But these preliminary results on the numerically determined α and η_t in isotropic turbulence simulations are quite encouraging.

Another potential problem with the mean field dynamo paradigm is that magnetic helicity conservation puts severe restrictions on the strength of the α-effect [38]. Magnetic helicity measures the linkages and twists in the magnetic field, and in many circumstances can be almost conserved, even when there is significant energy dissipation. The operation of any MFD automatically leads to the growth of linkages between the toroidal and poloidal mean fields and hence a mean field helicity. To satisfy total helicity conservation, this implies that there must be equal and oppositely signed helicity being generated in the fluctuating field. What leads to this helicity transfer between scales? It turns out that it is the turbulent electromotive force $\mathcal{E}$ that transfers helicity between the small and large scale fields. The large scale helicity is in the links of the mean poloidal and toroidal fields, while the small scale helicity is in what can be described as "twist" helicity of the small-scale field. Lorentz forces associated with the "twisted" small-scale field would like to untwist the field. This would lead to an effective magnetic α-effect, which opposes the kinetic α produced by the helical turbulence. The cancelation of the total α-effect leads to a catastrophic quenching of the dynamo. This quenching can be avoided if there is some way of transferring the twists in the small scale field out of the region of dynamo action, that is, if there are helicity fluxes out of the system [32, 243].

Such a helicity flux can result from the anisotropy of the turbulence combined with large-scale velocity shear [262] or the nonuniformity of the α-effect [146]. Another type of helicity flux is simply advection of the small scale field and its associated helicity out of the system [227]. This effect naturally arises in spiral galaxies where some of the gas is heated by SN explosions producing a hot phase that leaves the galactic disk, dragging along the small-scale part of the interstellar magnetic field. Large-scale dynamos may only work if such helicity fluxes are present.

Primordial Magnetic Fields from the Early Universe?

Could magnetohydrodynamics phenomena significantly influence the formation and evolution of cosmic structures, possibly affecting the current paradigm strongly based on gravitation? Do exist open problems of standard CDM model whose solutions could try benefit from progress in this field of astrophysics?

We have so far concentrated on the hypothesis that magnetic fields observed in galaxies and galaxy clusters arise due to dynamo amplification of weak seed fields. An interesting alternative is that the observed large-scale magnetic fields are a relic from the early Universe, arising perhaps during inflation or some other phase transition ([203, 254, 268] and references therein). It is well known that scalar (density or potential) perturbations and gravitational waves (or tensor perturbations) can be generated during inflation. Could magnetic field perturbations also be generated?

Indeed inflation provides several ideal conditions for the generation of a primordial field with large coherence scales [254]. First the rapid expansion in the inflationary era provides the kinematical means to produce fields correlated on very large scales by just the exponential stretching of wave modes. Also vacuum fluctuations of the electromagnetic (or more correctly the hypermagnetic) field can be excited while a mode is within the Hubble radius and these can be transformed to classical fluctuations as it transits outside the Hubble radius. Finally, during inflation any existing charged particle densities are diluted drastically by the expansion, so that the Universe is not a good conductor; thus magnetic flux conservation then does not exclude field generation from a zero field. There is, however, one major difficulty; the standard electromagnetic action is conformally invariant, and the Universe metric is conformally flat. One can then transform the evolution equation for the magnetic field to its flat space version locally for any FLRW Universe, and globally for a Universe with flat spatial sections. The field then decreases with expansion as $1/a^2$, where $a(t)$ is the expansion factor. (Interestingly slower decay can occur for modes comparable to the curvature scale, in open FLRW models due to coupling to the spatial curvature [16].)

Therefore, mechanisms for magnetic field generation need to invoke the breaking of conformal invariance of the electromagnetic action, which change the above behavior to $B \sim 1/a^{\epsilon}$ with typically $\epsilon \ll 1$ for getting a strong field. Such models can involve coupling of the electromagnetic action to the inflaton, the dilaton, the scalar curvature, etc. (cf. [102] and references therein). As $a(t)$ is almost exponentially increasing during slow roll inflation, the predicted field amplitude is exponentially sensitive to any changes in the parameters of the model that affects ϵ. Therefore, models of magnetic field generation can lead to fields as large as $B \sim 10^{-9}$ G (as redshifted to the present epoch) down to fields that are much smaller than that required for even seeding the galactic dynamo. Note that the amplitude of scalar perturbations generated during inflation is also dependent on the parameters of the theory and has to be fixed by hand. But the sensitivity to parameters seems to be stronger for magnetic fields than for scalar perturbations due to the above reason.

Another possibility is magnetic field generation in various phase transitions, like the electroweak transition or the Quantum Cromo Dynamics (QCD) transition due to causal processes. However, these generically lead to a correlation scale of the field smaller than the Hubble radius at that epoch. Hence very tiny fields on galactic scales obtain, unless helicity is also generated; in which case one can have an inverse cascade of energy to larger scales [13, 39, 50]. A number of mechanisms for helicity generation have been suggested involving parity violation during the electroweak phase transition [93, 221, 257].

If a primordial magnetic field with a present-day strength of even $B \sim 10^{-9}$ G and coherent on large scales is generated, it can strongly influence a number of astrophysical processes. An uniform field would, for example, select out a special direction, lead to anisotropic expansion around this direction, hence leading to a quadrupole anisotropy in the CMB. The degree of isotropy of the CMB then implies a limit of several nG on such a field [18]. Comparable limits may obtain, at least for the uniform component, from upper limits to the IGM Faraday rotation of high redshift quasars [33]. For tangled primordial fields, magnetically induced perturbations lead to both large and small angular scale anisotropies in the CMB temperature and polarization. The signals that could be searched for include excess temperature anisotropies (from scalar, vortical and tensor metric and fluid perturbations), B-mode polarization, and non-Gaussian statistics (cf. [103, 104, 163, 172, 223, 242, 271] and [240] for a review). A field at a few nG level and a nearly scale invariant spectrum produces, for example, rotational perturbations, leading to temperature anisotropies at the $5\,\mu$K level, and B-mode polarization anisotropies 10 times smaller [245]. Therefore, primordial magnetic fields of a few nG are potentially detectable via observation of CMB anisotropies.

After recombination, the Universe becomes mostly neutral, with a small but nonzero ionization fraction. This ionized component can still carry currents to sustain the fields. The Lorentz force drives then motions in the ionized component, whose energy can be dissipated into the IGM via ambipolar diffusion and, for small enough scales, by generating decaying MHD turbulence. Such processes significantly modify the thermal and ionization history of the post-recombination Universe. For example, a magnetic field of about 3 nG tangled on small sub Mpc scales leads to a reionization of the Universe, giving rise to Thomson scattering optical depths $\tau \sim 0.1$, although not in the range of redshifts needed to explain the WMAP polarization observations (the T-E cross correlation seen at low l) [224]. Future CMB probes like it Planck can potentially detect the modified CMB anisotropy signal from such partial reionization. This can be used to detect or further constrain small-scale primordial fields.

Potentially more exciting is the possibility that primordial fields induce the formation of subgalactic structures for $z > 10$. The compressional component of the Lorentz force induces inhomogeneities in the baryonic component. These grow due to gravitational instability for masses above the thermal and magnetic Jeans mass. And their gravitational influence also leads to the growth of DM perturbations. For nearly scale invariant spectra, fields as small as 0.1 nG could lead to the collapse of $10^6\,M_\odot$ DM halos, at high redshift $z > 0$, and hence naturally impact on the reionization of the Universe [224]. The enhanced electron abundance due to magnetic field dissipation in the IGM and in collapsing halos also leads to an enhancement of the formation of molecular hydrogen in these first halos, and so an enhanced cooling in these first objects [225]. Thus primordial magnetic fields of nG strength significantly modify structure formation in the Universe.

Furthermore, a 0.1 nG field in the IGM will also be sheared and amplified due to flux freezing, during the collapse to form a galaxy and lead to the few μG field observed in disk galaxies (cf. [154]). Of course, one may still need a dynamo to maintain such a field against decay and/or explain the observed global structure of disk galaxy fields [226]. Weaker primordial fields can also provide a strong seed field for the dynamo. Overall, it is interesting to continue to look for evidence of such a primordial field.

To Summarize

We have presented a brief overview of issues related to the origin of cosmic magnetic fields. Battery mechanisms produce only a small seed field, which needs to be amplified by a dynamo. How fluctuation and MFDs work is at present under intense scrutiny. Basic ideas of turbulent dynamos are in place. But a detailed understanding of the saturation of fluctuation dynamos, and the linear/nonlinear behavior of turbulent transport coefficients associated with MFDs, is still challenging. These dynamo generated fields can also be put back into the galactic and intergalactic medium via outflows induced by SNe and AGN [204], and can serve as stronger seed fields for subsequent small and large-scale dynamos. There is also increasing interest in finding natural mechanisms for primordial field generation in the early Universe, and their observational consequences. Our knowledge of galactic, cluster, and IGM fields has come from observing the synchrotron polarization and Faraday rotation in the radio wavelengths. The observational future is bright with planned instruments like the Square Kilometer Array (SKA) and its precursors [98, 99], which plan to extensively map the Faraday rotation over the whole sky. Indeed, the origin of cosmic magnetism is one of the key science projects of the SKA and one hopes for a rapidly growing understanding of the magnetic Universe.

Thanks a lot Kandaswamy.

References

1. Aalseth, C.E., et al.: Neutrinoless double-β decay of ^{76}Ge: First results from the International Germanium Experiment (IGEX) with six isotopically enriched detectors. Phys. Rev. **C 59**, 2108 (1999)
2. Abazajian, K.N., Beacom, J.F., Bell, N.F.: Stringent constraints on cosmological neutrino-antineutrino asymmetries from synchronized flavor transformation. Phys. Rev. **D 66**, 013008 (2002)
3. Adelman-McCarthy, J.K., et al.: The sixth data release of the sloan digital sky survey. Astrophys. J. Suppl. Ser. **175**, 297–313 (2008)
4. Aguilar-Arevalo, A.A., et al.: Search for electron neutrino appearance at the Δm^2 1 eV^2 scale. Phys. Rev. Lett. **98**, 231801 (2007)
5. Albrecht, A.: Cosmic inflation. Lectures presented at the NATO Advanced Studies Institute, Structure formation in the universe. Cambridge (1999)
6. Albrecht, A., et al.: Causality, randomness, and the microwave background. Phys. Rev. Lett. **76**, 1413 (1996)
7. Ambrosio, M., et al.: Measurement of the atmospheric neutrino-induced upgoing muon flux using MACRO. Phys. Lett. **B 434**, 451 (1998) [hep-ex/9807005]
8. Arnaboldi, C., et al.: New limit on the neutrinoless $\beta\beta$ decay of ^{130}Te. Phys. Rev. Lett. **95**, 142501 (2005) [hep-ex/0501034]
9. Astier, P., et al.: The supernova legacy survey: Measurement of Ω_M, Ω_λ and w from the first year data set. Astron. Astrophys. **447**, 31 (2006) [astro-ph/0510447]
10. Athanassopoulos, C.: Evidence for $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillations from the LSND experiment at the Los Alamos Meson physics facility. Phys. Rev. Lett. **77**, 3082 (1996) [nucl-ex/9605003]
11. Balbi, A., et al.: Constraints on cosmological parameters from MAXIMA-1. Astrophys. J. Lett. **545**, L1–L4 (2000)
12. Banday, A.J., et al.: Reappraising foreground contamination in the COBE-DMR data. Mon. Not. R. Astron. Soc. **345**, 897 (2003)
13. Banerjee, R., Jedamzik, K.: Evolution of cosmic magnetic fields: From the very early Universe, to recombination, to the present. Phys. Rev. **D 70**, 123003-1–123003-25 (2004)
14. Barger, V., et al.: Hiding relativistic degrees of freedom in the early Universe. Phys. Lett. **B 569**, 123–128 (2003)
15. Barkana, R., Loeb, A.: Identifying the reionization redshift from the cosmic star formation rate. Astrophys. J. **539**, 20 (2000)
16. Barrow, J.D., Tsagas, C.: Slow decay of magnetic fields in open Friedmann universes. Phys. Rev. **D 77**, 107302-1–107302-4 (2008)
17. Barrow, J.D., Juszkiewicz, R., Sonoda, D.H.: Universal rotation – How large can it be? Mon. Not. R. Astron. Soc. **213**, 917 (1985)
18. Barrow, J.D., Ferreira, P.G., Silk, J.: Constraints on a primordial magnetic field. Phys. Rev. Lett. **78**, 3610–3613 (1997)
19. Basko, M.M.: The thermalization length of resonance radiation with partial frequency redistribution. Astrofizika **17**, 69 (1981)
20. Baugh, C.M.: A primer on hierarchical galaxy formation: the semi-analytical approach. Report Prog. Phys. **69**, 3101 (2006)
21. Bean, R., Melchiorri, A., Silk, J.: Cosmological constraints in the presence of ionizing and resonance radiation at recombination. Phys. Rev. **D75**, 063505 (2007)

22. Beaudet, G., Goret, P.: Leptonic numbers and the neutron to proton ratio in the hot big bang model. Astron. Astrophys. **49**, 415–419 (1976)
23. Beaudet, G., Yahil, A.: More on big-bang nucleosynthesis with nonzero lepton numbers. Astrophys. J. **218**, 253–262 (1977)
24. Beck, R.: Magnetic fields in normal galaxies. R. Soc. of London Trans. Ser. **A 358**, 777–796 (2000)
25. Bennett, C.L., et al.: First-year Wilkinson microwave anisotropy probe (WMAP) observations: Preliminary maps and basic results. Astrophys. J. Suppl. **148**, 1–27 (2003)
26. Bevis, N., Hindmarsh, M., Kunz, M.: WMAP constraints on inflationary models with global defects. Phys. Rev. **D70**, 043508 (2004) [astro-ph/0403029]
27. Bevis, N., et al.: CMB power spectrum contribution from cosmic strings using field-evolution simulations of the Abelian Higgs model. Phys. Rev. **D 75**, 065015 (2007) [astro-ph/0605018]
28. Bevis, N., et al.: CMB polarization power spectra contributions from a network of cosmic strings. Phys. Rev. **D 76**, 043005 (2007)
29. Bevis, N., et al.: Fitting CMB data with cosmic strings and inflation. Phys. Rev. Lett. **100**, 021301 (2008)
30. Bhattacharjee, P.: Origin and propagation of extremely high energy cosmic rays. Phys. Rev. **327**, 109 (2000)
31. Biermann, L.: Über den Ursprung der Magnetfelder auf Sternen und im interstellaren Raum. Zeitschrift Naturforschung Teil **A 5**, 65–71 (1950)
32. Blackman, E.G., Field, G.B.: Constraints on the magnitude of α in dynamo theory. Astrophys. J. **534**, 984–988 (2000)
33. Blasi, P., Burles, S., Olinto, A.: Cosmological magnetic field limits in an inhomogeneous universe. Astrophys. J. **514**, L79–L82 (1999)
34. Bond, J.R., Efstathiou, G.: Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter. Astrophys. J. Lett. **285**, L45 (1984)
35. Boschan, P., Biltzinger, P.: Distortion of the CMB spectrum by primeval hydrogen recombination. Astron. Astrophys. **336**, 1 (1998)
36. Boughn, S., Crittenden, R.: A correlation between the cosmic microwave background and large-scale structure in the universe. Nature **427**, 45–47 (2004)
37. Brandenburg, A.: The inverse cascade and nonlinear alpha-effect in simulations of isotropic helical hydromagnetic turbulence. Astrophys. J. **550**, 824–840 (2001)
38. Brandenburg, A., Subramanian, K.: Astrophysical magnetic fields and nonlinear dynamo theory. Phys. Report **417**, 1–209 (2005) [astro-ph/0405052]
39. Brandenburg, A., Enqvist, K., Olesen, P.: The effect of Silk damping on primordial magnetic fields. Phys. Rev. **B 392**, 395 (1997)
40. Bucher, M., Moodley, K., Turok, N.: General primordial cosmic perturbation. Phys. Rev. **D 62**, 3508 (2000)
41. Bucher, M., Moodley, K., Turok, N.: Constraining isocurvature perturbations with CMB polarization. Phys. Rev. Lett. **87**, 191301 (2001)
42. Buchmuller, W., Di Bari, P., Plumacher, M.: Some aspects of thermal leptogenesis. New J. Phys. **6** 105 (2004) [hep-ph/0406014]
43. Burigana, C., et al.: CMB polarization constraints on radiative feedback. Mon. Not. R. Astron. Soc. **385**, 404–410 (2008)
44. Burles, S., Nollett, K., Turner, M.S.: Big bang nucleosynthesis predictions for precision cosmology. Astrophys. J. **552** L1 (2001)
45. Challinor, A., Lasenby, A.: Cosmic microwave background anisotropies in the cold dark matter model: A covariant and gauge-invariant approach. Astrophys. J. **513**, 1 (1999)
46. Chen, G., et al.: Looking for cosmological Alfvén waves in Wilkinson microwave anisotropy probe data. Astrophys. J. **611**, 655 (2004)
47. Chiang, L.-Y., et al.: Non-gaussianity of the derived maps from the first-year Wilkinson microwave anisotropy probe data. Astrophys. J. Lett. **590**, L65 (2003)
48. Chiang, L.-Y., Naselsky, P.D., Coles, P.: The robustness of phase mapping as a non-gaussianity test. Astrophys. J. Lett. **602**, L1 (2004)

49. Choudhury, T.R., Ferrara, A.: Updating reionization scenarios after recent data. Mon. Not. R. Astron. Soc. **371**, L55 (2006)
50. Christensson, M., Hindmarsh, M., Brandenburg, A.: Inverse cascade in decaying three-dimensional magnetohydrodynamic turbulence. Phys. Rev **E 64**, 056405-1–056405-6 (2001)
51. Chu, Y.Z., Cirelli, M.: Sterile neutrinos, lepton asymmetries, primordial elements: How much of each? Phys. Rev. **D 74**, 085015 (2006)
52. Cirelli, M., Strumia, A.: Cosmology of neutrinos and extra-light particles after WMAP3. J. Cosmol. Astropart. Phys. **13** (2006) [astro-ph/0607086]
53. Clarke, T.E., Kronberg, P.P., Bohringer, H.: A new radio-X-ray probe of galaxy cluster magnetic fields. Astrophys. J. Lett. **547**, L111–L114 (2001)
54. Cole, S., et al.: A recipe for galaxy formation. Mon. Not. R. Astron. Soc. **271**, 781 (1994)
55. Cole, S., et al.: The 2dF galaxy redshift survey: Power-spectrum analysis of the final data set and cosmological implications. Mon. Not. R. Astron. Soc. **362**, 505 (2005)
56. Coles, P., Chiang, L.-Y.: Characterizing the nonlinear growth of large-scale structure in the universe. Nature **406**, 376 (2000)
57. Coles, P., et al.: Phase correlations in cosmic microwave background temperature maps. Mon. Not. R. Astron. Soc. **350**, 989 (2004)
58. Colley, W.N., Gott, J.R. III.: Genus topology of the cosmic microwave background from WMAP. Mon. Not. R. Astron. Soc. **344**, 686 (2003)
59. Copeland, E.J., et al.: False vacuum inflation with Einstein gravity, Phys. Rev. **D 49**, 6410 (1994)
60. Copeland, E.J., Myers, R., Polchinski, J.: Cosmic F- and D-strings. J. High Energy Phys. **6**, 13 (2004)
61. Copeland, E.J., et al.: On the collision of cosmic superstrings (2007) [arXiv:0712.0808]
62. Courvoisier, A., Hughes, D.W., Tobias, S.M.: α effect in a family of chaotic flows. Phys. Rev. Lett. **96**, 034503-1–034503-4 (2006)
63. Cruz, M., et al.: A cosmic microwave background feature consistent with a cosmic texture. Science **318**, 1612 (2007) [arXiv:0710.5737]
64. Davis, M., et al.: The evolution of large-scale structure in a universe dominated by cold dark matter. Astrophys. J. **292**, 371 (1985)
65. de Bernardis, P., et al.: A flat universe from high-resolution maps of the cosmic microwave background radiation. Nature **404**, 955–959 (2000)
66. de Felice, A., et al.: Relaxing nucleosynthesis constraints on Brans-Dicke theories. Phys. Rev. **D 74** 103005 (2006) [astro-ph/0510359]
67. de Oliveira-Costa, A., et al.: Significance of the largest scale CMB fluctuations in WMAP. Phys. Rev. **D 69**, 063516 (2004a)
68. de Oliveira-Costa, A., et al.: The quest for microwave foreground X. Astrophys. J. Lett. **606**, L89 (2004b)
69. Dineen, P., Rocha, G., Coles, P.: Non-random phases in non-trivial topologies. Mon. Not. R. Astron. Soc. **358**, 1285 (2005)
70. Dolgov, A.D., et al.: Cosmological bounds on neutrino degeneracy improved by flavor oscillations. Nucl. Phys. **B 632**, 363 (2002) [hep-ph/0201287]
71. Dolgov, A.D., Hansen, S.H., Smirnov, A.Y.: Neutrino statistics and big bang nucleosynthesis. J. Cosmol. Astropart. Phys. **6**, 4 (2005) [astro-ph/0611784]
72. Doroshkevich, A.G., et al.: Ionization history of the cosmic plasma in the light of the recent cosmic background imager and future planck data. Astrophys. J. **586** 709 (2003)
73. Dunkley, J., et al.: Five-year Wilkinson microwave anisotropy probe (WMAP) observations: Likelihoods and parameters from WMAP data. Astrophys. J. Suppl. **180**, 306 (2009)
74. Durrer, R.: Gauge-invariant cosmological perturbation theory with seeds. Phys. Rev. **D 42**, 2533 (1990)
75. Durrer, R.: The cosmic microwave background. Cambridge University Press, Cambridge (2008)
76. Durrer, R., Zhou, Z.H.: Large-scale structure formation with global topological defects. Phys. Rev. **D 53**, 5394 (1996) [astro-ph/9508016]

77. Durrer, R., Howard, A., Zhou, Z.H.: Microwave anisotropies from texture-seeded structure formation. Phys. Rev. **D 49**, 681 (1994) [astro-ph/9311040]
78. Durrer, R., et al.: Cosmic microwave background anisotropies from scaling seeds: Fit to observational data. Phys. Rev. Lett. **79**, 5198 (1997) [astro-ph/9706215]
79. Durrer, R., Kunz, M., Melchiorri, A.: Reproducing the observed cosmic microwave background anisotropies with causal scaling seeds. Phys. Rev. **D 63**, 081301 (2001) [astro-ph/0010633]
80. Durrer, R., Kunz, M., Melchiorri, A.: Cosmic structure formation with topological defects. Phys. Rept. **364**, 1 (2002) [astro-ph/0110348]
81. Dvali, G., Tye, S.H.: Brane inflation. Phys. Lett. **B 450**, 72 (1999)
82. Dvali, G., Shafi, Q., Schaefer R.: Large scale structure and supersymmetric inflation without fine tuning. Phys. Rev. Lett. **73**, 1886 (1994)
83. Efstathiou, G.: A maximum likelihood analysis of the low cosmic microwave background multipoles from the Wilkinson microwave anisotropy probe. Mon. Not. R. Astron. Soc. **348**, 885 (2004)
84. Eidelman, S., et al.: Review of particle physics. Phys. Lett. **B 592**, 1 (2004)
85. Eisenstein, D.J., et al.: Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies. Astrophys. J. **633**, 560 (2005)
86. Eitel, K.: Direct neutrino mass experiments. Nucl. Phys. B Suppl. **143**, 197 (2005)
87. Elgaroy, O., Lahav, O.: Upper limits on neutrino masses from the 2dFGRS and WMAP: The role of priors. J. Cosmol. Astropart. Phys. **4**, 4 (2003)
88. Elliott, S.R., Vogel, P.: Double beta decay. Ann. Rev. Nucl. Part. Sci. **52**, 115 (2002) [hep-ph/0202264]
89. Eriksen, H.K., et al.: Asymmetries in the cosmic microwave background anisotropy field. Astrophys. J. **605**, 14 (2004)
90. Esposito, S., et al.: Non equilibrium spectra of degenerate relic neutrinos. Nucl. Phys. **B 590**, 539 (2000) [astro-ph/0005573]
91. Falcone, D., Tramontano, F.: Neutrino oscillations and neutrinoless double beta decay. Phys. Rev. **D 64** 077302 (2001) [hep-ph/0102136]
92. Fan, X., et al.: Constraining the evolution of the ionizing background and the epoch of reionization with $z \sim 6$ quasars. II. A sample of 19 quasars. Astron. J. **132**, 117 (2006)
93. Field, G.B., Carroll, S.M.: Cosmological magnetic fields from primordial helicity. Phys. Rev. **D 62**, 103008 (2000)
94. Fisher, N.I.: Statistical analysis of circular data. Cambridge University Press, Cambridge (1993)
95. Fraisse, A., et al.: Small-angle CMB temperature anisotropies induced by cosmic strings. Phys. Rev. **D 78**, 043535 (2008)
96. Freese, K., Kolb, E.W., Turner, M.S.: Massive, degenerate neutrinos and cosmology. Phys. Rev. **D 27**, 1689 (1983)
97. Fukuda, Y., et al.: Evidence for oscillation of atmospheric neutrinos (Super-Kamiokande Collab.) Phys. Rev. Lett. **81**, 1562 (1998)
98. Gaensler, B.M.: The square kilometre array: A new probe of cosmic magnetism. Astron. Nachr. **327**, 387–394 (2006)
99. Gaensler, B.M., Beck, R., Feretti, L.: The origin and evolution of cosmic magnetism. New Astron. Rev. **48**, 1003–1012 (2004)
100. Gallerani, S., Choudhury, T.R., Ferrara, A.: Constraining the reionization history with QSO absorption spectra. Mon. Not. R. Astron. Soc. **370**, 1401 (2006)
101. Gangui, A., Pogosian, L., Winitzki, S.: Cosmic string signatures in anisotropies of the cosmic microwave background. New Astron. Rev. **46**, 681 (2002) [astro-ph/0112145]
102. Gasperini, M.: Primordial magnetic seeds from string cosmology. Astron. Nachr. **327**, 399 (2006)
103. Giovannini, M.: Tight coupling expansion and fully inhomogeneous magnetic fields. Phys. Rev. **D 74**, 063002 (2006)
104. Giovannini, M., Kunze, K.E.: Magnetized CMB observables: A dedicated numerical approach, Phys. Rev **D 77**, 0630031–0630031-29 (2008)

105. Gnedin, N.Y.: Effect of reionization on structure formation in the universe. Astrophys. J. **542**, 535 (2000)
106. Gnedin, N.Y., Ferrara, A., Zweibel, E.G.: Generation of the primordial magnetic fields during cosmological reionization. Astrophys. J. **539**, 505–516 (2000)
107. Gopal, R., Sethi, S.: Generation of magnetic field in the pre-recombination era. Mon. Not. R. Astron. Soc. **363** 521–528 (2005)
108. Gunn, J.E., Peterson, B.A.: On the density of neutral hydrogen in intergalactic space: Astrophys. J. **142** 1633 (1965)
109. Guth, A.H.: Inflationary universe: A possible solution to the horizon and flatness problems. Phys. Rev. **D 23**, 347–356 (1981)
110. Hannestad, S.: Cosmological limit on the neutrino mass. Phys. Rev. **D 66**, 125011 (2002)
111. Hannestad, S., Raffelt, G.G.: Neutrino masses and cosmic radiation density: Combined analysis. J. Cosmol. Astropart. Phys., JCAP11 (2006) 016 [astro-ph/0607101]
112. Harrison, E.R.: Generation of magnetic fields in the radiation ERA. Mon. Not. R. Astron. Soc. **147**, 279–286 (1970)
113. Harrison, E.R.: Fluctuations at the threshold of classical cosmology. Phys. Rev. **D 1** 2726 (1970)
114. Haugen, N.E.L., Brandenburg, A., Dobler, W.: Is nonhelical hydromagnetic turbulence peaked at small scales? Astrophys. J. Lett. **597**, L141–L144 (2003)
115. Haugen, N.E.L., Brandenburg, A., Dobler, W.: Simulations of nonhelical hydromagnetic turbulence. Phys. Rev. **E 70**, 016308-1–016308-14 (2004)
116. Hawking, S.W.: Black hole explosions? Nature **248**, 30 (1974)
117. Hinshaw, G., et al.: First-year Wilkinson microwave anisotropy probe (WMAP) observations: Data processing methods and systematic error limits. Astrophys. J. **148**, 63 (2003)
118. Hinshaw, G., et al.: Three-year Wilkinson microwave anisotropy probe (WMAP) observations: Temperature analysis. Astrophys. J. Suppl. **170**, 288 (2007) [astro-ph/0603451]
119. Hinshaw, G., et al.: Five-year Wilkinson microwave anisotropy probe (WMAP) observations: Data processing, sky maps, and basic results (2008) [arXiv:0803.0732]
120. Hu, W.T.: Wandering in the background: A cosmic microwave background explorer. PhD thesis, UC Berkeley (1995) [astro-ph/9508126]
121. Hu, W.T.: CMB temperature and polarization anisotropy fundamentals. Ann Phys **303**, 203–225 (2003)
122. Hu, W., Sugiyama, N.: Anisotropies in the cosmic microwave background: An analytic approach. Astrophys. J. **444**, 489 (1995)
123. Hummer, D.G.: Total recombination and energy loss coefficients for hydrogenic ions at low density for $10 \simeq T_e/Z^2 \simeq 10^7$ K. Mon. Not. R. Astron. Soc. **268**, 109 (1994)
124. Ichikawa, K., Fukugita, M., Kawasaki, M.: Constraining neutrino masses by CMB experiments alone. Phys. Rev. **D 71**, 043001 (2005)
125. Ichikawa, K., Kawasaki, M., Takahashi, F.: Constraint on the effective number of neutrino species from the WMAP and SDSS LRG power spectra. J. Cosmol. Astropart. Phys. **5**, 7 (2007) [astro-ph/0611784]
126. Ichikawa, K., et al.: Increasing the effective number of neutrinos with decaying particles. J. Cosmol. Astropart. Phys. **5**, 8 (2007) [hep-ph/0703034]
127. Ichiki, K., Yamaguchi, M., Yokoyama, J.: Lepton asymmetry in the primordial gravitational wave spectrum. Phys. Rev. **D 75**, 084017 (2007) [hep-ph/0611121]
128. Ichiki, K., et al.: Magnetic field spectrum at cosmological recombination. Phys. Rev. **D 78**, 043535 (2008) [astro-ph/0701329]
129. Ivanchuk, A.V., Orlov, A.D., Varshalovic, D.A.: Effects of possible deviations of fundamental physical constants on primordial nucleosynthesis. Pis'ma Astronomicheskii Zhurnal **27**, 615 (2001)
130. Izotov, Y.I., Thuan, T.X.: Systematic effects and a new determination of the primordial abundance of ^4He and dY/dZ from observations of blue compact galaxies. Astrophys. J. **602** 200 (2004) [astro-ph/0310421]

131. Izotov, Y.I., Thuan, T.X., Stasinska, G.: The primordial abundance of ^4He: A self-consistent empirical analysis of systematic effects in a large sample of low-metallicity H II regions. Astrophys. J. **662**, 15 (2007) [astro-ph/0702072]
132. Jaffe, T., et al.: On the viability of bianchi type VII$_h$ models with dark energy. Astrophys. J. **644**, 701 (2006)
133. Jones, B.J.T., Wyse, R.F.G.: The ionisation of the primeval plasma at the time of recombination. Astron. Astrophys. **149**, 144 (1985)
134. Jones, N., Stoica, H., Tye, S.H.: The production, spectrum and evolution of cosmic strings in brane inflation. Phys. Lett. **B 563**, 6 (2003) [hep-th/0303269]
135. Jones, W.C., et al.: A measurement of the angular power spectrum of the CMB temperature anisotropy from the 2003 flight of BOOMERANG. Astrophys. J. **647**, 823 (2006) [astro-ph/0507494]
136. Kaiser, N., Stebbins, A.: Microwave anisotropy due to cosmic strings. Nature **310**, 391 (1984)
137. Kamionkowski, M., Kosovsky, A.: The cosmic microwave background and particle physics. Ann. Rev. Nucl. Part. Sci. **49**, 77 (1999) [astro-ph/9904108]
138. Kang, H.-S., Steigman, G.: Cosmological constraints on neutrino degeneracy. Nucl. Phys. **B 372**, 494 (1992)
139. Kasner, E.: Geometrical theorems on Einstein's cosmological equations. Trans. Amer. Math. Soc., **43**, 217–221 (1921)
140. Katrin Collaboration. KATRIN: A next generation tritium beta decay experiment with sub-eV sensitivity for the electron neutrino mass. (http://www-ik.fzk.de/katrin) (2001) [hep-ex/0109033]
141. Kauffmann, G., White, S.D.M., Guiderdoni, B.: The formation and evolution of galaxies within merging dark matter haloes. Mon. Not. R. Astron. Soc. **264**, 201 (1993)
142. Kazantsev, A.P.: Enhancement of a magnetic field by a conducting fluid. J. Exp. Theor. Phys. **26**, 1031–1034 (1968)
143. Kibble, T.W.B.: Topology of cosmic domains and strings. J. Phys. **A9**, 1387 (1976)
144. Kinney, W.H., et al.: Latest inflation model constraints from cosmic microwave background measurements. Phys. Rev. **D 78**, 087302 (2008) [arXiv:0805.1118]
145. Klapdor-Kleingrothaus, H.V., et al.: Latest results from the Heidelberg–Moscow double beta decay experiment. Eur. Phys. J. **A 12**, 147 (2001) [hep-ph/0103062]
146. Kleeorin, N., et al.: Helicity balance and steady-state strength of the dynamo generated galactic magnetic field. Astron. Astrophys. Lett. **361**, L5–L8 (2000)
147. Kneller, J.P., Steigman, G.: BBN for pedestrians. New J. Phys. **6**, 117 (2004)
148. Komatsu, E., et al.: First-year Wilkinson microwave anisotropy probe (WMAP) observations: Tests of gaussianity. Astrophys. J. Suppl. **148**, 119 (2003)
149. Komatsu, E., et al.: Five-year Wilkinson microwave anisotropy probe (WMAP) observations: Cosmological interpretation (2008) [arXiv:0803.0547v1]
150. Kraus, C., et al.: Final results from phase II of the mainz neutrino mass searchin tritium β decay. The Eur. Phys. J. **C 40**, 447 (2004)
151. Krause, F., Rädler, K.-H.: Mean-field magnetohydrodynamics and dynamo theory. Akademie-Verlag, Berlin (1980)
152. Krolik, J.H.: Further corrections to the theory of cosmological recombination. Astrophys. J. **353**, 21 (1990)
153. Kuiper, N.H.: Koninklijke Nederlandse Akademie Van Wetenschappen, Proc. Ser. A, LXIII, pp. 38–49 (1960)
154. Kulsrud R.M., Zweibel, E.G.: On the origin of cosmic magnetic fields. Report Prog. Phys. Reports Prog. Phys. **71/4**, 046901 (2008) [arXiv:0707.2783]
155. Kulsrud, R.M., et al.: The protogalactic origin for cosmic magnetic fields. Astrophys. J. **480**, 481–491 (1997)
156. Kuzmin, V., Rubakov, V., Shaposhnikov, M.: On anomalous electroweak baryon-number non-conservation in the early universe. Phys. Lett. **B 155**, 36 (1985)
157. La, D., Steinhardt, P.J.: Extended inflationary cosmology. Phys. Rev. Lett. **62**, 376 (1989)
158. Land, K., Magueijo, J.: Examination of evidence for a preferred axis in the cosmic radiation anisotropy. Phys. Rev. Lett. **95**, 071301 (2005)

159. Landau, S., Harari, D., Zaldarriaga, M.: Constraining nonstandard recombination: A worked example. Phys. Rev. **D 63**, 1303 (2001)
160. Lattanzi, M., Ruffini, R., Vereshchagin, G.V.: Joint constraints on the lepton asymmetry of the Universe and neutrino mass from the Wilkinson microwave anisotropy probe. Phys. Rev. **D 72** 063003 [astro-ph/0509079]
161. Lepp, S., P.C. Stancil, A. Dalgarno, A.: Chemistry of the early Universe. Mem. Soc. Astron. It. **69**, 331 (1998)
162. Lesgourgues, J., Pastor, S.: Massive neutrinos and cosmology. Phys. Report **429**, 307 (2006)
163. Lewis, A.: CMB anisotropies from primordial inhomogeneous magnetic fields. Phys. Rev. **D 70**, 043011 (2004)
164. Liddle, A.R., Lyth, D.H.: COBE, gravitational waves, inflation and extended inflation. Phys. Lett. **B 291**, 391 (1992)
165. Liddle, A.R., Lyth, D.H.: The cold dark matter density perturbations. Phys. Rept. **231**, 1 (1993)
166. Linde, A.: Eternal extended inflation and graceful exit from old inflation without Jordan–Brans–Dicke. Phys. Lett. **B 249**, 18 (1990)
167. Linde, A.: Axions in inflationary cosmology. Phys. Lett. **B 259**, 38 (1991)
168. Lobashev, V.M.: The search for the neutrino mass by direct method in the tritium beta-decay and perspectives of studies in the project CATRIN. Nucl. Phys. **A 719**, 153 (2003)
169. Lucchin, F., Matarrese, S.: Power-law inflation. Phys. Rev. **D 32**, 1316 (1985)
170. Luminet, J.-P., et al.: Dodecahedral space topology as an explanation for weak wide-angle temperature correlations in the cosmic microwave background. Nature **425**, 593 (2003)
171. Lyubarski, Y.E., Syunyaev, R.A.: The spectral features in the microwave background spectrum due to energy release in the early universe. Astrophys. Space Sci. **123**, 171 (1983)
172. Mack, A., Kahniashvili, T., Kosowsky, A.: Microwave background signatures of a primordial stochastic magnetic field. Phys. Rev. **D 65**, 123004 (2002)
173. Maltoni, M., et al.: Status of global fits to neutrino oscillations. New J. Phys. **6**, 122 (2004) [hep-ph/0405172]
174. Mangano, G., et al.: A precision calculation of the effective number of cosmological neutrinos. Phys. Lett. **B 534**, 8 (2002) [astro-ph/0111408]
175. Mangano, G., et al.: Relic neutrino decoupling including flavour oscillations. Nucl. Phys. **B 729**, 221 (2005) [hep-ph/0506164]
176. Mangano, G., et al.: Present bounds on the relativistic energy density in the universe from cosmological observables. J. Cosmol. Astropart. Phys. **3**, 6 (2007) [astro-ph/0612150]
177. Matarrese, S., et al.: Large-scale magnetic fields from density perturbations. Phys. Rev. **D 71**, 043502-1–043502-7 (2005)
178. McDonald, P., et al.: The linear theory power spectrum from the Lyα forest in the sloan digital sky survey. Astrophys. J. **635**, 761 (2005)
179. Moffatt, H.K.: Magnetic field generation in electrically conducting fluids. Cambridge University Press, Cambridge (1978)
180. Naselsky, P., Novikov, I.: The primordial baryonic clouds and their contribution to the cosmic microwave background anisotropy and polarization formation. Mon. Not. R. Astron. Soc. **334**, 137 (2002)
181. Naselsky, P.D., Polnarev, A.G.: Anisotropy and polarization of the microwave background radiation as a test of nonequilibrium ionization of the pregalactic plasma. Astrofizika **26**, 543 (1987)
182. Naselsky, P.D., Doroshkevich, A.G., Verkhodanov, O.: Phase cross-correlation of the Wilkinson microwave anisotropy probe internal linear combination map and foregrounds. Astrophys. J. Lett. **599**, L53 (2003)
183. Naselsky, P.D., et al.: Primordial magnetic field and non-gaussianity of the one-year Wilkinson microwave anisotropy probe data. Astrophys. J. **615**, 45 (2004)
184. Olive, K.A., Skillman, E.D.: A realistic determination of the error on the primordial helium abundance: Steps toward nonparametric nebular helium abundances. Astrophys. J. **617**, 29 (2004)

185. Olive, K.A., Steigman, G., Walker, T.P.: Primordial nucleosynthesis: Theory and observations. Phys. Report **333**, 389 (2000) [astro-ph/9905320]
186. Page, L., et al.: Three-year Wilkinson microwave anisotropy probe (WMAP) observations: Polarization analysis. Astrophys. J. Suppl. **170**, 335 (2007) [astro-ph/0603450]
187. Park, C.G., Park, C., Gott, J.R. III: Cleaned 3 year Wilkinson microwave anistropy probe cosmic microwave background map: Magnitude of the quadrupole and alignment of large-scale modes. Astrophys. J. **660**, 959 (2007)
188. Parker, E.N.: Hydromagnetic dynamo models. Astrophys. J. **122**, 293–314 (1955)
189. Peebles, P.J.E.: Recombination of the primeval plasma. Astrophys. J. Lett. **153**, 1–12 (1968)
190. Peebles, P.J.E.: The large-scale structure of the universe. Princeton University Press, Princeton (1980)
191. Peebles, P.J.E.: Origin of the large-scale galaxy peculiar velocity field: A minimal isocurvature model. Nature **327**, 210 (1987)
192. Peebles, P.J.E.: Cosmic background temperature anisotropy in a minimal isocurvature model for galaxy formation. Astrophys. J. **315**, L73 (1987)
193. Peebles, P.J.E.: An isocurvature cold dark matter cosmogony: I. A worked example of evolution through inflation. Astrophys. J. **510**, 523–530 (1999)
194. Peebles, P.J.E.: An isocurvature cold dark matter cosmogony: II. Observational tests. Astrophys. J. **510**, 531–540 (1999)
195. Peebles, P.J.E., Yu, J.T.: Primeval adiabatic perturbation in an expanding universe. Astrophys. J. **162**, 815 (1970)
196. Peebles, P.J.E., Seager, S., Hu, W.: Delayed recombination. Astrophys. J. **539**, L1 (2000)
197. Pequignot, D., Petitjean, P., Boisson, C.: Total and effective radiative recombination coefficients. Astron. Astrophys. **251**, 680 (1991)
198. Percival, W.J., et al.: The 2dF galaxy redshift survey: The power spectrum and the matter content of the Universe. Mon. Not. R. Astron. Soc. **327**, 1297–1306 (2001)
199. Perlmutter, S.J., et al.: Measurements of Ω and Λ from 42 high-redshift supernovae. Astrophys. J. **517**, 565–586 (1999)
200. Pontzen, A., Challinor, A.: Bianchi model CMB polarization and its implications for CMB anomalies. Mon. Not. R. Astron. Soc. **380**, 1387 (2007)
201. Popa, L.A., Vasile, A.: WMAP five-year constraints on lepton asymmetry and radiation energy density: Implications for *Planck*. J. Cosmol. Astropart. Phys. JCAP06 (2008) 028
202. Press, W.H., Schechter, P.: Formation of galaxies and clusters of galaxies by self-similar gravitational condensation. Astrophys. J. **187**, 425–438 (1974)
203. Ratra, B.: Cosmological seed magnetic field from inflation. Astrophys. J. **391**, L1–L4 (1992)
204. Rees, M.J.: The origin and cosmogonic implications of seed magnetic fields. Quart. J. R. Astron. Soc. **28**, 197–206 (1987)
205. Riess, A.G., et al.: Observational evidence from supernovae for an accelerating universe and a cosmological constant. Astron. J. **116**, 1009–1038 (1998)
206. Riess, A.G., et al.: Type Ia supernova discoveries at $z > 1$ from the hubble space telescope: Evidence for past deceleration and constraints on dark energy evolution. Astrophys. J. **607**, 665–687 (2004)
207. Ruffini, R., Song, D.J., Stella, L.: On the statistical distribution off massive fermions and bosons in a Friedmann universe. Astron. Astrophys. **125**, 265 (1983)
208. Ruffini, R., Song, D.J., Taraglio, S.: The "ino"-mass and the cellular large-scale structure of the Universe. Astron. Astrophys. **190**, 1 (1988)
209. Ruzmaikin, A.A., Sokoloff, D.D., Shukurov, A.M.: Magnetic fields of galaxies. Kluwer, Dordrecht (1988)
210. Rybicki, G.B., Dell'Antonio, I.P.: The time development of a resonance line in the expanding universe. Astrophys. J. **427**, 603 (1994)
211. Sarangi, S., Tye, S.-H.: Cosmic string production towards the end of brane inflation. Phys. Lett. **B 536**, 185 (2002) [hep-th/0204074]
212. Sarkar, S.: Big bang nucleosynthesis and physics beyond the standard model. Rept. Prog. Phys **59** 1493 (1996) [hep-ph/9602260]

213. Schekochihin, A.A., et al.: Simulations of the small-scale turbulent dynamo. Astrophys. J. **612**, 276–307 (2004)
214. Schekochihin, A.A., et al.: Plasma instabilities and magnetic field growth in clusters of galaxies. Astrophys. J. **629**, 139–142 (2005)
215. Schmidt, M.: Large redshifts of five quasi-stellar sources. Astrophys. J. **141**, 1295 (1965)
216. Schneider, R., et al.: Detectable signatures of cosmic radiative feedback. Mon. Not. R. Astron. Soc. **384**, 1525 (2008) [astro-ph/0712.0538]
217. Seager, S., Sasselov, D.D., Scott, D.: A new calculation of the recombination epoch. Astrophys. J. Lett. **523**, L1–L5 (1999)
218. Seager, S., Sasselov, D.D., Scott, D.: How exactly did the Universe become neutral? Astrophys. J. Suppl. **128**, 407 (2000)
219. Seljak, U. et al.: Cosmological parameter analysis including SDSS Lyα forest and galaxy bias: Constraints on the primordial spectrum of fluctuations, neutrino mass, and dark energy. Phys. Rev. **D 71**, 103515 (2005)
220. Seljak, U., Solsar, A., McDonald, P.: Cosmological parameters from combining the Lyman-alpha forest with CMB, galaxy clustering and SN constraints. J. Cosmol. Astropart. Phys., JCAP10(2006)014
221. Semikoz, V.B., Sokoloff, D.D.: Large-scale magnetic field generation by α effect driven by collective neutrino-plasma interaction. Phys. Rev. Lett. **92**, 131301 (2004)
222. Serpico, P.D., Raffelt, G.G.: Lepton asymmetry and primordial nucleosynthesis in the era of precision cosmology. Phys. Rev. **D 71**, 127301 (2005) [astro-ph/0506162]
223. Seshadri, T.R., Subramanian, K.: Cosmic microwave background polarization signals from tangled magnetic fields. Phys. Rev. Lett. **87**, 101301-1–101301-4 (2001)
224. Sethi, S., Subramanian, K.: Primordial magnetic fields in the post-recombination era and early reionization. Mon. Not. R. Astron. Soc. **356**, 778–788 (2005)
225. Sethi, S., Nath, B., Subramanian, K.: Primordial magnetic fields and formation of molecular hydrogen. Mon. Not. R. Astron. Soc. **387**, 1589–1596 (2008)
226. Shukurov, A.: Introduction to galactic dynamos. In: Mathematical aspects of natural dynamos. Dormy, E., Desjardins, B. (eds.) EDP Press (2004) [astro-ph/0411739]
227. Shukurov, A., et al.: Galactic dynamo and helicity losses through fountain flow. Astron. Astrophys. Lett. **448**, L33–L36 (2006)
228. Simha, V., Steigman, G.: Constraining the early-Universe baryon density and expansion rate. J. Cosmol. Astropart. Phys., JCAP06(2008)016
229. Smith, C.J., et al.: Light element signatures of sterile neutrinos and cosmological lepton numbers. Phys. Rev. **D 74**, 085008 (2006) [astro-ph/0608377]
230. Smith, K.M., Zahn, O., Doré, O.: Detection of gravitational lensing in the cosmic microwave background. Phys. Rev. **D 76**, 043510 (2007)
231. Smoot, G.F., et al.: Structure in the COBE differential microwave radiometer first-year maps. Astrophys. J. Lett. **396**, L1–L5 (1992)
232. Spergel, D.N., et al.: First-year Wilkinson microwave anisotropy probe (WMAP) observations: Determination of cosmological parameters. Astrophys. J. Suppl. **148**, 175–194 (2003)
233. Spergel, D.N., et al.: Three-year Wilkinson microwave anisotropy probe (WMAP) observations: Implications for cosmology. Astrophys. J. Suppl. **170**, 377 (2007)
234. Springel, V., Hernquist, L.: The history of star formation in a Λ cold dark matter universe. Mon. Not. R. Astron. Soc. **339**, 312 (2003)
235. Springel, V., Frenk, C.S., White, S.D.M.: The large-scale structure of the Universe. Nature **440**, 1137–1144 (2006)
236. Steigman, G.: Primordial nucleosynthesis in the precision cosmology era. Ann. Rev. Nucl. Part. Sci. **57**, 463 (2007)
237. Strumia, A., Visani, F.: Implications of neutrino data circa 2005. Nucl. Phys. **B 726**, 294 (2005) [hep-ph/0503246]
238. Subramanian, K.: Unified treatment of small- and large-scale dynamos in helical turbulence. Phys. Rev. Lett. **83**, 2957–2960 (1999)
239. Subramanian, K.: Hyperdiffusion in nonlinear large- and small-scale turbulent dynamos. Phys. Rev. Lett. **90**, 245003 (2003)

240. Subramanian, K.: Primordial magnetic fields and CMB anisotropies. Astron. Nachr. **327**, 403–409 (2006)
241. Subramanian, K.: Magnetizing the universe. In: from planets to dark energy: The modern radio universe. Beswick R., et al. (eds.) published by PoS [arXiv:0802.2804]
242. Subramanian, K., Barrow, J.D.: Microwave background signals from tangled magnetic fields. Phys. Rev. Lett. **81**, 3575–3578 (1998)
243. Subramanian, K., Brandenburg, A.: Magnetic helicity density and its flux in weakly inhomogeneous turbulence. Astrophys. J. Lett. **648**, L71–L74 (2006) [astro-ph/0509392v1]
244. Subramanian, K., Narasimha, D., Chitre, S.M.: Thermal generation of cosmological seed magnetic fields in ionization fronts. Mon. Not. R. Astron. Soc. Lett. **271**, L15–L18 (1994)
245. Subramanian, K., Seshadri, T.R., Barrow, J.D.: Small-scale cosmic microwave background polarization anisotropies due to tangled primordial magnetic fields. Mon. Not. R. Astron. Soc. **344**, L31–L35 (2003)
246. Subramanian, K., Shukurov, A., Haugen, N.E.L.: Evolving turbulence and magnetic fields in galaxy clusters. Mon. Not. R. Astron. Soc. **366**, 1437–1454 (2006)
247. Sur, S., Brandenburg, A., Subramanian, K.: Kinematic α-effect in isotropic turbulence simulations. Mon. Not. R. Astron. Soc. **385**, L15–L18 (2008)
248. Switzer, E.R., Hirata, C.M.: Primordial helium recombination III: Thomson scattering, isotope shifts, and cumulative results. (2007) [astro-ph/0702145]
249. Taylor, A.C., et al.: Clover - A B-mode polarization experiment. New Astron. Rev. **50**, 993 (2006) [astro-ph/0610716]
250. Tegmark, M., Silk, J., Blanchard, A.: On the inevitability of reionization: Implications for cosmic microwave background fluctuations. Astrophys. J. **420**, 484 (1994)
251. Tegmark, M., et al.: Cosmological parameters from SDSS and WMAP. Phys. Rev. **D 69**, 103501 (2004)
252. Tegmark, M., et al.: The three-dimensional power spectrum of galaxies from the sloan digital sky survey. Astrophys. J. **606** 702 (2004) [astro-ph/0310725]
253. Tremaine, S., Gunn, J.E.: Dynamical role of light neutral leptons in cosmology. Phys. Rev. Lett. **42**, 407 (1979)
254. Turner, M., Widrow, L.M.: Inflation-produced, large-scale magnetic fields. Phys. Rev. **D 37**, 2743–2754 (1988)
255. Turok, N.: A causal source which mimics inflation. Phys. Rev. Lett. **77**, 4138 (1996) [astro-ph/9607109]
256. Turok, N., Spergel, D.N.: Global texture and the microwave background. Phys. Rev. Lett. **64**, 2736 (1990)
257. Vachaspati, T.: Estimate of the primordial magnetic field helicity. Phys. Rev. Lett. **87**, 251302 (2001)
258. Varshalovich, D.A., Ivanchuk, A.V., Potehin, A.Yu.: Do the fundamental physical constants have the same values in different regions of space-time? Phys. J. Exp. Theor. Phys. **144** 1001 (1999)
259. Verner, D.A., Ferland, G.J.: Atomic data for astrophysics. I. Radiative recombination rates for H-like, He-like, Li-like, and Na-like ions over a broad range of temperature. Astrophys. J. Suppl. **103**, 467 (1996)
260. Vielva, P., et al.: Detection of non-Gaussianity in the Wilkinson microwave anisotropy probe first-year data using spherical wavelets. Astrophys. J. **609**, 22 (2004)
261. Vilenkin, A., Shellard, E.P.S.: Cosmic strings and other topological defects. Cambridge University Press, Cambridge (1994)
262. Vishniac, E.T., Cho, J.: Magnetic helicity conservation and astrophysical dynamos. Astrophys. J. **550**, 752–760 (2001)
263. Vogt, C., Enßlin, T.A.: A Bayesian view on Faraday rotation maps seeing the magnetic power spectra in galaxy clusters. Astron. Astrophys. **434**, 67–76 (2005)
264. Wagoner, R.V., Fowler, W.A., Hoyle, F.: On the synthesis of elements at very high temperatures. Astrophys. J. **148**, 3 (1967)
265. Watts, P.I.R., Coles, P.: Statistical cosmology with quadratic density fields. Mon. Not. R. Astron. Soc. **338**, 806 (2003)

266. Weinberg, S.: The cosmological constant problem. Rev. Mod. Phys **61**, 1–23 (1989)
267. White, S.D.M., Rees, M.J.: Core condensation in heavy halos – A two-stage theory for galaxy formation and clustering. Mon. Not. R. Astron. Soc. **183**, 341 (1978)
268. Widrow, L.M.: Origin of galactic and extragalactic magnetic fields. Rev. Mod. Phys. **74**, 775–823 (2002)
269. Wong, Y.Y.Y.: Analytical treatment of neutrino asymmetry equilibration from flavour oscillations in the early universe. Phys. Rev. **D 66** 025015 (2002) [hep-ph/0203180]
270. Yahil, A., Beaudet, G.: Big-Bang nucleosynthesis with nonzero lepton numbers. Astroph. J. **206**, 26–29 (1976)
271. Yamazaki, D.G., et al.: Effects of a primordial magnetic field on low and high multipoles of the cosmic microwave background. Phys. Rev. **D 77**, 043005 (2008)
272. Zabotin, N.A., Naselsky, P.D.: The neutrino background in the early universe and temperature fluctuations in the cosmic microwave radiation. Sov. Astron. **26**, 272 (1982)
273. Zel'dovich, Ya.B.: A hypothesis unifying the structure and the entropy of the Universe. Mon. Not. R. Astron. Soc. **160**, 1 (1972)
274. Zel'dovich, Ya.B., Novikov I.D.: The hypothesis of cores retarded during expansion and the hot cosmological model. Astronomicheskii Zhurnal **43**, 758 (1966)
275. Zel'dovich, Ya.B., Kurt, V.G., Sunyaev, R.A.: Recombination of hydrogen in the hot model of the Universe. JETF **28**, 146 (1969)
276. Zunckel, C., Ferreira, P.G.: Conservative estimates of the mass of the neutrino from cosmology. J. Cosmol. Astropart. Phys., JCAP08 (2007) 004

Chapter 4
From Galileo to Modern Cosmology: Alternative Paradigms and Science Boundary Conditions

Contributions by Carlo Burigana, Salvatore Capozziello, Cesare Chiosi, Mauro D'Onofrio, Malcolm Longair, Philip Mannheim, Paola Marziani, Moti Milgrom, Keith Olive, Thanu Padmanabhan, John Peacock, Francesca Perrotta, Luisa Pigatto, Rafael Rebolo, Luigi Secco, Jack W. Sulentic, Gerard t'Hooft, and Simon D.M. White

4.1 Outline of the Chapter

This chapter develops along two main lines. On the one hand, it is dedicated to those aspects of fundamental physics in tight relationship with modern cosmology: gravitation and dynamical theories, cosmological constant, and dark matter–energy content, early Universe phases, and the problem of the fundamental constants of physics. On the other hand, we have chosen to include a discussion of the various influences affecting modern astrophysics and cosmology. In fact, the development and the solutions to the fundamental problems of physics cannot be disjoined from the social conditions in which scientists operate.

We start with a tribute to Galileo, based on an interview of historic character with Luisa Pigatto (Sect. 4.2), revisiting his times and his relationship with the cultural environment of Venice, in honor of the memory of the man who contributed so much to the development of the "scientific method", and that, with his astronomical observations, prompted the first radical revolution of cosmological physics: the transition from the Ptolemaic paradigm to the Copernican view.

Such a beginning has at least two motivations. First, the influences of the Venetian cultural environment on Galileo are in general poorly known. Second, the man and the life of Galileo are here at the center of our curiosity for a particular aspect of science that sometimes could appear as taken for granted. How much does the structure and the sociological condition of our societies affect the development of modern science? Do we really live in a free society in which scientists can work without interferences? Could we have, possibly in different forms, another Galileo case today? Is the "scientific method" really used by modern scientists?

Before attempting an answer, and starting from Galileo, it was natural to first ask about the legacy of this scientist for modern cosmology. Malcom Longair was so kind in Sect. 4.3 to replay to this question linking his personal view on the Concordance Model to the basic tests of General Relativity (GR) (Sect. 4.4), the Einstein's theory that represents the basis of our modern interpretation of gravity. Of course, this leads to the problem of the cosmological constant, which Malcom has reviewed from a historical perspective in Sect. 4.5.1.

M. D'Onofrio and C. Burigana (eds.), *Questions of Modern Cosmology: Galileo's Legacy*, DOI 10.1007/978-3-642-00792-7_4,

The cosmological constant problem is indeed today the problem *of theoretical physics and cosmology. We have interviewed Thanu Padmanabhan in Sect. 4.5.2 with the aim of understanding if this unsolved problem will trigger, in the near future, a new shift of paradigm, or, in other words, a new revolution in physics with consequences of scientific relevance similar to those of the Copernican era. It is interesting in this respect to note that a scientific revolution may be prompted both by new observations (recall the discoveries of Galileo, such as Jupiter's satellites), and also by things that we are unable (at present) to see (e.g., the Dark Matter (DM) and Dark Energy (DE)). Thanu also presents his personal point of view about the possible alternatives to the cosmological constant in Sect. 4.7.*

As the problem of the cosmological constant is closely linked to that of DE, we decided to ask Francesca Perrotta to review the physical basis of cosmological scenarios involving DE and to discuss the possible candidates for such energy (see Sects. 4.6 and *4.6.1). She also comments on the question of possible analogies between the inflationary framework and the DE phenomenon in Sect. 4.6.2.*

On the basis of all these considerations, it emerges that gravity is still poorly understood. Despite the efforts of the last 30 years, its connection with the other fundamental forces of nature has not been realized yet. The DM and DE problems may simply be a consequence of our ignorance about gravity. Here, we will not enter specifically into such a big debate, but we will explore the alternative schemes of gravity that in recent years have claimed to solve both the DM and DE problems. Moti Milgrom reviews for us the pros and cons of his Modified Newtonian Dynamics (MOND) theory in Sect. 4.7.1, Salvatore Capozziello the advantages of $f(R)$ *theories in Sect. 4.7.2, and Philip Mannheim the properties of the conformal theory in Sect. 4.7.5.*

Gravity is also a problem for theoretical physics: the unification of the Quantum Mechanics with GR, the two big theories of the last century, is far from being completed. The Nobel Laureate for physics Gerard t'Hooft was very kind to replay us about the state-of-the-art of this process of unification of the fundamental forces. He also tells us what were the conditions of the very early Universe (see Sect. 4.8) and comments on the claimed crisis of theoretical physics that, according to Lee Smolin's book "The Trouble with Physics", is affecting the field, with enormous consequences for the new generations of astronomers and physicists.

A symptom of the aforementioned crisis of cosmology and theoretical physics is, according to several scientists, the so-called Anthropic solution or Anthropic landscape. *The idea of the string theory landscape was a concrete implementation of the Anthropic Principle, which claims that the fundamental constants of nature may have the values they have not for reasons of fundamental physics, but rather because their values are necessary for life. In the light of the recent paradigm change from Cold Dark Matter (CDM) to* Λ*CDM it is remarkable that in 1987 Steven Weinberg proposed that the observed value of the cosmological constant was so small because it is not possible for life to occur in a Universe with a much larger cosmological constant. We therefore asked Luigi Secco to review the predictive power of the Anthropic Principle in Sect. 4.10, after having addressed in this context the problem of the constants of physics, a fundamental aspect previously discussed from a more* *general point of view by Keith Olive in Sect. 4.9. Are they really constants and why they have the values they do?*

The "Why Now?" problem is an embarrassing question too, closely connected with the cosmological constant and the anthropic landscape solution. John Peacock will review for us these in Sect. 4.11 arriving at the concept of many universes.

We then move the discussion on the possible occurrence of a new revolution or shift of paradigm in modern physics. This is explicitly addressed in Sect. 4.12 and the related subsections, which are intended to delineate the role and the influences of our society in the development of science. Cesare Chiosi, Gerard t'Hooft, Paola Marziani, Thanu Padmanabhan, John Peacock, Jack Sulentic, and Simon White have expressed their opinions, giving a panorama of the different points of view. We will see from their discussion how many constraints operate today, limiting the freedom of scientists. Such conditioning has different faces, being sometime of economical character and sometime a more subtle form of segregation of the "heretics". In any case we believe that an ample reflection on these problems will be necessary in the astronomical community.

Finally, as an example of what we mean by speaking of direct influences of society on modern science, we present the case of Spain. Rafael Rebolo will remind us the big development of astrophysics in Spain along the last 20 years in Sect. 4.13.

Let's start with the tribute to Galileo.

4.2 Remembering Galileo

Dear Luisa *(Pigatto)***, it is well known that Galileo worked in Padova, but in a tight relationship with the Venice cultural environment. Could you please discuss the influence of such context on Galileo's activity and ideas?**

It is quite difficult to write something new on Galileo, considering the thousand of pages that have already been written about all the aspects of his life and scientific activity [30, 69, 75, 76, 78–80, 94]. Nonetheless, I will try to answer your questions, dear Mauro and Carlo, by taking into account different facts and documents that have already been analyzed by historians of the Serenissima and historians of science, albeit in a completely separate way. I hope that by doing so I may emphasize details by simply associating different pieces of information, particularly on the period concerning the invention of Galileo's telescope, and the place where the idea to make it was born, that is, the town of Venice.

Galileo first visited the splendid city of Venice on September 1592[1]. He went there to introduce himself to the Venetian patricians who retained power over the

[1] Information about Galileo's stay in the Venetian region can be found in Favaro's books [75, 78]. See also [53] and [173]. From now on, *Le Opere di Galileo Galilei* [76] will be cited as *Opere*, followed by a roman digit for the number of the volume, and a bold digit for the number of the document. An almost complete edition of *Le Opere* is available online (see Manuzio project in web page list). The original texts of quotations here translated into English are available following online edition (number of documents and pages).

appointment of chairs at the University of Padova. The chair *ad Mathematicam* had been vacant since 1588, when Giuseppe Moletti, greatly appreciated from Venetians, had passed away[2].

Galileo spent more than 20 days in Venice, except for a few trips to Padova, to meet Gianvincenzo Pinelli, a very illustrious man of his times and his fond supporter, who could recommended him to Venetian patricians, many of whom visited his house considered as "the refuge of Muses and the academy of all virtues at the times" [135, p. 19]. During his stay in Venice, Galileo was hosted by Giovanni Uguccioni, the Tuscan Resident in Venice, at "S.ta Iustina, in ca' Gradenigo"[3]. We can reasonably argue that the latter had introduced Galileo to

> Andrea Morosini's mezzanine, which had become very frequented by all those interested in literature, not only members of the nobility all of whom became great senators of the Serenissima Republic [...]; but also all sort of virtuous people, both secular and religious. More precisely, any man of letters happening to be in Venice, from Italy and other regions, would never have missed the opportunity of visiting this house, as it was one of the best known places consecrated to Muses [135, p. 18].

During his 18 years at the University of Padova, Galileo often visited this place, which "at time, 25, even 30 very virtuous people attended. [...] Everyone could talk about everything they liked, without any fear of changing subjects, provided they were always new, and debates aimed at knowing the truth" [135, p. 19].

Galileo was soon well known in this stimulating cultural environment – clearly described by Father Fulgenzio Micanzio – for his Tuscan eloquence and witty, intelligent remarks perhaps enlivened by extraordinary musical performances with lute he wonderfully played. As a child, Galileo had been instructed in music and had been taught to play this instrument by his father Vincenzio, a well known sixteenth century musician.

It was presumably at that time, on Morosini's mezzanine, that Galileo met the Servite friar, Paolo Sarpi, with whom he would be bound by great friendship and mutual esteem, and his circle of friends, who liked literary as well as scientific debates. And probably it was Father Paolo who took Galileo to Bernardo Sechini's shop at the "Golden Ship", "because a bunch of honorable, virtuous and polite men went to the "Golden Ship" in "Merceria" to exchange news. [...] Many foreign merchants also happened on that place who had travelled not only throughout Europe but also to the East and West Indies" [135, p. 19]. There, everyone could hear news about trade, wars, peace treaties, inventions, custom and usages of new peoples, etc., from anywhere in the world, so that in Venice, everyone knew everything. Father Paolo used to go there to hear interesting news that he then carefully arranged in his mind, endowed as he was with a prodigious memory, even "monstrous" memory, as his friend Father Fulgenzio said.

In those years, before constructing the telescope and revealing to the world his revolutionary celestial discoveries, Galileo frequently visited the Arsenal of Venice,

---

[2] Giuseppe Moletti was entrusted with the chair *ad Mathematicam* from 1577 to 1588 [77].

[3] Letter from Gio. Vincenzo Pinelli to Galileo, Padova 3 September 1592, *Opere*, X, **36** and letter from Giovanni Uguccioni to Grand Duke of Tuscany, Venice 26 September 1592, ivi, **40**.

which would have been so important in his scientific research. He had learned about its existence also from a prestigious literary source – Dante's *Divine Comedy* – at the time he had made a study about the geometrical structure of Hell [87][4].

He had dealt with the technical and practical problems faced at the Arsenal just 3 months after taking up his teaching position at the University of Padova. The *Provveditore all'Arsenale* (one of the Commissioners of the Arsenal), Giacomo Contarini, through the mediation of Gianvincenzo Pinelli, submitted him a question about the optimal position of oars in a vessel to obtain the most powerful thrust. With his scholarly analysis of March 1593, the young professor demonstrated the authoritative Commissioner his unquestionable scientific knowledge and clear ideas. At the same time, the scholarly reply that Contarini gave Galileo, as the latter had requested, demonstrated how theory and practice could differ in producing the required results ([75, II, p. 70, pp. 122–126], and *Opere*, X, **47** and **48**).

This was undoubtedly an important lesson that prompted Galileo to visit the Arsenal where he learnt so much during the 18 years he spent on duty for the Republic of Venice. The Arsenal was not very far from the house of Galileo's closest friend, Giovanfrancesco Sagredo – the famous Salviati's interlocutor in the *Dialogo* and *Discorsi* – and who occasionally hosted Galileo. The two friends presumably visit the Arsenal together, both amused and surprised at seeing as those masters' experience could produce admirable solutions, which were not immediately explicable in theoretical terms, but were confirmed by the experience handed down for centuries. It comes as no surprise, therefore, that the Arsenal, which was one of the largest dockyards in Europe at the time, and had provided Venice with its maritime power over the centuries, was celebrated by Galileo at the beginning of his last great scientific work *Discorsi e dimostrazioni matematiche intorno a due nuove scienze*. Beyond the practical problems faced in the book (see for instance [189]), we realize, from this opening paragraph, Galileo's high recognition of the lesson he had learnt there, where he could directly test the scientific method of the "experience of senses and certain demonstration" as he often recalled in his works, the very same method to apply to study nature in all its aspects:

> *Salviati* It seems to me, Venetian Lords, that the customary visit of the famous Arsenal of yours, particularly to that area where mechanics is practiced, provides speculative minds with a large amount of subjects for debate, given that there, all kind of instruments and machines are continually produced by a large number of craftsmen. Among them, there must certainly be artisans who have become very skilled and intelligent with experience they learned from their predecessors and the one they continually make themselves.
>
> *Sagredo* Your Lordship is not mistaken at all: I, too, curious by nature, visit that place for pleasure and to associate with the people whom we call *Proti*, for a certain predominance they have on the rest of the masters, and talking with them many times has helped me investigate the reason for effects which are not only magnificent, but also recondite and

[4] The famous Dante's verses say: "As in the Arsenal of the Venetians, during the winter the sticky pitch is boiling up to scrape and tar their ruined ships that cannot sail; on its place, some make a new boat, some bung up the ribs of vessels that sailed for many voyages; some hammer at the prow, some at the stern, some others make oars some others wind the shrouds, some others mend jib and mainsail", Dante, *Inferno*, XXI, 7–15.

> unimaginable. Of course, I was sometimes confused, and despaired of ever being able to understand events that was so beyond my mind and which, nonetheless, the experience proved to be true. [91, p. 7]

The relationship with Paolo Sarpi and his circle of friends, all of whom were interested in scientific news, was equally important to Galileo. Their correspondence shows that their debated subjects concerned magnetism, motion, mechanics[5]. On 1600, when *De Magnete* by Gilbert was published [95], besides debating about magnetism our friends made experiments in Venice upon the declination of magnetic "short and long, copper pointed needles both suspended and floating on water"[6], and at home in Padova, Galileo reinforced magnets he sold or gave friends as a present [75, I, pp. 237–243]. He was greatly requested by this small group of experimenters, both for his brilliant thinking and explanations.

If on one side Galileo acknowledged that his visits to the Venetian Arsenal gave him deeper insight into mechanics and taught him important lessons, on the other, he was reticent in describing the events that brought him to construct the telescope. This reticence gave rise to a series of wrong suppositions for which there is not any written proof, and for which Galileo was misunderstood and accused of taking credit for what he should have shared with others. We can get more reliable information from documented facts from which we can infer that the idea of making a telescope was conceived within the small circle of Paolo Sarpi's friends in Venice; we can also assume that the mentioned reticence was desired, perhaps necessary, if we bear in mind the events, well known to historians, which center around the Servite friar. Here is a short account of these events to better understand what I have just said.

In 1605, the Venetian Senate had passed two laws which forbade to build churches and monasteries without permission of public authority, and citizens of Venetian state had to be granted permission to bequeath goods to the Church. In the same period, two clergymen guilty of heavy common crimes had been arrested by Venetian authorities. Pope Paolo V considered both the new laws and the two arrests as violations of ecclesiastic law, and demanded that the Republic should repeal the decrees and hand the two clergymen over to ecclesiastic authorities, or else he would place Venice under an interdict. The Venetian Senate considered Paolo V's threat as interfering with the power of the Prince, and rejected his request. In April 1606, the Pope fulminated the interdict: the Doge and the Senate were excommunicated, and the clergy were prevented from celebrating religious services across the Venetian Territory. The Republic denounced the papal decree as illegitimate, and ordered its priests to carry out their ministry, otherwise they could leave the Venetian state, not without a previous permission from authorities. Instead, the Jesuits were expelled from Venice because of their ambiguous position on this situation. In this circumstance, the Senate appointed Paolo Sarpi as an expert in theology and Canon law, and at the time also as consultor *in iure* to the Republic. The quarrel with Pope Paolo

---

[5] Letter from Paolo Sarpi to Galileo Galilei, Venice 2 September 1602 and Venice 9 October 1604, in *Opere*, X, **83** and **104**; letter from Galileo to Sarpi, Padova 16 October 1604, ivi, **105**.

[6] Letter from Paolo Sarpi to Jacques Leschassier, Venice 3 February 1610 [200, p. 68].

V was settled by France in 1607: the Venetian Senate retained the two laws, and the two clergymen were handed over to papal authorities through the Frenchmen[7].

Paolo Sarpi's contribution was indeed very precious and absolutely relevant. Since then, the persecution against the Servite by Roman Curia never ended, taking on the characteristics of a spy story: Father Paolo became wanted, living or dead, better if dead. In the evening of 5 October 1607, as he was walking back to his Monastery, he was attacked by five hired assassins. They gave him more than 15 strokes with the dagger and three stabs hit him: two on the neck, one on the face. The dagger penetrated on the side of the right ear and exit near the nose. The assassins thought they had killed him, so they escaped and found refuge in the house of the Pope's nuncio. From here they fled by boat towards the papal territory. The Venetian Senate took urgent measures to protect the life of his consultor. Gerolamo Fabrici d'Acquapendente, the famous professor of anatomy at Padova University, was immediately summoned to Venice and ordered to stay permanently in the Monastery until the friar was recovered from the wounds. After that, Father Paolo had to change his habits: he had to take a gondola from Monastery to Rialto bridge, walk only across the crowded "Merceria" to Mark square and into the Ducal Palace, where he was to perform his duties. Sarpi spent the rest of the time in his Monastery, and here friends went to see him. He sent his numerous letters to important persons in France and other places through diplomatic courier, used a cryptographic system and after reading letters he received from his correspondents he destroyed them[8]. He did all this to prevent Roman spies from intercepting his letters and protect his friends [223, pp. xxxviii–lxxi]. A man in Paris did his best to slander Father Paolo, by accusing him of heresy as well as all the people who wrote to him: this person was the papal Nuncio, Maffeo Barberini [223, p. clxxxv], [135, p. 50], who became Pope Urban VIII in 1623, the Pope who later became a central figure in Galileo's trial and condemnation.

All these facts might explain why Galileo's name never appears in Sarpi's letters, as well as cast different light on the role played by Giacomo Badover, ex Galileo's pupil in Padova, immortalized in *Sidereus Nuncius* as one of the persons who had informed his teacher about some glasses by which "very distant objects became visible", and prompted him to make his first telescope. Already in March 1609, Badover had written to Father Paolo about "glasses from Hollanda"[9] while his letter to Galileo was probably sent the following July[10]. Why should not Sarpi have had to mention to his friend the news about glasses before Badover? I think he did, but Badover had not to know it, because he was a suspect person. He was born a Calvinist, but when he returned to France from Padova he had been converted to Catholicism by a Jesuit. Father Paolo considered Jesuits the great enemies of true

[7] Publications about the history of Paolo V's interdict are numerous and in different languages. The last in order of time is [201].

[8] Letter from Paolo Sarpi to Jérôme de l'Isle Groslot, Venice 14 September 1610 [199, ii, p. 115].

[9] Letter from Sarpi to Badover, Venice 30 March 1609, in [200, pp. 179–180]. Badover's letter to Sarpi is lost.

[10] In the *Sidereus Nuncius* [88], Galileo dates back Badover's lost letter to summer 1609.

faith, instigators, and accomplices of all the abuses the Roman Church committed because of its excessive power. There were good reasons for suspecting Badover, as appears from a "notice of the Nuncio in Paris, sent to Rome on July 1609", conserved in the secret Vatican Archive, which reports the contents of Sarpi's letter to Badover [223, p. CXCV]. Father Paolo could not know this notice, but a year later he was confirmed on his suspicion by a letter Martin Hasdale had written to Galileo from Prague, begging his former teacher to warn Father Paolo because somebody in Paris showed his letters[11].

I think we can now better understand the background of the small circle of Sarpi's friends. Galileo certainly knew all this, and was careful to avoid getting himself and his Venetian friends into trouble. From Florence, he thus wrote to Father Paolo through a common friend, Sebastiano Venier, Senator of the Republic[12].

Let us come to telescope. Before Badover's letter to Sarpi, information about the new device had arrived in Venice. On 6 January 1609 Father Paolo wrote: "I was informed of the news about 'glasses' more than a month ago, and I am so confident of this that I do not need to investigate the matter any further, or to philosophise about it, as Socrates forbids us to philosophise about events we have not experienced ourselves"[13]. We can assume that news had spread from the "Golden Ship" shop in "Merceria". The owner's son, Alvise Sechini, had studied in Leuven, and could translate what travellers and merchants from Flanders said. The young intelligent, educated man had made friends with Father Paolo, and translated news for him when needed [135, p. 19]. Our consultor could not image that within few months the Senate would have asked him for an opinion on that instrument offered on sale to Senate for one thousand sequins:

> As regards 'glasses' which were invented in Netherlands and are known as Galileo's 'glasses' in Italy, they were fully understood by Galileo when the Serenissima Seigniory were offered them for one thousand sequins; the Father was entrusted with the task of verifying what purpose they served, and give his opinion. As he was forbidden to disassemble them to see how they were made, he could image what they were. He told this to Mr. Galileo, who confirmed that Father Paolo had got it right [135, p. 51].

The attempt at selling the instrument occurred the summer 1609, but Sarpi expressed a negative opinion: "In Italy [...] those 'glasses' to see faraway things were issued; I admire them for the beauty of invention and respectability of art, but I consider them worthless in war, both at land and sea"[14].

Before going on, we should ask ourselves why the Venetian Senate asked his consultor *in iure* for an opinion. This is because senators knew about Sarpi's skill in scientific matters, especially in optics. Nine years earlier, Gerolamo Fabrici d'Acquapendente had mentioned Sarpi as the friend who had explained to him that the pupil both of cats and men narrows or widens in response to the brightness or dimness of the surrounding light: "This mystery was observed and explained to me

[11] Letter from Martin Hasdale to Galileo, Prague 4 June 1610, in *Opere*, X, **324**.

[12] Letter from Galileo to Paolo Sarpi, Venice 12 February 1611, in *Opere*, XI, **476**.

[13] Letter from Sarpi to Jérôme de l'Isle Groslot, Venice 6 January 1609, in [199, I, p. 181].

[14] Letter from Sarpi to Rossi, Venice 21 July 1609, in Sarpi [199, I, p. 279].

by Rev. Father Master Paolo, of the Order of friar Servants, a theologian and eminent philosopher, great scholar of mathematical disciplines, particularly optics, that I name here in order to honor him" [74, p. 93].

In conclusion, our Father had to manipulate the instrument for sale without looking inside it, although he certainly saw that the two lenses at each end of the tube were one convex and one concave. As Father Fulgenzio writes, Sarpi rushed to inform Galileo of this, urging him to construct a similar one. This circumstance is confirmed by Galileo himself, as years later, in *Il Saggiatore*, he wrote that after returning to Padova from Venice, he immediately made "the instrument, I gave account of it in Venice to the very same friends with whom I had discussed this subject the previous day". And after working very hard to improve the instrument, once he had obtained the good results we know, "I was advised by an affectionate patron of mine – as Galileo recalls – and presented it to the Prince before the whole College" [89, pp. 35–36]. Once again, I can infer that this advice came from Father Paolo who, about 1 month earlier, had been so negative about the instrument being sold by a foreigner. Unlike the latter, Galileo's telescope worked, and Galileo was right to present it almost immediately to Doge of Venice[15]. Unlike what is commonly believed, debates among friends on the new device certainly focused on the rules of refraction that Sarpi well knew, and although at the time the subject was known on a purely qualitative basis, that was enough to make them understand, for example, that the radius of curvature of the convex objective had to be much larger than that of concave eyepiece. Father Paolo himself provided these technical details in a letter to French jurisconsult Jacques Leschassier:

> As you know, this instruments is made up of two lenses (that you call "lunettes"), both shaped like spheres, one with a convex surface, and the other with a concave one. We have obtained the convex lens from a sphere, six feet in diameter, and the concave one from less than a digit in diameter[16]. These lenses form an instrument about 4-ft long, and enable us to see part of an object at an angle of 3 degrees, which the naked eye sees at an angle of 6 min[17].

These details clearly describes Galileo's telescope with 30 magnifications, which enabled him to make those extraordinary celestial discoveries we shall celebrate in 2009, the international year of astronomy. Sarpi explains magnification as the ratio between the visual angle under which the human eye sees the image and the object, respectively, which is also a modern definition of magnification[18].

---

[15] These facts are also reported in a letter from Giovanni Bartoli, Secretary of the Tuscan Resident in Venice, to Belisario Vinta, Venice 29 August 1609, in *Opere*, X, **233**.

[16] One Venetian foot is 34.7735 cm, one digit a tenth of a foot.

[17] Letter from Sarpi to Jacques Leschassier, Venezia 16 marzo 1610, in *Opere*, X, **272** and in [200, pp. 73–74].

[18] The magnification of an optical instrument can be defined as: $N = \tan\varphi_i / \tan\varphi_o$, φ_i and φ_o being the visual angles from which the rays, coming from the edges of the image and of the object respectively, converge into the optical center of the eye. The angular values by Sarpi, give a magnification of about 30, the same given by Galileo in the *Sidereus Nuncius* adopting the ratio of plane areas. In fact 900 magnifications, given there, correspond to the ratio of an area of a 30-side square respect to the 1-side one.

In this same letter, Father Paolo does not mention Galileo at all, and simply gives information about sky observations performed by means of the described instrument by "a mathematician of ours from Padova and some other our fellows experts in this art". He neither mentions the title nor the author of the booklet (the *Sidereus Nuncius* just off the press), that he was sending to his French correspondent, and which described the extraordinary discoveries "on the Moon, on the star of Jupiter, in the constellations of fixed stars". Therefore, Sarpi's friends were interlocutors, advisers, and privileged witnesses of the construction of Galileo's telescope and the subsequent unbelievable astronomical discoveries that followed. The place where Galileo repeatedly polished and adapted the lenses, made thousands observations of terrestrial objects to verify that instrument was not showing illusory things, where, at night, he continually observed the Moon, the stars, and Jupiter's satellites, was undoubtedly his house in Padova, its south windows opening on to the garden where he grew vines, so that a large portion of sky was visible. But the very place where he wished to exhibit the improvements of the telescope and talk about the new discoveries "with a friend able to understand and appreciate them" [90, p. 151] had to be the Monastery of friar Servants, given the circumstances mentioned above[19]. This fact is confirmed in a letter Galileo wrote to Father Paolo just 1 year later: "I think to remember that I spoke to you about the observation of Saturn last August"[20], and also by Father Fulgenzio many years later:

> I clearly remember that when Your Lordship made his first 'glasses' here, one of the things you observed were sunspots. I could indicate the precise spot where you using the 'glasses' showed them on a pale blue paper to the Father, of glorious memory. I also remember we talked about these sunspots, initially thinking they were a mistake derived from the instrument, or vapors in the atmosphere. Then, once we repeated the experience, we came to conclusion that the fact could be true to such an extent as to be reasoned out: and then you left[21].

Galileo spent in Padova 18 years. After his departure, what has been his legacy in the Venetian scientific community?

When Galileo left Padova for Florence in September 1610, did he leave behind him a fruitful scientific debate, mainly due to the revelation of a "new" sky that was now visible to the human eye, and which nobody had never seen before? Absolutely not! My statement is supported, for instance, by the famous letter Galileo wrote to Kepler in 1610[22]. The discovery of the four satellites of Jupiter had met with great scepticism, and Galileo was looking anxiously for witnesses to confirm the existence of these celestial bodies. Therefore, Galileo wrote to Kepler that, while the Grand Duke

---

[19] Letter from Sarpi to Jacques Leschassier, Venezia 16 marzo 1610, in [200, pp. 73–74] and in *Opere*, X, **272**.

[20] Letter from Galileo to Sarpi, Florence 12 February 1611, in *Opere*, XI, **476**.

[21] Letter from Micanzio to Galileo in Florence, Venice 7 September 1631, in *Opere*, XIV, **2210**. It seems to me very reliable thinking that Galileo had observed the Sun since the winter 1609, when foggy sky gives the opportunity to look at it naked eye, and successively had used a paper over which to project the solar image, as stated by Father Fulgenzio.

[22] Letter from Galileo to Kepler, Padova 19 August 1610, in *Opere*, X, **379**.

of Tuscany thanked Galileo just for discovering the four errant stars – the Medicean planets – that he had seen himself through the telescope, and had given to him a gift and a permanent position in Florence as a Mathematician and Philosopher, instead the Republic of Venice had confirmed his chair *ad Mathematicam* at the University of Padova for life with substantial pay rise, despite the new planets were illusory ("etiam illudentibus planetis"). The Venetians did not give him these prestigious acknowledgments for the astronomical discoveries, but for having provided them with an instrument for war. In spite of the alluring offer, Galileo wrote to Kepler: "I'm leaving anyway, and I'm going there, where I may endure the punishment for my dishonorable and miserly deceit". Galileo seemed offended that, despite his word, nobody believed in the existence of Jupiter's four satellites, although anyone could see (and had seen) them through his instrument. And what "about the primary philosophers of this University who, filled with snake obstinacy, never wanted to look at Planets, Moon, and telescope despite I became available to show them more than thousand times?" he wrote. And 1 year later, in recommending a successor to his vacant chair *ad mathematicam* he wrote to Sarpi and Sebastiano Venier that the new professor should be a person

> suitable to defend the dignity and excellence of such a noble profession against those who try to exterminate it; this kind of persons can be found in Padova, as you know very well. And I know that these people will arrange for such a man to be appointed only to dominate and frighten him, so that, if anything true and nice should ever be discovered, it will be suffocated by their tyranny[23].

However, after the Papal interdict, the political climate in Venice was changing. Sagredo and Sarpi, Galileo's main supporters, were engaged in other matters. On August 1608, Sagredo left Venice as Venetian ambassador to Siria, Sarpi, between late 1610 and early 1611, in addition to his duty as Consultor, was entrusted with the task "only him [...] of revising, examining, rearranging documents of the Archive concerning the dominion of this Adriatic Sea and boundaries with the Ecclesiastic State". This was a cumbersome task, for that Father Paolo was "very busy"[24].

History was taking its course: Father Paolo, in Venice, was completely absorbed by his duties and by political issues that particularly concerned him [53]; Galileo, in Florence and in Rome, was trying his best to have the Copernican system approved by the Catholic Church [207]. However, this does not mean that Sarpi did not worry about Galileo's vicissitudes, and his fatal attempts at making the Copernican truth accepted by the Roman Curia, dominated from untrustworthy Jesuits. In 1611 Sarpi had written in a note:

> Now, that the most Illustrious and Eminent Mr. Domenico Molino has informed me that Mr. Galileo Galilei is going to Rome, where he was invited to show his celestial discoveries, I fear that, if in these circumstances he tries to explain the intelligent reasons for which he prefer the theory of Canon Copernicus for our Solar system, the Jesuits and other friars will

[23] Letter from Galileo to Paolo Sarpi, Florence 12 February 1611, in *Opere*, XI, **476**.

[24] Letter from Domenico Molino to Jacques Leschassier, Venice 28 February 1611 [200, pp. 249–251].

> certainly disapprove of him. I imagine that the physic and astronomical issues will turn into theological ones, so that – I foresee, with my great trouble – to live in peace without being accused of heresy and having to face excommunication, he will be compelled to retract his beliefs about this subject. A day will come, I'm almost sure, when men enlightened by better studies will deplore the disgrace of Galileo and the injustice dealt so great man; in the mean time he has to face it and should not complain about it but in secret [99, II, pp. 70–71][25].

Although in 1611 Galileo's astronomical discoveries had been enthusiastically accepted by the Jesuits and the Roman Court, Father Paolo's predictions were soon confirmed. On 1616, the consultor himself was "compelled" to give his written opinion in favor of publishing in the Venetian territory the Holy Office decree forbidding, among others, the book *De Revolutionibus orbium caelestium* by Copernicus. Some Sarpi's considerations as exhibited to the Senate seem to have been taken, almost to the word, from a letter Galileo had written a year earlier to monsignor Dini, when news about a coming shortly prohibition of Copernicus's book and accusation that it followed a theory contrary to the Holy Scriptures, were circulating. These similarities, only partially reproduced below, once again confirm the close intellectual relationship between the two friends, the great esteem and affection with which Father Paolo followed Galileo's adventures on his Roman trips, and how he fully shared his view of the new Universe, displayed in the Venetian skies, thank also to his contribution:

> Nicolas Copernicus was not only a catholic, but also a religious and a canon; he was summoned in Rome at the time of Leone XIII, during the Lateranensis Council, when emendation of ecclesiastic calendar was debating, depending on him as a greatest astronomer [...] of whom doctrine every one followed, and finally the calendar was regulated in accordance with it. His hard work on the motion and constitution of celestial bodies was contained in six books that [...] he printed dedicating them to Pope Paolo III; since then, they have been sold publicly without any scruple [Galileo][26].

> Nicolas Copernicus has been a Catholic priest and public professor at the University of Rome. He had known Pope Paolo III, of holy memory, since he was cardinal and later when he was elected pope. His book was already printed about 100 years ago, was sold and read from all over the Europe, considering his author to be the most learned in astronomy in the world. Moreover, the correction of the year made by Pope Gregorio XIII is founded on his doctrine [Sarpi] [53, pp. 220–221].

The information is the same, including the wrong assumption that the Jesuit Christopher Clavius, who greatly appreciated the *Prutenic Tables* derived from Copernicus's astronomical tables, had actually used them for reforming the calendar. On the contrary, Clavius had used the mean value derived from the ancient *Alfonsine Tables* for the length of the tropic year of the Sun and the epacts criterion for the motion of the Moon ([46, p. 39 and pp. 96–97]; see also [65]).

---

[25] This note, reproduced also in [176, pp. XIII–XIV], is followed by a short but interesting comment about Galileo's and Sarpi's friendship and their figures as intellectuals at the time of Counter-Reformation.

[26] Letter from Galileo to Piero Dini, Florence 16 February 1615, in *Opere*, V, XII (X, **1081**).

Going back shortly to Father Paolo's written opinion, the greatly declaimed Venetian freedom was sacrificed to superior politic interests (peace with Pope and Roman Curia), and moreover – wrote Sarpi – "being here very few people who practice astronomy, there is no reason to worry about a scandal". And monsignor Antonio Querengo, whose house in Padova used to be open to anybody who loved literature included Galileo, exhibited himself with a very ungenerous letter that somewhat ridiculed his esteemed friend:

> Mr. Galileo's debates have dissolved like an alchemic smoke, since the Holy Office has stated that by supporting this opinion one openly dissents from infallible dogmas of the Church. Now we can finally rest assured that, although our brains are freewheeling, we may keep still in our place, without flying with the Earth like ants on a ball soaring into the air[27].

The ideas, works, and life of Galileo have been objects of a wide literature. In your opinion, do exist aspects and ideas that still deserve better historical investigations or possibly even unknown?

As I mentioned before, although I believe that the historians of science know and have debated almost all scientific aspects of Galileo's work and life, it seems to me that the inquisition trial and condemnation about "the immobility of Sun and motion of Earth" affected studies on galilean vicissitudes more than necessary. This gave Galileo's adversaries the opportunity of trying to conceal his true merits and the value of his scientific discoveries. Nonsense devoid of any scientific meaning has be written and published, perhaps willingly ignoring those details that reveal the scientific significance of these discoveries[28]. Consequently, except scholars, very few know about the revolution brought about by Galileo when he turned his telescope to the night sky opening the eyes of mankind to a new Universe. Both the public and school students know about Galileo's trial for his belief of the Earth's motion as an urban legend, and ignore or only vaguely know about his scientific discoveries, their meaning, and the greatness of a man who had the extraordinary courage of making a new vision of the sky accepted by a world of sceptic, suspicious, envious men.

Then, a more detailed analysis of Galileo's writings has to be given to make better known the scientific value of his astronomical observations and avoid falling into absurd disputes such as that of priority in discovering Jupiter's satellites that I mention here as an example.

In 1614, the German Simon Mayr published a booklet, *Mundus Jovialis*, where he alleged he had discovered (or seen) the four satellites of Jupiter before Galileo. The absurdity of mentioning this fact in a printed book 4 years after the *Sidereus Nuncius*, thus denying the value of written and documented communication, is immediately evident. Galileo realized that the four little stars near Jupiter were planets revolving around it, because he carefully observed them during all the clear nights, more than once each night, drawing their relative and very precise positions in a

[27] Letter from Antonio Querengo to Alessandro d'Este, Rome 5 March 1616, in *Opere*, XII, **1186**.

[28] I wish to remember here the important essay by Guglielmo Righini, *Contributo alla interpretazione scientifica dell'opera astronomica di Galileo* [192].

paper, writing down the time of his observations, marking their appearance and disappearance behind the planet, etc[29]. This was a very scientific method of observation, and this is how Galileo came to the true discovery. Quite unlike Mayr who – if he had actually observed the four small stars on January 1610 as he claimed – provided neither drawings nor times of observation. Actually, the circumstances described by German astronomer demonstrate that it was scientifically impossible for him to understand that the four stars were satellites revolving around Jupiter[30].

Which is your opinion concerning the cosmological view of Galileo beyond the Copernican model of the Solar system?

Galileo's profound belief in the truthfulness of the Copernican system drove him to facing the derision of his enemies and presumed friends, toilsome trips to Rome, a trial and consequent house arrest for life, all of which have been thoroughly discussed. As regards the constitution of the Universe as a whole, Galileo did not make a priori hypothesis. He was aware that nature, of which the sky was a part, was an open book written in mathematical language to read and learn from, and that the Creator, endowed with infinite wisdom, might have created a world with unimaginable characteristics. Many experiences were required to reason about the immensity of the Universe, its center – here or there – the unbelievable number of stars at different distance from the Earth, not fixed in a sole sphere, all things for which Galileo had immediately understood the potentiality of the telescope: "Other more extraordinary things will perhaps be discovered in the future, by me or others, by means of this instrument". [88, p. 6] But should have these discoveries never come to an end? Should have men never arrived to know the structure of the Universe completely, with absolute certainty? I think that Galileo's opinion about this complex problem is masterly expressed in a passage, not frequently quoted indeed, known as the "Fairy Tale of Sounds". This is why, dear Mauro and Carlo, I wish to answer your last question by quoting this passage from *Il Saggiatore* where Galileo was debating the constitution of comets with the Jesuit Orazio Grassi. The fairy tale of sounds quoted here, despite the limits of its translation, which prevent us from appreciating the formal beauty of the Italian language used by Galileo, is a clear, lucid, and extraordinary illustration of how the human mind proceeds gradually acquiring and widening its knowledge, in a process of apparently endless accumulation of "reasoned experiences".

The "Fairy Tale of Sounds"

A man, whom nature had endowed with keen intelligence and extraordinary curiosity, was born in a secluded place. To amuse himself, he bred several birds, whose song he greatly enjoyed. He marvelled at their ability to produce, at will, very different, sweet sounds with

[29] From 7 January to 2 March 1610, for the 45 clear nights, Galileo made 61 observations of Jupiter's satellites.

[30] About Simon Mayr's pretension see [75, I, pp. 340–348] and [69, pp. 220–221]. An English translation from Latin of *Mundus Jovialis* is available in The Observatory, **36**, 367–381, 403–412, 443–452, 498–503 (1916).

the very same air they used to breathe. One night, he heard a soft sound near his house and, thinking it could only be a small bird, went out to catch it. Once in the street, he found a small shepherd boy who produced sounds similar to the song of birds, albeit in a quite different way. The boy blew into a perforated piece of wood along which his fingers moved, closing and opening the holes on it. So surprised and naturally curious was the man, that he gave the shepherd a calf in exchange for his pipe. After collecting his thoughts and realizing that had the boy not come by he would never have known that in nature there actually were two different ways of creating sweet sounds, he decided to leave his house to experience other adventures. The following day, as he was walking by a hovel, he heard a similarly soft sound coming from within. He went inside to check if the sound came from a pipe or a blackbird, and found a child who moved a bow with his right hand as if was sawing the strings stretched over a concave piece of wood, while his left hand held the instrument, his fingers plucking the strings. The boy was capable of producing different, very sweet sounds without using any air. Knowing how intelligent and curious this man was, one can imagine how, after the initial surprise at discovering these two new, very unexpected ways of emitting sounds, the man convinced himself there might be others in nature. And imagine his bewilderment when, going into a church, he had to look behind the door to see who had produced that sound, and realized that it had come from hinges and small sheets of metal, in opening the door! Yet another time, attracted from curiosity, he went into a tavern, and believing to see someone with a bow touching lightly the chords of a violin, instead he saw a person who produced a very sweet sound by rubbing his fingertip on the rim of a glass. But when he realized that, unlike his birds, who emitted interrupted sounds by breathing, wasps, mosquitoes and big flies created perpetual sounds by beating their wings very quickly, the more his amazement increased the more his opinion decreased on his knowledge on how sound is produced. Not even all the experiences already acquired could have sufficed to make him understand or believe that crickets, without even flying, are able to emit so sweet and sonorous whistles not by breathing but shaking their wings. Thus, the man was almost convinced that there were no other possible ways to create sounds other than ones described above, and having he examined many organs, trumpets, pipes, strings instruments of every kinds, including the iron reed which, kept between someone's teeth, strangely uses the mouth cavity as a sound box and breath as a sound means; when, I say, he thought he had seen every thing, he was once again completely baffled and amazed when a cicada came in his hand and he couldn't lower its high shrieking sound neither by shutting its mouth nor by fastening its wings. Besides, he was not able to see neither scales nor other moving parts; but then as he lift thorax of the cicada, he saw some thin, albeit stiff cartilage underneath and, thinking that the noise was produced by its shaking, he decided to break it to silence the insect. But it was all in vain. And pushing his needle further inside and thus transfixing it, he took away with the voice its life, and the man had yet been incapable of verifying whether the sound actually came from that cartilage. Consequently, he became so insecure about his knowledge that, when he was asked how the sounds were produced, he generously answered he knew some ways, although he was confident that hundreds of unknown and unthinkable others existed.

I could provide many other examples to demonstrate how nature can produce its effects in a wealth of ways we cannot even envisage, had sense and experience not proved them to us, and to be true, experience may sometimes to be insufficient to make up for our inability. This is why, I shall be excused if I cannot explain the formation of this comet precisely, especially bearing in mind that I never claimed I could do so, as I knew that it might be formed in a ways beyond our imagination. Thus, our inability to understand how a cicada produces its sound, while it is singing in our hand, provides sufficient justification for me not to know how this comet is formed at so great distance [89, pp. 51–52].

Thank you very much Luisa for this very nice review of Galileo's time and relationships in Venice. Now, coming back to strictly scientific arguments, we will *explore the aspects of the legacy of Galileo for modern cosmology by investigating the observational tests that support the standard cosmological scenario. In the next interview Malcom Longair will be asked to discuss such tests.*

4.3 Galileo's Lesson Today

Dear Malcolm *(Longair)***, after Galileo we all have grown up with the idea that good science demands observational checks, and this is indeed the purpose of cosmological tests. Which are in your opinion the tests that better constrain the present Standard Cosmological Model? And why?**

Galileo's forging of the foundations of modern physics was a heroic achievement, which in my view is unparalleled at any other time in modern history. It is difficult for the scientist of the twenty first century to appreciate the full extent of the Galilean revolution. The tenets of Aristotelian physics were accepted without question as the way the world works. Replacing received dogma by science based upon experiment, precise measurement, and mathematics was a defining moment in history and one for which Galileo rightly deserves the very highest recognition.

In the cosmological dialogues developed in this volume, there is a common scientific methodology that scarcely merits justification or discussion. The vast majority of scientists agree about the roles of experiment, interpretation, and theory without the need to justify their positions. Science advances by a process of natural selection, constrained by the requirement of accounting for the results of observation and experiment and making predictions to new circumstances. We have it easy compared with Galileo on many counts.

There are important lessons to be learned from the Galileo case, which have been summarized by the recent publications of the books *Retrying Galileo* by Maurice Finocchiaro and *The Church and Galileo* edited by Ernan McMullin. The opening up of the papal archives clarified many of the issues involved. The subsequent reception history of the Galileo case reminds us of the need to be vigilant about assuming that modern perceptions are independent of political and sociological influence.

One of the more striking aspects of the story told in McMullin's book is the fact that, during Galileo's trials of 1616 and 1633, the Papal authorities did not investigate the experiments or observations on which his theories of the motion of the Earth, Sun, and planets were based. Their arguments concerning the issues of the motion and rotation of the Earth were taken to be largely philosophical and theological, rather than a subject for experimental investigation. I have an analogous concern about studies of the physics of the very early Universe, which involve applying theoretical ideas that cannot yet be supported by, and indeed might never be susceptible to observational and experimental validation. On the other hand, I will defend to the last the need to develop and understand physics far beyond what can be tested at the present time.

Another intriguing aspect of the Galileo case is that there were real physical problems in adopting Galileo's picture of the rotating Earth orbiting the Sun. For example, there was concern that objects would be thrown from the Earth's surface, just like objects on a rotating potter's wheel. In the absence of the laws of motion, the law of gravity and an understanding of centrifugal force, there was no way of countering that argument. It was this type of issue concerning the physical basis of the Copernican picture, which caused supporters of that picture, such as Didacus à Stunica, to concede that it was probably not correct. The parallel with present day cosmology is that there may well be unknown bits of physics that are crucial for understanding the origin of the Concordance Model. In what might be called *Donald Rumsfeld cosmology*, there are *unknown unknowns* – things which we do not know that we do not know. It would be disappointing if there were no new physics between the energies studied in the LHC at CERN and the 10^{19} GeV of the Planck era.

The catastrophe of the Papal Commission set up by Pope John Paul II in 1981 to re-examine the Galileo case, beautifully recounted by George Coyne in McMullin's book, is the story of how what should have been a balanced investigation based upon careful reading of the materials in the Vatican Archives turned out to be a travesty of *laissez faire*. One might argue that this could never happen in modern astrophysical cosmology. But in one's reflective moments, there might be a concern that the band-wagon of the standard concordance picture has become a relentless juggernaut. We should be beware of the standard model becoming received dogma, but rather a best-fitting model subject to observational and experimental validation like any other physical theory.

In my interpretation of the goals of the present Galilean Dialogue, we are not in the position of setting up Salviati, Galileo's *alter ego*, as the knowing cosmologist who has all the answers ready to counter the arguments advanced by Simplicius, the rather naive apologist for accepted dogma. The balance has to be much better sustained between those who are persuaded by the undoubted success of the standard concordance picture, and the sceptic who is aware of the dangers of being swept away by the tide of received opinion. I am reminded of Galileos's concern about Kepler's over-enthusiastic adoption of the Copernican world view:

> I do not doubt that the thoughts of Landsberg and some of Kepler's tend rather to the diminution of the doctrine of Copernicus than to its establishment as it seems to me that these (in the common phrase) have wished too much for it ...

We have to try to avoid being among those who "have wished too much for it ..". Personally, I find the concordance picture compelling and one I had not expected to witness during my lifetime. At the same time, it raises as many important physical problems as it resolves, a characteristic of all great physical theories. We have not yet obtained a proper understanding of the real significance of the observations and the conclusions to be drawn from them. Nonetheless, the encouraging aspect of this story is that, within the bounds of physical enquiry, there are real physics problems to be solved that can be addressed by observation and experiment.

I have addressed many of these issues in extenso in my recent book [125], but it is worthwhile reviewing briefly my position on these specific issues. Detailed references to all the works discussed here are to be found in my book.

What I find most compelling about the standard Concordance Model is not that any one piece of evidence provides the strongest constraints on the model, but rather that *all* the available evidence is consistent with a remarkable simple self-consistent cosmology. The long list of contemporary cosmological tests is remarkable:

- Isotropy, homogeneity, and spectrum of the Cosmic Microwave Background (CMB) radiation
- Hubble's Law determined by the Hubble Space Telescope and other physical methods
- Angular Power Spectrum (APS) of fluctuations in the CMB radiation and its polarization properties on a wide range of angular scales
- Power spectrum of baryon acoustic oscillations derived from massive redshift surveys
- Abundances of the light elements from primordial nucleosynthesis
- Solar system and laboratory tests of General Relativity
- Ages of the oldest objects in the Universe, such as globular clusters and white dwarfs
- Mass density on large physical scales from infall into supercluster systems and large-scale structures
- The redshift-apparent magnitude relation for type Ia Supernovae (SNe)

The important feature of this list is that each test provides an independent route to the determination of the best-fitting cosmological parameters. If any one of them disagreed with the expectation of the standard model, questions would undoubtedly be raised about how much credence should be given to it. But this has *not* happened. It is the first time in the history of cosmology that there is general agreement about the best-fitting cosmological model.

Most cosmologists would agree that the most impressive sets of observations are those of the spectrum and polarization of the CMB radiation provided by the Wilkinson Microwave Anisotropy Probe (WMAP) observatory. The predicted power spectrum and polarization of primordial adiabatic acoustic oscillations have been fully confirmed by these observations. Cosmologists have been squeezing every last piece of information from the WMAP observations and important estimates, or limits, have been found for a very large number of cosmological parameters. But it is equally impressive how well the simplest models can account for the observations without encountering any really serious conflict. I find the analysis of the first year WMAP data by Max Tegmark and his colleagues particularly revealing, particularly in showing how much information can be gleaned from the data. Furthermore, the mass of data can be accounted for by a cosmological model with only five free cosmological parameters. It is also pleasing that the physical origin of the predicted spectrum and polarization of the CMB radiation can be readily understood by simple physical arguments. It is very fortunate that most of the physics involved in the study of the temperature fluctuations is linear and so we can have confidence in the predictions of the models.

There are, however, concerns. We can reasonably claim that we know many of the key cosmological parameters with better than 10% accuracy, an extraordinary improvement over the situation 10 or 20 years ago. The objective of future observations and experiments is to test the models with higher and higher precision aiming for 1% accuracy and addressing key issues such as whether the cosmological constant really is a constant or whether it is a dynamical field. But this also means that we need to be sure of the physics to better than 1%. This is far from trivial. My own view is that the new physics will emerge as small deviations from the standard concordance picture, for example, through observations by the *Planck* satellite, but that also means that the predictions of the models must be robust at the 1% level.

4.4 Tests of General Relativity

Dear Malcolm *(Longair)*, **it is commonly believed that GR, well proved at many distance scales, fails at a length scale close to the Planck epoch. It seems therefore plausible to ask if GR may fail also at the largest scales, let say close to the scale of cosmology ($\sim 10^{28}$ cm), but cross check tests using different methods seems to exclude this possibility. Would you like to comment how robust is the theoretical background on which the Standard Model has been built?**

The robustness of the theoretical background necessarily depends upon the experimental and observational validation of the input physics. Let me review some aspects of the relevant theories.

The Concordance Model is based upon the general theory of relativity and so it is important to ask just how well that theory is supported by observations and experiment. For example, how good is *Einstein Equivalence Principle*? Following Clifford Will's splendid exposition, deviations from linearity of the relation between gravitational and inertial mass can be written

$$m_{\mathrm{g}} = m_{\mathrm{i}} + \sum_{\mathrm{A}} \frac{\eta^{\mathrm{A}} E^{\mathrm{A}}}{c^2}. \tag{4.1}$$

E^{A} is the internal energy of the body generated by interaction A, η is a dimensionless parameter that measures the strength of the violation of the linearity of the relation between gravitational mass m_{g} and inertial mass m_{i} induced by that interaction, and c is the speed of light. The internal energy terms include all the mass-energy terms that can contribute to the inertial mass of the body, for example, the body's rest energy, its kinetic energy, its electromagnetic energy, weak-interaction energy, binding energy, etc. If the inertial and gravitational masses were not exactly linearly proportional to each other, there would be a finite value η^{A}, which would be exhibited as a difference in the accelerations of bodies of the same inertial mass composed of different materials.

A measurement of, or limit to, the fractional difference in accelerations between two bodies yields the quantity known as the *Eötvös ratio*,

$$\eta = 2\frac{|a_1 - a_2|}{a_1 + a_2} = \sum_{\mathrm{A}} \eta^{\mathrm{A}} \left(\frac{E_1^{\mathrm{A}}}{m_1 c^2} - \frac{E_2^{\mathrm{A}}}{m_2 c^2} \right). \tag{4.2}$$

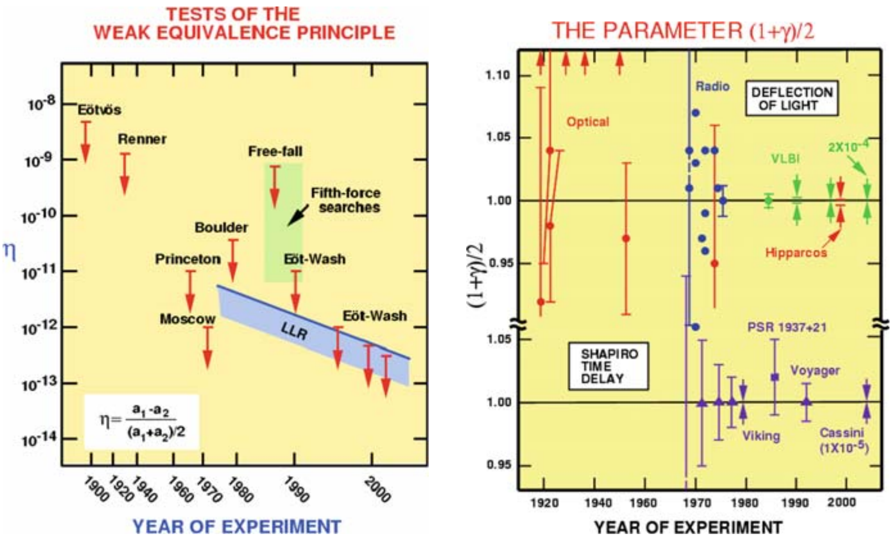

Fig. 4.1 (**a**) Tests of Einstein's equivalence principle, as parametrized by the Eötvös ratio η. (**b**) Solar system tests of the general theory of relativity as parametrized by the quantity γ. From [231]

According to the data shown in Fig. 4.1 (panel a), there is no deviation from a linear relation between gravitational and inertial mass at the level of about one part in 10^{13}. The NASA space mission STEP (Space Test of the Equivalence Principle) would be able to detect any nonlinearity at a level of one part in 10^{18} at which effects predicted by theories of elementary particles and extra dimensions might be observable. Any deviation from exact proportionality would have a profound impact for fundamental physics.

An elegant way of testing GR is to adopt *parametrized post-Newtonian* models for theories of relativistic gravity. To quote the words of Will,

> 'The comparison of metric theories of gravity with each other and with experiment becomes particularly simple when one takes the slow-motion, weak-field limit. This approximation, known as the post-Newtonian limit, is sufficiently accurate to encompass most solar-system tests that can be performed in the foreseeable future. It turns out that, in this limit, the space-time metric predicted by nearly every metric theory of gravity has the same structure. It can be written as an expansion about the Minkowski metric in terms of dimensionless gravitational potentials of varying degrees of smallness.'

The approach is to relax the powerful constraints implied by the Einstein Equivalence Principle, allowing a wider range of possible theories of relativistic gravity. In Table 4.1 we have two examples of extensions beyond GR, which illustrate the types of physics that can be tested:

Traditionally, there are four tests of GR:

- The *gravitational redshift* of electromagnetic waves in a gravitational field. Hydrogen masers in rocket payloads confirm the prediction at a level of about 5 parts in 10^5.

Table 4.1 Examples of parametrized post-Newtonian coefficients

Parameter	What it measures relative to general relativity	Value in general relativity	Value in conservative and semi-conservative theories
γ	How much space-curvature is produced by unit rest mass?	1	γ
β	How much "nonlinearity" in the superposition law for gravity?	1	β

- The *advance of the perihelion of Mercury*. Continued observations of Mercury by radar ranging have established the advance of the perihelion of its orbit to about 0.1% precision with the result $\dot{\omega} = 42.98(1 \pm 0.001)$ arcsec per century. GR predicts a value of $\dot{\omega} = 42.98$ arcsec per century. The corresponding limit to β is $\beta - 1 < 3 \times 10^{-3}$.
- The *gravitational deflection of light by the Sun* has been measured by VLBI and the values found correspond to $(1 + \gamma)/2 = 0.99992 \pm 0.00023$.
- The *time delay of electromagnetic waves* propagating through the spatially varying gravitational potential of the Sun. While en route to Saturn, the Cassini spacecraft found a time-delay corresponding to $(\gamma - 1) = (2.1 \pm 2.3) \times 10^{-5}$. Hence the coefficient $\frac{1}{2}(1 + \gamma)$ must be within at most 0.0012% of unity.

The historical improvement in the determination of the quantity $(1 + \gamma)/2$ from light deflection and time delay experiments is shown in Fig. 4.1 (panel b).

Another key test uses the binary pulsar PSR 1913+16 which is one of a pair of neutron stars in binary orbits about their common centre of mass. There is a displacement between the axis of the magnetic dipole and the rotation axis of the neutron star. The radio pulses are assumed to be due to beams of radio emission from the poles of the magnetic field distribution. Many parameters of the binary orbit and the masses of the neutron stars can be measured with very high precision by accurate timing measurements.

The binary pulsar emits gravitational radiation and this leads to a speeding up of the stars in the binary orbit. Figure 4.2 shows the change of orbital phase as a function of time for the binary neutron star system PSR 1913+16 compared with the expected changes due to gravitational radiation energy loss by the binary system. These observations enable many alternative theories of gravity to be excluded. These limits are expected to improve with the discovery of the remarkable system J0737-3039 in which both neutron stars are observed as pulsars. This pair of pulsars, discovered in 2003, is the most extreme relativistic binary system ever discovered with an orbital period of 2.45 h and a remarkably high value of its periastron advance, $\mathrm{d}\omega/\mathrm{d}t = 16.9^\circ$ year^{-1}. This system has the potential of providing very powerful constraints on alternative theories of gravity.

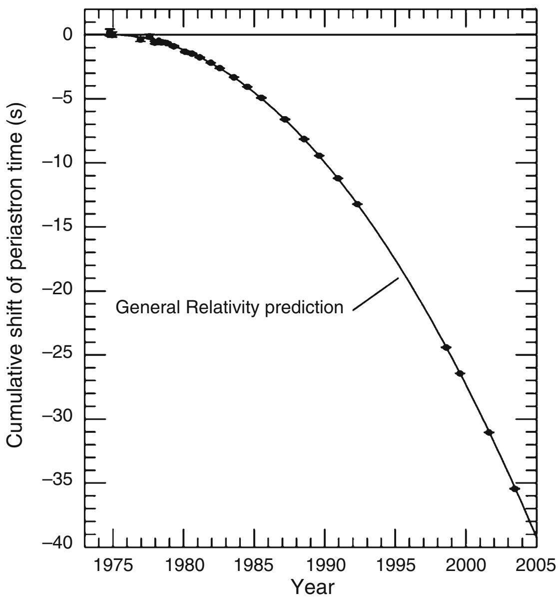

Fig. 4.2 The loss of energy by gravitational radiation by the binary pulsar results in an advance of the phase of the pulses over the years since it was discovered in 1976. From [231]

Table 4.2 Limits to the variability of the gravitational constant

Method	$(\dot{G}/G)/10^{-13}$ year^{-1}
Lunar laser ranging	4 ± 9
Binary pulsar PSR 1913+16	40 ± 50
Helioseismology	0 ± 16
Big bang nucleosynthesis	0 ± 4

In Table 4.2, various Solar system, astrophysical and cosmological tests are listed, which provide limits to the rate at which the gravitational constant could have varied over cosmological time-scales. It should be emphasized that the binary pulsar limit depends upon knowledge of the equation of state of neutron star matter at very high densities. Furthermore, the Big Bang nucleosynthesis argument assumes a power-law variation of the gravitational constant with cosmic epoch. Gary Steigman presents many more details of the latter argument[31] and I strongly commend his presentation to the reader. As he has pointed out, the last quoted value in Table 4.2 is on the optimistic side, about 10% variation being a safer upper limit.

Thus, there can have been little variation in the value of the gravitational constant over the last 10^{10} years.

Thank you Malcolm. The gravitational constant G is one of the two constants of GR: the other is the cosmological constant Λ. Could you also give us a historical overview of the cosmological constant, before we address the core of this scientific problem?

[31] He remarked also that while the binary pulsar limit is, in some sense, "here and now", Big Bang Nucleosynthesis (BBN) and any other cosmological limits generally require some model in order to extrapolate from "there and then".

4.5 Cosmological Constant

4.5.1 Historical Overview

Dear Malcolm *(Longair)*, **the history of the cosmological constant is very interesting, and today ad hoc interpretations have been proposed to justify its value. The anthropic principle is one of them. Would you like to summarize the history of the cosmological constant and express your opinion on such curious interpretation?**

I have discussed the tortuous history of the cosmological constant in some detail in my book [125]. It is a curious, but revealing story. Einstein introduced the cosmological constant into his field equations to create a static Universe which would incorporate Mach's Principle into GR. Almost immediately, De Sitter showed that Einstein had not achieved this aim as solutions of the equations existed even in the absence of matter in the Universe. In 1919, however, Einstein realized that the Λ-term would appear as a constant in his field equations quite independent of its cosmological significance.

For most of the twentieth century, cosmologists adopted ambivalent views about the Λ-term. In 1919, Einstein was not enthusiastic about the term, remarking that it "detracts from the formal beauty of the theory". Willem de Sitter had similar views and wrote in 1919 that the term

> detracts from the symmetry and elegance of Einstein's original theory, one of whose chief attractions was that it explained so much without introducing any new hypotheses or empirical constant.

Others regarded it as a constant that appears in the development of the GR and its value should be determined by astronomical observation.

In 1933, Lemaître suggested that the Λ-term could be interpreted in terms of a finite vacuum energy density. In his words,

> Everything happens as though the energy in vacuo would be different from zero.

This insight foreshadows the present interpretation of the cosmological constant, which associates it with DE. The same idea was revived by McCrea in 1951 in the context of providing a physical interpretation of the C-field introduced by Hoyle to describe the continuous creation of matter in Steady State cosmology. In his words,

> The single admission that the zero of absolute stress may be set elsewhere than is currently assumed on somewhat arbitrary grounds permits all of Hoyle's results to be derived within the system of GR theory. Also, this derivation gives the results an intellectual physical coherence.

A delightful sequel to this story is that, on the occasion of his 80th birthday in 1995, I invited Hoyle to lecture to the Cavendish Physical Society. He was delighted to accept this invitation because he had given his first lecture on steady state cosmology to the Cavendish Physical Society in 1948. Hoyle remarked wryly that his only mistake had been to call his creation field C rather than ψ. The exponential expansion of the early Universe according to the inflation picture involves a scalar field ψ, which performs exactly the same function as Hoyle's C field or Lemaître's cosmological constant Λ.

Throughout the twentieth century, the cosmological constant made regular reappearances in the literature in response the various cosmological problems but none of these arguments withstood detailed scrutiny until, in the last decade of century, compelling evidence for a nonzero value of the cosmological constant was found.

The key issue now is the significance of the cosmological constant, or DE, for cosmology. Is it a term that should appear in the geometric structure of GR? Or is it a field that might change with cosmic epoch? These issues can be thought of in terms of the properties of the form on the DE equation of state, $p = w\rho c^2$. Is $w = -1$, which would be precisely equivalent to the presence of the cosmological constant in the field equations, or does it differ from this value? Does w vary with redshift? It is now feasible to address these questions using very large samples of galaxies and weak gravitational lensing, as has been proposed in different variants of the DE space missions now being studied in the USA by NASA and in Europe by the ESA. These are very demanding missions, but they seem to me to be an essential part of unraveling what is undoubtedly one of the most remarkable cosmological and physical mysteries of twenty-first century science.

Thanks a lot Malcolm. In the following interview, we try to go deeper in the investigation of the cosmological constant problem with the aid of Thanu Padmanabhan.

4.5.2 The Problem of Theoretical Physics

Dear Thanu *(Padmanabhan)*, **the cosmological constant is a long standing, disturbing problem for physicists and cosmologists. It is today considered the crucial theoretical question of physics. Can you summarize the salient features of this problem?**

I share the view that the cosmological constant represents *the* problem of theoretical physics; in fact, it has been *the* problem of theoretical physics for several decades now and this problem – at a fundamental level – is quite independent of the observational results modern cosmology has brought in. Let me elaborate.

Einstein's theory of gravity, at least at low energies, is described by two constants G and Λ, where G is the Newtonian gravitational constant and Λ is the cosmological constant. Given these two constants, one can construct a dimensionless number $\lambda \equiv \Lambda(G\hbar/c^3)$ by introducing the Planck constant $\hbar$ and the speed of light c into the fray. It was known for decades that this dimensionless number λ is extremely tiny and less than about 10^{-123}. In the early days, this was considered puzzling but most people believed that this number is actually zero. The cosmological constant problem in those days was to understand why it is strictly zero. Usually, the vanishing of such a constant (which could have appeared in the low energy sector of the theory) indicates an underlying symmetry of the theory. For example, the vanishing of the mass of the photon is closely related to the gauge invariance of electromagnetism. No such symmetry principle is known to operate at low energies, which made this problem very puzzling. There is a symmetry – called supersymmetry – which does ensure that $\lambda = 0$ but it is known that supersymmetry is broken at sufficiently high energies and hence cannot explain the observed value of λ.

From late eighties onwards, the cosmological constant problem has acquired a new dimension. Early analysis of several observations [85] indicated that the expansion of the Universe is (partially) driven by a smooth, unclustered, component of energy (now-a-days called DE), which exerts negative pressure. This is confirmed dramatically by the SNe observations [183]. (For a critical look at the current data, see [170]; for recent reviews, see e.g., [162].) The rapid acceptance of DE by the community is partially due to the fact that – even before the SN data came up – there were strong indications for the existence of DE.

The simplest candidate for DE, consistent with all the observations today, is a cosmological constant with $\lambda \approx 10^{-123}$. So, in the simplest interpretation of the current observations, we need to explain why cosmological constant is nonzero and has this small value. In this sense, cosmological constant problem has got linked to the existence of DE in the Universe. I would, however, like to stress that these are logically independent issues. *Even if all the observational evidence for DE goes away, we still have a problem – viz., explaining why λ is zero.*

The literature dealing with cosmological constant also discusses what could be called the "why now" problem. That is, how come the energy density contributed by the cosmological constant (treated as the DE) is comparable to the energy density of the rest of the matter at the *current epoch* of the Universe. I personally do not believe this is a serious, *independent*, problem; if we have a viable theory predicting a numerical value for λ, then the energy density due to this cosmological constant will be comparable to the rest of the energy density at some epoch. So the real problem is in understanding the numerical value of λ; once that is solved the "why now" issue will take care of itself. In fact, we do not have a viable theory to predict the current energy densities of *any* component that populates the Universe, let alone the DE!. For example, the energy density of radiation today is computed from its temperature, which is an observed parameter – there is no theory which tells us that this temperature has to be 2.73 K when, say, galaxy formation has taken place for certain billion number of years.

It is also worth emphasizing another key aspect of the cosmological constant problem. Mathematically, the total Lagrangian density describing the interaction of gravity with matter has the form $L = (16\pi G)^{-1}[R - 2\Lambda] + L_{\text{matter}}$. What is really relevant is this *total* Lagrangian describing the gravity interacting with matter and hence, one could have thought of the cosmological constant as either part of gravitational sector or part of L_{matter}. No observation can distinguish between these two choices. But then, if you add another constant to the matter Lagrangian, that will also contribute as a cosmological constant! The addition of such a constant is a symmetry within the matter sector but gravity breaks this symmetry. This is because non-gravitational physics does not care about the absolute value of the energy and physical phenomena, which does not involve gravity dependent only on the *changes* in the energy. Gravity, however, couples to the absolute amount of energy and hence is sensitive to the shifting of the zero level of energy of the matter sector. This is, in fact, the crux of the cosmological constant problem.

Can you explain briefly the kind of approaches that have been attempted to solve the cosmological constant problem and why they have been unsatisfactory?

One nice possibility would be to postulate that $\lambda = 0$ and come up with a symmetry principle that will explain why this is the case. You probably have a greater chance of success in such an attempt than in coming up with an explanation for $\lambda \approx 10^{-123}$. But then, one needs to provide an alternative explanation for the DE observations.

One of the *least* esoteric ideas regarding the DE is that a cosmological constant term arises in the relevant equations because we have not calculated the energy density driving the expansion of the Universe correctly. The motivation for such a suggestion stems from the following fact: Einstein's equations tell you the kind of gravitational field that is produced by a particular distribution of matter in the Universe. We know that the distribution of matter in the Universe is very inhomogeneous. Rigorously speaking, we should calculate the gravitational field arising from this inhomogeneous distribution of matter and then average it over small scales to determine the large scale dynamics of the Universe. Unfortunately, this is a technically formidable task and is impossible to achieve today. What is usually done in cosmology is to first average over the distribution of matter in the Universe and then determine the dynamics of the Universe using the average energy density as the source. As Einstein's theory is nonlinear, these two procedures will lead to different kind of dynamics and there was a hope that this difference could mimic the DE.

In spite of some recent attention this idea has received [27] it is doubtful whether it will lead to the correct result when implemented properly. The reasons for my skepticism are the following:

- The effect described above, of course, exists. (For an explicit example, in a completely different context of electromagnetic plane wave, see [159]). The question that needs to be settled is how big is the correction compared to the average energy density in the Universe. It seems unlikely that when properly done, we will get a large effect for the simple reason that the amount of mass, which is contained in the nonlinear regions of the Universe today is subdominant.
- The key mathematical question that is to be addressed in this approach is that of identifying a suitable analogue of expansion factor for the Universe from an averaged geometry. This is nontrivial and it is not clear that the answer will be unique.
- This approach is too strongly linked to explaining the acceleration of the Universe as observed by type Ia SNe. Even if we decide to completely ignore all SNe data, we still have reasonable evidence for the existence of DE and it is not clear how this approach can tackle such evidence.

Another equally conservative explanation of the cosmic acceleration will be that we are located in a large underdense region in the Universe; so that, locally, the underdensity acts like negative mass and produces a repulsive force. While there has been some discussion in the literature [112] as to whether observations indicate such a local "Hubble bubble", this does not seem to be a tenable explanation that one can take seriously at this stage. Again, CMB observations indicating DE, for example, will not be directly affected by this feature though one does need to take into account the effect of the local void if it exists.

Finally, one should not forget that a *vanishing* cosmological constant is still a problem that needs an explanation. So even if all the evidence for DE disappears within a decade, we still need to understand why cosmological constant is zero. I stress this because there is a recent tendency to forget the fact that the problem of the cosmological constant existed (and was recognized as a problem) long before the observational evidence for DE, accelerating Universe, etc. cropped up. In this sense, cosmological constant problem has an important theoretical dimension, which is distinct from what has been introduced by the observational evidence for DE.

Moving on to more esoteric attempts for explaining DE, the approaches based on scalar fields have generated most of the publications. All these models essentially assume that the vanishing of cosmological constant will be explained by a future theory and goes on to produce the observed DE density in the Universe as due to an evolving scalar field. Unfortunately, these models are unsatisfactory for several reasons.

- They have no predictive power. It is always possible to construct a suitable model with a scalar field to reproduce *any* expansion history of the Universe!
- By and large, the fields used in the literature have no natural theoretical justification. All of them are non-normalizable in the conventional sense and have to be interpreted as a low energy effective potential in an ad hoc manner.
- One key difference between cosmological constant and scalar field models is that the latter lead to an equation of state, which varies with time. If observations have demanded this, or even if observations have ruled out the cosmological constant at the present epoch, then one would have been forced to take alternative models seriously. However, all available observations are consistent with cosmological constant and – in fact – the possible variation of the DE density is strongly constrained [111].
- The most serious problem with the scalar field models is the following: All the scalar field potentials require fine tuning of the parameters to be viable. These models, therefore, merely push the cosmological constant problem to another level, making it somebody else's problem!.

The second simplest possibility that has been attempted in the literature several times in different guises is to try and "cancel out" the cosmological constant by some process, usually quantum mechanical in origin. One can, for example, ask whether switching on a cosmological constant will lead to a vacuum polarization that will tend to cancel out the cosmological constant. A less subtle way of doing this is to invoke another scalar field (here we go again!) such that it can couple to cosmological constant and reduce its effective value [67]. Unfortunately, none of this could be made to work properly. The key point is that, in such attempts, there are two natural length scales that come into operation. First is the Planck length $L_{\mathrm{p}} = (G\hbar/c^3)^{1/2}$, which characterizes the quantum gravitational effects. The second is the length scale corresponding roughly to the size of the observed Universe determined by the current value of the cosmological constant, L_Λ. By and large, these approaches lead to an energy density that is either $\rho_{\mathrm{UV}} \propto L_{\mathrm{P}}^{-4}$ or to $\rho_{\mathrm{IR}} \propto L_\Lambda^{-4}$. The first one is too large while the second one is too small! What is required is an approach that will lead to the *geometric mean*, $\sqrt{\rho_{\mathrm{UV}}\rho_{\mathrm{IR}}}$, of these two energy densities that is hard to come by.

There are several other approaches, other than the two described above, which exist in the literature but it is fair to say that none of them gets even to the first base. So we really have a serious, open, question in our hands!

The Concordance Model is now so well established that it is difficult to formulate convincing alternatives. Several attempts have been performed up to now. Could you briefly review the most attractive ones?

While the Concordance Model has been remarkably successful as a template for describing cosmological observations, *one should not mistake precise observations for a fundamental understanding*. There are still several puzzling issues that needs to be understood about our Universe other than that of DE, which has grabbed, quite justifiably, most of the attention. Let me mention just three of them:

1. We need nearly 30% of energy density to be contributed by DM, which we hope will be a Weak Interacting Massive Particle (WIMP) – maybe the lightest supersymmetric partner. It is only a hope and until we see it in the lab, we need to keep an open mind regarding the nature of DM. For example, if LHC does not detect supersymmetry, one may have to re-look at the DM paradigm or at least become more sceptical.
2. We do not have a viable, predictive, model for understanding the energy density contributed by the baryons today. The most natural Universe one could think of will be baryon–antibaryon symmetric and the life forms as we know today will not exist.
3. There is no natural candidate from particle physics for producing a viable inflationary scenario. For example, we cannot today *predict* either the amplitude or the index of the initial spectrum of inhomogeneities generated by inflation. The best we can do is to claim consistency of cosmological observations with the inflationary paradigm, which is not a very satisfactory state of affairs.

I will argue, based on the above facts, that we are far from *understanding* the structure of our Universe but have only succeeded in parametrizing our ignorance in terms of a set of parameters, which have been measured with unprecedented accuracy. While this a triumph for observational cosmology, it will be dangerous to accept the theoretical backdrop at an equally firm footing. So, in principle, there is always scope for alternative theoretical paradigms.

The trouble, however, is that *we do not have today a single viable alternative* to the concordance cosmology. The same goes for inflationary paradigm – while I consider it unsatisfactory, there is no other game in town!

A way to escape the present impasse of theoretical physics and cosmology may come from a revolution in our understanding of the current physics. Do you think that we are close to such an event? Do we really need a deep revision of our current understanding of the Universe?

Absolutely! I think *cosmological constant problem essentially has to do with our misunderstanding of the nature of gravity*. To solve this problem comprehensively, we need a major paradigm shift as regards gravity and have to view gravity as an emergent phenomenon [167] – like for example, elasticity. Let me explain the features which strongly suggest such a point of view.

To begin with, it is obvious that in terms of the energy scales, the cosmological constant problem is an infrared problem par excellence. At the same time, the occurrence of $\hbar$ in $\lambda = \Lambda(G\hbar/c^3)$ shows that it is a relic of a quantum gravitational effect (or principle) of unknown nature. One is envisaging here a somewhat unusual possibility of a high energy phenomenon leaving a low energy relic and an analogy will be helpful to illustrate this idea [169]. Suppose we solve the Schrodinger equation for the Helium atom for the quantum states of the two electrons $\psi(x_1, x_2)$. When the result is compared with observations, we will find that only half the states – those in which $\psi(x_1, x_2)$ is antisymmetric under $x_1 \longleftrightarrow x_2$ interchange – are realized in nature. But the low energy Hamiltonian for electrons in the Helium atom has no information about this effect! Here is a low energy (IR) effect, which is a relic of relativistic quantum field theory (spin-statistics theorem) that is totally nonperturbative, in the sense that writing corrections to the Hamiltonian of the Helium atom in some $(1/c)$ expansion will *not* reproduce this result. I suspect the current value of cosmological constant is related to quantum gravity in a similar spirit. There must exist a deep principle in quantum gravity, which leaves its nonperturbative trace even in the low energy limit that appears as the cosmological constant.

Second, as I have already stressed, the addition of a constant to the energy (or Lagrangian) is a symmetry of the matter sector (at least at scales below the scale of supersymmetry breaking). The matter equations of motion do not care about constant that is added. But, in the conventional approach, gravity breaks this symmetry. *This is the root cause of the cosmological constant problem.* As long as gravitational field equations are of the form $E_{\mathrm{ab}} = \kappa T_{\mathrm{ab}}$, where E_{ab} is some geometrical quantity (which is G_{ab} in Einstein's theory) the theory cannot be invariant under the shifts of the form $T^{\mathrm{a}}_{\mathrm{b}} \rightarrow T^{\mathrm{a}}_{\mathrm{b}} + \rho\delta^{\mathrm{a}}_{\mathrm{b}}$ arising from an addition of a constant to the matter Lagrangian. Since such shifts are allowed by the matter sector, it is very difficult to imagine a definitive solution to cosmological constant problem within the conventional approach to gravity.

So we need a formalism in which gravity does not couple to the bulk energy density and only responds to energy differences. But one can prove that (1) If a geometrical variable – like the metric – represents the gravitational degree of freedom that is varied in the action and (2) we demand full general covariance, we cannot make gravity immune to the bulk energy density. Clearly a new, drastically different approach to gravity is required. We have to abandon the usual picture of treating the metric as the fundamental dynamical degrees of freedom of the theory and treat it as providing a coarse grained description of the space–time at macroscopic scales, somewhat like the density of a solid – which has no meaning at atomic scales [195]. The unknown, microscopic degrees of freedom of space–time (which should be analogous to the atoms in the case of solids) will play a role only when space–time is probed at Planck scales (which would be analogous to the lattice spacing of a solid [213]).

Moreover, in the study of ordinary solids, one can distinguish between three levels of description. At the macroscopic level, we have the theory of elasticity that has a life of its own and can be developed purely phenomenologically. At the other extreme, the microscopic description of a solid will be in terms of the statistical mechanics of a lattice of atoms and their interaction. Both of these are well known; but interpolating between these two limits is the thermodynamic description of a solid at finite temperature, *which provides a crucial window into the existence of the corpuscular substructure of solids.* As Boltzmann has emphasized, heat is a form of motion and we will not have the thermodynamic layer of description if matter is a continuum all the way to the finest scales and atoms did not exist! *The mere existence of a thermodynamic layer in the description is proof enough that there are microscopic degrees of freedom.* Let us translate these lessons from a solid to the space–time. Again we should have three levels of description. The macroscopic level is the smooth space–time continuum with a metric tensor $g_{\mathrm{ab}}(x^i)$ and the equations governing the metric have the same status as the phenomenological equations of elasticity. At the microscopic level, we expect a quantum description in terms of the "atoms of space–time" and some associated degrees of freedom q_{A}, which are still elusive. But what is crucial is the existence of an interpolating layer of thermal phenomenon associated with null surfaces in the space–time. Just as a solid cannot exhibit thermal phenomenon if it does not have microstructure, *thermal nature of horizons, for example, cannot arise without the space–time having a microstructure.*

In such a picture, we normally expect the microscopic structure of space–time to manifest itself only at Planck scales or near singularities of the classical theory. However, in a manner which is not fully understood, the horizons – which block information from certain classes of observers – link [160] certain aspects of microscopic physics with the bulk dynamics, just as thermodynamics can provide a link between statistical mechanics and (zero temperature) dynamics of a solid. The reason is probably related to the fact that horizons lead to infinite redshift, which probes *virtual* high energy processes; it is, however, difficult to establish this claim in mathematical terms.

The above paradigm, in which the gravity is an emergent phenomenon, is anchored on a fundamental relationship between the dynamics of gravity and thermodynamics of horizons [164] and the following three results are strongly supportive of the above point of view:

- There is a deep connection between the dynamical equations governing the metric and the thermodynamics of horizons. An explicit example was provided in [161], in the case of spherically symmetric horizons in four dimensions in

which it was shown that, Einstein's equations can be interpreted as a thermodynamic relation $T\,\mathrm{d}S = \mathrm{d}E + P\,\mathrm{d}V$ arising out of virtual radial displacements of the horizon. Further work showed that this result is valid in *all* the cases for which explicit computation can be carried out – like in the Friedmann models [233] as well as for rotating and time dependent horizons in Einstein's theory [119].
- The Hilbert Lagrangian has the structure $\mathcal{L}_{\mathrm{EH}} \propto R \sim (\partial g)^2 + \partial^2 g$. In the usual approach the surface term arising from $\mathcal{L}_{\mathrm{sur}} \propto \partial^2 g$ has to be ignored or canceled to get Einstein's equations from $\mathcal{L}_{\mathrm{bulk}} \propto (\partial g)^2$. But there is a peculiar (unexplained) relationship between $\mathcal{L}_{\mathrm{bulk}}$ and $\mathcal{L}_{\mathrm{sur}}$, which makes the gravitational action "holographic" with the same information being coded in both the bulk and surface terms and one of them is sufficient. One can indeed obtain Einstein's equations from an action principle, which uses *only* the surface term and the virtual displacements of horizons [165, 166]. Since the surface term has the thermodynamic interpretation as the entropy of horizons, this establishes a direct connection between space–time dynamics and horizon thermodynamics.
- Most importantly, recent work has shown that *all the above results extend far beyond Einstein's theory.* The connection between field equations and the thermodynamic relation $T\,\mathrm{d}S = \mathrm{d}E + P\,\mathrm{d}V$ is not restricted to Einstein's theory alone, but is in fact true for the case of the generalized, higher derivative Lanczos-Lovelock gravitational theory in D dimensions as well [1, 172]. The same is true [146] for the holographic structure of the action functional: the Lanczos-Lovelock action has the same structure and – again – the entropy of the horizons is related to the surface term of the action. *These results show that the thermodynamic description is far more general than just Einstein's theory* and occurs in a wide class of theories in which the metric determines the structure of the light cones and null surfaces exist blocking the information.

The conventional approach to gravity fails to provide any clue on these results just as Newtonian continuum mechanics – without corpuscular, discrete, substructure for matter – cannot explain thermodynamic phenomena.

Taking a cue from the above arguments, let us suppose there are certain microscopic – as yet unknown – degrees of freedom q_{A}, analogous to the atoms in the case of solids, described by some microscopic action functional $A_{\mathrm{micro}}[q_{\mathrm{A}}]$. In the case of a solid, the relevant long-wavelength elastic dynamics is captured by the *displacement vector field*, which occurs in the equation $x^{\mathrm{a}} \to x^{\mathrm{a}} + \xi^{\mathrm{a}}(x)$. In the case of space–time, we no longer want to use metric as a dynamical variable, and so we need to introduce some other degrees of freedom, analogous to ξ^{a} in the case of elasticity, and an effective action functional based on it. Normally, varying an action functional with respect certain degrees of freedom will lead to equations of motion determining *those* degrees of freedom. But we now make an unusual demand that varying our action principle with respect to some (nonmetric) degrees of freedom should lead to an equation of motion *determining the background metric*, which remains nondynamical.

Based on the role expected to be played by surfaces in space–time, we shall take the relevant degrees of freedom to be the normalized vector fields $n_i(x)$ in the space–time [168]. That is, just as the displacement vector ξ^a captures the macro-description in case of solids, the normalized vectors (e.g., normals to surfaces) capture the essential macro-description in case of gravity in terms of an effective action. More formally, we expect the coarse graining of microscopic degrees of freedom to lead to an effective action in the long wavelength limit. It turns out that one can indeed develop a full theory along these lines and obtain the equations of motion for the background metric. Unlike the conventional approach, the resulting equations show that gravity is immune to the bulk energy density and only responds to the shifts in the energy density! *Thus this approach can explain quite easily why the classical value for the cosmological constant should be zero.*

The description of gravity using the action principle mentioned earlier provides a natural back drop for gauging away the bulk value of the cosmological constant, as it decouples from the dynamical degrees of freedom in the theory. Once the bulk term is eliminated, what is observable through gravitational effects, in the correct theory of quantum gravity, should be the *fluctuations* in the vacuum energy. These fluctuations will be nonzero if the Universe has a DeSitter horizon, which provides a confining volume. In this paradigm, the vacuum structure can readjust to gauge away the bulk energy density $\rho_{\rm UV} \simeq L_{\rm P}^{-4}$, while quantum vacuum *fluctuations* generates the observed value $\rho_{\rm DE}$.

If the cosmological constant arises due to the fluctuations in the energy density of the vacuum, then one needs to understand the structure of the quantum gravitational vacuum at cosmological scales. Quantum theory, especially the paradigm of renormalization group, has taught us that the concept of the vacuum state depends on the scale at which it is probed. The vacuum state that we use to study the lattice vibrations in a solid, say, is not the same as vacuum state of the Quantum Electro Dynamics (QED) and it is not appropriate to ask questions about the vacuum without specifying the scale. If the space–time has a cosmological horizon that blocks information, the natural scale is provided by the size of the horizon, L_Λ, and we should use observables defined within the accessible region. The energy inside a region bounded by a cosmological horizon will then exhibit fluctuations and it can be shown that [163] this leads to a *surviving* energy density due to the cosmological constant, which is the geometric mean: $\rho_{\rm DE} = \sqrt{\rho_{\rm IR}\,\rho_{\rm UV}}$. This is precisely what is observed.

I stress that the computation of energy fluctuations is completely meaningless in the conventional models of gravity in which the metric couples to the bulk energy density. Once a UV cut-off at Planck scale is imposed, one will always get a bulk contribution $\rho_{UV} \approx L_{\rm P}^{-4}$ with the usual problems. It is *only because* we have a way of decoupling the bulk term from contributing to the dynamical equations that, we have a right to look at the subdominant term $L_{\rm P}^{-4}(L_{\rm P}/L_\Lambda)^2$. Approaches in which the sub-dominant term is introduced in an ad hoc manner are technically flawed as the bulk term cannot be ignored in these usual approaches to gravity. Getting the correct value of the cosmological constant from the energy fluctuations is not as difficult as understanding why the bulk value (which is larger by 10^{120}!)

can be ignored. Our approach provides a natural backdrop for ignoring the bulk term – and as a bonus – we get the right value for the cosmological constant from the fluctuations. It is small because it is a purely quantum effect.

Thank you Thanu. The DE is certainly the other face of the cosmological constant problem. We will now discuss this in an interview with Francesca Perrotta.

4.6 Dark Energy Models

Dear Francesca *(Perrotta)***, as the evidence of a non-negligible cosmological constant contribution and the difficulties to explain why its density parameter is of the order of unit today, it strongly emerged the idea of introducing a corresponding energy content to the energy–momentum side of the Einstein equation, in the form of the so-called dark energy, able to generate an accelerated expansion of the Universe. Could you describe the physical basis of the various ideas proposed in this context?**

An expanding Universe is a hallmark of the Big Bang theory of the Universe. Over the last century, incontrovertible evidence has accumulated to support that scenario, corroborated by the detection of the relic radiation from the Big Bang and from the detailed observations of the abundances of the light elements. The standard model of cosmology has been completely upset in 1998, when astronomers found that distant type Ia SNe were dimmer than expected in a decelerating Universe [181, 182, 190]. Data from SNe could be explained in a Friedmann model with a positive cosmological constant: assuming that the Universe is dominated by matter and a possible vacuum energy component (or cosmological constant), the observed magnitude–redshift diagrams were converted into limits on the cosmological density parameters $\Omega_{\rm m}$ and Ω_Λ, ruling out both a flat CDM model ($\Omega_{\rm m} = 1$) and an open Universe with zero cosmological constant. The need for a revival of the cosmological constant was confirmed by subsequent analysis of the CMB and large scale structure datasets. For example, Percival et al. [178] performed a joint analysis of the CMB datasets and of the 2-Degree Field Galaxy Redshift Survey (2dFGRS), finding a strong evidence for a positive cosmological constant. Independent observations (lensing of quasars, cluster abundance, galaxy peculiar velocities, cluster dynamics, etc.) indirectly confirmed this finding. The main point is that, while the CMB observations clearly point to a flat geometry, the content of matter on cosmological scales is too low to justify the flatness.

In general, accelerated expansion could be achieved by adding a new component to the cosmic fluid, whose energy–momentum tensor corresponding to negative pressure. Modern cosmology refers to this (yet unknown) fluid as "dark energy"; in this framework, the cosmological constant is regarded as just one candidate to the DE podium. To match the acceleration deduced from type Ia SNe and the observed flatness, the energy density of DE should be 70% of the total: thus, DE would be the dominant component of the Universe, determining its ultimate fate and evolution. But the nature of DE is still matter of pure speculation, and many models have been proposed. What we know is that, to drive cosmic acceleration, it must have negative pressure. We know that it is a very homogeneous component, as DE cluster have never been observed on sub-horizon scales. We also know that its current energy density is extremely low of the order of the critical density $10^{-29}\,\mathrm{g\,cm^{-3}}$: nonetheless, it can (and in fact does) have a macroscopic impact on the Universe evolution, as it uniformly fills the space.

The simplest explanation would be that of a cosmological constant Λ. However, it historically suffers a serious theoretical problem: its energy density today has to be of the order of the critical density, $10^{-47}\,\mathrm{GeV}^4$. On the other hand, particle physics interprets the cosmological constant as the energy density of the ground state of a theory. A quantum field is thought as a collection of an infinite number of harmonic oscillators in the momentum space: the ground state of such a system does not diverge, thanks to the introduction of a cutoff applied on the energies past, which we no longer trust our field theory. The Planck Energy ($E = 10^{19}\,\mathrm{GeV}$) plausibly fixes this threshold, so that the expected value of the vacuum energy density is of the order $\rho_{\mathrm{vac}} = 10^{72}\,\mathrm{GeV}^4$, a discrepancy of about 120 orders of magnitude with respect to its observed value. This impressive discrepancy is known as the "cosmological constant problem". One could postulate that a "bare", geometrical cosmological constant (in the lhs of the Einstein's equations of GR), opposite in sign and exactly equal in magnitude to the zero-point quantum energy (sourcing the rhs of the Einstein's equation), conspires to cancel out the "net" vacuum energy. But, given the large number of elementary particles, this would be an extreme case of fine tuning. For this reason, one generally adopts a different point of view, namely that the cosmological constant is zero due to some mechanism that is not yet understood, and that the Λ-like behavior should be ascribed to a different source.

There is another perhaps philosophical question we can ask, namely why the cosmic acceleration began only at so recent times: noticeably, if it began earlier in the Universe evolution, collapsed structures would never form and we would never had a chance to exist. While this evidence may be interpreted as a realization of the Anthropic Principle (AP), it poses another problem to the cosmological constant model (the "coincidence problem"), as its energy density ρ_Λ does not vary with time: another fine tuning is thus required to obtain $\rho_\Lambda \sim \rho_{\mathrm{m}}$ just today (already at $z \sim 2$ the cosmological constant is subdominant).

All these difficulties motivated the proposal of several alternative scenarios and candidates for DE; amongst them, probably the most widely studied in the literature are the so-called *quintessence* models, a class of scalar field implementations of DE. In such scenarios, the "missing energy" resides in the potential and kinetic energy of a dynamical scalar field (Quintessence field): if the potential is sufficiently flat at late times, the consequent slow-rolling motion can mimic a cosmological constant (a fluid with negative pressure) that decreases with time. For particular choices of the field potential, the dynamics of the scalar field, governed by the Klein–Gordon equation, gives a natural explanation of the current smallness of DE density, thus alleviating the cosmological constant problem. The differences between Quintessence models and Λ also results in many cosmological observables: while the cosmological constant is perfectly homogeneous on all scales, this time varying component can develop perturbations, which may leave an observable imprint on the CMB and large scale structure, although, in general, DE does not clump on sub-horizon scales. Furthermore, while the cosmological constant equation of state w = pressure/density is $w = -1$ at all times, the equation of state of a dynamical scalar field is generally time dependent, resulting in different predictions for the expansion rate and for the distances.

A simple minimally coupled scalar field is generally affected by the same initial fine tuning problem of Λ: the DE density needs to lie in a very narrow range in order for its present value to match the observed values. However, some classes of quintessence potentials can alleviate the coincidence problem by virtue of a so-called tracker behavior in which the evolution of the scalar field is independent on the initial conditions and can naturally set the low energy scale of the DE.

In tracking models, the energy density of the field closely tracks (but is less than) the radiation density until matter–radiation equality, which triggers quintessence to start behaving as DE, eventually dominating the Universe (the conditions for such behavior have been worked out by several authors [122, 185, 219, 229]). An exponential potential for the Quintessence field is able to generate such scaling behavior, but it has been shown that it cannot come to dominate the cosmic energy density.

Some special cases of quintessence are phantom energy, in which the energy density of quintessence actually increases with time, and k-essence (kinetic quintessence), which has a nonstandard form of kinetic energy.

A perhaps more physically motivated model constructs a Quintessence scalar field coupled to neutrinos, where the scalar field potential naturally incorporates the small mass scale. Indeed the required "missing energy" density can be written as $\rho_{\mathrm{DE}} = (10^{-3}\,\mathrm{eV})^4$, suggestive of a light neutrino mass. Such model would generate a slowly decaying cosmological "constant", which eventually relaxes to zero.

In the past years, an interesting class of models was considered as a DE candidate, the so-called Chaplygin gas: this is a perfect fluid with equation of state $p = -A/\rho$, being A a positive constant. The interest around this model was motivated by its connection to string theory and Supersymmetry. From a cosmological point of view, its striking feature is that it allows for a very elegant unification of DM and DE. Despite the appeal of its features, this model was not able to match many observational data, and it was soon abandoned.

Extended models of Quintessence differs from the canonical ones since they allow for a nonminimal coupling of the scalar field to the Ricci curvature.

Such cosmologies have the appealing feature that the same field causing the time (and space) variation of the cosmological constant is the source of a varying Newton constant in the manner of Jordan–Brans–Dicke. In addition, they avoid the problem of fine tuning, thanks to a specific field dynamics instaured by virtue of gravity coupling. An interesting prediction of these models is that coupled quintessence may cluster on some sub-horizon scale, with potentially detectable effects on galaxies and clusters.

In completely different approaches, DE is not needed at all to explain cosmic acceleration: they call into the game possible failures of GR on very large scales (hundreds of millions light years, larger than superclusters); this is the case, for example, of the braneworld model [71]. However, most of these models are inconsistent with the observation, or they are equivalent to quintessence on sub-horizon scales.

A controversial but fascinating idea is that the accelerated expansion could originate from a "backreaction" process [118]. The way GR implements the cosmic evolution is based on the assumption of a smooth background supporting the growth of linear perturbations up to the nonlinear regime. However, it has been noticed that the two mechanisms (smoothing and evolution) do not commute, which should introduce an extra term on the rhs of the Einstein's equations. Instead of introducing some exotic, unknown form of DE, such conservative model has the advantage of making a refined use of a well established theory: the perturbations that create the current acceleration are just the familiar Newtonian ones that lead to the formation of large scale structure. It is still matter of debate whether the sub-horizon backreaction can actually produce such a macroscopic effect.

4.6.1 Dark Energy Candidates

Dear Francesca *(Perrotta)*, **to your opinion, what are the most plausible candidates for the DE and what are the most crucial kinds of observations able to distinguish between different models?**

In synthesis, excluding few alternative ideas, the current trend is to represent DE as a slowly rolling "Quintessence" scalar field or by a cosmological constant. Although there is no shortage of models, most Quintessence options are lacking in motivation and require significant fine tuning of initial conditions or the introduction of a fine tuned small scale into the fundamental Lagrangian. An attempt to distinguish between this huge class of ideas will necessarily rely on observational cosmology: the only strategy to follow to restrict the list of candidates, is to achieve a good knowledge of any imprint it can leave, and try to identify such signatures through dedicated observations.

A critical clue to understand the underlying physics would be the measurement of the equation of state of DE at different times. It is possible to give a parametrical representation of the DE equation of state, which includes its dynamics in a model independent way, to provide a clear comparison between predictions from different models: while each model of DE has its own form for $w(z)$, the parametrization proposed by Linder [123] is very general:

$$w(z) = w_0 + w_{\mathrm{a}}(1-a) = w_0 + w_{\mathrm{a}}z/(1+z). \tag{4.3}$$

The time variations of the DE equation of state are represented by the parameter w_a: to understand how important this parameter is, one can simply notice that the eventual measurement of a non zero value of w_{a} would unambiguously rule out the cosmological constant as a DE candidate. This parametrization is enough to evince basic properties such as the evolution of $\rho_{\rm DE}$, how much of it is there, in terms of its present energy density $\Omega_{\rm DE} = \rho_{\rm DE} \times 8\pi G/(3H^2)$, and how does it cluster, which is characterized by the DE sound horizon.

At present, there are no tight constraints on the current DE equation of state w_0, mainly because of a strong degeneracy between cosmological parameters. Assuming a constant equation of state, data are consistent with $w_0 = -1$ (to 10% precision), that is, compatible with a cosmological constant, but not ruling out Quintessence models for which w can be very close to -1 on average at recent times. However, thanks to the many high precision experiments expected in the next few years, the situations is rapidly improving, as the combination of complementary cosmological probes can remove this degeneracy: indeed, while classical methods for detecting DE are well established (based on SNe, CMB, weak lensing and cluster abundances), new methods are being proposed to improve the current constraints.

Let us review the main observational impacts of a DE component. First of all, we can notice that $w(z)$ can be interpreted as an effective equation of state: it can be introduced in the Friedmann equation for the expansion rate, independently on our knowledge of the physical mechanism leading to the cosmic acceleration. In a flat Universe, the expansion rate is then represented as [124]:

$$H^2(z)/H_0^2 = \Omega_{\rm m}(1+z)^3 + (1-\Omega_{\rm m})\exp\left[3\int_0^z d\ln(1+z')[1+w(z')]\right]. \quad (4.4)$$

It is clear how a DE component will affect distances. The comoving distance is defined as

$$r(z) = \int_a^1 {\rm d}t/a(t) = \int_0^z {\rm d}z/H(z). \quad (4.5)$$

Thus, most importantly, the DE will affect, through the expansion rate, also the angular diameter distances and the luminosity distances. Those are fundamental quantities in observations such as gravitational lensing surveys, type Ia SNe luminosities, and CMB acoustic peaks features. Indeed, probably the most powerful tool for investigating the evolution of DE (e.g., for estimating the parameters w_0 and w_a) will come from a dedicated experiment such as the proposed satellite SNAP[32], based on the systematic observation of spectra and calibrated light-curves of over 2000 distant ($z > 1$) type Ia SNe.

In addition, the expansion rate described by the Friedmann equation enters into the linear perturbation growth of the matter component,

$$\ddot{\delta} + 2H\dot{\delta} - \frac{3}{2}H^2\Omega_m\delta = 0, \quad (4.6)$$

where δ is the fractional matter perturbation density.

Analyzing the evolution of δ, one finds that the DE should affect the matter perturbations and the large scale structure in three different ways. First of all, the linear

[32] See SNAP in web page list.

matter power spectrum $P(k)$ turns over at a wavenumber $k_{\rm eq}$ (corresponding to the matter–radiation equivalence scale), which is proportional to $\Omega_{\rm m}$. Assuming a flat Universe, the main effect of DE is that of reducing $\Omega_{\rm m}$, thus increasing the scale of the turnover in the matter power spectrum. In fact, this was one of the first evidences for DE. The second effect is also related to the reduced matter content in most of the DE models: as a result of the Poisson equation, the matter overdensities are inversely proportional to $\Omega_{\rm m}$ for a fixed potential. Therefore, the amplitude of the power spectrum increases as the matter decreases (or if the DE goes up, for a fixed curvature). If we normalize the amplitude to the largest angular scales probed by CMB experiments (such as COBE or WMAP), which fix the amplitude of the primordial gravitational potential on such scales, a DE cosmology will imply a higher normalized matter power spectrum. The third effect on large scale structure is due to the growth factor of linear perturbations: for a fixed curvature, a DE Universe has earlier structure formation than a matter-dominated Universe. Whatever structure we observe today, it should have been in place at earlier times. Future deep galaxy surveys will certainly give a fundamental contribution to the assessment of the DE impact on the matter structure growth; to date, the first (and only) attempt to measure the DE equation of state at high redshifts, or its derivative, has been done by Seljak et al. [206]. That analysis required a combination of different datasets, the Lyα forest analysis of the SLOAN Digital Sky Survey (SDSS), the SDSS galaxy bias analysis, the constraints from SDSS galaxy clustering, the updated SNe data, and the first year WMAP data. The Lyα forest and the CMB allow to measure the amplitude of fluctuations at high redshifts, while galaxy clustering gives a measure at low redshifts: by comparing the different datasets, one can infer the DE impact on the structure growth and constrain the DE equation of state at different redshifts. They found that, up to a redshift as large as $z = 1$, $w(z)$ stays very close to the cosmological constant value, $w(z = 1) = -1.03 \pm 0.28$.

The next few years will also be characterized by the high-precision imaging of the CMB by the *Planck*[33] satellite. Naively, we might not expect DE affecting significantly the CMB properties: it is a late Universe effect, dominating the cosmic fluid at recent times only, while the CMB mostly probes the recombination epoch. However, the CMB can be a useful tool to constrain DE: it provides an inventory of the energy density of everything else in the Universe; it acts as a standard ruler on the last scattering surface, so that the geometry can be accurately determined; some CMB anisotropies are created very recently, so that DE can leave a direct imprint (e.g., the Integrated Sachs–Wolfe (ISW) effect and the Sunyaev–Zel'dovich Effect (SZE)); in some DE models (e.g., tracking models) DE may be relevant already at the recombination epoch [29, 68].

How does the CMB weight the DE content? From the Friedmann equation, one has

$$\rho_{\rm DE} + \rho_{\rm m} + \rho_{\rm b} + \rho_\gamma + \rho_\nu - 3k/(8\pi G a^2) = 3H^2/(8\pi G) = \rho_{crit}. \qquad (4.7)$$

[33] See *Planck* in web page list.

In the above formula, knowing the expansion rate, the CMB gives direct information about several terms, with the exception of ρ_{DE}, which can thus be determined as the "missing" energy density to explain the present expansion rate.

The observed CMB temperature constraints the photon density to a level of 0.004%. The height of the CMB Doppler peaks constraints the total matter density (DM plus baryons): the most recent observation by the satellite WMAP [215] give $\Omega_{\mathrm{m}}h^2 = 0.127 \pm 0.009$. Baryon density is constrained by the light element abundance, as well as by the CMB acoustic peaks ratio $\Omega_{\mathrm{b}}h^2 = 0.0223 \pm 0.0007$. Neutrino density amounts to less than 1% of the total density, as established by the small scale damping in Large Scale Structure (LSS) and overall neutrino mass limits. The angular sizes of the CMB structures constrain the curvature term to be less than 2%: the limit in this case depends on the DE assumptions. Finally, the critical density (rhs of the Friedmann equation above), measured through the Hubble constant, is the biggest source of possible systematic errors. Assuming $h = 0.72$ it follows, by difference,

$$\Omega_{\mathrm{DE}} \approx 1 - 0.13h^{-2} = 0.75. \tag{4.8}$$

The property of the CMB to act as a cosmic yardstick relies on the fact that it is imprinted with the scale of the sound horizon at last scattering. This characteristic scale is shown by the Doppler peaks. Any physical scale on the last scattering surface will subtend an angle on the sky, whose value depends on the distance to that surface and on the overall geometry. In particular, the physical size of the acoustic sound horizon at recombination,

$$\lambda = \int_0^{t_{\mathrm{rec}}} c_{\mathrm{s}}(t)\,\mathrm{d}t \tag{4.9}$$

is a well known quantity (the time of recombination and the sound speed are well determined); the angular distance to last scattering surface is given by

$$\int_0^{z_{\mathrm{rec}}} \mathrm{d}z' \frac{1}{\sqrt{\rho_{\mathrm{DE}}(z') + \rho_{\mathrm{m}}(z') + \rho_{\mathrm{b}}(z') + \rho_{\mathrm{rad}}(z')}}. \tag{4.10}$$

Roughly, the angular size of a feature imprinted on the last scattering surface is the ratio between its physical size and this angular distance; however, a small amount of curvature would also change the observed angular size, an effect hardly distinguishable from changes of the angular distance.

In summary, both the curvature and the DE can change the angular size of the Doppler peaks. Assuming a cosmological constant, we get a constraint on curvature; conversely, if we assume a flat Universe, we can find a constraint on the DE density and its evolution on time. It is important, however, to notice that the CMB fixes a single integrated quantity and has a fundamental degeneracy between the DE density and its equation of state, which could only be broken by other observations.

Another CMB indicator for DE comes from the ISW effect, by virtue of which additional anisotropies in the CMB are generated at low redshifts ($z < 4$ and large, super degree angular scales, corresponding to multipoles $\ell < 10$). When the Universe becomes dominated by DE, a late time transition occurs in the total equation of state, causing the potential depth associated to a density fluctuation to decay as photons are passing through. A direct detection of the ISW is difficult because the effect is biggest where the cosmic variance is largest. However, it can be identified through a nonzero correlation between the galaxy number density and the CMB temperature: this correlation has been observed by Scranton et al. [204] and by Granett et al. [98] between the luminous red galaxies from the SDSS and the WMAP temperature map of the CMB, providing significant ($> 90\%$) physical evidence for DE. This correlation is a promising line of research for DE, also in view of the incoming *Planck* data.

Another consequence of the reduced $\Omega_{\rm m}$ for a fixed curvature in DE (or cosmological constant) models is that of changing the expected cluster number counts. The physical principle is quite simple: since a DE component is smooth on sub-horizon scales, it does not contribute to clumping. Thus, for a fixed curvature, the DE just decreases the available matter for gravitational collapse: small density fluctuations cease to grow once the DE (or the curvature) begins to dominate. In such models, then formation of structures ceased before than they would in higher matter density models. The difference is significant: normalizing to the rich cluster abundance seen today at $8\,h^{-1}$ Mpc, the probability of finding a rich cluster at redshift 0.7 is about 100 times smaller for a flat CDM model than for a flat ΛCDM.

The redshift distribution of the number of cluster is one of the most straightforward observables: it is a function of the cluster mass function, $\mathrm{d}^2 n/(\mathrm{d}M\,\mathrm{d}z)$ and of the redshift evolution of the comoving volume $\mathrm{d}V/\mathrm{d}\Omega(z)$. The mass function is obtained from the growth rate of density fluctuations $\delta(z)$. Since both $\delta(z)$ and $\mathrm{d}V/\mathrm{d}\Omega(z)$ depend on the cosmological model, useful constraints can be obtained by comparing the observed $N(z)$ with the predicted values.

Clusters can be found and weighted via their notable observational properties, such as gravitational lensing on background objects, optical emission by member galaxies, X-ray emission by hot intracluster medium, and the SZE on the CMB. Counting clusters of galaxies as a function of redshift allows to extract the combined information about the structure growth and the geometry of the Universe. If cluster masses could be measured accurately, the degeneracy between growth and curvature effects would be removed: Press–Schechter approach or N-body calculations are used to determine the expected mass function, which is then compared with the observations. The main issue with the X-ray, optical, and Sunyaev-Zel'dovich cluster detection is that the relation between the cluster mass and the corresponding observable (X-ray temperature, flux, or integrated SZE flux) may undergo nonstandard redshift evolution, making difficult to constraint efficiently the DE parameters. For this reason, one of the most promising DE probe comes from gravitational lensing, where the mass is measured through its gravitational effects, thus providing a direct measurement of the mass distribution in the Universe as opposite to the distribution of light. In particular, we do expect important hints from weak lensing: weak gravitational lensing detects the gravitational mass of an overdensity (in particular, a cluster) via the slight distortion imprinted on the images of the far background galaxies. The observable associated to the mass is the shear field, encoding the magnitude and direction of these distortions. In a weak lensing survey, clusters are identified, where the S/N peaks above some threshold; the result is a filtered shear map, which is unambiguously associated to the underlying mass distribution.

Another new and promising approach to constrain DE is based on BAO in the galaxy power spectrum: baryons are responsible for oscillations-like features in the shape of matter power spectrum and two-point galaxy correlation function, which are the imprints of the acoustic oscillations of the early Universe plasma.

The "wavelength" of BAO is determined by the comoving sound horizon at recombination, thus it is known from CMB measurements. A measurement of apparent size of BAO in redshift or angle space leads to purely geometric quantities such as the Hubble parameter or the angular diameter distance. The mechanism that makes these features a probe of DE is similar to that occurring in the CMB acoustic peaks: since the DE parameters affect the angular diameter distances (through the expansion rate), they will also shift the positions of the peaks in the projected galaxy power spectrum. (A more subtle effect is due to the DE effect on the overall amplitude of the galaxy power spectrum, as DE also enters the growth factor of matter perturbations.) Therefore, these features can be used as CMB-calibrated standard rulers for determining the distances to a given redshift and constraining DE. The main difference with respect to the CMB as a DE probe is that, in principle, we can observe these baryonic features at different redshifts, considering galaxy surveys over different redshift ranges: as a consequence, we can get information not only on the total amount of DE density needed to match the observed angular scales, but also on the time-dependent DE parameters.

BAO have been detected in modern low-redshift surveys, such as 2dF [49] and SDSS [73,179,180]. It will be greatly advantageous to probe higher redshifts, where the linear regime (showing BAO) extends to smaller scales: the most efficient way to measure BAO will rely on deep photometry and large coverage. Future planned and ongoing surveys, such as the ground-based Large Synoptic Survey Telescope (LSST)[34], Visible and Infrared Survey Telescope for Astronomy (VISTA)[35], Subaru [116] and the space-based SuperNova Acceleration Probe (SNAP), would be able to put significant constraints on the DE parameters, which can be improved when combined with CMB constraints expected from the *Planck* mission. DE is also being addressed by two ESA proposals, merged under the name *Euclid*. These are DUNE, a weak gravitational lensing survey in the visible and near infrared, and SPACE, a near infrared imaging and spectroscopic mission to detect BAO in the galaxy power spectrum. NASA and the US department of Energy will support a probe focused on investigating DE, the Joint Dark Energy Mission (JDEM), which will observe type Ia SNe to track the expansion rate evolution.

All these many complementary techniques are required to constrain DE, all of them will play a key role in the coming decade.

[34] See LSST in web page list.

[35] See VISTA in web page list.

4.6.2 Dark Energy and Inflation Analogies?

Dear Francesca *(Perrotta)*, **both inflation and DE involve scalar fields or, more in general, energy components at the basis of an accelerated expansion of the Universe, but at very different cosmic times and energy scales. Do exist possible analogies between these two frameworks?**

Scalar fields are ubiquitous in cosmology; both the cosmology of weakly coupled scalar fields and their theoretical motivations have been much studied since the advent of the idea of inflation. And it is natural to notice the similarity between two processes of expansion rate acceleration: the inflationary phase in the early Universe, occurring at a temperature scale of 10^{19} GeV, and the current DE dominated epoch, taking place in a cold (10^{-12} GeV) Universe, which may be thought as a "late-time inflation". Both of these processes can be realized, under specific conditions, by a scalar field, that is, the inflaton and a Quintessence field. One may be tempted to ask whether they are simply the different faces of the same medal, one revealing the dynamics of the Universe at the smallest physical length scales (the Planck scale), one telling us about the behavior if the Universe on the horizon scales: it is indeed the horizon that poses a limit below which the Quintessence field would be as homogeneous as a cosmological constant. Despite the analogy between the inflaton and a Quintessence field, the last needs not to participate to the inflationary phase. It may well not to be the same field as the inflaton: we can only plausibly assume, based on the observations, that this field should have an energy density that survives the inflationary phase and that comes to dominate the whole energy density at recent times. If Quintessence (or DE, in general) were to dominate early enough, it would suppress the baryonic structure growth on small scales, because the Universe would expand faster than the perturbation can collapse. Unfortunately, we do not have any physical motivation to relate the two fields, or in general to relate the early inflation to the late inflation phase in the cosmic evolution: in part, we are limited by the difficulty in testing theories based on energy scales well beyond the reach of accelerators, which leaves a question mark on the processes actually occurred in the very early Universe. In part, we are still in trouble when trying to relate the macroscopic scales (on which gravity dominates) to the microscopic ones: a complete and satisfactory quantum gravity theory has still to be understood. So for the moment we must drop any pretension to close the circle, until some new hint may eventually come from theorists to make light on this crucial problem. Nonetheless, there are very promising perspectives of "making light on the dark side", thanks to future observations and cosmological probes: if, in the future, we will be able to discern the dynamical DE from a cosmological constant, for example, through an accurate and reliable detection of time variations of the DE equation of state, this would open new perspectives both in cosmology and in particle physics, possibly opening the doors for a new revival of the cosmological constant.

Thank you very much Francesca.
Here below we start our investigation of the proposed alternatives to the standard theories of gravity. First, we ask again Thanu to comment on some curious coincidences that follow from the value of the cosmological constant.

4.7 Alternatives to Standard Gravity Theories

Dear Thanu *(Padmanabhan)***, it has been noticed that the current value of the cosmological constant leads to several numerical coincidences like, for example, the acceleration parameter in MOND is comparable to $a_0 \equiv c^2/L_\Lambda$. What is your opinion regarding these coincidences?**

There are indeed several numerical coincidences that one could come up with in the study of cosmological constant. In addition to the coincidence with the MOdified Newtonian Dynamics (MOND) parameter, it has been pointed out that (1) This acceleration $a_0 \sim$ Å sec^{-2} is also comparable to the Pioneer anomaly – which is an unexplained discrepancy observed in the acceleration of the Pioneer space probe. (2) The length scale corresponding to cosmological constant $L_\Lambda \approx 0.04$ mm – which is macroscopic – is also close to the Compton wavelength associated with a neutrino of mass about 10^{-2} eV.

It is difficult to judge categorically whether these coincidences have any deep significances. Personally, I believe they do *not*. In the case of MOND, there are several well known difficulties in accepting it as a viable alternative to the standard Einstein's theory of gravity. So it is not easy to see how the cosmological constant can in any way be connected with the MOND. Similarly, it is not possible to use the coincidence between the value of the cosmological constant and the Pioneer anomaly and come up with the model for the latter. As regards the neutrino mass, it is just one of the many low energy parameters in the standard model of particle physics. While we do not understand the mass spectrum of fermions at any level of satisfaction, it is nevertheless true that the neutrino mass is not expected to play any deeper role than, say, the electron mass.

Another difficulty in making a viable connection in all these numerical coincidences is the following. In many of the approaches to understand such numerical coincidences, we are not talking about cosmological constant as a strict constant but instead view it as a parameter linked to the current size of the Universe. If one links the cosmological "constant" to the size of the Universe, then it becomes an evolving parameter in the theory. This leads to several deep conceptual and technical difficulties and unless we assume that other parameters (like the neutrino mass) also evolve in time, the numerical coincidence at the present epoch requires special fine tuning. (This is closely related to the question "why now" problem of the cosmological constant mentioned before.) If you want to connect the value of the cosmological constant to other physical parameters and treat *all* of them as time dependent, then one creates further problems for us to solve!

Comment by Moti Milgrom: I would like to take issue with the statement of Padmanabhan. He dismisses the relevance of the coincidence of the MOND acceleration constant a_0 with the cosmological acceleration on the basis of his own disbelief in MOND as a basic theory. In matter of fact, whether MOND is a fundamental theory or not, the very special and central role the constant a_0 plays in galaxy dynamics is well established and is here to stay. For instance, you find it everywhere in the data itself. All round systems, from giant molecular clouds, through globular clusters and elliptical galaxies, to clusters of galaxies lie, in the mass-radius plane, near the line with constant M/R^2. The value of this ratio when multiplied by G gives $a_ø$. Another example: the baryonic Tully–Fisher relation agrees well (over many orders of magnitude in mass) with a relation of the form $\alpha M = V^4$. The proportionality constant α has dimensions of G times acceleration, and when divided by G gives a_0 (this is independent of the previous appearance as it refers to asymptotic regions in disc galaxies). Yet another instance, when one plots the mass discrepancy in disc galaxies as a function of observed acceleration one finds that data from all galaxies lie on the same line (when plotting vs. radius, or other attribute the data points are scattered all over). This discrepancy line has the value 1 for high accelerations and equals a/a_0 for low acceleration. The transition occurs at a_0 so this is also the value of the slope. I have not mentioned the word MOND in all this; but clearly MOND has caused us to uncover all these facts. Many facets of observed galactic dynamics are thus controlled by an acceleration constant that coincides with the cosmic acceleration. What one makes of this is up to ones own way of thinking.

Thank you again Thanu. As you can see, Moti Milgrom promptly reacted to your comments about MOND. So we now ask him to introduce the reader to the pros and cons of this alternative scheme of gravity.

4.7.1 MOND

Dear Moti *(Milgrom)***, your theory MOND has received an increasing attention in the astronomical literature during the last years. Could you focus on the fundamental aspects of the theory and on its successes and failures in explaining astronomical observations?**

MOND [137, 138] is an alternative to Newtonian dynamics that aims at explaining the observations of galactic systems without DM. It greatly departs from standard physics in systems with low accelerations. It posits the appearance of a new "constant of nature" with the dimensions of acceleration, a_0. This appears in a conceptually similar way to that of the Planck constant in quantum theory (or to the speed of light, c, in relativity). So, for example, a_0, like $\hbar$, marks the boundary between the old, classical regime, and the modified regime. The classical regime can be seen as a limit of the full quantum theory when one formally pushes the boundary $\hbar$ to zero in all the equations of physics (or pushing c to infinity to get the classical limit of relativity). In a similar vein, a MOND theory has to approach the classical (non-MOND) theory when a_0 is sent to 0 in all the equation (or in other words, when all the relevant accelerations in the system are much larger than a_0 itself). To implement the desired MOND phenomenology, we require from the theory that in the opposite limit – namely when we send a_0 to infinity in all the equations of physic – we should obtain equations that, for purely gravitational systems, contain a_0 and G only in the combination a_0G. The essence of this is that when accelerations are much smaller than a_0 the relation between mass, distance, and acceleration is not the usual Newtonian expression, $a = MG/r^2$, but $a = \sqrt{MGa_0}/r$. More generally, a simplistic rule that captures the essence of MOND is as follows: If g_N is the acceleration of a test particle calculated at some position in a galactic system with Newtonian dynamics, then the acceleration, g, that MOND predicts is related to it by $\mu(g/a_0)g = g_N$. The function $\mu(x)$ is an interpolating function. Its exact form is not known yet (and is a subject of recent studies), but we know that it has to go to 1 in the limit $x \gg 1$, and that $\mu(x) \approx x$ for $x \ll 1$. Because $\mu < 1$, and sometimes $\mu \ll 1$, the MOND acceleration is larger than the Newtonian value for the same mass and distance. This is how MOND enables us to do away with DM: even with only the baryonic matter that is observed in galaxies, the predicted accelerations from MOND can be larger then the Newtonian value, as is observed. DM is invoked to achieve a similar effect; but it works, of course, in a completely different way.

The above requirements can be incorporated in various MOND theories. For example, in the nonrelativistic regime one can modify the Poisson equation for the gravitational field [14], which would count as modifying gravity, or one can modify the kinetic action of particles leading to modification of Newton's second law, or to modified inertia [140]. Ultimately, one would like to incorporate these principles in a relativistic extension, which is a MOND version of GR. The state of the art of this effort is Bekenstein's Tensor–Vector–Scalar (TeVeS) theory [12].

To further use the analogy with quantum theory, just as $\hbar$ appears in many relations and phenomena in the deep quantum mechanical regime (the Black Body (BB) spectrum, the photoelectric effect, the Hydrogen atomic spectrum, the quantum Hall effect, etc.) so does a_0 appear in many relations that are predicted by MOND, concerning the dynamics of galactic systems (galaxies of all sorts, galaxy groups, clusters, and super-clusters). These relations are mutually independent in the sense that they do not follow from each other in the context of the DM paradigm. In fact, some of them blatantly contradict the predictions of DM, and those that do not would each require a separate explanation in the framework of the DM paradigm. In MOND, the underlying theory unifies them as consequences of the same principle. Most of these laws were looked for and noticed in the data only because MOND predicted them. So MOND has already achieved this one important role of a theory: to direct the eye to previously undiscovered regularities in nature. Some examples: MOND predicts that the velocity on a circular orbit around a mass M becomes independent of the radius of the orbit for large radii (asymptotic flatness of rotation curves) and that the asymptotic velocity, V_∞, depends only on M through $V_\infty^4 = MGa_0$ (see Fig. 4.3). It also predicts that the onset of a mass discrepancy (in a disc galaxy say) always occurs at a radius where $V^2/R \approx a_0$. Likewise, there are

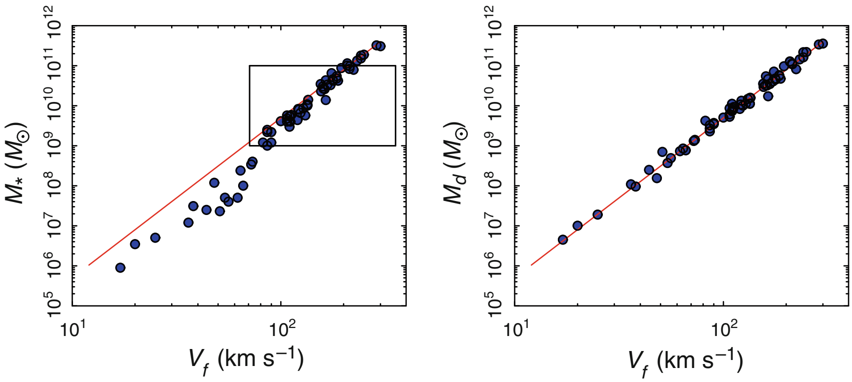

Fig. 4.3 The mass of galaxies in a large sample plotted against the plateau velocity of their rotation curves. (*Left*) The traditional Tully–Fisher plot with mass in stars only (taken proportional to the K-luminosity). (*Right*) The total mass including that of gas as is dictated by MOND's prediction. The *solid line* has the log–log slope of 4, predicted by MOND, and is not a fit. Adapted from [134]

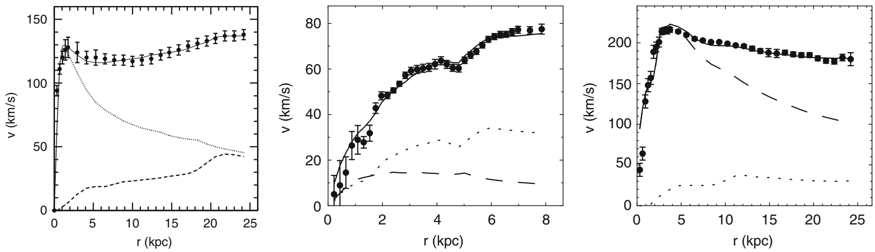

Fig. 4.4 The observed rotation curves (*points*) and MOND curve (*in solid lines*) for NGC 3657 (*left*), NGC 1560 (*center*), and NGC 2903 (*right*). The first from [196], the last two from [197]). The other lines are the Newtonian curves calculated for the stars and gas alone (they add in quadrature to give the full Newtonian curve)

many other MOND laws of galactic motion, analogous to Kepler's laws of planetary motion. These laws are only extractions from the general prediction that MOND makes of the full acceleration field around any object based on only the baryonic matter in the object without need for DM. In particular, MOND predicts correctly the rotation curves of disc galaxies based only on the distribution of their baryons (see Fig. 4.4). For reviews see [13, 144, 197, 202].

MOND is enormously predictive, much more so then Newtonian dynamics with DM. This is because MOND predicts the full gravitational potential of a galactic system from only the distribution of baryonic matter, which can be observed directly. MOND is thus also a much more easily falsifiable paradigm. The flagship of MOND testing concern its predictions for the rotation curves of individual galaxies (over a hundred tested so far), which have been confirmed with great success. I show three examples in Fig. 4.4. With very few exceptions, DM is, of course, not capable of making such predictions for individual galaxies as these would depend on the unknowable history of the object. (What we see in the literature for DM are always best fit results with two parameters for the halo corresponding to the halo size and mass; the MOND curves are basically predictions if we know the baryon distribution.) Most of the major MOND predictions do not depend on the formation and evolution history of the system. In the DM paradigm the relative quantities and distributions of baryons and DM in a system should strongly depend on history, as the two types of matter suffer different influences. This makes it impossible to make predictions for individual system, only perhaps general approximate rules.

MOND makes a large number of predictions in the form of laws of galactic motions (e.g., [144]). I already mentioned three such laws in the introduction: The asymptotic flatness of rotation curves, the mass-asymptotic-speed relation (see Fig. 4.3 for observational confirmation), and the onset of discrepancy at a fixed acceleration, a_0. Others I can mention are a mass–velocity dispersion relation for round systems, underlying the well established Faber–Jackson relation for elliptical galaxies; an upper bound of order a_0/G on the mean surface density of quasi-isothermal spheroidal systems, as indeed is observed for all such astronomical systems from globular clusters, via elliptical galaxies, to clusters of galaxies; and many more such laws, as listed, for example, in [144]. The constant a_0 appears in many of these Kepler-like laws and was determine from them with consistent results.

The only galactic systems where MOND systematically does not fully explain away the mass discrepancy are clusters of galaxies. MOND does reduce substantially the amount of required DM in the cluster at large, by about a factor of 5, but there is a remaining discrepancy typically of a factor of 2 for the whole cluster. We then have to say that galaxy clusters contain some yet undiscovered baryons, about as much as the baryonic mass already observed. This would be only a small fraction of the baryonic budget in the Universe, much of which is still unaccounted for, anyway. These surmised, yet undetected baryons are distributed in a rather more centrally concentrated manner and hence they actually dominate the very core of clusters (within a few hundred kiloparsecs). This fact, which we have known for many years from studies of isolated clusters, also explains the recent observations of the "bullet cluster", which has been claimed to be a counter example to MOND [47].

Can MOND be tested at different astronomical scales, from Solar system to clusters of galaxies? Recently, it has been suggested by Scarpa et al. that Globular Clusters may provide a further test for MOND. The DM content in these objects has always been considered almost negligible. If so, their observed velocity dispersion profiles disagree with the expected Newtonian fall-off behavior but it seems to be in agreement with MOND's prediction. Is this test so robust to prove MOND or several kinds of complications, such as tidal effects or possible DM clumps, may significantly affect it? What is your opinion on this point?

Yes, in principle, MOND can be tested on all these scales. The effects of MOND in the realm of the galaxies are very strong. The predicted discrepancies with Newtonian dynamics can be as large as a factor of $50 \div 100$ in the very outskirts of galaxies now tested by weak lensing, or in sparse galaxy groups. The predicted discrepancies accurately probed with rotation curves of disc galaxies, or in dwarf spheroidal galaxies, routinely reach a factor of 10. In galaxy clusters they reach a factor of 5 approximately. In the Solar system, or on earth (laboratory experiments), the expected effects are very small, as the accelerations are very large. Despite this, possible approaches to testing MOND locally have been discussed. In the solar system, the Pioneer anomaly and its like [152] seem to offer a tool. In fact, the claimed anomaly is just of the order of the MOND acceleration a_0 and there are versions of MOND that can explain the anomaly naturally, without conflicting with the lack of anomaly in the planetary motion. Space-time locations for MOND testing in low acceleration regimes in the Solar system and on earth have also been discussed.

In fact, MOND as it is presently formulated, predicts that we should not see appreciable departures from standard physics for globular clusters at small galactocentric distances. This is because the field of the galaxy – which in these regions is not in the deep MOND regime – erases intrinsic MOND effects. So, I think that the findings of Scarpa et al. are indeed due possibly to tidal effects or other such influences, and do not really reflect MOND effect. But having said that, maybe we are missing something in our knowledge of MOND, since the effect found by Scarpa et al. does seem to start at the MOND acceleration.

Can MOND be included in the actual cosmological context?

MOND is not yet developed enough to deal fully and adequately with cosmology and structure formation, and, in particular, fully address the issue of cosmological DM. But there have been many preliminary works in this direction including suggestions of ways in which relativistic formulations of MOND are capable of eliminating cosmological DM [66, 209].

But having said that, it is exciting to note that MOND itself very strongly smells of cosmology: a very interesting observation, which possibly hints at the deepest and most revolutionary consequence of MOND is that $a_0 \approx cH_0/2\pi \approx c\sqrt{\Lambda/3}/2\pi$, where H_0 is the expansion rate of the Universe – the Hubble constant, and Λ is the measured cosmological constant. This coincidence can be described pictorially by noting that accelerating with a_0 takes one from zero speed to c roughly in the time since the big bang. This means that a_0 coincides with some accelerations of clear cosmological significance.

To understand some of the implications of this coincidence, we recall that as a prelude to quantum gravity one can combine $\hbar$ with G and c to form the so-called Planck length and Planck mass. These indicate where we can expect combined effects of quantum physics with strong gravity. The analogous quantities in MOND are $\ell_0 \equiv c^2/a_0 \approx 10^{29}$ cm, and $M_0 \equiv c^4/Ga_0 \approx 6 \times 10^{23}$ $M_\odot$, which are of the order of the Hubble distance, and the mass within the horizon, respectively. This tell us that combined effects of strong gravity and MOND only appears in connection with the Universe at large, and that there are no local Black Holes (BHs) with surface acceleration as low as a_0.

These coincidences may also point to some deep and strong connection between local MOND physics and cosmology: either the same new physics explain MOND and the DE effects, or MOND is an expression in local physics of the DE, or cosmological constant (in the sense that if there was no DE we would have had $a_0 = 0$, for which MOND reduces to standard dynamics).

I personally believe that the understanding of cosmology in the framework of MOND will come together with, or after, the understanding of how MOND emerges from, or related to, cosmology.

Does there exist a possible physical explanation of why gravity should change its apparent behavior below a certain field strength?

Well, do we understand why quantum effects enter below a certain value of the action as given by Planck's constant?

But actually, in the case of MOND we can bring up possible reasons for the transition occurring at some field strength (or, rather, at a given acceleration scale, a_0). These reasons all have to do with the coincidence that I already mentioned between a_0 and the cosmological acceleration parameters. From different physical points of view, an acceleration of a system, a, carries with it some length scale defined by $\ell = c^2/a$. For instance, this is the transition radius from the near field to the radiation zone for a radiating, accelerated charge. It is the size of the region within which we can erect a locally flat coordinate system for an accelerating observer. It is the typical wavelength of the Unruh radiation[36] for an accelerated observer. Also, an observer with a uniform acceleration carries with it an event horizon of this typical size. Now, because of the cosmological connection I mentioned, for an observer with $a \gg a_0$ this length ℓ is much smaller than the radius of the cosmological horizon, while for $a \ll a_0$ the opposite is true. The former observer may be probing within his ℓ only the local Universe; so in this sense it is oblivious to some details of cosmology; this is not true for the observer having $a \ll a_0$. While we have not been able yet to construct a full theory based on this kind of arguments, they do point to some possibilities. For example, I pointed out [141] the possible relevance of the fact that in a de Sitter Universe controlled by a cosmological constant Λ, a uniformly accelerated observer (of acceleration a), sees an Unruh temperature $\propto \sqrt{a^2 + c^2\Lambda/3}$. This indeed gives a transition in the behavior with a at some special acceleration $\hat{a}_0 \sim c\sqrt{\Lambda/3}$, which may underlie some transition in the dependence of inertia on acceleration at an acceleration scale that is connected with cosmology the way MOND has it. (See more on this in my discussion of the possible origin of MOND below.)

What do you think is the origin of MOND?

Galileo discovered that bodies near the earth's surface fall with a constant acceleration, g_0. This became a new law of physics – incorporating g_0 as a new physical constant – a law that, as we now understand, is relevant for measurements near the Earth's surface. Had Galilei also known the escape speed from the Earth's

[36] An observer in noninertial (accelerated) motion through the quantum vacuum in which we are, finds itself immersed in a field of radiation, known as the Unruh radiation, that is a transformed form of the particle content of the quantum vacuum. This effect is analogous to the Hawking radiation, which is, in a sense, pumped out of the quantum vacuum due to the presence of a black hole. The typical frequency of the Unruh radiation is determined by the magnitude of the acceleration. For example, for an observer that moves on a constant acceleration trajectory the Unruh radiation is thermal with a temperature that is proportional to the acceleration.

surface, V_{es}, or the velocity of a satellite on an Earth grazing orbit, V_{orb} (both measurable, in principle, from the Earth surface) he could have wondered why it is that these are numerically related to g_0 and the Earth's radius, $R_\oplus$. As we know $V_{\text{orb}}^2 = g_0 R_\oplus$ or $V_{\text{es}}^2 = 2g_0 R_\oplus$. Newtonian gravity explained it all, of course, and showed that Galilei's law is an emergent one, and g_0 is not a constant in the more fundamental theory. As I explained earlier, the MOND acceleration a_0, the speed of light, c, and the "radius of the Universe", R_{H}, are related in a very similar way: $c^2 \sim a_0 R_{\text{H}}$. The analogy can be pushed even further: if we are oblivious to Earth gravity (i.e., we do not know, or ignore, the fact that free fall results from the gravitational pull of the Earth) we can conclude that the law of motion near the Earth's surface, relating force $\mathbf{f}$ to acceleration $\mathbf{a}$, is not quite Newton's law, but instead: $\mathbf{f} = m(\mathbf{a} - \mathbf{g}_0) = m\mathbf{a}\mu(\mathbf{a}/g_0)$, where $\mu(\mathbf{a}/g_0) \equiv 1 - \mathbf{n} \otimes (\mathbf{a}/g_0)/(a/g_0)^2$ (Here μ is a tensor, and $\mathbf{n}$ is a unit vector in the direction of $\mathbf{g}_0$). This would appear as a modification of Newton's second law involving the constant g_0 that is formally similar to MOND.

I view this as a hint that MOND as formulated today is only a tip of some iceberg reflecting the workings of a deeper theory; in other words, MOND is an emergent paradigm. Not only that, but the above connection says that the MOND may be rooted in cosmology, and that it results from the influence of cosmology and the state of the Universe at large on local dynamics; or, that there is an agent that enters both. I have been considering the idea that the quantum vacuum is responsible for this connection. The idea that local dynamics is determined by an interaction of local systems with some universal agent has been widely discussed under the general name of Mach's principle. MOND could provide a concrete link of this sort.

In the end, this may turn out to be the deepest, most significant aspect of MOND, above and beyond its introducing new dynamics, and eliminating DM.

Interestingly, this possible connection raises the possibility that a_0 varies with cosmic time. If it is always related to cH_0 in the same way, or if it is related to Λ but this changes, then a_0 would follow suit. It is also possible, however, that Λ is a constant and so is a_0. The notion of a variable a_0 opens up interesting possibilities, as such variations may, in turn, induce secular evolution in galactic systems. Such variations in a_0 could also be in the basis of the cosmological coincidence $\Omega_\Lambda \sim 1$ via anthropic considerations. Also, the general connection of MOND with cosmology, and, in particular, the possibility of a_0 varying may provide a mechanism for connecting the local arrow of time with the cosmological one.

Another feature of existing MOND theories that brings to mind its tentative, effective nature is the appearance of the interpolating function. Its role in MOND is similar to that of the BB function in quantum physics, or the Lorentz factor in relativity. They all interpolate some piece of physics between the old, classical regime and the new, modified regime. They where all introduced first as phenomenological tools. The BB function and the Lorentz factor where later derived from basic theories, and I expect the same to happen with μ within a future underlying theory of MOND.

What is your general opinion about MOND and its significance for the DM paradigm?

The successes of MOND argue against DM simply because they support MOND as a competing paradigm. But, I believe that such success – namely the fact that galactic systems conform to very strict rules and regularities as predicted by MOND – in itself, and without relation to MOND, speak strongly against DM as the explanation of the mass discrepancies. What is the meaning of these successes if we interpret them in terms of DM? They mean that the baryons in galactic systems determine the full gravitational field of these systems. Since, supposedly, DM by far dominates these systems, it follows that MOND's predictions tell us that the baryons alone can be used to determine everything about the amount and distribution of DM, not only as general rules but actually system by system. In other words, given the baryon distribution in a galaxy, one can, using the MOND formula, determine the DM distribution in this particular galaxy. Now, this is quite absurd: the expected relation between baryons and DM in the DM paradigm is very haphazard and strongly dependent on the particular history of formation and evolution of a system. These involve collapse, mergers, energy loss by dissipation, interactions with magnetic fields, cannibalism of dwarf galaxies, ejection of baryons by supernova wind, and by ram pressure stripping, accretion of gas, etc. Baryons and DM respond very differently to these processes; so the end result and the relation between the two components should be strongly dependent on the individual history. In fact, we know now, in support of this conclusion, that the typical baryon-to-DM ratio in galaxies is today an order of magnitude smaller than the cosmic ratio, with which incipient galaxies should have started. The processes that caused 90% of the baryons to be lost have surely caused large scattering in the relative properties of the two components. It is inconceivable, in my opinion, that the many and varied tight connections between the two, which constitute the MOND laws of galactic motion, can be reproduced in a dark matter scenario. Can one imagine, in analogy, that knowledge of the properties of a star uniquely determines all the properties of its planetary system: number of planets and their sizes and orbital radii? Clearly not; the characteristics of the planetary system depend on those of the parent star, but they also depend strongly, as is the case for galactic systems, on the unknown, and unknowable, complex history of the system. MOND makes predictions for individual systems; Newtonian dynamics with DM will never be able to do this, except perhaps for very few systems whose history can be pinpointed. This, in fact, brings me to one example where the CDM paradigm (but not DM in general) does make a definite prediction for individual galaxies. These are the debris dwarf galaxies discussed by Bournaud et al. [25]. As explained by them, these galaxies where formed from the debris of a collision between two galaxies. All past history of the relative baryon to DM is then erased, and one can safely predict in the CDM paradigm that hardly any DM should be found in these dwarfs. Yet, all three dwarfs that were analyzed show significant mass discrepancies, contrary to the predictions of CDM. They do follow MOND's predictions very well [93, 143].

Beside the MOND relations there are very few instances of astronomical correlations satisfied so faithfully by astronomical objects. Those that come to mind such as the (zero-age) main sequence for nonrotating, nonmagnetic stars of a given composition, or the Hubble law, like the MOND laws, are inevitable consequences of laws of physics, not the consequences of complex histories like those of planetary systems or galaxies in the DM scenario. In MOND, you cannot conceive of a galaxy that does not satisfy the MOND laws, just as you cannot conceive of an incipient star (nonrotating, of standard composition, etc.) that does not sit on the main sequence. But you can easily construct Newtonian galaxies with DM that totally disobey these laws.

I feel that all these are severe problem for the CDM paradigm that should be addressed in earnest by its advocates.

In which directions you suggest to focus the efforts of young astronomers that want to face the MOND challenge or similar alternative ideas, both theoretically and observationally?

MOND still leaves a vast range of problems and open questions for people to work on. Perhaps the most exciting are those connected with constructing more fundamental theories of MOND. I believe MOND is an effective theory that must emerge from a more fundamental theory at a deeper stratum. In particular, it will be exciting to understand its possible connection with cosmology. Finding such theories and connections is a great challenge. Even at a more modest level, there is still much progress to be made in developing and understanding relativistic versions of MOND in the spirit of TeVeS and the like, and also developing in full their observational implications. These are theories that are constructed to reproduce MOND phenomenology and to incorporate the usual relativistic principles, but that do not necessarily aim at a deeper understanding of the origins of MOND. People presently working on MOND have made important progress, but we may need people with different backgrounds, different insights, perhaps different ways of thinking, to advance the topic even more.

There is, of course, also still a great volume of phenomenological and observational headway to be made. One strategy may be to concentrate on phenomenological aspects where the predictions of MOND and DM are very distinct. There are in fact specific observations on which the predictions of MOND and DM differ greatly; and these should perhaps get the utmost attention. Some examples:

- I already mentioned the very interesting, if somewhat overlooked, observations of debris dwarf galaxies by Bournaud et al. [25], for which the CDM paradigm predicts no mass discrepancies, while MOND predicts large mass discrepancies if these galaxies have low accelerations. Existing studies point very clearly in favor of MOND on this issue, but we need more cases to study.
- Another sort of observations concerns the degree of flaring of galactic discs. MOND predicts that in addition to the effects similar to those produced by a spheroidal halo, disc galaxies should show a mass discrepancy in the disc itself [142]. CDM does not produce discs and predicts no such disc discrepancy. Existing studies show that to explain the observations of disc flaring one needs either hefty DM discs, or highly oblate DM halos, neither of which is consistent with CDM or with other observations (e.g., [114]). MOND handles these rather well (e.g., [198]). More studies of this problem can be decisive.

- The question of whether galaxy cluster indeed contain some yet undetected baryons, as required by MOND, is also a crucial question to be investigated.
- Continued observations and analysis of galaxy rotation curves is also very important. Failure of MOND in enough clear cut cases can falsify it. On the other hand, continued success will help buttress MOND as an alternative paradigm. Such analysis can also help constrain the exact form of the relevant extrapolating function, and thus constrain theories of MOND.

Thanks a lot Moti for this very objective review of MOND and its perspectives for the future.

We now start to bring the discussion on the proposed alternatives to GR that may solve the problem of invoking a DM and DE contribution in our Universe. Salvatore Capozziello and his group will now explain to us the advantages of the $f(R)$ theories. In this approach, DM becomes a curvature effect. Let us see what they are.

4.7.2 $f(R)$ Theories

Dear Salvatore *(Capozziello)*, **you and your collaborators recently claimed the possibility of a breakdown of the GR theory that may explain the cosmic acceleration of the Universe and other astrophysical problems, such as the flat rotation curves of spiral galaxies, without invoking any form of DE and DM. Can you describe your idea and its main implications? Do exist other observable predictions of your model that are not "ad hoc", that is, expected on the basis of the fact that the theory has been built to explain them?**

Astrophysical observations are pointing out huge amounts of DM and DE needed to explain the observed large scale structures and cosmic accelerating expansion. Up to now, no experimental evidence has been found, at fundamental level, to explain such mysterious components. The problem could be completely reversed considering DM and DE as "shortcomings" of GR and claiming for the "correct" theory of gravity as that derived by matching the largest number of observational data. As a result, accelerating behavior of cosmic fluid and rotation curves of spiral galaxies is reproduced by means of "curvature effects", being gravity an interaction depending on scale. This could be a straightforward explanation why the M/L ratio is increasing with the size of astrophysical self-gravitating systems.

The impressive amount of good quality data of last decade has shed new light on the effective picture of the Universe. Type Ia SNe, anisotropies in the CMB, and matter power spectrum inferred from large galaxy surveys represent the strongest evidences for a radical revision of the Cosmological Standard Model. In particular, the Concordance ΛCDM Model predicts that baryons contribute only $\sim$4% of the total matter–energy budget, while the exotic CDM represents the bulk of the matter content ($\sim$25%) and the cosmological constant Λ plays the role of the so called DE ($\sim$70%).

Although being the best fit to a wide range of data, the ΛCDM model is affected by strong theoretical shortcomings that have motivated the search for alternative models [52]. DE models mainly rely on the implicit assumption that Einstein's GR is the correct theory of gravity.

Nevertheless, its validity on the larger astrophysical and cosmological scales has never been tested, and it is therefore conceivable that both cosmic speed up and DM represent signals of a breakdown in our understanding of the gravitational interaction. Following this line of thinking, the choice of a generic function $f(R)$ as the gravitational Lagrangian, where R is the Ricci scalar, can be derived by matching the data and by the "economic" requirement that no exotic ingredients have to be added. This is the underlying philosophy of what are referred to as $f(R)$ gravity [32, 35]. From a theoretical standpoint, different issues suggest that higher order terms must necessarily enter the gravity Lagrangian. In fact, such terms come out as one-loop corrections in field quantization on curved space–times and they seem unescapable in any perturbation approach to achieve a self-consistent theory of quantum gravity [22, 26].

It is worth noting that Solar system experiments show the validity of GR at these scales so that $f(R)$ theories should not differ too much from GR at this level [157]. In other words, the parameterized post-Newtonian limit of such models must not violate the experimental constraints on Eddington parameters. A positive answer to this request has been recently achieved for several $f(R)$ theories [34], nevertheless it has to be remarked that this debate is far to be definitively concluded. Although higher order gravity theories have received much attention in cosmology, as they are naturally able to give rise to the accelerating expansion (both in the late [43] and in the early [218] Universe), it is possible to demonstrate that $f(R)$ theories can also play a major role at astrophysical scales [39]. In fact, modifying the gravity action can affect the gravitational potential in the low energy limit.

Provided that the modified potential reduces to the Newtonian one on the Solar system scale, this implication could represent an intriguing opportunity rather than a shortcoming for $f(R)$ theories. In fact, a corrected gravitational potential could offer the possibility to fit galaxy rotation curves without the need of DM. In addition, one could work out a formal analogy between the corrections to the Newtonian potential and the usually adopted DM models. To investigate the consequences of $f(R)$ theories on both cosmological and astrophysical scales, let us first remind the basics of this approach.

4.7.3 DE as a Curvature Effect

From a mathematical viewpoint, $f(R)$ theories generalize the Hilbert–Einstein Lagrangian $\mathcal{L}_{\rm HE} = \sqrt{-g}R$ as $\mathcal{L} = \sqrt{-g}f(R)$, without assuming a priori the functional form of Lagrangian density in the Ricci scalar. The field equations are obtained by varying with respect to the metric components to get [32]:

$$f'(R)R_{\alpha\beta} - \frac{1}{2}f(R)g_{\alpha\beta} = f'(R)^{;\mu\nu}\left(g_{\alpha\mu}g_{\beta\nu} - g_{\alpha\beta}g_{\mu\nu}\right) + T^{\mathrm{M}}_{\alpha\beta}, \tag{4.11}$$

where the prime denotes derivative with respect to the argument and $T^{\mathrm{M}}_{\alpha\beta}$ is the standard matter stress–energy tensor. Defining the *curvature stress–energy tensor* as

$$T^{curv}_{\alpha\beta} = \frac{1}{f'(R)}\left\{\frac{1}{6}g_{\alpha\beta}\left[f(R) - Rf'(R)\right] + f'(R)^{;\mu\nu}(g_{\alpha\mu}g_{\beta\nu} - g_{\alpha\beta}g_{\mu\nu})\right\}. \tag{4.12}$$

Equation (4.11) may be recast in the Einstein-like form as

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R = T^{\mathrm{curv}}_{\alpha\beta} + T^{M}_{\alpha\beta}/f'(R) \tag{4.13}$$

where matter non-minimally couples to geometry through the term $1/f'(R)$. The presence of term $f'(R)_{;\mu\nu}$ renders the equations of fourth order, while, for $f(R) = R$, the curvature stress–energy tensor $T^{\mathrm{curv}}_{\alpha\beta}$ identically vanishes and (4.13) reduce to the standard second-order Einstein field equations. As it is clear, from (4.13), the curvature stress–energy tensor formally plays the role of a further source term in the field equations so that its effect is the same as that of an effective fluid of purely geometrical origin.

However, the metric variation is just one of the approaches towards $f(R)$ gravity: in fact, one can face the problem also considering the so-called Palatini approach (see e.g., [127,214]) where the metric and connection fields are considered independent. Apart from some differences in the interpretation, one can deal with a fluid of geometric origin in these case as well.

The scheme outlined above provides all the ingredients we need to tackle with the dark side of the Universe. Depending on the scales, such a curvature fluid can play the role of DM and DE. From the cosmological point of view, in the standard framework of a spatially flat homogenous and isotropic Universe, the cosmological dynamics is determined by its energy budget through the Friedmann equations. In particular, the cosmic acceleration is achieved when the right hand side of the acceleration equation remains positive (in physical units with $8\pi G = c = 1$):

$$\frac{\ddot{a}}{a} = -\frac{1}{6}\left(\rho_{tot} + 3p_{tot}\right), \tag{4.14}$$

where a is the scale factor, $H = \dot{a}/a$ the Hubble parameter, the dot denotes derivative with respect to cosmic time, and the subscript tot denotes the sum of the curvature fluid, and the matter contribution to the energy density and pressure. From the above relation, the acceleration condition, for a dust dominated model, leads to

$$\rho_{\mathrm{curv}} + \rho_M + 3p_{\mathrm{curv}} < 0 \rightarrow w_{\mathrm{curv}} < -\frac{\rho_{\mathrm{tot}}}{3\rho_{\mathrm{curv}}} \tag{4.15}$$

so that a key role is played by the effective quantities

$$\rho_{\rm curv} = \frac{8}{f'(R)} \left\{ \frac{1}{2} \left[f(R) - R f'(R) \right] - 3H\dot{R} f''(R) \right\} , \tag{4.16}$$

and

$$w_{\rm curv} = -1 + \frac{\ddot{R} f''(R) + \dot{R} \left[\dot{R} f'''(R) - H f''(R) \right]}{\left[f(R) - R f'(R) \right]/2 - 3H\dot{R} f''(R)} . \tag{4.17}$$

As a first simple choice, one may neglect ordinary matter and assume a power-law form $f(R) = f_0 R^n$, with n a real number, which represents a straightforward generalization of the Einstein GR in the limit $n = 1$. One can find power-law solutions for $a(t)$ providing a satisfactory fit to the type Ia SNe data and a good agreement with the estimated age of the Universe in the range $1.366 < n < 1.376$ [36]. On the other side, one can develop the same analysis in presence of the ordinary matter component, although in such a case, one has to numerically solve field equations. Then, it is still possible to confront the Hubble flow described by such a model with the Hubble diagram of type Ia SNe.

The data fit turns out to be significant (see Fig. 4.5) improving the χ^2 value and, it fixes the best fit value at $n = 3.46$ when it is accounted only the baryon contribute $\Omega_b \approx 0.04$ (according with BBN prescriptions). It has to be remarked that considering DM does not modify the result of the fit, supporting the assumption of no need for DM in this model. From the evolution of the Hubble parameter in term of redshift one can even calculate the age of the Universe. The best fit value $n = 3.46$ provides $t_{univ} \approx 12.41$ Gyr. It is worth noting that considering $f(R) = f_0\, R^n$ gravity represents only the simplest generalization of Einstein theory.

In other words, it has to be considered that R^n-gravity represents just a working hypothesis as there is no overconfidence that such a model is the correct final gravity theory. In a sense, we want only to suggest that several cosmological and

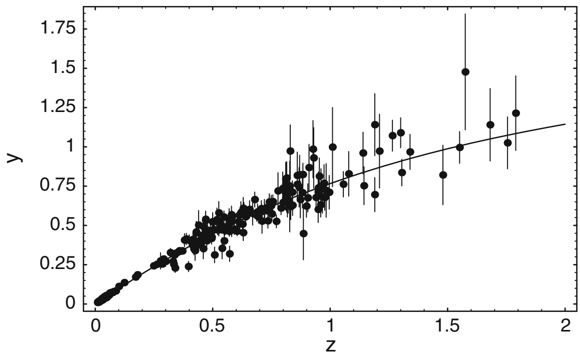

Fig. 4.5 The Hubble diagram of 20 radio galaxies together with the "gold" sample of type Ia SNe, in term of the redshift as suggested in [57]. The best fit curve refers to the $f(R)$-gravity model without DM

astrophysical results can be well interpreted in the realm of a power law extended gravity model. As matter of fact, this approach gives no rigidity about the value of the power n, although it would be preferable to determine a model capable of working at different scales. Furthermore, we do not expect to be able to reproduce the whole cosmological phenomenology by means of a simple power law model, which has been demonstrated to be not sufficiently versatile [38].

For example, we can easily demonstrate that this model fails when it is analyzed with respect to its capability of providing the correct evolutionary conditions for the perturbation spectra of matter overdensity [232]. This point is typically addressed as one of the most important issues, which suggest the need for DM. In fact, if one wants to discard this component, it is crucial to match the observational results related to the LSS of the Universe and the CMB which show, respectively, at late time and at early time, the signature of the initial matter spectrum. As important remark, we note that the quantum spectrum of primordial perturbations, which provides the seeds of matter perturbations, can be positively recovered in the framework of R^n-gravity. In fact, $f(R) \propto R^2$ can represent a viable model with respect to CMB data and it is a good candidate for cosmological Inflation. To develop the matter power spectrum suggested by this model, we resort to the equation for the matter contrast obtained in [232] in the case of fourth order gravity. This equation can be deduced considering the conformal Newtonian gauge for the perturbed metric [232]:

$$ds^2 = (1 + 2\psi)\,dt^2 - a^2(1 + 2\phi)\,\Sigma_{i=1}^{3}(dx^i). \tag{4.18}$$

In GR, it is $\phi = -\psi$, since there is no anisotropic stress; in extended gravity, this relation breaks, in general, and the $i \neq j$ components of field equations give new relations between ϕ and ψ. In particular, for $f(R)$ gravity, due to nonvanishing $f_{R;i;j}$ (with $i \neq j$), the $\phi - \psi$ relation becomes scale dependent. Instead of the perturbation equation for the matter contrast δ, we provide here its evolution in term of the growth index $f = d\ln\delta/d\ln a$, that is the directly measured quantity at $z \sim 0.15$:

$$f'(a) - \frac{f(a)^2}{a} + \left[\frac{2}{a} + \frac{1}{a}E'(a)\right] f(a) - \frac{1-2Q}{2-3Q} \cdot \frac{3\Omega_m\, a^{-4}}{n\, E(a)^2\, \tilde{R}^{n-1}} = 0, \tag{4.19}$$

$E(a) = H(a)/H_0$, $\tilde{R}$ is the dimensionless Ricci scalar, and

$$Q = -\frac{2 f_{RR}\, c^2\, k^2}{f_R\, a^2}. \tag{4.20}$$

For $n = 1$ the previous expression gives the ordinary growth index relation for the Cosmological Standard Model. It is clear, from (4.19), that such a model suggests a scale dependence of the growth index which is contained into the corrective term Q so that, when $Q \to 0$, this dependence can be reasonably neglected. In the most general case, one can resort to the limit $aH < k < 10^{-3} h\,\mathrm{Mpc}^{-1}$, where (4.19) is

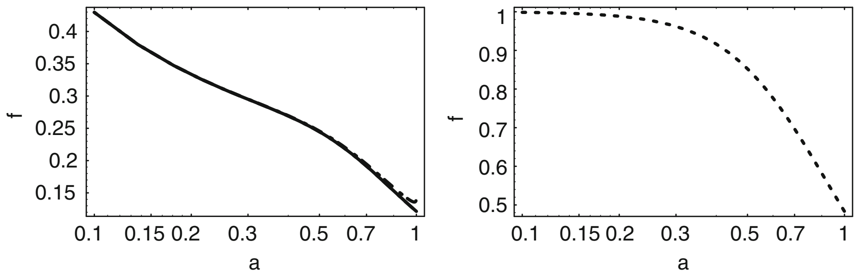

Fig. 4.6 Scale factor evolution of the growth index f: (*left*) modified gravity, in the case $\Omega_{\rm m} = \Omega_{bar} \sim 0.04$, for the type Ia SNe best fit model with $n = 3.46$, (*right*) the same evolution in the case of a ΛCDM model. In the case of R^n-gravity it is shown also the dependence on the scale k. The three cases $k = 0.01,\ 0.001,\ 0.0002$ have been checked. Only the latter case shows a very small deviation from the leading behavior

a good approximation, and nonlinear effects on the matter power spectrum can be neglected.

Studying numerically (4.19), one obtains the growth index evolution in term of the scale factor; for the sake of simplicity, we assume the initial condition $f(a_{ls}) = 1$ at the last scattering surface as in the case of matter-like domination. The results are summarized in Fig. 4.6, where we show, in parallel, the growth index evolution in R^n-gravity and in the ΛCDM model.

In the case of $\Omega_{\rm m} = \Omega_{bar} \sim 0.04$, one can observe a strong disagreement between the expected rate of the growth index and the behavior induced by power law fourth order gravity models. This negative result is evidenced by the predicted value of $f(a_{z=0.15})$, which has been observationally estimated by the analysis of the correlation function for 2.2×10^5 galaxies in 2dFGRS dataset sample at the survey effective depth $z = 0.15$. The observational result suggests $f = 0.58 \pm 0.11$ [120], while our model gives $f(a_{z=0.15}) \sim 0.117\ (k = 0.01),\ 0.117\ (k = 0.001),\ 0.122 (k = 0.0002)$. Although this result seems frustrating with respect to the underlying idea to discard the dark components from the cosmological dynamics, it does not give substantial improvement in the case of R^n-gravity model plus DM. In fact, it is possible to show that, even in this case, the growth index prediction is far to be in agreement with the ΛCDM model and again, at the observational scale $z = 0.15$, there is not enough growth of perturbations to match the observed LSS. In such a case one obtains $f(a_{z=0.15}) \sim 0.29\ (k = 0.01),\ 0.29\ (k = 0.001),\ 0.31\ (k = 0.0002)$, which are quite increased with respect to the previous case but still very far from the experimental estimate. It is worth noting that no significantly different results are obtained if one varies the power n, of course in the case of $n \to 1$ one recovers the standard behavior if a cosmological constant contribution is added. These results seem to suggest that an extended gravity model that considers a simple power law of Ricci scalar, although cosmologically relevant at late times, is not viable to describe the evolution of Universe at all scales. In other words, such a scheme seems too simple to give account for the whole cosmological phenomenology. In fact, in [232] a gravity Lagrangian considering an exponential correction to the Ricci scalar $f(R) = R + A \exp(-B\ R)$ (with A, B two constants) gives more interesting results and displays a grow factor rate, which is in agreement with the observational results at least in the DM case. To corroborate this point of view, one has to consider that when the choice of $f(R)$ is performed starting from observational data (pursuing an inverse approach) as in [37], the reconstructed Lagrangian is a non-trivial polynomial in term of the Ricci scalar. A result that directly suggests that the whole cosmological phenomenology can be accounted only with a suitable nontrivial function of the Ricci scalar rather than a simple power law function. As matter of fact, the results obtained with respect to the study of the matter power spectra in the case of R^n-gravity do not invalidate the whole approach, since they can be referred to the too simple form of the model.

4.7.4 DM as a Curvature Effect

The results obtained at cosmological scales motivates further analysis of $f(R)$ theories. In a sense, one is wondering whether the curvature fluid, which works as DE, can also play the role of effective DM thus yielding the possibility of recovering the observed astrophysical phenomenology by the only visible matter. It is well known that, in the low energy limit, higher order gravity implies a modified gravitational potential. Therefore, in our discussion, a fundamental role is played by the new gravitational potential descending from the given fourth order gravity theories we are referring to. By considering the case of a pointlike mass m and solving the vacuum field equations for a Schwarzschild-like metric, one gets from a theory $f(R) = f_0 R^n$, the modified gravitational potential [39]:

$$\Phi(r) = -\frac{Gm}{2r}\left[1 + \left(\frac{r}{r_c}\right)^{\beta}\right], \tag{4.21}$$

where

$$\beta = \frac{12n^2 - 7n - 1 - \sqrt{36n^4 + 12n^3 - 83n^2 + 50n + 1}}{6n^2 - 4n + 2}, \tag{4.22}$$

which corrects the ordinary Newtonian potential by a power-law term. In particular, this correction sets in on scales larger than r_c, value of which depends essentially on the mass of the system. The corrected potential, see (4.21), reduces to the standard $\Phi \propto 1/r$ for $n = 1$ as it can be seen from the relation in (4.21).

The result in (4.21) deserves some comments. As discussed in detail in [39], we have assumed the spherically symmetric metric and imposed it into the field (4.11) considered in the weak field limit approximation. As a result, we obtain a corrected Newtonian potential, which accounts for the strong nonlinearity of gravity related to the higher-order theory. However, we have to notice that Birkhoff's theorem does not hold, in general, for $f(R)$ gravity (see [40] for a demonstration) but other spherically symmetric solutions than the Schwarzschild one can be found in these extended theories of gravity [41].

The generalization of (4.21) to extended systems is achieved by dividing the system in infinitesimal mass elements and summing up the potentials generated by each single element. In the continuum limit, we replace the sum with an integral over the mass density of system taking care of eventual symmetries of the mass distribution (see [39] for details). Once the gravitational potential has been computed, one may evaluate the rotation curve $v_c^2(r)$ and compare it with the data. For extended systems, one has typically to resort to numerical techniques, but the main effect may be illustrated by the rotation curve for the pointlike case, that is:

$$v_c^2(r) = \frac{Gm}{2r}\left[1 + (1-\beta)\left(\frac{r}{r_c}\right)^{\beta}\right]. \tag{4.23}$$

Compared with the Newtonian result $v_c^2 = Gm/r$, the corrected rotation curve is modified by the addition of the second term in the right hand side of (4.22). For $0 < \beta < 1$, the corrected rotation curve is higher than the Newtonian one. As measurements of spiral galaxies rotation curves signals a circular velocity higher than those which are predicted on the basis of the observed luminous mass and the Newtonian potential, the above result suggests the possibility that our modified gravitational potential may fill the gap between theory and observations without the need of additional DM.

It is worth noting that the corrected rotation curve is asymptotically vanishing as in the Newtonian case, while it is usually claimed that observed rotation curves are flat (i.e., asymptotically constant). Actually, observations do not probe v_c up to infinity, but only show that the rotation curve is flat within the measurement uncertainties up to the last measured point. This fact by no way excludes the possibility that v_c goes to zero at infinity. In order to observationally check the above result, we have considered a sample of low surface brightness galaxies with well measured HI + Hα rotation curves extending far beyond the visible edge of the system. Low surface brightness galaxies are known to be ideal candidates to test DM models since, because of their high gas content, the rotation curves can be well measured and corrected for possible systematic errors by comparing 21 cm HI line emission with optical Hα and [NII] data. Moreover, they are supposed to be DM dominated so that fitting their rotation curves without this elusive component is a strong evidence in favor of any successful alternative theory of gravity.

Our sample contains 15 low surface brightness galaxies with data on both the rotation curve, the surface mass density of the gas component and R-band disk photometry extracted from a larger sample selected by de Blok and Bosma [60]. We assume the stars are distributed in an infinitely thin and circularly symmetric disk with surface density $\Sigma(r) = \Upsilon_\star I_0 exp(-r/r_d)$, where the central surface luminosity I_0 and the disk scalelength r_d are obtained from fitting to the stellar photometry. The gas surface density has been obtained by interpolating the data over the range probed by HI measurements and extrapolated outside this range. When fitting to the theoretical rotation curve, there are three quantities to be determined, namely the stellar mass-to-light (M/L) ratio, $\Upsilon_\star$ and the theory parameters (β, r_c). It is worth stressing that, while fit results for different galaxies should give the same β, r_c is related to one

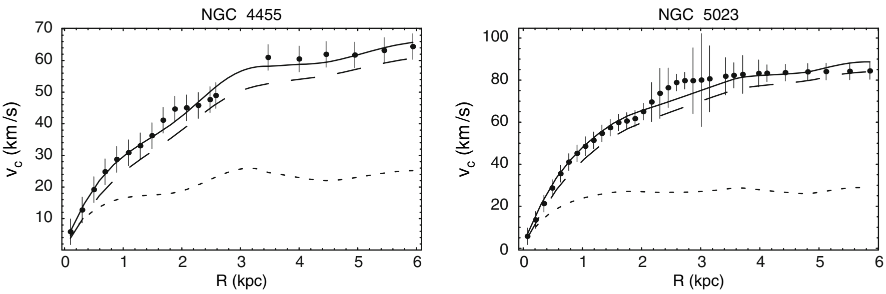

Fig. 4.7 Best fit theoretical rotation curve superimposed to the data for the low surface brightness galaxy NGC 4455 (*left*) and NGC 5023 (*right*). To better show the effect of the correction to the Newtonian gravitational potential, we report the total rotation curve $v_c(r)$ (*solid line*), the Newtonian one (*short dashed*) and the corrected term (*long dashed*)

of the integration constants of the field equations. As such, it is not a universal quantity and its value must be set on a galaxy-by-galaxy basis. However, it is expected that galaxies having similar properties in terms of mass distribution have similar values of r_c so that the scatter in r_c must reflect somewhat the scatter in the circular velocities. In order to match the model with the data, we perform a likelihood analysis determining for each galaxy, using, as fitting parameters β, $\log r_c$ (with r_c in kpc) and the gas mass fraction[37] f_g. As it is evident considering the results from the different fits, the experimental data are successfully fitted by the model (see [39] for details). In particular, for the best fit range of β ($\beta = 0.80 \pm 0.08$), one can conclude that R^n gravity with $2.3 < n < 5.3$ (best fit value $n = 3.2$ which well overlaps the above mentioned range of n fitting type Ia SNe Hubble diagram) can be a good candidate to solve the missing matter problem in low surface brightness galaxies without any DM.

At this point, it is worth wondering whether a link may be found between R^n gravity and the standard approach based on DM haloes as both theories fit equally well the same data. As a matter of fact, it is possible to define an *effective DM halo* by imposing that its rotation curve equals the correction term to the Newtonian curve induced by R^n gravity. Mathematically, one can split the total rotation curve derived from R^n gravity as $v_c^2(r) = v_{c,N}^2(r) + v_{c,corr}^2(r)$ where the second term is the correction. Considering, for simplicity a spherical halo embedding a thin exponential disk, we may also write the total rotation curve as $v_c^2(r) = v_{c,disk}^2(r) + v_{c,DM}^2(r)$ with $v_{c,disk}^2(r)$ the Newtonian disk rotation curve and $v_{c,DM}^2(r) = GM_{DM}(r)/r$ the DM one, $M_{DM}(r)$ being its mass distribution. Equating the two expressions, we get

$$M_{DM}(\eta) = M_{vir}\left(\frac{\eta}{\eta_{vir}}\right)\frac{2^{\beta-5}\eta_c^{-\beta}(1-\beta)\eta^{\frac{\beta-5}{2}}\mathcal{I}_0(\eta) - \mathcal{V}_d(\eta)}{2^{\beta-5}\eta_c^{-\beta}(1-\beta)\eta^{\frac{\beta-5}{2}}\mathcal{I}_0(\eta_{vir}) - \mathcal{V}_d(\eta_{vir})}. \tag{4.24}$$

[37] This is related to the M/L ratio as $\Upsilon_\star = [(1-f_g)M_g]/(f_g L_d)$ with $M_g = 1.4 M_{HI}$ the gas (HI + He) mass, $M_d = \Upsilon_\star L_d$ and $L_d = 2\pi I_0 r_d^2$ the disk total mass and luminosity.

with $\eta = r/r_d$, $\Sigma_0 = \Upsilon_\star i_0$, $\mathcal{V}_d(\eta) = I_0(\eta/2)K_0(\eta/2) \times I_1(\eta/2)K_1(\eta/2)$[38] and

$$\mathcal{I}_0(\eta, \beta) = \int_0^\infty \mathcal{F}_0(\eta, \eta', \beta) k^{3-\beta} \eta'^{\frac{\beta-1}{2}} e^{-\eta'} d\eta' \tag{4.25}$$

with $\mathcal{F}_0$ only depending on the geometry of the system and "vir", indicating virial quantities. Equation (4.24) defines the mass profile of an effective spherically symmetric DM halo whose ordinary rotation curve provides the part of the corrected disk rotation curve due to the addition of the curvature corrective term to the gravitational potential. It is evident that, from an observational viewpoint, there is no way to discriminate between this dark halo model and R^n gravity.

Having assumed spherical symmetry for the mass distribution, it is immediate to compute the mass density for the effective dark halo as $\rho_{DM}(r) = (1/4\pi r^2) dM_{DM}/dr$. The most interesting features of the density profile are its asymptotic behaviors that may be quantified by the logarithmic slope $\alpha_{DM} = d\ln \rho_{DM}/d\ln r$, which can be computed only numerically as function of η for fixed values of β (or n). As expected, α_{DM} depends explicitly on β, while (r_c, Σ_0, r_d) enter indirectly through η_{vir}. The asymptotic values at the center and at infinity denoted as α_0 and α_∞ result particularly interesting. It turns out that α_0 almost vanishes so that in the innermost regions the density is approximately constant. Indeed, $\alpha_0 = 0$ is the value corresponding to models having an inner core such as the cored isothermal sphere and the Burkert model [28].

Moreover, it is well known that galactic rotation curves are typically best fitted by cored dark halo models. On the other hand, the outer asymptotic slope is between −3 and −2, that are values typical of most dark halo models in literature. In particular, for $\beta = 0.80$ one finds $(\alpha_0, \alpha_\infty) = (-0.002, -2.41)$, which are quite similar to the value for the Burkert model (0, −3). It is worth noting that the Burkert model has been empirically proposed to provide a good fit to the low surface brightness and dwarf galaxies rotation curves. The values of $(\alpha_0, \alpha_\infty)$ we find for the best fit effective dark halo therefore suggest a possible theoretical motivation for the Burkert-like models. Because of the construction, the properties of the effective DM halo are closely related to the disk one. As such, we do expect some correlation between the dark halo and the disk parameters. To this aim, exploiting the relation between the virial mass and the disk parameters, one can obtain a relation for the Newtonian virial velocity $V_{vir} = GM_{vir}/r_{vir}$

$$M_d \propto \frac{(3/4\pi\delta_{th}\Omega_m\rho_{crit})^{\frac{1-\beta}{4}} r_d^{\frac{1+\beta}{2}} \eta_c^\beta}{2^{\beta-6}(1-\beta)G^{\frac{5-\beta}{4}}} \frac{V_{vir}^{\frac{5-\beta}{2}}}{\mathcal{I}_0(V_{vir}, \beta)}. \tag{4.26}$$

We have numerically checked that (4.26) may be well approximated as $M_d \propto V_{vir}^a$, which has the same formal structure as the Baryonic Tully–Fisher (BTF) relation $M_b \propto V_{flat}^a$ with M_b the total (gas + stars) baryonic mass and V_{flat} the circular velocity on the flat part of the observed rotation curve. To test whether the BTF can

[38] Here I_l and K_l, with $l = 1, 2$ are the Bessel functions of first and second type.

be explained, thanks to the effective DM halo we are proposing, we should look for a relation between $V_{\rm vir}$ and $V_{\rm flat}$. This is not analytically possible since the estimate of $V_{\rm flat}$ depends on the peculiarities of the observed rotation curve such as how far it extends and the uncertainties on the outermost points. For given values of the disk parameters, we therefore simulate theoretical rotation curves for some values of $r_{\rm c}$ and measure $V_{\rm flat}$ finally choosing the fiducial value for $r_{\rm c}$ that gives a value of $V_{\rm flat}$ as similar as possible to the measured one. Inserting the relation thus found between $V_{\rm flat}$ and $V_{\rm vir}$ into (4.26) and averaging over different simulations, we finally get

$$\log M_{\rm b} = (2.88 \pm 0.04) \log V_{\rm flat} + (4.14 \pm 0.09) \tag{4.27}$$

while a direct fit to the observed data gives [133]

$$\log M_{\rm b} = (2.98 \pm 0.29) \log V_{\rm flat} + (3.37 \pm 0.13). \tag{4.28}$$

The slope of the predicted and observed BTF are in good agreement, thus leading further support to our approach. The zeropoint is markedly different with the predicted one being significantly larger than the observed one. However, it is worth stressing that both relations fit the data with similar scatter. A discrepancy in the zeropoint can be due to our approximate treatment of the effective halo, which does not take into account the gas component. Neglecting this term, we should increase the effective halo mass and hence $V_{\rm vir}$ which affects the relation with $V_{\rm flat}$ leading to a higher than observed zeropoint. Indeed, the larger is $M_{\rm g}/M_{\rm d}$, the more the points deviate from our predicted BTF thus confirming our hypothesis. Given this caveat, we can conclude, with confidence, that R^n gravity offers a theoretical foundation even for the empirically found BTF relation.

Although the results outlined along this paper are referred to a simple choice of fourth order gravity models ($f(R) = f_0 R^n$), they could represent an interesting paradigm. In fact, even if such a model is not suitable to provide the correct form of the matter power spectra, and this suggests that a more complicated Lagrangian is needed to reproduce the whole dark sector phenomenology at all scales, we have shown that considering extensions of GR can allow to explain some important issues of cosmological and astrophysical phenomenology. We have seen that extended gravity models can reproduce type Ia SNe Hubble diagram without DM, giving significant predictions even with regard to the age of Universe. In addition, the modification of the gravitational potential which arises as a natural effect in the framework of higher order gravity can represent a fundamental tool to interpret the flatness of rotation curves of low surface brightness galaxies. Furthermore, if one considers the model parameters settled by the fit over the observational data on low surface brightness rotation curves, it is possible to construct a phenomenological analogous of DM halo whose shape is similar to the one of the Burkert model. As the Burkert model has been empirically introduced to give account of the DM distribution in the case of low surface brightness and dwarf galaxies, this result could represent an interesting achievement since it gives a theoretical foundation to such a model.

By investigating the relation among dark halo and the disk parameters, we have deduced a relation between $M_{\rm d}$ and $V_{\rm flat}$, which reproduces the baryonic

Tully–Fisher. In fact, exploiting the relation between the virial mass and the disk parameters, one can obtain a relation for the virial velocity, which can be satisfactorily approximated as $M_\mathrm{d} \propto V_\mathrm{vir}^a$. Even such a result seems very intriguing as it gives again a theoretical interpretation for a phenomenological relation. As a matter of fact, although not definitive, these phenomenological issues on $f(R)$ can represent a viable approach for future investigations and in particular support the quest for a unified view of the dark side of the Universe.

In summary, these results motivate a careful search for a fundamental theory of gravity able to explain the full cosmic dynamics with the only main ingredients which we can directly experience, namely the background gravity, the baryonic matter, the radiation, and the neutrinos. However, the results outlined along this paper are referred to simple choices of $f(R)$, while it is likely that a more complicated Lagrangian is needed to reproduce the whole dark sector phenomenology at all scales. Nevertheless, although not definitive, these achievements represent an intriguing issue for more exhaustive, future investigations in the same track of GR which, a part some shortcomings, is an excellent theory and a well defined standpoint of modern physics. In particular, exploiting such models can reveal a useful approach to motivate a more careful search for a fundamental theory of gravity working not only at ultraviolet scales (quantum gravity) but also at infrared scales (Galaxy, large scale structure and cosmology).

What are the consequences of your model for particle physics framework(s) and for current experiments in this direction?

Another consequence of such an approach is for particle physics and experiments in this direction. As we said, one of the main conundrums of recent particle physics is that no direct evidence of DM and DE candidates has been found out at a fundamental level, also if the *macroscopic* evidences of such ingredients are well known and well tested. This lack could be, as in the case of ether at the beginning of last century, due to the fact that such candidates cannot be revealed simply because they do not exist. If this is the situation, people should seriously face the problem to revise the theory of gravity and try to explain astronomical observations only considering particles and interactions which are *really* observed.

On the other hand, it is interesting to point out that several other observable predictions exist, which are not "ad hoc" for the models and could be the *experimentum crucis* to definitively retain or rule out extended theories of gravity. In particular, the massive modes of gravitational waves, the neutrino oscillations and the anomalous acceleration of spacecrafts outside the Solar system.

In the first case, extending GR, for example assuming $f(R)$-fourth order gravity, naturally leads to extra polarizations and massive modes for gravitational waves as soon as the post-Minkowskian limit of the theory is considered. If the forthcoming space and ground-based interferometric experiments, as *Virgo*, LIGO (Laser Interferometer Gravitational-Wave Observatory) or LISA (Laser Interferometer Space Antenna), will reveal such new features, this will be a very strong evidence to support such theories vs. GR [42].

Secondly, the mechanism of neutrino oscillations could be not an intrinsic feature, but it could be induced by gravitational coupling. In this case, stringent constraints on oscillations could be used to select the true theory of gravity [33].

A third direct evidence could come from the anomalous acceleration revealed by spacecrafts outside Solar system [2]. Such a feature cannot be framed by the experimental errors so then a natural explanation could come from searching for theories of gravity, which do not give the standard Newtonian potential in the weak field limit but correct it by polynomial or exponential (Yukawa-like) scale-dependent terms. This could be a natural way to address such a phenomenon [21, 40].

What is the difference between your approach and MOND?

This approach is completely different from MOND as it is related to a field theory and only claims for a generalization of GR, which should be better tested at other scales. More precisely, the differences with MOND consist on the facts that our modified potentials come directly from the post-Newtonian limit of the theory, exactly as in GR, and no phenomenological parameter is introduced. Essentially, we ask for an action of gravity, which could be something else with respect to the Hilbert-Einstein one, but the foundations of the theory are the same (see [35] for details). On the other hand, MOND and its generalizations do not come from a well posed field theory and several assumptions are often "ad hoc". In this sense, the paradigm of $f(R)$ gravity is the same of GR while it is not so for MOND. Besides, the effective action of any unification scheme can be recast as an extended theory of gravity, in particular an effective $f(R)$-theory [35], so the approach is consistent with any field theory.

To conclude, we think that such an approach is fully consistent with the Galilean spirit since only data and self-consistent theories should be retained without introducing further "ad hoc" ingredients[39].

Thank you very much Salvatore. Now, we will review the nice properties of the conformal theory of gravity, which according to Philip Mannheim can solve both the DM/DE problems and the quantum gravity question. In the conformal theory, gravity has an additional symmetry that manifests itself with an invariance under local conformal transformations of the metric.

4.7.5 Conformal Gravity

Dear Philip *(Mannheim)***, the nature of dark matter and dark energy is the actual nightmare of present day theorists. Why do we believe in dark matter and dark energy – and do we have to?**

Here we discuss the basis for believing in DM and DE, and identify its possible shortcomings. We show that while the standard Newton–Einstein theory of gravity

[39] Following also the Occam razor prescriptions, we can say that: *Entia non sunt multiplicanda praeter necessitatem.*

is sufficient to describe gravitational phenomena, on Solar system distance scales it is not necessary; with the case for DM and (meV scale) DE resting entirely on the viability of the extrapolation of the standard Solar system wisdom to the much larger astrophysical and cosmological scales. It is this lack of necessity of the standard theory that permits one to consider alternatives to it. We discuss the merits of one such alternative, conformal gravity, and show that it can solve the DM, DE and quantum gravity problems within one comprehensive framework. We discuss whether we might be approaching a paradigm shift in gravitational theory, and consider what form it might take.

4.7.5.1 Why We Believe in Dark Matter

While the development of the standard Newton–Einstein theory of gravity represents one of the most celebrated achievements in physics, as such the standard theory stands on two totally different types of cornerstones. One of them is a broad covariance principle, namely that the space–time metric describes the gravitational field, while the other is the purely phenomenological requirement that the equations of motion for the metric be second order differential equations. While one should not contemplate modifying the covariant, metric description of gravity, the emergence of the DM problem, the DE or cosmological constant problem, and the quantum gravity problem invite consideration of whether one could relax the phenomenological requirement, and whether theories based on alternate, but equally covariant theories of gravity, might enable us to resolve these troubling concerns.

In trying to analyze how one could possibly make any changes to the standard theory and determine what is actually mandated by observation, we need to distinguish between two quite separate issues, namely what particular gravitational field a given source is to set up, and how a test particle is to respond to it. The first of these issues relates to the validity of Newton's law of gravity and the fundamentality of Newton's gravitational constant $G_{\rm N}$, while the second relates to the principle found by Galileo, namely that the response of a particle to an external gravitational field is independent of its mass. Following his development of special relativity, Einstein was faced with the problem that Newtonian gravity was not compatible with the relativity principle, and would thus need to be modified in some way, just as Newton's laws of motion had to be modified in order to make them compatible with special relativity. In finding a successful modification to Newtonian gravity, it was Galileo's principle and its Eötvös experiment refinement that served as Einstein's principle guide, leading him to generalize Galileo's principle to the equivalence principle. Specifically, by generalizing the gravitational potential ϕ to the metric $g_{\mu\nu}$ and by requiring that the metric couple to matter covariantly, Einstein was able to deduce that test particles would then move on the geodesics associated with the metric:

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{d\tau}\frac{dx^\sigma}{d\tau} = 0, \quad \Gamma^\mu_{\nu\sigma} = \frac{1}{2}g^{\mu\lambda}\left(\frac{dg_{\nu\lambda}}{dx^\sigma} + \frac{dg_{\sigma\lambda}}{dx^\nu} - \frac{dg_{\nu\sigma}}{dx^\lambda}\right), \tag{4.29}$$

with the geodesic equation immediately being independent of the mass of the particle, and with the Christoffel symbol $\Gamma^\mu_{\nu\sigma}$ describing the background gravitational field in which the test particle is to move.

As such, the Christoffel symbols which appear in (4.29) need not make any specific reference to Newton's law of gravity per se or to its associated fundamental G_N as the form of (4.29) is generic. And indeed, this should be the case since Galileo's principle was obtained long before Newton's law of gravity, and is logically independent of it. Indeed, the validity of (4.29) does not require the specification of any particular theory of gravity at all other than that it be covariant. Specifically, for any covariant metric theory of gravity in which the pure gravitational piece of the action I_{GRAV} is a general coordinate scalar function of the metric, the quantity $\Lambda^{\mu\nu} = 2(-g)^{-1/2}(\delta I_{GRAV}/\delta g_{\mu\nu})$ will, simply because of covariance, automatically obey $\Lambda^{\mu\nu}{}_{;\nu} = 0$, and through the gravitational equation of motion $\Lambda^{\mu\nu} = T^{\mu\nu}$ then automatically lead to the covariant conservation of $T^{\mu\nu}$ and thus to geodesic motion and the equivalence principle for test particles. As such then, the fact that motion in an external gravitational field is geodesic only implies that gravity is a metric theory, and in itself does not select out any particular one.

To determine what the form of the gravitational equations of motion might be, Einstein did not follow the approach which led him to the equivalence principle by appealing to some further fundamental principle, but instead appealed solely to a purely phenomenological requirement, namely that the theory reduce to the second order Poisson equation $\nabla^2\phi = 4\pi G_N\rho$ and its Newtonian potential solution $\phi = -MG_N/r$ in the weak gravity limit. With this requirement Einstein was then led to base the theory on the geometric Ricci tensor $R_{\mu\nu}$ and Ricci scalar $R^\alpha{}_\alpha = g^{\mu\nu}R_{\mu\nu}$, since they are both second derivative functions of the metric; with covariance then leading to the Einstein–Hilbert action $I_{EH} = -(1/16\pi G_N)\int d^4x(-g)^{1/2}R^\alpha{}_\alpha$ and the Einstein equations

$$-\frac{1}{8\pi G_N}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R^\alpha{}_\alpha\right) = T_{\mu\nu} \tag{4.30}$$

for a source with energy–momentum tensor $T_{\mu\nu}$, and with dimensional analysis requiring the introduction of the dimensionful G_N as an overall coefficient. Given (4.30), a specification of any particular $T_{\mu\nu}$ would then fix the Christoffel symbols needed for the geodesics. With the application of the Einstein equations to the Solar system not only recovering Newton's Law of Gravity, but also leading to general relativistic corrections to it which were then spectacularly confirmed (the three classic Solar system tests), a consensus in the gravity community emerged that the issue of what the correct theory of gravity is had once and for all been settled.

Despite this, it is important to note that the three classic tests only test the structure of the geometry exterior to a static, spherically symmetric source such as the sun or the earth, a region where the geometry can be described by the Ricci flat Schwarzschild solution in which $R_{\mu\nu} = 0$. However, since these tests are only applied in the region where $T_{\mu\nu} = 0$, they are not sensitive to the structure of the geometry in the $T_{\mu\nu} \neq 0$ region interior to the source, and thus do not tell us what $R_{\mu\nu}$ is to be equal to in the interior region. Thus while the Einstein equations imply that $R_{\mu\nu}$ is zero if $T_{\mu\nu}$ is zero, confirming this solution in the $T_{\mu\nu} = 0$ region does not, and in fact cannot, establish the validity of the Einstein equations in the region where $T_{\mu\nu}$ is nonzero.

Indeed, as noted by Eddington not all that long after Einstein's development of GR, if the Einstein equations were to be replaced by some alternate set of equations in which $T_{\mu\nu}$ was instead to be related to some derivative function of $R_{\mu\nu}$ (we give a specific example of this below), such a theory would still have the Schwarzschild solution as an exact exterior region solution since the vanishing of $R_{\mu\nu}$ secures the vanishing of its derivatives as well. Since in such a case we would still recover the Schwarzschild solution while bypassing the Einstein equations altogether, we see that the Einstein equations are only sufficient to give the three classic tests, but not at all necessary. It is this lack of necessity which opens the standard Newton–Einstein theory to challenge, with the standard theory still awaiting a principle, which would make it both necessary and sufficient. Moreover, it is the exploitation of this very lack of necessity, which can lead to alternatives to DM.

However, if we ignore these issues of principle, and simply take the Einstein equations as a given, on extrapolating them to distances much larger than the Solar system ones on which the theory was first established, we immediately run into phenomenological problems, with it being found impossible to describe observations in galaxies and clusters of galaxies using (4.30) with a $T_{\mu\nu}$ which consists of established luminous sources alone (see e.g., [130]). To bring the theory into agreement with observation on these large distance scales one introduces DM, and one not only introduces DM, one introduces it in copious amounts, with galaxies and clusters of galaxies needing altogether more DM than luminous matter; and not only that, in spiral galaxies the needed DM has to be located predominantly in regions where the luminous matter itself is not located. We thus see the circularity of the argumentation: one assumes the standard theory to be correct, one finds that with luminous matter alone it fails, and so one introduces however much DM is needed and in whichever locations it would be required to make things work. Thus either DM does indeed exist, or the extrapolation of the standard Solar system wisdom to larger distance scales is unreliable.

Apart from the circularity of the above reasoning, we also note that dynamical DM models of galaxy formation have yet to explain some of the key aspects of the systematics of the rotational velocity curves which are detected in spiral galaxies. Specifically, because DM is taken to be both nonluminous and weakly interacting, it can be dissipation-less and thus produce stable DM halos, and one can nicely produce various classes of such halos from fluctuations in the standard cosmology (viz. the cosmology which is obtained when the standard Einstein theory is extrapolated to cosmological distance scales) [150, 151]. However, as yet, this same fluctuation theory does not prescribe which specific DM halo is to go with which specific luminous distribution of matter. Consequently, from a measurement of the luminous matter content of a galaxy alone, DM theory cannot yet predict what the associated galactic velocity rotation curve would look like. Rather, all it can do is tell us which particular halo to use once a rotation curve has been measured. One of the key features which would support alternate, non-DM, theories would be an ability to predict the systematics of galactic rotation curves from a knowledge of the luminous content of the galaxies alone. Three specific candidate alternate theories which satisfy this criterion are the MOND (modified Newtonian dynamics) theory of Milgrom [137–139], the MOG (modified gravity) theory of Moffat [145], and the conformal gravity theory, which is being explored by the present author [130]. Moreover, each of these theories makes a different prediction regarding the behavior of the rotation curves at distances larger than those currently available. MOND requires the rotation curves to all be asymptotically flat, MOG requires them to eventually fall, and conformal gravity requires them to rise. With standard DM theory also requiring asymptotic flatness, the asymptotic behavior of rotation curves can serve as a useful diagnostic.

To summarize, we see that the wisdom that it is actually mandated from studies of the Solar system is embodied in the statements that gravity is a metric theory, that the coupling of the metric to matter is general coordinate invariant, that test particles move on geodesics, and that on Solar system distance scales the metric exterior to a static spherically symmetric source is the Ricci flat Schwarzschild metric. Any theory which incorporates all of these features preserves the key features of the Einstein theory, doing so even if the metric shows departures from the Schwarzschild geometry on larger distance scales, departures which might be able to account for galactic distance scale issues without the need for DM. Consequently, what is needed to eliminate the need for DM is to generalize not Newton but Schwarzschild.

4.7.5.2 Why We Believe in Millielectronvolt Scale Dark Energy

Problems for the standard theory are also encountered when it is extrapolated even further, namely to cosmological distance scales. In the standard theory cosmology is described by the Friedmann evolution equations

$$\dot{R}^2(t) + kc^2 = \dot{R}^2 \left[\Omega_m(t) + \Omega_\Lambda(t)\right], \tag{4.31}$$

$$-q(t) = \frac{R(t)\ddot{R}(t)}{\dot{R}^2(t)} = -\frac{1}{2}\Omega_m(t) + \Omega_\Lambda(t), \tag{4.32}$$

where $R(t)$ is the expansion radius of the Universe and k is its spatial 3-curvature. In (4.31) and (4.32), we have introduced a matter[40] source $\Omega_{\mathrm{m}}(t) = 8\pi G_{\mathrm{N}}\rho_{\mathrm{m}}(t)/3c^2H^2(t)$ and a cosmological constant source $\Omega_\Lambda(t) = 8\pi G_{\mathrm{N}}\Lambda/3cH^2(t)$, with $H(t) = \dot{R}(t)/R(t)$ being the Hubble expansion rate parameter. As an evolution

[40] Here the term “matter” refers to all gravitational sources in the energy–momentum tensor, whether they be massive or massless.

equation, in the absence of any Λ-dependent term, (4.32) would lead to an expected negative value for the quantity $R(t)\ddot{R}(t)/\dot{R}^2(t)$ because the galaxies gravitationally pull on each other (i.e., $\Omega_{\mathrm{m}}(t)$ is positive). However, when the current era $-q(t_0) = R(t_0)\ddot{R}(t_0)/\dot{R}^2(t_0)$ was measured through the use of Type Ia SNe as standard candles, $-q(t_0)$ was actually found to be positive [183,190], implying that there has to be some cosmologically repulsive gravitational force at work that is causing galaxies to accelerate away from one another. To account for this repulsive effect in the standard theory one has to add an additional gravitational source known generically as DE, as just like DM, the source for this new effect has to be non-luminous as there is no apparent visible manifestation of it. A particular candidate for DE would be a cosmological constant since a positive Λ would provide a positive contribution to $R(t)\ddot{R}(t)/\dot{R}^2(t)$ in (4.32). On taking both the SNe data and the anisotropy structure of the cosmic microwave background into consideration [59, 222], a best fit using (4.31) is obtained with current era parameters $\Omega_{\mathrm{m}}(t_0) = 0.3$, $\Omega_\Lambda(t_0) = 0.7$ (to require $-q(t_0) = 0.55$), and these are now regarded as the numbers which cosmology supplies and which fundamental theory is thus obliged to explain.

As such, both of the above values for $\Omega_{\mathrm{m}}(t_0)$ and $\Omega_\Lambda(t_0)$ create problems for the standard theory. With the cosmological contribution of visible matter as determined from galaxy counts giving an $\Omega_{\mathrm{m}}(t_0)$ of only 0.01 or so, we see that to produce an $\Omega_{\mathrm{m}}(t_0)$ as large as 0.3 we would, just as with galaxies, again need copious amounts of DM, this time for cosmological reasons. While elementary particle theory permits of many candidate particles which could serve as this DM (such as axions or supersymmetric partners of the established elementary particles), for the moment there has been no direct detection of any such particles despite the extensive searching for them which has been going for many years now. The direct detection of DM particles is critical for the viability of the standard theory so that the existence of such particles could be confirmed by non-gravitational means, with use of the large hadron collider providing perhaps the best near term opportunity to search for them.

While the DM problem is the problem of having an $\Omega_{\mathrm{m}}(t_0)$ which is too large (i.e., too large to be accountable for by luminous matter alone), the DE problem is one of having an $\Omega_\Lambda(t_0)$ which is too small. Specifically, the same elementary particle theory which can yield Higgs or supersymmetric particles with masses in the TeV ($= 10^{15}$ degrees) region that the large hadron collider can explore, will at the same time simultaneously yield a contribution to $\Omega_\Lambda(t_0)$ of order 10^{60} or so (viz. an equivalent BB with a temperature of 10^{15} degrees). Consequently, some mechanism has to be found which can quench this value by 60 orders of magnitude, since the measured $q(t_0)$ is of order one, not of order 10^{60}. For the moment, and despite years of vigorous effort, no mechanism has yet been found which could provide for so mammoth a quenching, and the belief in DE in the way it is used in the standard theory is not so much that of believing that $\Omega_\Lambda(t_0)$ is greater than $\Omega_{\mathrm{m}}(t_0)/2$ (since the measured value of $q(t_0)$ is undeniably negative), but rather in believing that $\Omega_\Lambda(t_0)$ is of order one rather than of order 10^{60} and that Λ itself is accordingly to be associated with a miniscule meV energy scale.

So dire is the situation regarding a value for $\Omega_\Lambda(t_0)$ of order one, that appeal is now being made to the AP to provide an explanation. As such, the use of the AP to fix the values of physical parameters such as Λ requires that three conditions hold: first, that we have reached a brick wall in trying to explain the value of Λ by conventional means, and second, that there is a theory which admits of an entire range of allowed values for Λ, and thirdly that included within this range there is only a very narrow range of values for Λ for which intelligent observers could emerge (for instance, for galaxies to be able to form in the standard theory does not permit the gravitationally repulsive $\Omega_\Lambda(t_0)$ to be all that much bigger than the value of order one that is now being suggested for it [227]). As regards the first condition, it is almost impossible to demonstrate that we have indeed reached a brick wall or that we have reached a limit to knowledge. Rather all we can say is that we have run out of ideas. As regards the second condition, it turns out that there actually is a theory with such a plethora of values for Λ, namely the string theory landscape picture with its of order 10^{500} solutions, so string theory does at least provide some range of allowed values for Λ (with the landscape converting what had been a manifest vice for string theory (viz. too many solutions) into a possible virtue). However, whether any single one of these 10^{500} solutions looks anything like the Universe we observe remains to be seen, so it is not at all clear if the third condition is actually met anywhere in the landscape – and even if there might be a region in the landscape where Λ is indeed small enough, one would still need to show that in that region one can reconcile so small a value for Λ with the very large mass scale for supersymmetric particles (not less than at least a TeV scale) that is needed to account for their lack of detection to date.

However, trying to explain how Λ could be so small is not the only problem facing the standard cosmology. Of concern is not just the need to understand the current era where one tries to explain why the current era $-q(t_0) = -\Omega_m(t_0)/2 + \Omega_\Lambda(t_0)$ is a positive quantity of order one, but also of reconciling current era observations with the evolution required by the Friedmann equations over the entire history of the Universe. Specifically, with the standard cosmology being initiated at an initial time $t = 0$ (of order the 10^{-43} s Planck time) by a big bang singularity in which $\dot{R}(t = 0)$ is infinite, (4.31) implies that at the initial time the quantity $\Omega_m(t = 0) + \Omega_\Lambda(t = 0)$ must have been equal to one, no matter what the value of the spatial 3-curvature k. With Λ being constant and with $\rho_m(t)$ redshifting continuously for 10^{10} years from an enormous initial 10^{32} degree Planck temperature T_{PL} until the current few degree temperature $T(t_0)$ of today, the required closeness in magnitude of $\Omega_m(t_0)$ and $\Omega_\Lambda(t_0)$ today implies that at time $t = 0$ the matter density $\Omega_m(t = 0)$ would have had to have been incredibly close to one (within one part in $T_{PL}^4/T^4(t_0) \sim 10^{126}$ or so), while equally, the cosmological constant term $\Omega_\Lambda(t = 0)$ would have had to have been incredibly close to zero (within one part in 10^{126}). And, not only that, any small change in the initial conditions would cause the Universe to evolve into some totally different configuration today. The Friedmann evolution equations thus need an incredible amount of early Universe fine tuning in order to subsequently produce the current Universe we see today, with the required values of $\Omega_m(t_0)$ and $\Omega_\Lambda(t_0)$ just not being natural values for it. What would be natural for the Friedmann evolution equations would be an $\Omega_m(t)$ which fell to zero in the early Universe itself and an $\Omega_\Lambda(t)$, which rose to one in the same epoch.

To underscore the fact that the fine-tuning required for the Friedmann equations is different from that required of Λ, we note that this very same problem would exist for a current era $\Omega_{\rm m}(t_0)$ of magnitude 0.3 even if the Λ term were to be absent altogether. Indeed, this is the original flatness fine-tuning problem which inflation [100] was invented to resolve (and which would have done so had $\Omega_{\rm m}(t_0)$ been equal to one and $\Omega_\Lambda(t_0)$ been zero). However, with $\Omega_{\rm m}(t_0)$ now needing to be less than one, the Friedmann evolution equations are once again stricken with an early Universe fine-tuning problem, a problem for which there is no known solution. Moreover, while an anthropic argument has been advanced for Λ, there is no apparent anthropic argument that would require the initial $\Omega_\Lambda(t = 0)$ to be zero to one part in 10^{126}. Indeed, one's ability to appeal to the Anthropic Principle (AP) at all requires that there be a whole range of values of Λ for which humans could exist, and another range for which they could not. One cannot readily reconcile anthropism with the Friedmann evolution equations and one cannot readily reconcile anthropism with fine-tuning. With the fine-tuning problems lying with the Friedmann equations themselves, one way to avoid such problems would be to replace the Friedmann equations by a different set of equations, which would give the Universe a very different evolution history; and as we shall see below, conformal gravity will precisely yield an evolution equation which is free of fine-tuning problems, and will do so while simultaneously solving the cosmological constant problem as well.

In cosmology, the cosmological constant has had a long and checkered history. It was first introduced by Einstein himself in an attempt to obtain a static Universe, with the fine-tuned choice of $\Omega_\Lambda(t) = \Omega_{\rm m}(t)/2 = k/3$ in (4.31) and (4.32) leading to a static Universe with $\dot{R}(t) = 0$, $\ddot{R}(t) = 0$. With the discovery of Hubble that $\dot{R}(t)$ was nonzero, Einstein abandoned the idea of a static Universe, discarded Λ, and even regretted that he had introduced it at all. However, while it was known since Hubble's time that $\dot{R}(t)$ was nonzero, it was not known whether $\ddot{R}(t)$ was positive or negative, and had Einstein known that $\ddot{R}(t)$ was not only nonzero but actually positive, he would again have had to consider introducing a positive Λ. And indeed, such a Λ was in fact introduced into the standard theory following the modern era discovery that the Universe actually was accelerating. What underlies this historical turn of events is the fact that our ability to add or remove Λ from (4.31) and (4.32) at will is because there is no principle in standard gravity which speaks to its presence or absence. Indeed, we had noted earlier that there was no fundamental principle which even led Einstein to the second order (4.30) in the first place, with it being the very absence of any such principle which has engendered the cosmological constant problem. The requirement that the gravitational equations be general coordinate invariant only requires the gravitational action to be a general coordinate scalar and does not select any one particular general coordinate scalar action over any other, to thus leave the issue of the cosmological constant not only unaddressed but essentially unaddressable. In addition, on top of these gravitational considerations, modern elementary particle physics has even made the situation yet more severe since it naturally provides the Universe with a vacuum energy which would typically be associated with a TeV region scale as that is the electroweak interaction symmetry breaking scale as determined by experiment. Nonetheless, regardless of whether or not there actually might be a specific mechanism, which would make Λ as small as the standard theory would like it to be, there is actually a direct, model-independent observational test to see if nature actually favors a value for Λ this low. Specifically, because $\rho_{\rm m}(t)$ redshifts while Λ does not, given their current era values, $\Omega_\Lambda(t)$ will only dominate over $\Omega_{\rm m}(t)$ at late time redshifts of order $z < 1$ or so. Above a redshift of $z = 1$, there should thus be a switch over from acceleration to deceleration, with such a switch over in the Hubble plot thus serving as a diagnostic test for the standard cosmology, and especially for any anthropic basis for an meV scale Λ.

In assessing the problems which the standard cosmology faces, we recall that the Friedmann equation early Universe fine tuning problem is the difficulty in reconciling the early Universe constraint $\Omega_{\rm m}(t=0) + \Omega_\Lambda(t=0) = 1$ with current era values for $\Omega_{\rm m}(t_0)$ and $\Omega_\Lambda(t_0)$ that are of the same order of magnitude. One way to resolve this problem would be to eliminate the Big Bang singularity (viz. $\dot{R}(t = 0) = \infty$) as it is the reason that there is an early Universe constraint in the first place. To do this one could consider the possibility that instead of being controlled by the standard positive Cavendish $G_{\rm N}$, cosmology would instead be controlled by an effective $G_{\rm eff}$ that is negative, since with a repulsive $G_{\rm eff}$ there would simply be no initial singularity. As regards the cosmological constant problem, we note that it arises due to a mismatch between standard cosmology and elementary particle physics, with particle physics yielding a large Λ and standard cosmology needing a small one. With efforts to quench the magnitude of Λ having so far failed, and rather badly so, one should ask whether Λ might in fact not be quenched at all, and whether the energy–momentum tensor might really contain a TeV scale cosmological constant. At first this might seem to be impossible. However, it is not actually Λ itself which is measured in cosmology, but rather $\Omega_\Lambda(t)$, that is, not Λ but $G_{\rm N}$ times Λ. Thus, cosmology measures not Λ but the amount by which it gravitates. Hence, if cosmologically $G_{\rm N}$ could be replaced by a $G_{\rm eff}$ which is altogether smaller in magnitude than $G_{\rm N}$, then $8\pi G_{\rm eff}\Lambda/3cH^2(t)$ could potentially be of order one without Λ itself needing to be quenched. For this to happen, just like the analog dimensionful Fermi coupling constant $G_{\rm F}$ that controls low energy weak interactions, the dimensionful $G_{\rm N}$ would have to no longer be a fundamental input parameter in the gravitational action, but would instead need to emerge as a dynamically induced parameter that would only control local physics and not control global cosmology at all. To solve the problems that the standard model faces then, we can consider the possibility that $G_{\rm N}$ is not fundamental, and that cosmologically it is replaced by a $G_{\rm eff}$ that is both small and negative. And as we shall now show, this precisely occurs in conformal gravity, and not only does it occur, it occurs quite naturally with no fine-tuning of parameters being needed. Moreover, since $G_{\rm eff}$ is negative, conformal cosmology is naturally repulsive in all epochs, with acceleration thus occurring at both low and high redshift. In the conformal theory then there is no switch over between acceleration and deceleration at a redshift of one. Monitoring the Hubble plot above a redshift of one should thus enable one to discriminate between the standard cosmology and its conformal competitor, in a way that could be definitive for both.

4.7.5.3 Conformal Gravity

As a candidate theory of gravity, conformal gravity is motivated by the requirement that just like the other three fundamental interactions, gravity too is to be a renormalizable theory that is built on a fundamental action whose coupling constants are all dimensionless, and whose mass scales are all generated dynamically in the vacuum by spontaneous symmetry breaking. To achieve this, in addition to general coordinate invariance, conformal gravity also endows gravity with an additional symmetry, namely local conformal invariance, viz., invariance of the theory under any and all local conformal transformations on the metric of the form $g_{\mu\nu}(x) \to e^{2\alpha(x)} g_{\mu\nu}(x)$. With the imposition of such a symmetry, the pure gravitational piece of the action is then fixed to uniquely be of the Weyl form:

$$I_{\mathrm{W}} = -\alpha_{\mathrm{g}} \int \mathrm{d}^4 x (-g)^{1/2} C_{\lambda\mu\nu\kappa} C^{\lambda\mu\nu\kappa}, \tag{4.33}$$

where $C_{\lambda\mu\nu\kappa}$ is the Weyl conformal tensor and α_{g} is a dimensionless gravitational coupling constant. The great appeal of the conformal symmetry is not only that it actually uniquely specifies the form of the gravitational action in the first place, in addition it does not permit the presence of any fundamental cosmological constant term in the action since such a term would violate the underlying conformal invariance of the theory. Conformal gravity thus has a control on the cosmological constant which the standard theory lacks.

In the presence of a matter action I_{M}, variation with respect to the metric of the total $I_{\mathrm{W}} + I_{\mathrm{M}}$ yields gravitational equations of motion of the form [130]

$$4\alpha_{\mathrm{g}} W^{\mu\nu} = 4\alpha_{\mathrm{g}} \left[W^{\mu\nu}_{(2)} - \frac{1}{3} W^{\mu\nu}_{(1)} \right] = T^{\mu\nu}, \tag{4.34}$$

where $W^{\mu\nu}_{(1)}$ and $W^{\mu\nu}_{(2)}$ are the fourth order derivative functions

$$W^{\mu\nu}_{(1)} = 2g^{\mu\nu} (R^{\alpha}{}_{\alpha})^{;\beta}{}_{;\beta} - 2(R^{\alpha}{}_{\alpha})^{;\mu;\nu} - 2R^{\alpha}{}_{\alpha} R^{\mu\nu} + \frac{1}{2} g^{\mu\nu} (R^{\alpha}{}_{\alpha})^2, \tag{4.35}$$

$$W^{\mu\nu}_{(2)} = \frac{1}{2} g^{\mu\nu} (R^{\alpha}{}_{\alpha})^{;\beta}{}_{;\beta} + R^{\mu\nu;\beta}{}_{;\beta} - R^{\mu\beta;\nu}{}_{;\beta} - R^{\nu\beta;\mu}{}_{;\beta} - 2R^{\mu\beta} R^{\nu}{}_{\beta} + \frac{1}{2} g^{\mu\nu} R_{\alpha\beta} R^{\alpha\beta}. \tag{4.36}$$

With $R_{\mu\nu} = 0$ thus being a solution to the theory in regions where $T_{\mu\nu} = 0$, the Schwarzschild solution is thus an exact exterior solution to the theory, with conformal gravity thus satisfying the three classic tests of GR. Conformal gravity thus emerges as a theory, which is able to meet the three classic tests while bypassing the second order derivative Einstein equations altogether. Conformal gravity thus provides an explicit realization of the point noted above, namely that the standard Newton–Einstein theory is only sufficient to give the standard Solar system wisdom but not necessary.

Now while the vanishing of $R_{\mu\nu}$ entails the vanishing of $W_{\mu\nu}$, the converse does not follow, as $W_{\mu\nu}$ could vanish without $R_{\mu\nu}$ vanishing. In fact the general solution exterior to a static spherically symmetric source is found [131] to be of the form

$$ds^2 = -[1 - 2\beta/r + \gamma r]\,dt^2 + [1 - 2\beta/r + \gamma r]^{-1}\,dr^2 + r^2 d\Omega_2{}^{41}. \quad (4.37)$$

The Schwarzschild solution with its $1/r$ Newtonian limit is thus recovered for small enough $r \ll (2\beta/\gamma)^{1/2}$, while the departures from Schwarzschild at large r, where the potential rises rather than falls thus modify Newtonian gravity in precisely the region where the standard Newton–Einstein theory needs to resort to DM. Moreover, with this rising potential conformal gravity transits from a local to global theory. Specifically, Newton's theorem that for spherically distributed matter one can ignore the material outside the region of interest only holds for $1/r$ potentials. It does not hold for any other potential, and so in essentially all alternate theories which seek to modify the law of force in order to eliminate the need for DM, one cannot restrict to integrating the potential over the matter within the source region alone. Rather, one must integrate over both the matter in the local source and over the rest of the Universe as well; and it is precisely when one takes this global effect into consideration that conformal gravity is then found capable of accounting for galactic rotation curve systematics without recourse to DM [130]. Invoking DM can thus be viewed as an attempt to describe a global effect in purely local terms. Apart from the fact that conformal gravity is actually able to solve the galactic DM problem (to thus suggest that the origin of the DM problem may indeed lie in the lack of reliability of the extrapolation of standard gravity beyond its Solar system origins), its successful fitting should be regarded as being of significance since conformal gravity was not initially proposed for galactic rotation curve purposes at all. Rather, it was first studied in order to address the cosmological constant problem [128], with the solution in (4.37) only being found much later. Since its original objective was to address the cosmological constant problem via use of a symmetry principle, we shall now describe how it has fared on it.

4.7.5.4 Conformal Cosmology

In conformal gravity the matter action also has to be conformal invariant, and for the generic model of a massless fermion conformally coupled to a symmetry breaking scalar field[42], the most general conformal invariant action takes the form

$$I_M = -\int d^4x(-g)^{1/2}\left[\tfrac{1}{2}S^{;\mu}S_{;\mu} - \tfrac{1}{12}S^2R^{\mu}{}_{\mu} + \lambda S^4 + i\bar{\psi}\gamma^{\mu}[\partial_{\mu} + \Gamma_{\mu}]\psi - hS\bar{\psi}\psi\right], \quad (4.38)$$

[41] Here $d\Omega_2 = d^2\theta + \sin^2\theta d\phi^2$.

[42] The scalar field should be understood as a Ginzburg–Landau c-number order parameter produced by a fermion condensate, and as such should not observed in a collider experiment.

and we note the presence of the back reaction term $-S^2 R^{\mu}{}_{\mu}/12$ of the scalar field on the geometry, with its expressly negative coefficient. For such a matter action, on modeling the fermion contribution by a perfect fluid and giving the scalar field a non-zero vacuum expectation value S_0, the energy–momentum tensor takes the form

$$T_{\mu\nu} = (\rho_{\rm m} + p_{\rm m})U_\mu U_\nu + p_{\rm m} g_{\mu\nu} - \frac{1}{6}S_0^2\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R^{\alpha}{}_{\alpha}\right) - g_{\mu\nu}\lambda S_0^4. \quad (4.39)$$

Because of its high symmetry, in a Robertson–Walker geometry the Weyl tensor vanishes identically. Consequently from (4.34) we find that in conformal cosmology the energy–momentum tensor is zero, with conformal cosmology thus knowing where the zero of energy is, a key need for solving the cosmological constant problem. Since for conformal cosmology $T_{\mu\nu} = 0$, from (4.39) we see that conformal cosmology is described by

$$\frac{1}{6}S_0^2\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R^{\alpha}{}_{\alpha}\right) = (\rho_{\rm m} + p_{\rm m})U_\mu U_\nu + p_{\rm m} g_{\mu\nu} - g_{\mu\nu}\lambda S_0^4. \quad (4.40)$$

We recognize (4.40) as being identical in form to the standard gravity cosmic evolution equation save only that $G_{\rm N}$ has been replaced by an effective cosmological $G_{\rm eff}$ of the form

$$G_{\rm eff} = -\frac{3}{4\pi S_0^2}. \quad (4.41)$$

Conformal cosmology thus acts just like a standard cosmology except that matter sources now gravitate with a strength which is not fixed by the Cavendish experiment, but by a $G_{\rm eff}$ that is negative rather than positive and that in addition becomes smaller the larger the value of S_0, that is, the same scalar field, which causes the cosmological constant to be big simultaneously causes the amount by which it gravitates to be small. Conformal cosmology thus naturally self-quenches the gravitational effect of the cosmological constant without needing to quench Λ itself; and with $G_{\rm eff}$ being negative, the cosmology also has no initial singularity and thus no early Universe fine-tuning problem [129]. Additionally, with a negative $G_{\rm eff}$, cosmological gravity is naturally repulsive, to thus automatically cause the Universe to accelerate without fine-tuning. Conformal cosmology thus naturally resolves the primary difficulties we identified above for the standard cosmology and its specific approach to the problem could be instructive for other attempts to address these same issues.

Beyond these general issues of principle one also has to ask how well conformal cosmology does in practice. To this end we note that in the conformal theory (4.31) and (4.32) are replaced the analogous

$$\dot{R}^2 + kc^2 = \dot{R}^2\left[\bar{\Omega}_m(t) + \bar{\Omega}_\Lambda(t)\right], \quad -q(t) = -\bar{\Omega}_m(t)/2 + \bar{\Omega}_\Lambda(t), \quad (4.42)$$

where $\bar{\Omega}_m(t) = 8\pi G_{\rm eff}\rho_m(t)/3c^2H^2(t)$, $\bar{\Omega}_\Lambda(t) = 8\pi G_{\rm eff}\Lambda/3cH^2(t)$. On identifying Λ with a TeV region temperature T_V and associating $\rho_{\rm m}(t_0)$ with the current temperature of the Universe, we see that the ratio $\bar{\Omega}_{\rm m}(t_0)/\bar{\Omega}_\Lambda(t_0) = \rho_{\rm m}(t_0)/c\Lambda \sim T^4(t_0)/T_{\rm V}^4 = O(10^{-60})$ is completely negligible today, with ordinary matter not making any substantial contribution to current era cosmic expansion (this would also be the case in the standard theory if one were to use an $\Omega_\Lambda(t_0)$ of order 10^{60}). Rather, in the conformal theory matter is only of importance to cosmic expansion in the early Universe before $\rho_m(t)$ redshifts away. In the conformal theory one can also show that the spatial 3-curvature k is negative [130], since the need to have a zero $T_{\mu\nu}$ in the presence of a positive energy density perfect fluid is achieved by having negative energy density in the gravitational field, and thus negative spatial 3-curvature. In consequence of this pattern of values for $\bar{\Omega}_{\rm m}(t)$, $\bar{\Omega}_\Lambda(t)$ and k, from (4.42) and without any fine-tuning at all, one finds the bounds

$$0 \le \bar{\Omega}_\Lambda(t) \le 1, \quad -1 \le q(t) \le 0 \tag{4.43}$$

at all times for which $\bar{\Omega}_\Lambda(t) \gg \bar{\Omega}_m(t)$. Thus no matter how big Λ might be, its contribution to cosmic expansion is tamed. For the cosmology one can also solve for the luminosity distance, to obtain a dependence on redshift of the form

$$d_{\rm L} = -\frac{c}{H(t_0)}\frac{(1+z)^2}{q(t_0)}\left(1 - \left[1 + q(t_0) - \frac{q(t_0)}{(1+z^2)}\right]^{1/2}\right) \tag{4.44}$$

as expressed in terms of one free parameter, the current era value of the deceleration parameter $q(t_0)$. In Fig. 4.8, we have plotted the expectation of this formula for the currently available high redshift SNe Hubble plot data to obtain [130] a best conformal cosmology fit with $q(t_0) = -0.37$ (i.e., nontrivially right in the middle of the range allowed by (4.43)), which is every bit as good in quality as an $\Omega_{\rm m}(t_0) = 0.3$, $\Omega_\Lambda(t_0) = 0.7$ standard cosmology fit. In the figure, we have included the nine $z > 1$ data points of the so-called gold standard data of [191], and for the moment one cannot discriminate between the conformal gravity expectation of continuing acceleration at higher redshift ($q(t)$ remains negative in (4.43)) and the switch over to deceleration required of the standard cosmology. With the upcoming DE searches targeting the high redshift Hubble plot, it should be possible to test both standard and conformal cosmology in the near future and discriminate between them. While much work still needs to be done on the conformal theory (for instance, its predictions for the anisotropy of the cosmic microwave background have yet to be worked out), it does exhibit some novel generic features that one might want to incorporate in trying to solve the DE and cosmological constant problems.

4.7.5.5 Quantum Gravity

Troubling as the cosmological constant problem is for the standard theory, the problem of making the standard theory compatible with quantum mechanics is even

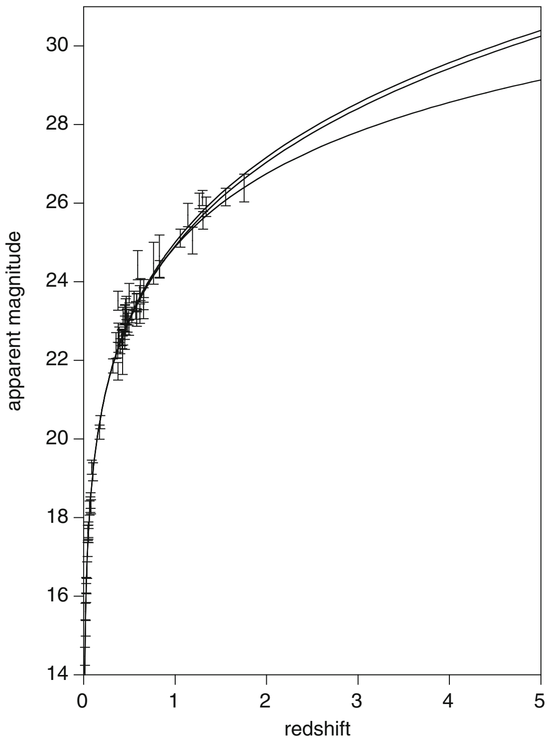

Fig. 4.8 Hubble plot expectations for $q(t_0) = -0.37$ conformal gravity (*highest curve*), for the $\Omega_m(t_0) = 0$, $\Omega_\Lambda(t_0) = 0$, $q(t_0) = 0$ empty Universe baseline (*middle curve*), and for $\Omega_m(t_0) = 0.3$, $\Omega_\Lambda(t_0) = 0.7$ standard gravity (*lowest curve*)

more disturbing. Moreover, the problem of quantum gravity cannot be considered as being a purely microscopic one, which is detached from macroscopic cosmology. Specifically, as the standard Einstein gravitational theory is not renormalizable, its quantum radiative corrections are not small but infinite. Hence, quantum corrections to the standard cosmology cannot be ignored. To try to rectify this one generalizes gravity to string theory, as the very act of spreading out points into strings leads to controllable radiative corrections, corrections that, however, are found to only be unitary if string theory is formulated in ten spacetime dimensions[43]. But then, if one starts with such a ten-dimensional string theory as input, one now has to show that when one descends to four dimensions one precisely recovers the standard Friedmann cosmological evolution (4.31) and (4.32) as the output which is relevant on cosmological scales. Thus it is only if standard macroscopic classical gravity emerges from a consistent microscopic quantum gravity theory that one can legitimately use standard classical gravity at all. Since string theory possesses far more fields than just the graviton (and especially scalar ones), and since the ensuing output classical theory would have to live with whatever value for the cosmological constant string theory supplies, it is not clear that one can claim that an

[43] For an appraisal of this and other approaches to quantum gravity, see, for example, [211].

$\Omega_m(t_0) = 0.3$, $\Omega_\Lambda(t_0) = 0.7$ cosmology is indeed the cosmological expectation of string quantum gravity.

Moreover, the problem of quantum gravity is not only a problem for the standard theory, it is also a problem for almost all candidate alternate gravitational theories as well. Specifically, since alternate theories only seek to modify classical gravity on astrophysical or cosmological distance scales, the easiest way to achieve this one very specific objective is to modify the standard Einstein theory by adding terms to it (typically by introducing additional fields), terms which are to be negligible on Solar system distance scales. As the equations of motion of such *Einstein plus* theories reduce to the standard Einstein equations of (4.30) on Solar system distance scales, they automatically recover its solutions on those distance scales too, and thus meet the standard Solar system tests. However, as these alternate theories do contain the Einstein–Hilbert term, as classical theories they suffer from the same cosmological problems as the standard theory, while as quantum theories they are not renormalizable. For such alternate theories one thus either has to find some solution to this problem, or look to string theory for a cure. However, if the standard Einstein theory is to be the four-dimensional limit of string theory, none of the Einstein-plus theories could be its limit as well. Consequently, for Einstein plus theories the problem of quantum gravity might be even more severe than it is for Einstein gravity itself.

In string theory, the connection between microscopic physics and macroscopic physics is quite indirect. However, for theories such as electrodynamics the connection is straightforward. Specifically, because quantum electrodynamics is a renormalizable theory, the primary effect of radiative corrections is to turn bare quantities into dressed ones, with the output classical theory (matrix elements of the quantum operators) then (roughly) obeying the same Maxwell equations as the dressed quantum fields. For renormalizable theories then the microscopic q-number fields and the macroscopic c-number fields obey equations of motion of the same general form. Thus in a renormalizable theory of gravity, one will also be able to use dynamical equations of motions of any given form for both the microscopic and macroscopic fields. With conformal gravity being a renormalizable theory, the quantum corrections to its classical cosmology are thus small and under control, and its behavior in the deep Euclidean region far off the mass shell is fully acceptable. Now we recall that to meet the standard Solar system tests, the equations of a given theory do not need to reduce to those of the Einstein theory on Solar system distance scales. Rather, it is only the solutions that need to reduce to those of Einstein. Theories such as conformal gravity which are not Einstein plus can still meet the Solar system constraints. Thus by not being in the Einstein plus category, conformal gravity does not contain the Einstein–Hilbert term at all, and it is for this very reason that its radiative corrections are fully under control. As with both the DM and the cosmological constant problems then, a possible solution to the renormalization problem is to not include the Einstein–Hilbert term in the gravitational action at all.

While its renormalizability has always been recognized as being one of conformal gravity's most attractive features far off the mass shell, on the debit side its quantum corrections have long been thought to have a severe problem, namely not to be unitary on the mass shell. The problem can be summarized in the generic fourth-order scalar field theory whose action is

$$I = \frac{1}{2}\int \mathrm{d}^4x \left[\partial_\mu\partial_\nu\phi\partial^\mu\partial^\nu\phi - (M_1^2 + M_2^2)\partial_\mu\phi\partial^\mu\phi + M_1^2M_2^2\phi^2\right] \tag{4.45}$$

(the scalar field ϕ represents a typical component of the metric fluctuation $h_{\mu\nu}$ around a flat background $\eta_{\mu\nu}$) and whose equation of motion is

$$(\partial_t^2 - \nabla^2 + M_1^2)(\partial_t^2 - \nabla^2 + M_2^2)\phi(\bar{x}, t) = 0. \tag{4.46}$$

With $k^2 = E^2 - \bar{k}^2$, the associated propagator has the form

$$D^{(4)}(k^2) = \frac{1}{(k^2-M_1^2)(k^2-M_2^2)} = \frac{1}{M_1^2-M_2^2}\left(\frac{1}{k^2-M_1^2} - \frac{1}{k^2-M_2^2}\right). \tag{4.47}$$

Unlike the standard $1/k^2$ propagator of the Einstein theory, the conformal gravity propagator behaves as the far more convergent $1/k^4$ at large k^2 to thereby secure renormalizability. However, this good convergence appears to come at a price, namely the relative minus sign in the propagator, as it would appear to indicate that some of the poles in the propagator have residues which are negative. Such negative residues would be associated with negative-norm ghost states, and would lead to a violation of unitarity.

Because of the seriousness of this unitarity problem, higher derivative field theories have largely been abandoned. However, it turns out that there is a flaw in the above expectation, namely its assumption that the appropriate inner product for the theory is the standard Dirac one, something that cannot actually be assured until the Hilbert space needed for the theory has actually been constructed. And when it was actually constructed in a particular case, the Pais–Uhlenbeck quantum-mechanical oscillator model [171] (a model that serves as a prototype for fourth-order derivative field theories), it was found [18] that in the relevant Hilbert space the Hamiltonian of the theory was not Hermitian but was instead of the PT symmetric type studied by Bender and collaborators [17]. For such theories one has to define a different norm, the PT norm, and with respect to the PT inner product all the states in the theory are found to have positive norm and the time evolution operator is unitary. Because of the change in norm, the relative minus sign in (4.47) does not signal the presence of ghost states. Consequently, renormalizability and unitarity of higher derivative theories need not be in conflict, with conformal gravity potentially being a fully consistent theory of quantum gravity in four space–time dimensions.

4.7.5.6 Are We Approaching a Paradigm Shift?

At the present time, gravity theory stands at a critical juncture, just as it did in Galileo's time. On the one hand, the standard gravity fits to astrophysical data and especially to the cosmic microwave background data are extremely impressive and compelling, and yet, on the other hand, the standard theory has a large number of quite severe problems. While these problems are theoretical rather than observational, they actually would be observational if either an $\Omega_{\rm m}(t_0) = 0.3$ worth of DM did not in fact exist or if particle physics actually supplied standard cosmology with an $\Omega_\Lambda(t_0)$ of order 10^{60}. Nonetheless, despite these concerns, by and large the community finds the standard theory fits to be so compelling that it takes the view that even if it is not currently understand why the numbers are the way they are, it must be the case that in the real world $\Omega_{\rm m}(t_0)$ is equal to 0.3 and $\Omega_\Lambda(t_0)$ is equal to 0.7; with the task of fundamental theory being only to derive these numbers but not to challenge them. Then, if these numbers do describe the real world, and if it is the AP which explains why $\Omega_\Lambda(t_0)$ is not of order 10^{60}, and if it is the use of extra dimensions which makes quantum gravity consistent, the current standard picture itself would represent a paradigm shift. In fact it would actually represent not one but two types of paradigm shift. The existence of more than four space–time dimensions is a paradigm shift regarding the structure of the Universe, while the use of the AP is a paradigm shift regarding what we believe physics is ever able to accomplish, with there being certain things that we could not hope to ever be able to explain at all.

By the same token, any modification of standard gravity away from its Newton–Einstein form would also represent a paradigm shift, but one of a quite different nature. Rather, here we would not be changing any of the rules of physics themselves (so long as we maintain covariance), but only some of the equations of physics, a much milder step. While this would be true of any Einstein plus type modification to standard gravity, for the case of the conformal gravity theory there would also be a departure from the way paradigm changes have previously occurred in physics. Specifically, while equations do get revised from time to time, ordinarily they are revised in a way which then recovers previously established equations in a specific limit. With conformal gravity the previous equations are not in fact to be recovered at all. Rather, it is only previous solutions which are to be recovered in the limit.

Conformal gravity also makes a few other shifts in paradigm as well. First, it does not treat Newton's constant as being fundamental – though this is not that severe a shift as the equally dimensionful weak interaction Fermi constant is not considered to be fundamental either. Second, conformal gravity distinguishes between local gravity and global gravity in a way which is not the case for the standard theory. In conformal gravity, local physics and global physics are not only controlled by differing strengths ($G_{\rm eff}$ of (4.41) for cosmology and β of (4.37) for local physics), they are even controlled by equations of different order. Specifically, with the Weyl tensor vanishing identically for the homogeneous and isotropic Robertson–Walker geometry, conformal cosmology is controlled by the vanishing of second-order $T_{\mu\nu}$ given in (4.39). But, with the Weyl tensor being nonzero in the presence of any inhomogeneity (such as a mass source which is localized in some region), its dynamics is controlled by the fourth order (4.34). In such a case by matching the interior and exterior solutions to (4.34) at the surface of the source [130] the parameter β in (4.37) can be related to the energy–momentum tensor of the source with a strength that depends on the value of the dimensionless gravitational coupling constant $\alpha_{\rm g}$, which appears in (4.33) and (4.34). Since this same α_g parameter decouples in highly symmetric cases such as cosmology where $W_{\mu\nu} = 0$, the scales of the cosmological G_{eff} and the local β are thus completely unrelated. Conformal gravity thus naturally releases cosmology from having to be controlled by a local physics scale, to show that in principle local and global physics (viz. homogeneous ($C_{\mu\nu\sigma\tau} = 0$) and inhomogeneous ($C_{\mu\nu\sigma\tau} \neq 0$) physics scales) can be independent.

Finally, while all of the ideas advanced so far to deal with the DM, cosmological constant and quantum gravity problems will involve some sort of paradigm shift, should they all in the end prove to be of insurmountable difficulties, we would then have to face a paradigm shift that might be even bigger still.

Thanks a lot Philip. We take advantage of your introduction to quantum gravity to introduce the following interview. Now Gerard t'Hooft, the Nobel Laureate for physics, will tell us about the process of unification of the fundamental forces of nature, that is not finished yet. Later, we will ask his opinion about the "Anthropic Landscape" and the prospects for these scientific studies in the future.

4.8 Early Universe: Connecting Particle Physics and Cosmology

Dear Gerard *(t'Hooft)*, **what is your opinion on the most recent ideas of unification of gravity with the other fundamental interactions?**

The unification process takes many steps. First, we have the unified theories of all forces except gravity. The idea is to identify a large Yang-Mills gauge group that has all known gauge groups, U(1), SU(2), and SU(3), as subgroups. It appears that SU(5) and SO(10) have the right algebraic properties, and SO(10) appears to agree best when the observed stability of the proton is taken into account. However, even this relatively simple first step is as yet far from certain; we only have a few numbers, such as the relative strengths of the coupling parameters, that seem to point in this direction, but the hierarchy problem is not solved. This is the problem of the vastly different scales that appear to separate the masses of the particles that emerge from this scheme; we fail to understand the origin of these wide scale separations. The most likely answer to this mystery is often assumed to involve supersymmetry, the symmetry between fermions and bosons, but this is not enough. Already here, new ideas are needed.

On top of that, we now try to add the gravitational force. Again, the scale at which this interaction takes place is different, and we are eager to get a better understanding of the origin of this difference. Naive theories tend to yield only small differences in scales. Here also, we are not waiting for radical flashes of insight that overthrow everything we know. The clashes in our understanding are not in the grand picture, but in some details such as the scale hierarchies. And here also, I do expect a gradual paradigm shift.

I have what could be characterized as a minority's view on the interpretation of quantum mechanics. I refuse to believe that quantum mechanics will be at the foundation of a new theory in the same way as it is in present day atomic and particle physics. A fundamental theory should show a deterministic law of evolution. This is the direction that I feel to be drawn to, whenever I attempt to imagine fundamentally finite and comprehensive theories: quantum mechanics is "emergent", it is a feature that controls the statistics of the outcomes of experiments at scales much bigger than the Planck scale. It is as if we wish to make a theory of sand grains. Individual grains of sand obey very precisely defined dynamical laws, but when we want to build a castle out of sand, what we need is the statistical properties of sand grains. I like to compare atoms, molecules, elementary particles and such with sand castles. To understand these, you need to know how sand behaves as a material, rather than the properties if individual grains. I think that physics at the Planck scale will be like the dynamical laws of individual sand grains. The reason why determinism is needed at that scale I cannot explain in these pages [220, 221], but is closely related to the breakdown of space-time continuity that many of us expect at the Planck scale.

What do you think of the crisis of theoretical physics in explaining our early Universe? Do you agree with Lee Smolin about the failure of String theory? Do you recognize in the Anthropic landscape of Leonard Susskind a symptom of such a crisis?

I do not at all agree that there is a crisis. I would not call it a crisis if our progress is a bit slower than expected. There are differences of opinion on the direction to go. This is usually considered as a healthy situation in science.

Whether or not string theory is to be regarded as a failure depends very much on what your expectations of this theory were. I vividly remember the early 1980s, when indeed string theory was the new hype. There were quite a few enthusiasts – I would not name them now – who strongly believed that string theory would solve all our extant problems; the algebraic features of the Standard Model, the values of the coupling parameters such as the fine structure constant, the mass ratios of all elementary particles, and all other fundamental parameters would soon be explained in terms of the new theories. Some proclaimed in public that "physics would be over by the end of the century", and yes, they meant the twentieth century. I was ridiculed when I protested against so much naivety. I said that such claims were wrong three times over: they were wrong, unwise and unfair. First, I was convinced that the Standard Model could never be explained by string theory; the ingredients for such an explanation were totally absent. Second, it was unwise to make such claims, as they would obviously be debunked, and such an exposure would be damaging to our reputation as theoretical physicists. Third, the claims were unfair; one day a young fellow might discover the real explanations for our present mysteries, and quite possibly some methods now used in string theory might be included in his/her arguments. It would be unfair then to claim that it was all foreseen by the string theorists of the 1980s, because clearly they had no idea how to derive such results.

Having had no such high expectations ever of string theory, I am not overly disappointed by today's more modest views about the status of these ideas. Quite to the contrary, string theory has survived better than I had expected. It is showing more internal consistency, in particular with regard to BHs, it revealed remarkable relations among different schemes and models through relations called "duality" and "holography", and it appears to provide for new technical possibilities to do certain complex calculations in other arias of field theory, such as Quantum Cromo Dynamics (QCD). As string theory is fundamentally quantum mechanical, I do not expect that it will ever serve as the "final theory", but it is quite conceivable that string theory will help us along the way towards such a theory. I am convinced that string theory, at best, will serve as a very special mathematical instrument to help us understand to categorize the effects of the deeper theory that we are still searching for. In short, string theory is an effective, emergent theory just as what we say today about the Standard Model.

Lee Smolin did not do a good service to the science community by writing this book [212]. The public at large might conclude that string theorists were dishonest, imposing, and a disgrace to science. They were none of those things. String theorists were acting in response to a strong conviction which now has been put better in perspective. They were right in suspecting that this theory would have a beautiful and suggestive inner structure and that it might have a lot to do with reality, although the exact way in which this relation works still escapes us. Many of the string theorists were too optimistic about the theory's potential, but they should not be blamed for that as nobody could exactly foresee which results would come next. In fact, I am more annoyed by the second part of Smolin's book, where he advertises "loop quantum gravity" while dropping all of the critical and rigorous demands with which he did attack string theory in the first part. Loop quantum gravity is an interesting and promising approach but it is not at all superior to string theory.

String theory came with the remarkable concept of the "landscape", where nothing but the AP could be used to identify our place in it. The problem with the landscape is that it is clearly a very ugly picture of our world, while it cannot be dismissed off hand as being wrong. It can be defended as being a natural next step after Copernicus removed planet Earth from the center of the solar system, and after Hubble's expanding Universe made it clear that neither the solar system, nor the entire galaxy are anywhere close to the center of the Universe. But, as experimental verification might never be possible, most of us would rather dismiss the landscape as being a nasty aberration ensuing from taking string theory too literally. I believe that it is too early to tell whether our theories indeed imply the existence of a landscape of Universes. I take it that what the landscape really means is that our present theoretical insights cannot exclude the emergence of a couple of hundred apparently arbitrary digits as "constants of nature". The only way to establish the values of these digits is by experiment. Well, today's Standard Model displays the same situation: there are constants of nature whose values we cannot derive from first principles. So, string theory has given us not as much progress as was hoped for, but did not make the situation much worse than what it was. Please do keep in mind, however, that in my conviction string theory is not the end of the road, and so our views concerning freely adjustable parameters may well change in the future.

What is your opinion about the status and the current perspectives of the studies on the very early Universe? Can you discuss the major advances and the limits of the present day theories concerning the physical conditions of the Universe at these epochs? Can we hope to experimentally prove some of these theories in a near future?

Some 30 years ago, cosmology, in my view, was not much more than Science Fantasy, a myth about the creation of times without any solid backing in real science. Now, this has changed. We have data, solid data that can be used to select one theory out of many others. I think this is a marvelous new situation. I am astonished about the successes of the inflationary Universe, a theory I do have some difficulties with, as it requires a very unnatural type of interactions. I do realize, of course, that something like inflation must have occurred during the earliest times; otherwise, the present Universe could not have grown as large as it is today. Apparently, theoreticians did manage to capture some of its details correctly; otherwise the agreement with the observed fluctuations in the cosmic background radiation could not have been so good.

So, to answer your question: we have already seen many experimental, or rather, observational tests of these theories. They are not totally exclusive of course, but they strongly indicate that theorists may well be on the right track. More of such tests will be possible in the near future, when better telescopes and other vital detectors will become available. New generations of telescopes are to be expected, since the technological limits for telescopes are still far away. In the not so distant future, I expect, colossal light-gathering machines can be put in outer space with angular resolutions that today can only be dreamed of. They will give us sharp images of what the Universe was like when it was very young. This is important information, because we need to know how inhomogeneities and other structures came about in the Universe, and how DM is distributed. The cosmic background measurements can be improved by measuring polarizations, and detection of gravitational waves will add new dimensions to observational cosmology. At the other end of the scale, particle accelerators such as LHC and its successors, probably in the form of linear particle colliders, will tell us more about the physics of the elementary particles that we urgently need to understand the beginning of time.

Yet the biggest advances should come from our theories. We have only a very rudimentary understanding of the dynamics of the early Universe. The basis of the very early beginnings of the cosmos is as difficult to understand as the principles of elementary particle theory. In fact, they go hand in hand. There are prohibitive limitations to our human ability to do logic thinking, but nevertheless I do have a lot of confidence in humanity; I think we will be able to overcome many more hurdles in the near future. We would not resolve the most difficult issues very soon, but I am sure that much more progress can be expected.

The question whether there are limits to what we can do, is difficult to answer. Are there limits to our abilities in any science? When it comes to our abilities to construct theories, it is hard to imagine that we are bounded by limits. As for cosmology, it is often claimed that the ultimate moment when the Universe got started, and Nature's laws were "switched" on, in short, the "initial state", can only be explained by theologians. Also our Universe's "reason of existence" can surely not be examined by theorists. However, even here I am not so sure; one can imagine that theories answering such questions might be formulated one day. Some say that the Universe might not have a beginning. Others say that the AP explains our "raison d'être", so perhaps it is fair to suspect that there might be no limit for theories.

Are there limits to our capacities to observe things? How precisely will we be able to determine the Universe's age? Could error margins of just a few years be possible? Today, that would be difficult to imagine, but one never knows where human ingenuity will lead us.

Thank you Gerard. Indeed the quantum nature of the early Universe is today very poorly known and a lot of surprises will certainly come up in the future.

An old question, still connected with the primeval quantum condition of the Universe, is that concerning the values of the constants of physics. We will read below the opinion of Keith Olive on this.

4.9 Constants in Physics?

Dear Keith *(Olive)*, **since Dirac, the time coincidence of the age of the Universe with the product of the electron radius with the ratio of the electric and gravitational forces, has found any possible explanation. This is only one of the many curiosities coming from combinations of the physics constants. Would you please introduce the reader to this very interesting subject and give us your opinions on such coincidences? Do you believe that such things may find an explanation when the quantum nature of our Universe will be much clear to us?**

Several alternatives ideas in cosmology are based on the possibility that the constants of physics vary along the age of the Universe. An example for all is that of a much higher velocity of light at the beginning, invoked to avoid the difficulties coming from inflation. Would you like to comment about the robustness of such alternative scenarios?

The notion of time-varying constants goes back to Dirac and his large number hypothesis [63]. Dirac noticed that the ratio of the electromagnetic interaction between a proton and an electron to their gravitational interaction, $e^2/G_N m_p m_e \sim 10^{40}$, is roughly the same as the ratio of the size of the Universe to the "size" of the electron, $m_e c^3/e^2 H_0 \sim 10^{40}$, where $H_0 \simeq 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ is the present-day Hubble parameter. Furthermore, both ratios are roughly the square root of the total number of baryons in the observable Universe, $c^3/m_p G_N H_0 \sim 10^{80}$. Dirac supposed that if the similarity in these ratios is not coincidence, then they should remain constant over time. Noting that $H_0 \sim t^{-1}$ is not constant, Dirac proposed a time variation in Newton's constant $G_N \sim t^{-1}$. However, the desired result could have been achieved by taking $e^2/m_e \sim t^{1/2}$. The choice simply depends on one's choice of units. Dirac's choice of units would naturally fix as constants the quantities e, c, and m_e. However, in Planck units, one may rather choose G_N, c, and $\hbar$ as fixed. Of course, the large number hypothesis has been excluded by experiments as the predicted variation of $\dot{G}_N/G_N \sim -10^{-10}$ year^{-1} is about two orders of magnitude larger than the limits from the Viking landers on Mars which gave $\dot{G}_N/G_N = (2\pm4)\times10^{-12}$ year^{-1} [103].

At first glance, there would appear to be a long list of potentially nonconstant constants: the fine-structure constant, α; the speed of light, c; Newton's constant, G_N; Botzmann's constant, k_B; Planck's constant, $\hbar$; the electric permitivity and magnetic permeability of space, ϵ_0 and μ_0; Fermi's constant, G_F, and other gauge coupling constants; Yukawa coupling constants which fix the masses of quarks and leptons; etc. However, in this list we must first distinguish what may be called a fundamental parameter of the theory vs. a fundamental unit. Variations in the latter (or in any dimensionful quantity) simply reflect a change in our system of units and as such are not unambiguously observable. That is not to say the Universe with a variable speed of light is equivalent to one where the speed of light is fixed, but that any observable difference between these two Universes cannot be uniquely ascribed to the variation in c. In contrast, variations in dimensionless parameters represent fundamental and observable effects. As such, it becomes operationally meaningless to talk about "measuring" the variation in the speed of light or whether a variation in α is due to a variation in c or $\hbar$. It is simply a variation in α. A discussion on the number of fundamental units in physics is the subject of a trialogue by Duff, Okun, and Veneziano [70].

Okun [154] provides a nice example based on the hydrogen atom which illustrates our inability to detect the variation in c despite the physical changes such a variation would inflict. Lowering the value of c would lower the rest mass energy of an electron $E_e = m_e c^2$. When $2E_e$ becomes smaller than the binding energy of the electron to the proton in a hydrogen atom, $E_b = m_e e^4/2\hbar^2$, it becomes energetically favorable for the proton to decay to a hydrogen atom and a positron. Clearly this is an observable effect providing evidence that *some* constant of nature has changed. However, the quantity of interest is really the ratio $E_b/2E_e = e^4/4\hbar^2 c^2 = \alpha^2/4$. Therefore, one cannot distinguish which constant among e, $\hbar$, and c other than the dimensionless quantity $\alpha = e^2/\hbar c$ is changing.

The prospect that the fundamental constants of nature vary in time has been piqued by the indication that the fine structure constant was smaller at cosmological redshifts $z = 0.5 \div 3.5$ as suggested by observations of quasar absorption systems [148, 226]. The statistically significant result of $\Delta\alpha/\alpha = (-0.54 \pm 0.12) \times 10^{-5}$, where $\Delta\alpha$ is defined as the past value minus the present one, is based on the many-multiplet method which makes use of the α dependence of the relativistic corrections to atomic energy levels and allows for sensitivities which approach the level of 10^{-6}. This method compares the line shifts of elements which are particularly sensitive to changes in α with those that are not. At relatively low redshift ($z < 1.8$), the method relies on the comparison of Fe lines to Mg lines. At higher redshift, the comparison is mainly between Fe and Si. At all redshifts, other elemental transitions are also included in the analysis.

More recent observations taken at VLT/UVES using the many multiplet method have not been able to duplicate the previous result [45,184,216]. The use of Fe lines in [184] on a single absorber found $\Delta\alpha/\alpha = (-0.05 \pm 0.17) \times 10^{-5}$. However, since the previous result relied on a statistical average of over 100 absorbers, it is not clear that these two results are in contradiction. In [45], the use of Mg and Fe lines in a set of 23 systems yielded the result $\Delta\alpha/\alpha = (-0.06 \pm 0.06) \times 10^{-5}$ and therefore represents a more significant disagreement and can be used to set very stringent limits on the possible variation in α. The latter analysis has been recently criticized [149] and defended [217].

The result found in [45] and in the statistically dominant subsample of 74 out of the 128 low redshift absorbers used in [148] are sensitive to the assumed isotopic abundance ratio of Mg. In both analyses, a Solar ratio of ^{24}Mg:^{25}Mg:^{26}Mg = 79:10:11 was adopted. However, the resulting shift in α is very sensitive to this ratio. Furthermore, it is commonly assumed that the heavy Mg isotopes are absent in low metallicity environments characteristic of Quasi Stellar Object (QSO) absorption systems. Indeed, had the analyses assumed only pure ^{24}Mg is present in the QSO absorbers, a much more significant result would have been obtained. The Keck/Hires data [148] would have yielded $\Delta\alpha/\alpha = (-0.98 \pm 0.13) \times 10^{-5}$ for the low redshift subsample and $\Delta\alpha/\alpha = (-0.36 \pm 0.06) \times 10^{-5}$ for the VLT/UVES data [45].

The sensitivity to the Mg isotopic ratio has led to a new possible interpretation of the many multiplet results [4, 5]. The apparent variation in α in the Fe–Mg systems can be explained by the early nucleosynthesis of ^{25,26}Mg. A ratio of (^{25}Mg + ^{26}Mg)/^{24}Mg = 0.62 ± 0.05 (0.30 ± 0.01) is required by the data in [148] ([45]).

The heavy Mg isotopes are efficiently produced in the asymptotic giant branch phase of intermediate mass stars, particular in stars with masses 4–6 times the mass of the sun. During hot-bottom burning, these stars become sufficiently hot ($T > 7 \times 10^7$ K) so that proton capture processes in the Mg–Al cycle become effective. Proton capture on ^{24}Mg then leads to the production of ^{25}Mg (from the decay of ^{25}Al) and to ^{26}Al (which decays to ^{26}Mg). A second contributing process occurs deeper in the star during thermal pulses of the helium-burning shell. During these thermal pulses, ^{22}Ne is produced by α-capture on ^{14}N, which itself is left over from the CNO cycle. Heavy magnesium isotopes are then produced via the ^{22}Ne(α,n)^{25}Mg and ^{22}Ne$(\alpha,\gamma)^{26}$Mg reactions.

Whether or not the claimed variations in α are real, there exist various sensitive experimental checks that constrain the variation of coupling constants (see e.g., [224]). Limits can be derived from cosmology (from both big bang nucleosynthesis and the microwave background), the Oklo reactor, long-lived isotopes found in meteoritic samples, and atomic clock measurements.

The success of BBN relies on a fine balance between the overall expansion rate of the Universe and the weak interaction rates which control the relative number of neutrons to protons at the onset of nucleosynthesis. Changes in the expansion rate, which is proportional to $\sqrt{G_{\rm N} N}$, where N is the number of relativistic particles, or changes in the weak rates, which may result from changes in fundamental parameters, affect the neutron to proton ratio and ultimately the ^{4}He abundance, Y. Thus one can use the concordance between the theory and the observational determination of the light element abundances to constrain new physics [54].

Changes in the fine structure constant affect directly the neutron–proton mass difference which can be expressed as $\Delta m_{\rm N} \sim a\alpha\Lambda_{\rm QCD} + bv$, where $\Lambda_{\rm QCD} \sim \mathcal{O}(100)$ MeV is the mass scale associated with strong interactions, $v \sim \mathcal{O}(100)$ GeV determines the weak scale, and a and b are numbers which fix the final contribution to $\Delta m_{\rm N}$ to be -0.8 and 2.1 MeV, respectively. As one can see, changes in α directly induce changes in $\Delta m_{\rm N}$, which affects the neutron to proton ratio. The relatively good agreement between theory and observation, $|\Delta Y/Y| \lesssim 5\%$ allows one to set a limit $|\Delta\alpha/\alpha| \lesssim 0.05$ ($\Delta Y/Y$ scales with $\Delta\alpha/\alpha$) [19, 31, 109, 117, 153]. Since this limit is applied over the age of the Universe, we can obtain a limit on the rate of change $|\dot{\alpha}/\alpha| \lesssim 4 \times 10^{-12}$ year^{-1} over the last 13 Gyr.

In the context of unified or string-inspired theories, it is possible to derive significantly stronger limits on the variation of α if all gauge couplings are related at some unification scale. In this case, a change in the fine structure constant would imply a change in other couplings and mass scales (if Yukawa couplings are also changed in a correlated manner) [31]. The dominant effects are found in induced variations of the QCD scale $\Lambda_{\rm QCD}$ and the weak scale Higgs expectation value v. The variations $\Delta\Lambda/\Lambda \simeq 30\Delta\alpha/\alpha$ and $\Delta v/v \sim 80\Delta\alpha/\alpha$ [31] are translated into variations in all low energy particle masses. In short, once we allow α to vary, virtually all masses and couplings are expected to vary as well, typically much more strongly than the variation induced by the Coulomb interaction alone. For example, the nucleosynthesis bound improves by about two orders of magnitude in this case.

Coupled changes in α can have effects on the light elements, which go beyond altering the ^{4}He abundance [48,61,64,110,147]. In addition to affecting the neutron–proton mass difference (as well as the neutron lifetime), changes in fundamental parameters will affect the deuteron binding energy [48, 64]. Unfortunately, there is considerable model dependence in the relation between quantities such as $\Delta B_{\rm D}/B_{\rm D}$ and $\Delta\alpha/\alpha$ although reasonable estimates can be made. In addition to its effect on the deuterium abundance, changes in the binding energy of deuterium has a large effect on the late-time abundance of ^{7}Be, which after decay alters the primordial ^{7}Li abundance. Indeed, a change in the binding energy of only a few percent, corresponding to a change in α of only a few parts per 10^5, can lower the ^{7}Li abundance to match observations, while at the same time leaving the other elements compatible with their respective observed abundances [48].

One can also derive cosmological bounds based on the microwave background. Changes in the fine-structure constant lead directly to changes in the hydrogen binding energy, $E_{\rm b}$. As the Universe expands, its radiation cools to a temperature at which protons and electrons can combine to form neutral hydrogen atoms, allowing the photons to decouple and free stream. Measurements of the microwave background can determine this temperature to reasonably high accuracy (a few percent) [215]. Decoupling occurs when $\eta^{-1}\exp(-E_{\rm b}/T) \sim 1$, where $\eta \sim 6 \times 10^{-10}$ is the ratio of the number density of baryons (protons and neutrons) to that of photons.

Thus, changes in α of at most a few percent can be tolerated over the time scale associated with decoupling (a redshift of $z \sim 1100$) [193].

Interesting constraints on the variation of α can be obtained from the Oklo phenomenon concerning the operation of a natural reactor in a rich uranium deposit in Gabon approximately two billion years ago. The observed isotopic abundance distribution at Oklo can be related to the cross section for neutron capture on ^{149}Sm [72, 208]. This cross section depends sensitively on the neutron resonance energy E_r for radiative capture by ^{149}Sm into an excited state of ^{150}Sm. The observed isotopic ratios only allow a small shift of $|\Delta E_r| \sim E_r$ from the present value of $E_r = 0.0973$ eV. This then constrains the possible variations in the energy difference between the excited state of ^{150}Sm and the ground state of ^{149}Sm over the last two billion years. Assuming that the energy difference is due to the α-dependence of the Coulomb energy alone, a limit $|\Delta\alpha/\alpha| \lesssim 10^{-7}$ can be obtained [72,84]. However, if all fundamental couplings are allowed to vary interdependently, a much more stringent limit $|\Delta\alpha/\alpha| < (1-5) \times 10^{-10}$ may be obtained [155].

Bounds on the variation of the fundamental couplings can also be obtained from our knowledge of the lifetimes of certain long-lived nuclei. In particular, it is possible to use precise meteoritic data to constrain nuclear decay rates back to the time of Solar system formation (about 4.6 Gyr ago). Thus, we can derive a constraint on possible variations at a redshift $z \simeq 0.45$ bordering the range ($z = 0.5 \div 3.5$) over which such variations are claimed to be observed. The pioneering study on the effect of variations of fundamental constants on radioactive decay lifetimes was performed by Peebles and Dicke [177] and by Dyson [72]. The isotopes which are most sensitive to changes in α are typically those with the lowest β-decay Q-value, Q_β. The isotope with the smallest Q_β value (2.66 ± 0.02 keV) is ^{187}Re. A precise age of 4.558 Gyr for angrite meteorites can be determined by the ^{207}Pb-^{206}Pb method [126]. The data on ^{187}Re and ^{187}Os in iron meteorites formed within 5 Myr of the angrite meteorites [210] then give a limit $-8 \times 10^{-7} < \Delta\alpha/\alpha < 24 \times 10^{-7}$ [156].

Finally, there are a number of present-day laboratory limits on the variability of the fine-structure constant using atomic clocks. Three recent experiments have led to marked improvement in the limit on $\Delta\alpha/\alpha$. An experiment comparing hyperfine transitions in ^{87}Rb and ^{133}Cs over a period of about 4 years yields $\Delta\alpha/\alpha < 6 \times 10^{-15}$ [132]. Another experiment comparing an electric quadrupole transition in $^{199}Hg^{+}$ to the ground-state hyperfine splitting in ^{133}Cs yields $\Delta\alpha/\alpha < 6 \times 10^{-15}$ over a 3 year period [23], and an experiment comparing the 1S-2S transition frequency in atomic hydrogen to ^{133}Cs yields $\Delta\alpha/\alpha = (1.1 \pm 2.3) \times 10^{-15}$ over a 4 year period [83]. The combined results from the latter two give $\Delta\alpha/\alpha = (-0.9 \pm 4.2) \times 10^{-15}$ or $\dot{\alpha}/\alpha \lesssim 10^{-15}$ year^{-1} [83].

Thank you very much Keith. Now following the flavor of the Anthropic Principle we will enter in the question of the predictive power of such a principle, which in turn is connected with the problem of why the constants of physics have the values they do. Luigi Secco will introduce us to this field.

4.10 On the Anthropic Principles

Dear Luigi *(Secco)***, there are several flavors of the "Cosmological Principle", such as the "Copernican", the "Perfect" and the "Weak Anthropic" and "Strong Anthropic" ones. Would you like to stress the differences and the most important assumptions underlying each one?**

Before discovery of CMB, thanks to and Wilson, 1965 (e.g., [9] p. 368), two opposite alternatives were possible regarding cosmological evolution. One was related to the *Perfect Cosmological Principle*, that is, *The Universe is, on average, the same everywhere, in all directions and at all times.* On this Principle, Bondi, Gold, Hoyle and Narlikar [24, 105] grounded their *Steady State theory* of cosmic description. To cure the problem of lack of constant density due to the discovery by Hubble [108] that the Universe expands with time, they were available to assume that the creation of only one hydrogen atom/liter each 5×10^{11} years had resolved the problem [24]. In Hoyle's words [105]: *Using continuous creation of matter, we shall attempt to obtain, within the framework of the general theory of relativity, but without introducing a cosmical constant, a Universe satisfying the wide cosmological principle* (i.e., the perfect one) *that shows the required expansion properties.* The other alternative was that the Universe did not be characterized by the monotony imposed by the Perfect Cosmological Principle but, on the contrary, by a high level of variety and fantasy owing to the continuous change of density and temperature. Penzias's and Wilson's discovery proved that it was in this second way that expansion occurs. Indeed, perfect cosmological behavior was ruled out by the perfect BB energy distribution of CMB, which tells us that matter and radiation were in thermodynamic equilibrium in the remote past even if they are not now. At present cosmological density (of about 3 protons/m^3) one photon actually needs more than the Universe age to interact with an electron. The conclusion was that density must have changed relentlessly during cosmic evolution and its temperature as well. A fantastic scenario appeared in the history of the Universe marked by a sequence of different physical phenomena, exactly the opposite of the Perfect Cosmological Principle's depiction. The latter subtracted to the Universe all the inventiveness and fantasy, relegating its story to a flat picture albeit obeying our mental needs of *Ockham's Razor Principle*[44].

4.10.1 Cosmological Principle

In the meantime, the large surveys (e.g., that of 157320 galaxies using the Anglo Australian Telescope (e.g., [174]), the other which contains around 10^6 galaxies collected in Shane's and Wirtanen's catalogue ([51] p. 293) and the APM survey

---

[44] Proposed by William of Ockham in the fourteenth century: *Pluralitas non est ponenda sine neccesitate*, which translates as *entities should not be multiplied unnecessarily*.

with over two million galaxies (e.g., [125])) began to map the Universe structure on the large scale. At the end of 1989, the COBE satellite was launched. It measured the BB temperature of the CMB ($\sim$3 K) and revealed its very small fluctuations (of the order of 10^{-5}) at the Universe age of about 3×10^5 years after the Big-Bang [50]. So the texture of the young and old Universe appeared to be characterized by a high degree of *homogeneity* and *isotropy* ratified in the *Cosmological Principle*: *On the large scale (greater than 200 Mpc) the Universe is, at a good extent, homogeneous and isotropic at any one time.* To describe the cosmological evolution one then needs to set up a metric for the space-time, which takes into account this Principle. From the point of view of a physical observer, it means that each hyper-surface which describes the Universe in a four-dimensional space has to appear to him at all times to be *homogeneous* and *isotropic* and to be the same even if the observer changes. In other words, the Cosmological Principle implies the *Copernican Principle* (e.g., [175]) that is: *We are not privileged observers of the Universe*. In turn, if we find that distant galaxies are heading away from us in all directions, that does not mean that we have the privilege of being at the fixed center of an expanding Universe, but rather that the nature of cosmic expansion is such that the recession of distant galaxies is what would be observed by any other observer located in any other point in the Universe (see [113]). The immediate consequence is: a distant galaxy has to move away in all directions because there is not a special one, but all directions means no direction at all, i.e., *it is the reference frame which expands*.

4.10.2 Modern Cosmology and Center of the Universe

After the conclusion of the Great Debate, which occurred at the National Academy of Sciences in Washington (April of 1920) (e.g., Chap. 1 of [136] and references therein) between Shapley and Curtis, the sequence of shifts related to the center of Universe arrived at its end. This center had been moved from Earth to Sun by Copernicus (1543) and Galileo (1609). After them, at the end of the eighteenth century, Herschel concluded that the Sun lay roughly at the center of his first Milky Way map, and until Kapteyn's Universe (e.g., [115]) the heliocentric position was strongly sustained. It was Shapley who discovered the Sun's off-center of about 15 kpc (between 1915 and 1919) using globular cluster distribution in galactic longitude. But *the last act of the Copernican revolution* was the conclusion of the Great Debate: neither the Earth, nor the Sun, nor even our Galaxy was special. Modern Cosmology will very soon conclude that the center of the Universe does not exist. This result since the beginning of the twentieth century has had and nowadays continues to have an enormous impact on the cultural level. If the Life site inside the Universe texture was without any relevance, the Life itself appeared to be like mould on a lost planet in the Universe's immensity. Man and his Life are absolutely marginal in the cosmical context, to such a point that, according to the Copernican Principle, the human view point itself is not of any importance because each observer located anywhere in the Universe would see the Universe in the same way. On one side, Man, by accepting this tremendous humiliation, has been able to construct cosmological models actually based on the Copernican Principle, which on the other side, have brought again to the scene the connection between Cosmos and Life.

4.10.3 The Large Numbers Puzzle

The story began in 1923 with Eddington, who suggested that the large total number of protons inside the Universe's horizon, $N \simeq 10^{80}$, might play a part in determining some fundamental constants of Nature ([9] Chap. 4). There are indeed some Large Numbers coincidences: e.g., the *coupling constant of the gravitational force* is, in dimensionless units:

$$\alpha_G = G m_N^2 / \hbar c \simeq 10^{-39}$$

where m_N is the nucleon mass, G the gravitational constant, $\hbar$ the Planck's constant/2π, and c the light speed. α_G turns out to be about $N^{-1/2}$. Moreover the ratio of electric force between proton and electron to their gravitational force is:

$$N_1 = e^2 / G m_p m_e \simeq 10^{39}$$

that is about $N^{1/2}$ (m_p, m_e proton and electron mass, respectively and e the elementary charge). Even the ratio between the age of Universe, t_o, and the light crossing time of the classical electron radius, r_e, is of the same order of magnitude:

$$N_3 = \frac{t_o}{r_e / c} \simeq 6 \cdot 10^{39}.$$

Since 1937 Dirac [63] took into account very seriously these coincidences, concluding that gravity must weaken at increasing cosmic time (*see the previous contribution of Keith Olive in this book*). To understand the deep meaning of these kinds of apparently strange speculations, we have to look back at what we know now by means of modern cosmological models. The large numbers relationships showed indeed that there was a mysterious connection between the *macro-cosmos* and the *micro-cosmos* as now has been proved. The increasing of microscopic complexity at the beginning of Universe (let us think, for example, of the *recombination epoch* when the atoms form) occurred indeed in order to allow for the formation of LSS, which in turn would allow microscopic complexity growth to continue by the nuclear synthesis inside stars formed as *sub-units* of these structures. One appears related to the other in an extraordinarily beautiful link. To these interesting speculations in the 1930s one should add the relationship underlined by Dicke in 1961 between Universe age and the time requested to form carbon, necessary to Life, inside the stars, which again was another macro–micro cosmos link.

4.10.4 Anthropic Principle

Yet it was Carter in 1974 who for the first time pointed out that the key to previous links was Life. It is due to him that we have the first formulation of the so-called Anthropic Principle: *what we can expect to observe must be restricted by the conditions necessary for our presence as observers* ([44] p. 291). Barrow and Tipler [9] took up again Carter's idea by reformulating the AP in two ways, in the Weak form as follows: *the observed values of all physical and cosmological quantities are not equally probable but they take on values restricted by the requirement that there exist sites where carbon-based life can evolve and by the requirement that the Universe be old enough for it to have already done so*. In the Strong form it is as follows: *the Universe must have those properties which allow life to develop within it at some stage in its history*. What is the relevance of these two formulations? They open the door to a large spectrum of interpretations. I think many controversies arise from the different meanings people assign to the same statements. It seems to me that the most important thing lies in this: *for the first time a physical connection between histories of the Universe and Life is clearly enunciated*. To ground these Principles, a very long chain of physical phenomena was collected by [9] with the impressive *fine tunings* Life requests in order to appear. We will refer to them later in this interview.

We wish first of all to underline the deep difference between Weak AP and Strong AP. It should be noted that the Weak AP formulation changes into the Strong AP one essentially by substituting *can* by *must*. That means the *possibility* of Life development in the Weak AP version becomes *necessity* in the Strong AP one. That implies the existence of a metaphysical finality in the whole cosmological evolution. A metaphysical component appears indeed in the Strong AP formulation. On the contrary, when we analyze the different characteristics of some cosmos evolutionary phases or the chemical–physical properties of some substances involved in the processes of Life building up, we do not pass over the Science domain. Indeed we simply consider Life itself as a complex phenomenon, looking at it as a function of many parameters and testing how sensitive it is to a small variation of one of its variables. Nevertheless, taking into account the probability of the *monstrous sequence*[45] [106] of compatible and independent events of which Life needs, it is manifest how *the probability of Life not appearing is great*. But *we are*! Then the question arises: *why?* In my opinion it is important to be as precise as possible about the limit between the two domains, that of Science and that of all which lies above the Science. We can refer to the latter as Metaphysics (following the original meaning related to the order of Aristotle's works, *metà tà fusiká*, that is, *all that is beyond the Physics*) or Trascendence, which are used as synonyms in this context. The problem of Science–Trascendence relationship goes back to the origin of modern Science thanks to Galileo. Even if the debate is outside the aim of

[45] *I do not believe that any scientist who examined the evidence would fail to draw the inference that the laws of nuclear physics have been deliberately designed with regard to the consequences they produce inside the stars. If this is so, then my apparently random quirks have become part of a deep-laid scheme. If not then we are back again at a monstrous sequence of accidents.* From [106].

this contribution, let me stress which are in my opinion some of main lines of the problem in order to locate Weak AP and Strong AP correctly. Science refers to the domain of quantity, of corporeity; Metaphysics refers to the domain of what lies beyond them. The two methodologies are completely different (see [55]). The two domains develop along two orthogonal axes: the horizontal one, Science, the vertical one, Trascendence. Then, inside the Science domain are soundly located the questions of the incredible fine tuning connections between Cosmos and Life which Weak AP underlines. On the contrary, when we wish to add to the *whys* the possible answers to them as Strong AP claims, we are going beyond Science. But I think the requests Weak AP asks us to answer are so strong and so unavoidable that it is wrongfully considered weak. Then hereafter we will reformulate Weak AP as follows: *the collection of facts which connect the factors (see next answer) – constraining the main features of cosmos and its evolution – with Life, by which it is possible to infer how strongly the Life phenomenon is depending on these factors*. We refer to it as WRAP.

What is the interplay between the present cosmological scenarios and the development of life?

To answer this question, we have first to stress the principal factors which constrain the main features of the Universe. We can briefly group them into the three following sectors (see [56]):

A. The values of the main constants in Fundamental Physics
B. The global properties of the Universe and its history
C. The space dimensionality

Related to each sector, a huge collection of relationships proposed by many authors (e.g., [8, 9, 56, 96, 187]) may prove the interplay between the present cosmological scenarios and the development of Life. To keep it short we will take into account only few exemplifications for the sectors (A) and (B). To understand how special is the spatial dimensionality equal 3 of our Universe, we invite the reader to refer at the references cited above.

A. *The values of the main constants in Fundamental Physics*. We will refer only to the coupling constant values of the four forces. At the very early phases after the Big Bang there are no doubts that the Universe was the highest energy laboratory of particle Physics. Even if the reliability of the physical phenomena is decreasing, going back toward singularity, the electro-weak unification proved at about 100 GeV at Ginevra CERN, allow us to take seriously into account the other possible unification as depicted by the Grand Unification Theory (GUT) at the time of about $\sim 10^{-35}$ s. At this time, when the energy density of the Universe had to be about 10^{15} GeV, according to GUT the three forces: the strong and the electro-weak could be unified. The spontaneous breaking of the high level symmetry corresponding to this unified force, due to the Universe's expansion allows the strong force to separate from the electro-weak by assuming a coupling constant which now is equal to: $\alpha_S \simeq 15$. At about $\simeq 10^{-11}$ s the electro-weak symmetry breaks too and the corresponding unified force splits into the electromagnetic force (with a typical coupling constant value of,

$\alpha \simeq 1/137$) and into the weak one (with a typical coupling constant of, $\alpha_W \simeq 10^{-5}$). These values with which the forces detached one from the other are crucial in the primordial nucleosynthesis epoch ($\simeq$200 s). Indeed if α_S increases only by 0.3%, dineutron binds and with, $\Delta\alpha_S/\alpha_S$, increasing by 3.4% the diproton is bound too. But if, $\Delta\alpha_S/\alpha_S$, decreases less than 9%, the deuterium nucleus fails to be bound (see Davies in [9] Chap. 5). These little changes might have catastrophic consequences for Life. For example, if the deuteron was unbound, the consequences for the nucleosynthesis of elements necessary for development of Biology are strong because *a key link in the chain of nucleosynthesis would be removed* ([9] Chap 5). If, on the contrary, the strong interaction was a little stronger, the diproton stable bound state would have the consequence that *all the hydrogen in the Universe would have been burnt to* ^{2}*He (diproton) during the early stages of the Big Bang and no hydrogen compounds or long-lived stable stars would exist today* ([9] Chap. 5). Indeed the reactions to form ^{4}He would find a channel about 10^{18} times faster in comparison with those without diproton formation. The hydrogen reserve would have been quickly consumed without allowing, for example, the water formation. Moreover the stability of a nucleus of a mass number A and atomic number Z hinges on a fine link between the strengths of electromagnetic and strong forces as follows:

$$\frac{Z^2}{A} \leq 49\left(\frac{\alpha_S}{10^{-1}}\right)^2\left(\frac{1/137}{\alpha}\right)$$

Thus, if the electromagnetic interaction was stronger (increased α*) or a stronger interaction a little weaker (decreased* α_S*), or both, then biologically essential nuclei like carbon would not exist in Nature* ([9] Chap. 5).

These are only very few exemplifications of a huge collection given in the cited book.

B. *The global properties of Universe and its history.*

4.10.5 Growth of Complexity

There is an impressive trend in the whole history of the Universe beginning from the very early phases at which the high level symmetries break. For example, when the Higgs mechanism is at work, the electro-weak force differentiates into weak and electromagnetic ones by distinguishing between, $W^{\pm}$, and, Z_o, massive mediator bosons for the first and the massless photons for the second. The Universe's aim appears to differentiate, to articulate itself by increasing its complexity or, in other words, to grow its microscopic thermodynamical information[46], I, without violating the second thermodynamic Principle, that is increasing in the meantime its total entropy. The same paradigm appears to be followed as well at the *recombination epoch* when atoms form. After this time, the large structures form by enriching the

[46] According to Layzer [121], $I = S_{max} - S$ where S_{max} means the maximum value the entropy of a system may have as soon as the constraints on it, which fix its entropy value to S, are relaxed.

Universe with extraordinarily beautiful complexity on the macroscopic scale too. It appears as if a *large carving was in progress*. But all this is exactly what the end-product, Life, needs. Indeed Life requests a very high level of structuration, of differentiation, of local thermodynamical information. The Universe indeed appears in some sense bend to these needs.

4.10.6 The Fine Tuned Expansion

Before the introduction of inflation mechanism thanks to Guth [100] (e.g., [50]), one of the problems to be solved in modern cosmology was the *flatness problem*. It appeared very strange indeed that in whichever Universe we were (*open*, *critical* or *closed*) the corresponding model describing it had the same limit of the parameter density Ω when time goes toward singularity[47]. Actually all the models not only had to tend to the critical one, that is, with $\Omega = 1$, but in a way surprisingly tuned: for example, at the Planck epoch, $t_P \simeq 10^{-43}$ s, an open Universe had to have: $\Omega(t_P) = 1 - 10^{-62}$, and a closed one, $\Omega(t_P) = 1 + 10^{-62}$. Moreover at the present age, t_o, that is, after about 14×10^9 years the density parameter has not to be much different from 1. The acceptable models indeed have to be characterized by a density parameter inside these limitations: $0.03 < \Omega(t_o) < 2$. It means the Universe has now to differ from a critical one of less than a factor 100, when its density variation along the whole evolution is of 123 orders of magnitude. The Ω constraints are not trivial because they are strictly related to the age of Universe and to the possibility to form structures. Actually, without this small shift which tunes the expansion the Universe had collapsed very soon or expanded so fast that neither stars neither galaxies might form. In both cases the Life *had not the necessary conditions to develop itself*. The inflation seems to take care of this very peculiar initial condition of the Universe even if the previous fine tuning problem transfers into the not completely resolved inflationary modulation (see [50] Chap. 7). Moreover most of the inflationary models does predict an almost flat Universe that is a very special value of $\Omega_o = \Omega(t_o) \simeq 1$. But dynamical estimates about ordinary and DM amount yield typical value of $\Omega(t_o) \simeq 0.3$. In 1998 the analysis of the distant-redshift relationship using high redshift type Ia SNe, has led to the discovery that the Universe expansion from about the time of Sun's birth ($z \simeq 0.5$) is accelerating [190]. According to this extraordinary scientific new the dominant contribution to the present-day energy budget is a component called DE. *Dark* because we ignore its origin and because its equation of state is[48]: $p = wc^2\rho$, with $w < -1/3$ that is an ingredient with an *anti-gravitational character*. The most straightforward candidate to produce it is a positive *cosmological constant* Λ with an equation of state parameter $w = -1$. Immediately a fine-tuning problem again arises. A cosmological constant Λ seems to have caused the inflationary accelerated expansion at the GUT

[47] Ω is the ratio of matter and energy density to the value of the same quantities for the critical Universe. Open Universe corresponds to $\Omega < 1$, critical to $\Omega = 1$, closed to $\Omega > 1$.

[48] The relation between pressure, p, and energy density, ρ; c is the light speed.

epoch if its energy density was $\rho_\Lambda \simeq (10^{15}\text{GeV})^4$ (in natural units). At the present the energy density to obtain $\Omega_{\Lambda o} \simeq 0.7$ (so that $\Omega_o = \Omega_o(matter) + \Omega_{\Lambda o} \simeq 1$) has to be $\rho_{\Lambda o} \simeq 10^{-47}\,\text{GeV}^4$. It means that the DE density had to be decreased of about 123 orders of magnitude from the Planck's scale. The mistery is *why* it had now to become so small but with the right amount in order that the total budget of *ordinary matter+DM+DE* yields properly $\Omega_o \simeq 1$. An interesting answer is given by Rees [187] who includes as one of possible solutions for Λ mistery that *its tuning ... may be a fundamental request for our existence*. Actually an higher value of Λ before or after galaxy formation would cause catastrophic consequences by preventing or destroying their formation owing to its repulsion effect against that of gravitational condensation. Such kind of Universe *would be sterile*.

4.10.7 Carbon and Oxygen Nucleosynthesis

Let us remember at least another of the many incredible bottle-necks through which the Universe's evolutive story passes in order to produce the large amount of carbon that Life needs. That appears when nucleosynthesis inside the first generation of stars transforms three helium nuclei into one of carbon as follows: $3^4\text{He} \rightarrow {}^{12}\text{C}$. At first two helium nuclei collide producing the nucleus of ${}^8\text{Be}$. But this nucleus is unstable and would decade in $\simeq 10^{-7}$ s unless it captures a third helium nucleus in order to change into ${}^{12}\text{C}$. But this chain at the reaction temperature of about $10^8 K$ would not produce enough carbon for Life unless the last reaction was resonant, that means there would exist a level of ${}^{12}\text{C}$ nucleus about equal to the intrinsic energy of the two nuclei ${}^8\text{Be} +{}^4\text{He}$ plus the mean typical kinetic energy of collision at $10^8 K$, so that the reaction rate would increase strongly . The resonance level indeed exists and it corresponds to 7.6549 MeV as Hoyle predicted since 1954. This resonance channel was soon verified by Fowler in the laboratory (e.g., [158, 188]). This is also a good example of the prediction capability the WRAP indeed has owing to its nature of a real physical principle.

It should be noted that the next reaction of carbon burning by which oxygen is produced has also to be tuned but in the opposite way. Indeed the following reaction: ${}^{12}\text{C} +{}^4\text{He} \rightarrow {}^{16}\text{O} + \gamma$ has not to be resonant. If it does, all the carbon would be transformed into oxygen. Luckily it does not even if there is a resonant level for the oxygen nucleus but at a little bit too low energy, 7.1187 MeV. So comparable quantities of carbon and oxygen are produced to make the CO molecule a common one. As consequence the formaldehyde H_2CO is the association of two of the most common molecules (H_2 and CO) in the Universe so that:

$$(\text{H}_2\text{CO})_n \rightarrow \ \textit{sugars and carbohydrates}$$

are easily built up [107]. In turn the carbon and oxygen energy nuclear levels are strictly depending on the values α and α_S properly have. If α would vary more than 4% or α_S more than 0.4% the carbon or oxygen production will change of a factor in the range 30–1,000 [8].

My Conclusions

From this brief glance at the cosmology, we may conclude that the Universe appears to be built up in an extraordinary way. It does not look like a chaotic muddle of things but as something of which intelligibility continues to excite in us as in Einstein[49] some time ago deep surprise. We feel as we were in front of a great cathedral about which science allows us to understand the constructive ingredients but nothing about its ultimate meaning. The question is: *What is the key to trying to understand it?* My answer is: Beauty. Cosmos structure is permeated by Beauty. But Beauty is an *indecipherable cipher* which does not enter into the scientific domain. Nevertheless, it is the key to reading the whole Reality [55]. According to all the great traditional Religions, it tell us something related to an inner tension to offer, to a gratuitous gift, which is a distinguishing feature of Love. Then the Universe and Life inside the Universe appear by this light to be a benevolent, completely free offer. The deepest secret inside the cosmos turns out then not to be Life but Life as Gift [205].

Thanks a lot Luigi.
Now, John Peacock will review for us the "Why Now?" problem, a question that seems to contradict one of the fundamental principles of our cosmology: the fact that we cannot be privileged observers of the Universe.

4.11 Many-Universes

Dear John *(Peacock)***, the existence of a non zero vacuum density has raised two important cosmological questions related with the energy scale associated with the vacuum density and with the problem commonly known as the "Why now?" issue. Would you please comment and explain these questions and express your opinion about the anthropic point of view?**

The scale problem concerns the energy scale corresponding to the vacuum density. If we adopt the values $\Omega_{\rm v} = 0.75$ and $h = 0.73$ for the key cosmological parameters, then

$$\rho_{\rm v} = 7.51 \times 10^{-27}\,\mathrm{kg\,m^{-3}} = \frac{\hbar}{c}\left(\frac{E_{\rm v}}{\hbar c}\right)^4,$$

where $E_{\rm v} = 2.39\,\mathrm{meV}$. The vacuum density should receive contributions from the zero-point fluctuations of all quantum fields, and one would expect a net value for $E_{\rm v}$ of order the scale at which new physics truncates the contributions of high-energy virtual particles: anything from 100 to 10^{19} GeV. The why-now problem further asks

[49] It was Einstein who asked the question: *Why is the world comprehensible?* He could say only: *The eternal mystery of the world is its comprehensibility* (see [86]).

why we are observing the Universe at just around the time when this strangely small vacuum density first comes to dominate the cosmic density.

I think every physicist would start off with the hope that the scale problem can be solved directly, by figuring out how to calculate the vacuum density. There is a tendency to be suspicious of the anthropic line of reasoning, in which we seek to explain the value of the vacuum density as some kind of selection effect. This sounds like taking the easy way out (which I don't think it is), and abandoning the ideal of hoping to calculate everything from first principles. I can sympathize with this attitude, but it certainly can't apply to the why-now problem: this is a question that involves the existence of observers, so observers must necessarily be involved in the answer, and anthropic reasoning cannot be evaded. One might be able to get away with what might be called one-Universe anthropic arguments. Here, we note that observers will inevitably exist only at special times in this Universe: complex structures cannot form until temperatures reach $T \ll 1,000\,\mathrm{K}$ and the Universe becomes matter dominated. If we could tie the onset of vacuum domination to this radiation-matter transition, we would have a convincing package. The quintessence programme aims to find a dynamical contribution to the vacuum energy where the change in cosmological expansion history at matter-radiation equality prompts the effective vacuum density to change from a sub-dominant contribution at early times, to something that resembles Λ by the present. This would be a pleasing solution to the why-now problem, but it is generally agree that the model does not work unless the potential is tuned by hand.

This leads to consideration of the more radical many-Universe mode of anthropic reasoning. Here, one envisages making many copies of the Universe, allowing the value of the vacuum density to vary between different versions. Although most members of the ensemble will have large vacuum densities comparable in magnitude to typical particle-physics scales, rare examples will have much smaller densities. Since large values of the vacuum density will inhibit structure formation, observers will tend to occur in models where the vacuum density falls in a small range about zero – thus potentially solving both the scale and why-now problems. This solution was outlined by Weinberg [228], who actually went further and turned the argument into a *prediction* of a non-zero cosmological constant. Weinberg's view was that the natural value of Λ was large in magnitude, and that the observed value was only prevented from being above some limit by anthropic selection effects – but "there is no reason for it to be smaller". The stunning success of Weinberg's prediction should impress critics who claim that anthropic ideas have no predictive power.

Today, anthropic reasoning has become more respectable through the "landscape" of string theory, driven by the recognition that there will be many possible different vacuum states and low-energy phenomenology. This is an encouraging development if you think the basic reasoning is sound, and there is an analogy with Darwin here. Having enunciated the mechanism of natural selection, a microscopic mechanism (genes and DNA) was needed to make evolution work. But the anthropic argument stands irrespective of the correctness or otherwise of landscape ideas: it is the only argument I know that deals with the why-now question. The longer this state of affairs goes on, the longer we will be driven to think more seriously about an ensemble of Universes and what this means physically. Do the other Universes really exist? Where are they? Could we ever detect them? I suspect the answer to the last question is "no", but we can perhaps still infer their existence with confidence. Our Universe is in the position of having had a stroke of great luck with its unusually low value of Λ, rather like someone who finds they have won the national lottery. Now, every lottery winner knows they are not special: someone has to win. But if you won the lottery and were then told that no-one else had entered on that day, you would be skeptical. The only way for there to be an unusually lucky winner is for there to be a lot of losers out there somewhere. It's a strange and disturbing vision to imagine so many Universes, each appallingly hostile to life, in order to understand just the one we inhabit; but I think it's the best answer we have at present.

Thanks again John, and sorry if we take advantage of you again to start the following discussion. Everybody can see that today there are many influences acting on scientists and constraining their work: the sociological, economical, ethical and religious condition of our society may act in different ways on the work of scientists, exerting various kind of pressures on them. Many things have changed since Galileo, except maybe the fact that human society may sometimes behave very badly towards its scientists.

Let us start to see why.

4.12 Science and Society and Self-Organization of Astrophysical Community

4.12.1 Comments on Sociological and Economical Influences

Dear John *(Peacock)***, science today is affected by "sociological" and "economical" conditions that seem to have a relevant influence on its progress. What kind of possible solutions may be adopted to guarantee scientific freedom and the development of ideas and investigation methods alternative or complementary to most accepted ones?**

I am more aware of financial constraints in science than sociological ones. I know that some people complain of a herd instinct in cosmology, and that standard thought can be blinkered, but I think this is in large part misguided. Many cosmologists that I know are iconoclasts by nature, and would like nothing better than to disprove the standard model. Indeed, I know for a fact that this was a strong motivation for many of the leading CMB experimentalists, who were horrified that their measurements turned out to agree exactly with ΛCDM orthodoxy, rather than opening the door to new and unfamiliar territory. I am reminded of something I think Steven Weinberg said: that the problem with cosmologists is not that they are forever having wild ideas, but that they often do not take the simplest predictions of their theories seriously enough. This is not to justify a complacent certainty that the standard model must be right: but critics should accept that it has had many successes, so any critique must grow out of a deep familiarity with these successes.

The financial limits are also sociological in a way, however, since society has to decide the level at which it wants to support fundamental science. There is no easy answer to this, and it has changed with time. If we are honest, astronomers and other physicists should acknowledge that we are still living off the fat of the glory days of the Second World War, when governments realized that physicists were useful for the Defence of the Realm. I think this connection with the military machine is something that most of us would prefer to forget, but it is difficult to deny. Gradually, this dependence of the state on physicists has declined, and the question is how much further it has to fall.

Thank you John. The possibilities of a forthcoming revolution in physics are certainly attractive, but how close is this possibility and how much is this linked to the current sociological conditions? Let us listen the opinion of Paola Marziani.

Dear Paola *(Marziani)*, **do you think that a scientific revolution is possible in the current sociological condition? What kind of astrophysical observations could trigger it?**

I wish to point out two aspects that may sound pretty obvious since they come from a laywoman astronomer and not from a social scientist. First, a scientific revolution needs people who are so outstanding (in the literal meaning of the word) to be able to carry out new observations of revolutionary impact and interpret them. A social mechanism must allow those outstanding people to become scientifically productive. And it may be a blessing that neither the involved researchers nor others fully realize the importance of their discoveries. I am saying this because I cannot take out of my mind the biography of some of the most prominent scientists and thinkers of the twentieth century: Einstein, Freud, Turing. All of them faced threats that jeopardized their success in science. The power of deterrence works well for positive achievements: intimidate one Galileo and you may have an intellectual desert for a long, long time. A generation of German and Italian physicists and astronomers was disbanded by dictatorship and war. Are we sure that our present-day society is so free, and free of discrimination and conditioning, to allow outstanding individuals to become outstanding scientists with the knowledge, the will and the courage to go outside mainstream science?

A community of scientists must then be receptive of the new discoveries that may be scoffed at in one place but appreciated in another. There must be an "else where" where a paradigm shift can propagate. Equalization of science policies, work habits, and polarization of astronomy toward big science carried with a limited number of extra-powerful instruments may be deleterious to the development and affirmation of alternative theory [230]. After all, almost everyone with a past in mainstream science is posed to lose from revolutionary discoveries.

If one restricts your question to Cosmology, there is one discovery that will really revolutionize the field: non-cosmological redshifts that is, redshifts not due to the large scale expansion of the Universe. Non-cosmological redshifts would invalidate the use of redshift as a distance indicator, and undermine the foundation of much of the present astrophysical work way beyond cosmological issues proper: supermassive BHs in distant quasars, most of quasar physics, galactic evolution, the CMB, and many more technical issues would lose their meaning. It would be a truly Copernican revolution: the impact would be immensely larger than if, for example, only the CMB were found to be of noncosmological origin, as discussed by Dr. Robitaille in this book.

However, the probability of such changes actually occurring seems small at present. Champions of the idea of non-cosmological redshifts (present since quasars were discovered) were eventually ostracized seemingly on non-scientific ground. In recent years, there has been a revival of papers pointing out oddities and inconsistencies of the cosmological interpretation of redshift [3, 15, 16, 194, just to cite a few references]. There has even been a manifesto invoking attention and funding for alternative cosmologies (which I signed since it did not imply that standard big-bang cosmology is necessarily wrong). The issue is still somewhat open (especially for redshifts larger than 1), even if (I repeat here my opinion) there is no strong evidence foreshadowing a major revolution.

Perhaps another possibility is the contact with alien intelligent beings who may force humankind to rethink much of their physical understanding of the Universe, and even to confront with their past (and present) of genocide and wars.

Thank you very much Paola. Along the same vein, we can suspect that the scientific work of young scientists is strongly influenced by the environment in which they live. Cesare Chiosi has sent us his opinion on this point.

Dear Cesare *(Chiosi)*, **what do you think about the present way of doing astrophysical research? Do you believe that for a young astronomer is it possible today to pursue his research interests in an environment free from constraints limiting his scientific creativity? To what extent boundary conditions influence creativity in science?**

What to say without disappointing anybody? Often, I think that Globalization has infected Science. I may be wrong, but it seems that pressure of success, affirmation, long lists of published papers, leadership in specific areas have pushed away the simple pleasure of curiosity, the investment of time and efforts just to widen the horizon of personal knowledge. This tendency begins very soon in a young researcher's carrier. He/her is in fact expected to publish paper after paper, possibly in journals of high impact factor to increase his/her chances of a job (if any, at least in my own country!). Monochromatism of expertise is the obvious result of it. In addition to this, is the growing need of large teams to realize big, very expensive instrumentations, in which young fellows are likely lost in the crowd. No time to do anything else but the specific task assigned to him/her, with a great damage of personal scientific growing and initiative. Nothing against big instrumental projects, undoubtedly necessary to deepen our experimental information, but this is the price to pay. Same considerations would apply to large groups intending to attack specific problems from all view angles (from observations to interpretation). There are several subtle poisons in the air: monopolization of information and ideas, self-referencing, and over-production of papers. New information is often out of reach for long time.

Self-referencing, an obvious consequence of this sort of monopolization, makes the rest. To express opinions out of the main stream may be hard if not impossible. In this context, most of papers either confirm, sustain, share, agree with someone else results. Only a minority dare to say "things may be different". Other opinions are simply ignored. Over-production of papers is good to justify the economical cost and to rise more funds. However, the reverse side of the coin is that often single researchers are formally authors (in collaboration with many others) of an incredibly large number of refereed papers per year. Now, a year is made of 52 weeks, 7 days each, 24 h per day ... the rest to your imagination. It is like an assembly chain! Often, very expensive instrumentations, large legacies on single projects impose to rediscover what in reality was already long known, with much less detail of course but the leading ideas already in place. This implies a short decaying time scale of any result and progress. All this can be easily understood as due first to marketing laws, second to the simple lack of time and information. Too a pessimistic view? How to cure it? Sorry, I do not know!

Thanks a lot Cesare for having remarked the relevance of this problem and the difficulty to solve it. These observations raise the question of whether even in the scientific ambient new forms of ostracisms towards heretic views may still appear, producing a new "Galileo case". Let us continue in this discussion with the following interview.

Dear Thanu *(Padmanabhan)*, **remembering the Galileo experience, do exist today the political and sociological conditions for a scientific revolution?**

Science is practiced by scientists who are human and hence one cannot expect a perfect, objective, progress. The currently established scientific norms (peer reviewed publications, wide access to and dissemination of information, peer reviewed funding procedures ...) are probably the best we can hope for. Within this context, one often notices that a generation of reputed scientists devote their time, effort and energy in an initially promising direction of research (say, e.g., in a particular approach to quantum gravity) only to slowly realize that they were wrong all along. But it is unrealistic to hope that they will admit this debacle and move over to another paradigm. The emotional involvement (not to mention the need to sustain positions, groups etc.) will make these scientists as closed group to keep nearly dead ideas on life-supporting machines for a long time! This is inevitable and we need to accept it as a fact of life.

But the younger generation will be able to see through this and can easily adapt to newer paradigms, rejecting the once popular ideas which have outlived their utility. The older generation, in spite of their important early contributions and intellectual prowess, will eventually be sidelined unless they abandon the ideas which were initially attractive but have lead nowhere over a period of time. I am positive that this will happen in the case of the present day views regarding gravity as well but one cannot predict the timescale over which it will take place. Given the current practices followed in science, it will happen at a significantly shorter timescale compared to in the days of Galileo – but one cannot ignore the fact that sociological effects have a strong bearing on the progress of science even today.

Thank you very much Thanu.

We will now enter more specifically on the questions that are more closely linked to the way in which the astrophysical community is organized today. Jack Sulentic was very kind to address some tricky questions concerning the way in which researchers may work today. He expressed also his opinion about the behavior of scientists themselves at the beginning of the propagation of a new paradigm.

4.12.2 Comments on Astrophysical Community Self-Organization

Dear Jack *(Sulentic)***, science is made by men and women. It is therefore possible that their characters and beliefs influence their approach to science. This occurred at the Galileo's epoch as well as it does today. Do you believe that this could partially contribute to the success of the standard cosmological vision? What is your personal experience and opinion about this?**

That brings us to the sociology of science. Long experience in an esoteric field like astronomy, and especially cosmology, provides insights into the sociology that underlies human involvement in these activities. Ones initial naivete about how science is conducted and how ideas are exchanged can be quickly lost. It seemed to me – as a grad student – that investigating controversial areas would be an exciting thing to do. So why then the hostility and suspicion, if those are the correct words, that one actually experiences after working on a controversial problem? Chip Arp actually warned me about these reactions – he said people might be "put off" by an association with him – when he invited me to Pasadena to work with him in the area of non-Doppler redshifts and alternatives to the Big Bang paradigm. Another manifestation of this loss of naivete involves surprise at the lack of true intellectual discourse in the field. Maybe it has always been that way and I had some idealistic vision of how science was conducted.

After some time in the field, I began to realize that there were different kinds of people doing science and these people had quite different motivations and goals. Everything became clearer to me when I came across a beautiful tribute to Max Planck written by Albert Einstein (see e.g., Ferris [82]). It turns out that Einstein already realized there were different kinds of people in physics (he called it the temple of science). At the risk of oversimplification he was saying and we can say that there are careerists and truth seekers doing science. Careerists are motivated by the desire to advance their careers and truth seekers by more complex and unrealistic goals (e.g., "the love of science"). The latter tend to have their feet planted less firmly on the ground. In order to remove any good vs. evil connotations (so prevalent in puritan societies) let us henceforth refer to them as Baconian(s) (after Sir Francis Bacon often referred to as a father of the scientific method or as the ultimate empiricist). Of course Bacon was a contemporary of Galileo (e.g., [20]) whose anniversary we commemorate! In Pasadena I was lectured more than once that one could no longer be a Baconian in (1970s) astrophysical science – I remember my incredulity upon receiving this lecture from a senior person.

Einstein obviously belonged to the latter class and the point of his comments was that Max Planck was also a Baconian. Certainly not all Baconians were/are as clever as Einstein or Planck but they do share similar motivations. This is not to suggest that careerists cannot be able scientists. In fact they often show higher technical adeptness than many truth seekers. A major departure occurs because careerists build their career on their technical adeptness within the existing paradigm. Threats to the paradigm are often perceived as threats to their continued rise in the field – money, awards, memberships and access to large telescopes. While oversimplified the realization of the existence of such different kinds of scientists removed the scales from my eyes and from that point onward I could understand why people behaved the way that they did. I never felt to be advocating any particular replacement for the standard paradigm. Advocacy is not part of the game and nor is caring which answer is the correct one. I simply believed that testing and questioning all aspects of the paradigm was a fundamental part of being a scientist. I learned that this philosophy was not shared by careerists.

Thus the careerist response to most of the controversial questions in cosmology will be along the lines of "this is nonsense ... there is nothing to discuss ... it has all been discredited". In fact it is now well understood that, far from advocacy, even investigation of such areas is deemed inappropriate. It all becomes understandable. While we may be building ever larger and grander instruments, they are not intended as vehicles for testing the standard model – they are symbols of power and glory. They will be tied up in endless surveys lest access falls into the hand of "unreliable" investigators. Endless surveys where curiously nothing new is ever found. Well even if something slightly new is discovered, the last line of the abstract reporting it will always read "supports the standard picture". A jaundiced view for sure – but sadly for the scientific enterprise too often true. It is easy to get time on a big telescope to obtain low S/N observations of high redshift sources (the higher the better) but very difficult to get high S/N observations of a low redshift source. The former results are uncertain and therefore unlikely to be constraining (i.e., scientifically falsifiable) while the latter can be a threat. And applying for telescope time to test unpopular ideas is virtually impossible. Students know this but are, in any case, trained to believe rather than question the standard paradigm.

One can easily find case studies that illustrate the strength of careerist devotion to the standard model. The uncritical acceptance of any new evidence that supports it and rejection of evidence that does not. It is easy too see if you keep yourself out of social circles created by the careerists (its called "networking" in the business world) and almost impossible to see if you are inside. Careerists can be credited for "socializing" astronomy – Baconians are often loners. As an example of evidence embraced uncritically we might consider the "Baldwin effect" (discussed by Paola Marziani in Chap. 2) involving an apparent anti-correlation between a line strength measure (equivalent width of broad CIV1549) and source luminosity for quasars. The discovery paper [7] involved a very modest sample of quasars and a surprisingly strong correlation. The author expressed an appropriate skepticism about its reality. In subsequent years the Baldwin effect was lionized even as the strength of the correlation decreased with every subsequent quasar sample used to test it.

Questionable data manipulation was sometimes used to confirm it and soon it was being “found” everywhere. I would say that the Baldwin effect fervor became so strong that few insiders would have had the courage to question it – they would have found their popularity in sudden decline. We now know a little more about quasars (spectroscopically) and we know that they are not all the same although that impression persists. It now looks like the Baldwin effect is an intrinsic effect and that the extreme values observed in the correlation represent quasars at different evolutionary stages. It is not a tool that could enable us to use quasars as standard candles. Or to put it more simply – we find a Baldwin effect in a sample of quasars with a small spread of redshift and/or luminosity – implying that it is intrinsic and has nothing to do with cosmology [6, 10]. Richard Feynman said that “science is a culture of doubt”. A little more doubt was needed over more than 20 years after that the Baldwin effect was enshrined to the point of inspiring a meeting titled “Quasars as Standard Candles” [81].

Thanks Jack. We are not sure at all that the distinction between careerists and Baconians is so sharp as it appears from your discussion. Indeed, careerists could have a Baconian part too, and viceversa. Often, the balancing between these two attitudes depends on many circumstances and opportunities and, as you may recognize, the most important thing is the quality of the contributions of scientists, that can eventually derive from both of them.

To your opinion, is there still space for a thoroughly unbiased, empirical study in astrophysics? Or, are some theoretical and effective, yet still basically unproven paradigms so deeply entrenched in our scientific discourse, that we are forced into biased analysis?

I remember the first time – while observing at Palomar as a postdoc – that I was instructed on this subject. As mentioned earlier, I was instructed why the answer to the question is “no”. It certainly surprised me then but it does not anymore. Almost all careerists would say “no” and that there is no need for such studies anymore. I think the answer is “yes” and will always be “yes”. I reject the notions: (1) that we live in a special time and (2) that we know all or most of the laws of nature. A “no” answer would support both of these notions. I suppose the “no” response would be delivered most emphatically in the US. When the answer to this question becomes “yes” then science will again triumph over ideology. My response is necessarily philosophical and I am not holding my breath.

It seems that we are indeed forced into biased analyzes. One sees it all the time. It would be very difficult for someone entrenched in the paradigm to see them. One must understand that. A biased analysis would be called a correct analysis by many careerists. Not all but many. The earlier story about the Baldwin effects provides a case study in what can happen. How could anyone believe that a very small sample of sources might be representative of the quasar phenomenon? It would be close to miraculous for that to happen. But it still goes on today. Why did no one say “look – we examined this effect with a larger sample and the correlation gets weaker”? How could workers selectively omit sources from their sample – which did not show the effect – in order to get the effect. One must blame some of this lack of skepticism, and eagerness to embrace paradigms, on poor graduate school training. Grad training in astronomy is not supposed to be like seminary training – it should be the opposite.

Can, in your opinion, a scientific revolution of the weightiness of the Copernican revolution take place today? In the terminology coined by Kuhn, can a paradigm shift propagate in scientific sub-communities and spread to the scientific community at large?

The answer is "of course yes". But it will not be easy because the way science is conducted in the modern world has changed so much. The time of Copernicus and Galileo saw a small number of truth seekers making observations and trying to explain them. It also saw the emergence of the Copernican paradigm that has been amply verified. That was largely empiricism until Newton came into the world, at the death of Galileo, to develop the physical laws and mathematical tools needed to explain Copernican/Galilean empiricism.

One can argue that trends that shape future astronomical research along the lines followed in high energy physics are motivated by a wish to forestall future revolutions. If true it is likely a subconscious process. A recent polemic against the high energy approach in astronomical research [230] struck a deep chord with me. The high energy approach is something that only a careerist could love. I have seen many clever young people leave astronomical research in the past 30 years. The above polemic argues – I think correctly – that the growth/dominance of large team research – drives creative individuals away from science. This trend will grow with the advent of the giant telescopes. A single 100 m telescope is in no way better than ten or twenty 4–6 m telescopes. The latter would maximize use of brain power by allowing many individuals and groups to conduct experiments. How to decide who merits access to a 100 m telescope? Only a large team and only for confirmation of the standard model. I think other creative individuals have left the file (or entered astronomical bureaucracy) because frankly boring.

Modern science was strongly shaped in the US which dominated the twentieth century. This is a culture where (exceptionalism) fosters career and especially acquisition of money. It is also hyper-competitive. Money equals power and acquisition of more money is facilitated by power. Things were basically OK before second world war when science was largely done in a small number of elite schools. But Sputnik changed everything and thousands of people entered the field. More careerists than Baconians. They have created a large and powerful scientific edifice. Little room for free thinkers who are easily pushed aside. They cannot easily survive in this new environment. Some adapt and do good work in a trendy area that they adopt in order to survive. But free inquiry suffers and is now almost dead. If a new revolution comes it will likely arise in Europe (if they resist the impulse to mimic the US in areas other than pure excellence) or the rest of the world (India/ China? Latin America?).

Four hundred years after Galileo first used a telescope to observe the heaven we see large, giant and monster telescopes coming on-line or in the development stages. If the current state persists none of them will be used (wasted?) for exploratory research. They will be dedicated to supporting the paradigm which for many is now a religion. It is difficult to imagine any observation that could kill the Big Bang paradigm now. It is seductive – even I am drawn to it against my instincts! It is a complex edifice – a citadel of complexity – one must accept it or dig very deep to have any hope of understanding it well enough to challenge it. I know from experience that, if you try to dig into any one brick of the edifice you will begin to encounter problems and inconsistencies. But it is a thankless enterprize – only a Baconian would be crazy enough to try. If the situation improves we can look forward to seeing many new and fundamental discoveries in the next 100 years. The before mentioned surveys that will continue to monopolize big telescopes on Earth and in space do have a bright side. They data is being archived, made public and linked through a "virtual observatory". It is already possible for people to enter this archive and (carefully) use this data. There are no doubt many nuggets of gold that have been overlooked by the careerists!

In the present term I think the most disappointing trends involve: (1) the one-sidedness of so many published papers and (2) declining seminar attendance world-wide. Careers are not advanced by admitting, much less discussing, evidence against ones thesis. This goes far beyond the lack of dialog over controversial issues raised above. Similarly ones career is only fostered by attending seminars very closely related to ones work. All of this suppresses creative dissent and makes a scientific revolution, large or small, very unlikely.

Thanks a lot Jack.

The relationship between scientific communities is sometimes influenced by basic problems that need to be solved, such as the economical question of how research funds are shared among the communities themselves. Now Simon White will defend his recent paper on the comparison between particle physics and astrophysics communities. They have worked together in a fruitful way, but sometimes they have not shared the same view about best strategy to address relevant cosmological problems.

Dear Simon *(White)*, **recently you presented an interesting comparison between the astronomical and particle physics approaches to cosmology. You suggested that the concentration of theoretical and experimental efforts on the problem of DE, and more generally on themes strongly related to fundamental physics, may be a danger for the astronomical community. Would you like to summarize here your point of view?**

Particle physicists and astrophysicists have long interacted productively on topics of mutual interest. Important examples include the discovery of helium in the Sun at the end of the nineteenth century, the elucidation of nucleosynthesis in stars and in the Big Bang, the use of neutron stars to explore the equation of state of ultra-dense matter, the solar neutrino problem and its relation to the discovery of neutrino masses, astrophysical constraints on the nature and mass of axions, limits on the parameters of the Minimal Supersymmetric Model from searches for annihilation radiation from clumps of DM, and the use of statistical and data-processing techniques from particle physics to analyze the very large datasets produced by searches for gravitational microlensing. This cross-fertilization has been very fruitful and has stimulated new research directions in both fields. It will undoubtedly continue.

The DE problem fascinates both cosmologists and high energy physicists, but it does not fit the pattern of previous cross-disciplinary collaborations. In my view, a failure to appreciate and take proper account of the differences will seriously weaken astrophysics, reducing its extraordinary current vitality, breadth and versatility, as well as its appeal both to the general public and to the next generation of bright and ambitious young scientists.

It is over 90 years since Einstein introduced the cosmological constant in order to construct a static Universe. Over this period, it has been revived repeatedly as a possible explanation for various cosmological puzzles. The recent demonstration of accelerated expansion was, nevertheless, a great surprise. Suggested explanations include a modification of Einsteinian gravity, an unexpected effect of nonlinearities within it, and an effective scalar field which may reflect the unification of gravity and quantum mechanics in a higher dimensional string theory. These all require new physics which is manifest only in the global evolution of the Universe.

Current observations are consistent with a cosmological constant to within about 10% in the equation of state parameter w, that is, they suggest $w = -1 \pm 0.05$ (1σ). There are plenty of ad hoc theories for DE in which w differs measurably from -1, but none has strong independent theoretical underpinning. Many high-energy theorists argue that additional fine-tuning, independent of that already needed to explain the unnaturally small value of the current density of DE, is required for $w + 1$ to be measurably different from zero. Attempting to constrain DE by precise measurements of the cosmic expansion and linear growth histories is thus analogous to searching for lost keys under a street light. Just as the drunk searches in the only place where he could find his keys, rather than in the place where he believes he lost them, so planned DE experiments probe possibilities we can constrain, rather than possibilities which are generally agreed to be plausible.

The precision of current measurements of cosmological parameters is such that uncertainties in the evolution of the cosmic scale factor and the amplitude of linear fluctuations no longer limit our understanding of how galaxies and galaxy clusters form. The complexities of strongly nonlinear processes such as star formation, BH formation and feedback are much more important. Achieving the primary measurement goals of planned DE surveys is thus unlikely to shed light on structure formation issues.

The danger here is clear. Optimizing a survey to obtain, for example, precise measurements of the baryon oscillation signal may require lowering the signal-to-noise of individual spectra in order to maximize the number of redshifts obtained. A successful result would be a more precise measure of the expansion history, but there is a substantial a priori probability that this would simply be a narrowing of the error bound around $w = -1$. This will not help significantly in understanding the nature of the DE, nor will it help us to understand galaxy formation since the quality of the spectra will be too low to provide much useful information about individual objects. In my view, this would be a meagre return for the effort and money invested.

My conclusion is that we must design observatories to explore DE as one of many issues, rather than experiments tuned specifically for optimal constraints on w and its derivatives. We must ensure that the data returned by our instruments are of the quality and type needed to address a broad range of astronomical issues, not just the expansion and linear growth histories of the Universe.

On the contrary, it is also possible that exactly the need of an extremely precise knowledge of astrophysical phenomena, required to disentangle fundamental physics signatures, further stimulate astrophysical studies, and viceversa. For example, this happened to some extent in the context of the *Planck* mission project, the logical step forward after WMAP. Do you believe that such a strong interaction between these two different approaches to the Universe science might have also positive consequences for both?

The *Planck* mission has not yet flown (at the time of writing). My group has been heavily involved in trying to provide computing infrastructure for joint analysis of the data from the two instruments. Our experience over the last decade has not been positive and has reinforced my view that large and diverse collaborations of this kind are an inefficient way to do science. We all hope, of course, that the mission will be successful and will produce great results, but its organization is in dramatic contrast to that of WMAP which involved a small and tightly knit team. While the different national and scientific communities involved in *Planck* (space experimentalists and instrument-builders, cosmologists from astronomical and high-energy backgrounds, microwave astronomers ...) certainly bring differing expertise to the table, they also bring different working habits and expectations. Time will tell whether all this can be brought together to a productive conclusion.

In the DE context, I think the dramatic success story of microwave background experiments over the last decade is a significant factor leading many cosmologists to the unrealistic expectation that surveys to measure the cosmic expansion and structure growth histories will be limited by statistics rather than by systematic uncertainties. The CMB is unique in astrophysics, in that all the information is contained in linear perturbations of an extremely simple system. In addition, confusion from contaminating foregrounds has turned out, fortuitously, to be negligible, at least for the temperature fluctuations. We are unlikely to be this lucky when we use SNe, galaxies or galaxy clusters as cosmic tracers. High-energy physicists are, of course, trained to search large amounts of accelerator data for small signals hidden among a haystack of confusing effects. This expertise will undoubtedly be very helpful when analyzing DE survey data.

To your opinion, comes the major problem you underlined from the necessity of an appropriate balance of available funding resources or is it intrinsic to the two different methodological approaches?

The issue which most concerns me is intrinsic to the two different approaches. For some years the experimental focus of high-energy physics has narrowed to concentrate on fewer and fewer issues. For example, even the internal structure of the proton is now considered by many as “nuclear physics”. The community has organized itself into ever larger teams to build ever bigger instruments for an ever smaller number of accelerators. Programmes are organized around a small number of “Big Questions”. It is unclear whether the world will be able to afford a successor to the LHC, and each team working on an LHC instrument already involves more than 1,000 physicists. Organization at this level requires an industrial approach, and it is unclear how the field can continue in the future. Concern about such trends is in part responsible for the increasing number of physicists switching to "astroparticle" topics like DM searches.

In contrast, astrophysics has always been opportunistic, making progress simultaneously on many fronts. Its division of labor differs from that of high energy physics, separating those who build observatories and instruments from those who use them, thereby allowing much smaller teams to address problems of forefront interest. Even under today's imperative of single page executive summaries, astronomy has a very diverse set of primary objectives: DM and DE; the origin and evolution of galaxies, stars and planetary systems; the nature of BHs; physics under extreme conditions; the very early Universe; gravitational wave and neutrino astronomy; the origin of life ... This diversity has allowed ambitious young scientists to find individual niches and to establish themselves as independent, internationally recognized researchers while still graduate students or postdocs. It is also partially responsible for the popularity of astronomy with the general public – compare the number of popular or amateur astronomy journals with the number of similar journals in high-energy physics, or the number of newspaper articles addressing the two areas.

The rise of "survey astronomy" has resulted in a dramatic expansion in the number of large teams active in our field. Although typically more loosely structured than in high-energy physics, these teams share the general ethos of accelerator instrument teams. The large scale of DE surveys and the active participation of high-energy physicists in them has emphasized and accelerated this trend. Its effects are very clear, for example, in current citation statistics, where many highly cited scientists acquired their status through citations to papers with long author lists where their individual role is invisible. Such teams emphasize hierarchy. Senior scientists determine the careers of their juniors by writing references which provide the sole means for outsiders to judge quality, and in addition they take credit for "team science" where their role is often managerial rather than creative. This is far less egalitarian and transparent than the traditional astronomy system where authorship lists show who is primarily responsible for the content of a research article.

In my view, such "large team science" is less attractive for the best young scientists than traditional astronomical research, where small teams propose science programmes for forefront instrumentation at national or international observatories. Some projects, DE surveys perhaps, do require the large team approach, but astronomy will be impoverished if such projects come to dominate our field. Astronomical advances have typically come from inspired and creative individuals, rather than from planned programmes by large teams. I believe that our understanding of DE, like that of the perihelion advance of Mercury 100 years ago, is more likely to be advanced by a new and revolutionary insight than by an industrial-strength campaign of "precision" measurements.

Thank you Simon.

A key question for the future of science is certainly how new generations of scientists are educated and how much space will be left for their creativity. Gerard t'Hooft will now tell us his opinion on this point.

Dear Gerard *(t'Hooft)*, **during the last 20 years the way of working in physics and astrophysics is changed a lot. New generations of astronomers and physicists are growing within a world in which competition and working possibilities are the very big challenge of this new century. In several fields only large and well-organized research teams could have a real impact on science evolution. This often requires a great specialization of individual research activity. Do you believe that this could have a negative impact for the future of our science?**

I do not see the problem of large and well-organized research teams. Many of today's questions in science can only be answered by well-orchestrated joint efforts that require the utmost of our abilities and resources. I am delighted that we get support from the public and its politicians to realize our big adventures such as putting telescopes and other detectors in outer space and building gigantic particle accelerators that cost billions. It is our duty as scientists to show our gratitude for this support by informing the public as well as we can about our findings. Only if no new findings of any interest can be reported about we have a problem; we will then lose our support.

Fortunately, not all of science has developed this way. We still have small scale methods, cheap and modest experiments, calculations and theoretical methods of analysis that cost not much more than the salaries of a couple of devoted scientists. What makes science proceed further, is a combination of all procedures that we can manage to perform. Indeed, this implies that many of us will be extremely specialized. Some of us will continue to be generalists however, and all of these people together ensure that science in general, and ours in particular, will flourish.

Thank you very much Gerard. Finally, for this chapter, we want to present a positive example of a strong progress of astrophysics, supported by a parallel development of a society. Spain is this example. Since the start of democracy we have seen in this country an explosion of possibilities for science and scientists. Here, Rafael Rebolo will show us how this was possible.

4.13 Boundary Condition for Astrophysics Development: A Modern Example

Dear Rafael *(Rebolo)*, **during the recent decades the Canaries began a very important site for ground based instrumentation at various spectral domains and with remarkable international scientific collaborations. Do you like to discuss the evolution and grow up of such facilities and their impact for the Spanish astrophysical community? How much the general opening and social, cultural and economical growing of the Spain society is responsible for the**

good conditions for such scientific development? Can you see some parallelism between the modern Spain society and the Venice ambient at the epoch of Galileo?

Astronomy is a well developed branch of science in Spain that contributes to the world production with more than 5% of the total number of research papers. This was not the case 30 years ago when only a few spanish scientists were dedicated to this field. Today, more than 400 PhDs conduct astronomical research in Spain, mainly in several large research institutes and astronomy departments, but also a significant number are part of small groups in physics departments in universities distributed across the whole country. Spain is leading the construction of one of the largest optical/infrared telescopes, the 10.4 m Gran Telescopio Canarias, is a full member of the ESO and actively participates in the astronomy missions of the ESA. In these last 30 years major changes took place in Spain, most remarkably the insaturation and consolidation of democracy which has led the nation to a rather unique period of development in many areas including education, culture and science. Astronomy has benefitted from this major achievement of the Spanish society.

Research is an essential element of human development that needs economic, human and material resources. In Spain the fraction of the Gross Domestic Product dedicated to Research and Development is still significantly below that of USA, Japan, Germany, or France. While in the last decades there has been an objective growth of the resources dedicated to Research and Development, that has favored a very positive evolution of science in the country, the dramatic expansion in Spanish Astrophysics was largely due to special circumstances which favored a growth above the average level of many other fields of science. The installation of several large astronomical observatories in Spain has played a major role in this development. For optical and infrared observations the observatories at the Canary Islands (Tenerife and La Palma) and Calar Alto (Almera), and at millimeter wavelength the observatory at Sierra Nevada have offered exceptional research opportunities to European astronomers including the Spanish Astronomy groups that were emerging at the end of the 1970s. These groups were the seeds of research institutes and University departments later consolidated thanks to the leading role of key astronomers and the vision of academic and political authorities. In that respect, we could find some parallelism with the situation faced by Galileo at a time where Venice was one of the most powerful territories in the Mediterranean. The will to support science and astronomy led to an exceptional set of discoveries at the beginning of the seventeenth century which left imprinted Galileo's name in the history of science. We possibly face now an epoch of similarly exciting discoveries that range from detection of other "earths" and discovery of life elsewhere, to find direct evidence on the nature of the dominant forms of matter and energy in the Universe. Fortunately, the evolution of astronomy in Spain led to a situation where Spanish astronomers can contribute significantly to these and many other important research topics in the near future.

4.13.1 *Astronomy in the Canaries*

There are few places in the world meeting the demanding conditions required for frontier astronomical observations. We have been fortunate in Spain to have several very good astronomical sites, and in particular, two exceptional ones at the higher reaches of the islands of Tenerife and La Palma. The outstanding conditions at Observatorio del Roque de los Muchachos (La Palma, 2,400 m.) indeed makes this an extraordinary site for the exploration of the Universe which competes among the best three in the world.

The Istituto de Astrofisica de Canaria (IAC) was conceived as to make the best use of the scientific and technical possibilities offered by the sky of the islands. Nowadays, near 400 people directly work for astronomy in the Canaries, about 300 develop activities under the umbrella of the IAC and the other 100 work for collaborating institutions like the Isaac Newton Group of Telescopes or the Foundation Telescopio Nazionale Galileo. The remarkable conditions of the observatories for solar and night sky observations have attracted many advanced telescopes to these islands and prompted the construction of our own 10.4 m telescope. Spain is aware of the very exceptional sky conditions and decided to protect them by a national law that preserves the quality of the sky for astronomical observation limiting light pollution among other things.

Modern astrophysics started in the Canaries in the early 60s when Francisco Sánchez started observations at Teide Observatory. Astronomy in the Canaries begun much earlier, possibly in 1856 with the astronomical campaign organized by the scottish astronomer Charles Piazzi Smyth. Following the suggestions by Isaac Newton that telescopes should be installed in high summits where the atmosphere is clear and stable. Smyth decided to organize an observing campaign in Tenerife and installed a telescope first at 2,717 m in Guajara mountain (Tenerife) and later at a higher altitude in Altavista (3,250 m). The campaign took place over 65 days performing observations of the Moon, planets, binary stars, zodiacal light and even ultraviolet solar radiation. He published in 1858 the book "Tenerife, an astronomer's experiment" where details of his work can be found. In particular, it is remarkable the use of stereoscopic photography.

In 1910, Jean Mascart conducted astronomical observations in Tenerife as part of an international team aimed to obtain images of Halley comet among other things. He installed the observatory in Guajara, as Smyth did, and his notes reveal how impressed he was by the transparency of the sky. The installation of the first professional telescope (according to modern standards) at Teide Observatory took place in 1964. It was owned by the Université de Bordeux in France where the astrophysics group was interested in zodiacal light studies. The first solar telescope will be installed in 1969, and this will lead to the formation of the first solar physics group in Spain, currently one of the largest in Europe. Since those years the number of telescopes at Teide Observatory only grows, the 1.5 m infrared flux collector from UK is installed and becomes one of the largest infrared dedicated telescopes in the world. This would be a precursor of the 3.8 m UK Infrared Telescope (UKIRT) in

Fig. 4.9 The 3.5 m Telescopio Nazionale Galileo at Observatorio del Roque de los Muchachos

Hawaii. Today this telescope now owned by IAC still performs routine astronomical observations mainly imaging in the near infrared with the instrument CAIN (CAmera INfrarroja), and in the optical with the lucky-imaging system (FastCam). The largest Solar telescopes were also installed at Teide Observatory, remarkably the german Gregory Coudé and Vacuum Tower Telescopes and the italo–french telescope THEMIS (Télescope Héliographique pour l'Étude du Magnétisme et des Instabilités Solaires). In addition, this observatory hosts laboratories for Helioseismology and for CMB studies with several pioneer experiments in these fields.

In the 1970s, the advantages of building a new observatory in La Palma, at the summit known Roque de los Muchachos became obvious to many groups in Europe. The seeing campaigns seemed to indicate a potentially exceptional site for both solar and night sky observations as it would be later confirmed by more than 30 years of observations. Sweden, UK, Netherlands and later a consortium of Nordic countries (Finland, Norway, Denmark, and Sweden) and also Italy and Belgium decided to install the most advanced telescopes in this observatory. In 1985, the 2.5 m Isaac Newton Telescope and 1 m Jacobus Kapteyn telescopes were inaugurated and by the end of the 1980s the 4.2 m William Herschel Telescope started operation. In the 90s this initial set of telescopes (which included a solar tower from the Royal Academy of Sciences in Sweden and the Carlsberg meridian telescope) is enlarged with the 2.5 m Nordic Telescope, the 3.6 m Telescopio Nazionale Galileo and the Dutch solar telescope. In the last decade, the Mercator and the Liverpool telescopes completed the suit of optical facilities in operation. After the precursor experiment on High Energy Gamma Rays (HEGRA), two new large telescopes (16 m each) have

Fig. 4.10 The 10.4 m Gran Telescopio Canarias at Observatorio del Roque de los Muchachos

been built by a large European consortium (MAGIC: Major Atmospheric Gamma-ray Imaging Cherenkov Observatory) aimed to measure the Cherenkov radiation produced by extremely energetic gamma rays. The first MAGIC telescope started operation several years ago and the second will start very soon.

Finally, Spain decided to build a 10.4 m segmented telescope in 1998 for operation at Roque de los Muchachos Observatory. This is now a project led by Spain in collaboration with México (Universidad Nacional Autónoma de México and Instituto Nacional de Astronomía, Optica y Electrónica) and the University of Florida. The telescope consists of 36 segments (each approximately 1.9 m across). The active optics system of the primary mirror will correct distortions caused by changes in temperature and mechanical stresses. Several instruments have been built for first light, including an optical imager and multiobject spectrograph (OSIRIS: Optical System for Imaging and low/intermediate-Resolution) and a mid-infrared camera (CanariCam) for operation in the range 8–20 μm. Currently, a multiobject near-infrared spectrograph (EMIR: Espectrógrafo Multiobjeto InfrRrojo) and an adaptive optics system for correction of the distortions introduced by atmospheric turbulence are under development. The telescope is expected to begin scientific operation by the end of 2008.

The agreements of cooperation on Astrophysics signed by Spain with more than 19 countries and 50 scientific institutions from all over the world have ensured an efficient installation, operation and scientific exploitation of a large number of forefront facilities at the observatories of the IAC. This set of telescopes conform in fact the European Northern Observatory (ENO) where each scientific institution retains ownership of the facilities it has contributed and IAC is responsible for management of common services. During the last decades, ENO has played a major role in European astronomy and has been crucial for the development of astrophysics in Spain.

4.13.2 Cosmology in the Canaries

Dear Rafael *(Rebolo)***, what are the recent most crucial experiments for cosmology in Canaries?**

The optical and infrared telescopes at the Canary Observatories have produced a large number of important results on Cosmology. Among them, it is remarkable the contribution to the study of distant SNe which led to the discovery of accelerated expansion of the Universe, the spectroscopic observations of very high redshift quasars and the abundance measurements of light elements (He, Li) that set constrains on BBN. Observations of clusters of galaxies and of the spatial distribution of galaxies have also provided constraints on the growth of large scale structures in the Universe.

In addition, a number of CMB experiments have been conducted at Teide Observatory since 1984 which produced as most remarkable results the first independent confirmation of the level of anisotropy detected by COBE (Tenerife CMB experiment), the detection of acoustic peaks in the APS (VSA: Very Small Array) and the detection of anomalous microwave emission (COSMOSOMAS: COSMOlogical Structures On Medium Angular Scales). Aimed to measure the level of anisotropy in the CMB at large angular scales ($5° \div 20°$), the Tenerife CMB experiment led by Rod Davies and his group at University of Manchester in collaboration with the IAC and Cambridge University, started a long term programme to measure the microwave sky in the frequency range 10–30 GHz [58]. The idea was to measure the synchrotron and free-free emission at high Galactic latitude regions using the receivers at 10 and 15 GHz and the CMB signal at 30 GHz. We demonstrated first that at high Galactic latitude the foreground emission was such that at 30 GHz would produce an rms below 10 μK at the angular scales of interest (larger than the horizon at the recombination). By the end of the 80s, we installed a new receiver at 30 GHz aimed to measure the CMB anisotropy with a precision better than this rms. It turned out that COBE found in 1992, the level of anisotropy was at 30 μK rms and we confirmed it by the end of 1993 with measurements at 10, 15, and 33 GHz [101] that showed 5σ detections of primordial spots in the CMB.

After the Tenerife Experiment, the goal was to measure the first peak of the APS and we set up two initiatives in the mid-1990s, the IAC-Bartol millimetric experiment (led by Piccirillo) and the two-element interferometer to search for anisotropies in scales of 1°–2°. This latter interferometer [102] was the precursor of the VSA, an interferometer at 33 GHz conformed by 14 antennas [225] which was built by the Universities of Cambridge, Manchester, and the IAC and started scientific operation in 2001 at Teide Observatory. The VSA has adopted several configurations, the so-called compact, extended and superextended (the current one) which produced detection of the first, second, and third acoustic peaks in the APS of

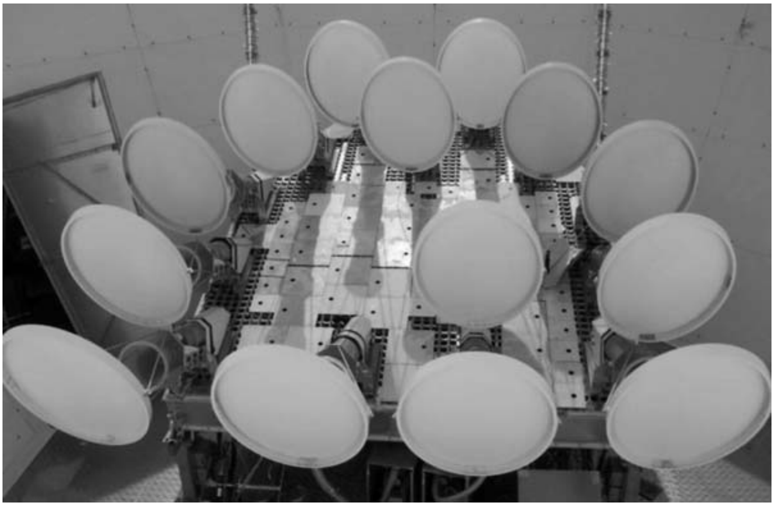

Fig. 4.11 The superextended configuration of the microwave interferometer very small array at Teide observatory

CMB fluctuations [62, 97, 186, 203]. VSA is now observing the CMB at an angular resolution of 6 arcmin in order to determine the amplitude of fluctuations at multipoles higher than $\ell = 1{,}500$. The most recent result obtained with this experiment is the detection of a cold spot towards the Corona Borealis supercluster of galaxies which could be caused by the so-called warm Inter Galactic Medium (IGM) [92] (see also Genova-Santos et al., in preparation). This would be the first detection of a Sunyaev-Zel'dovich effect towards a supercluster of galaxies.

The other experiment currently in operation is COSMOSOMAS, built by the IAC to explore the microwave sky ($10 \div 17$ GHz) with an angular resolution of 1°, this experiment has achieved the first unambiguous detection of anomalous microwave emission in the molecular complex of Perseus [225] and evidence for this new foreground, possibly associated to electric dipole radiation of carbon-based molecules, at high Galactic latitude [104].

At present, we are developing a new experiment (QUIJOTE-CMB), which will start operation in October 2008 aimed to measure the polarization of the CMB at angular scales larger than 1°. A consortium led by IAC, with the Instituto de Fisica de Cantabria, the engineering company IDOM, and the Universities of Cantabria, Manchester and Cambridge as partners is aiming to build several microwave telescopes to obtain polarization maps over a region of 10^4 square degree at five frequencies in the range 10–30 GHz. The goal is to reach a precision better than 1 μK per degree at 30 GHz and a few μK for the rest. These maps will provide an unprecedented view on the polarization of the synchrotron emission and of anomalous microwave emission. The latter is expected to be very weakly polarized at these angular scales (see [11]) but synchrotron is known to be significantly polarized. We intend to measure the synchrotron polarization at the lowest frequencies and correct its contribution at high Galactic latitude in the 30 GHz map to a level that should show the imprint of B-modes in this map if their amplitude to scalar modes is $r = 0.1$. The presence of B-modes in the polarization of the CMB is a robust prediction of inflationary models and result from the generation of primordial gravitational waves. The amplitude of the B-modes provides a direct measurement of the energy scale of Inflation.

The QUIJOTE (Q, U, I JOint TEnerife CMB) data will complement the higher frequency maps that the *Planck* satellite will obtain in the next years. The combination of data from both experiments should lead to stringent constraints on the amplitude of the B-modes.

Thank you Rafael for your passionate review of the salient steps of the progress of astrophysics and cosmology in Spain and in the Canaries, so introducing us to the discussion on future projects which constitutes the object of the next chapter.

References

1. Akbar, M., Cai, R.G.: Thermodynamic behavior of field equations for f(R) Gravity. Phys. Rev. Lett. **B 648**, 243 (2007)
2. Anderson, J.D., et al.: Study of the anomalous acceleration of Pioneer 10 and 11. Phys. Rev. **D 65**, 082004 (2002)
3. Arp, H.: Catalogue of discordant redshift associations. Apeiron, Montreal, Canada, p. 234, ISBN 0968368999 (2003)
4. Ashenfelter, T.P., Mathews, G.J., Olive, K.A.: The chemical evolution of Mg isotopes vs. the time variation of the fine structure constant. Phys. Rev. Lett. **92**, 041102 (2004)
5. Ashenfelter, T.P., Mathews, G.J., Olive, K.A.: The fine-structure constant as a probe of chemical evolution and AGB nucleosynthesis in damped Lyman-alpha systems. Astrophys. J. **615**, 82 (2004)
6. Bachev, R., et al.: Average ultraviolet quasar spectra in the context of eigenvector 1: A Baldwin effect governed by the eddington ratio? Astrophys. J. **617**, 171–183 (2004)
7. Baldwin, J.A.: Luminosity indicators in the spectra of quasi-stellar objects. Astrophys. J. **214**, 679 (1977)
8. Barrow, J.D.: The constants of nature: From α to Ω. In: Johnatan Cape (ed.) London (2002); Transl: I numeri dell'universo, Oscar Mondadori (2003)
9. Barrow, J.D., Tipler, F.J.: The anthropic cosmological principle, Clarendon Press, Oxford (1986)
10. Baskin, A., Laor, A.: On the origin of the CIV Baldwin effect in active galactic nuclei. Mon. Not. R. Astron. Soc. **350**, L31–L35 (2004)
11. Battistelli, E.S., et al.: Polarization observations of the anomalous microwave emission in the perseus molecular complex with the COSMOSOMAS experiment. Astrophys. J. Lett. **645**, L141–L144 (2006)
12. Bekenstein, J.D.: Relativistic gravitation theory for the modified Newtonian dynamics paradigm. Phys. Rev. **D 70**, 083509 (2004)
13. Bekenstein, J.D.: The modified Newtonian dynamics – MOND and its implications for new physics. Contemp. Phys. **47**, 387 (2006)
14. Bekenstein, J.D., Milgrom, M.: Does the missing mass problem signal the breakdown of Newtonian gravity? Astrophys. J. **286**, 7 (1984)
15. Bell, M.B.: Distances of quasars and quasar-like galaxies: Further evidence that quasi-stellar objects may be ejected from active galaxies. Astrophys. J. **616**, 738 (2004)

16. Bell, M.B.: Further evidence that the redshifts of AGN galaxies may contain intrinsic components. Astrophys. J. Lett. **667**, L129 (2007)
17. Bender, C.M.: Making sense of non-Hermitian hamiltonians. Report Prog. Phys. **70**, 947–1018 (2007)
18. Bender, C.M., Mannheim, P.D.: No-ghost theorem for the fourth-order derivative Pais-Uhlenbeck oscillator model. Phys. Rev. Lett. **100**, 110402 (2008)
19. Bergstrom, L., Iguri, S., Rubinstein, H.: Evolution of cosmic magnetic fields: From the very early Universe, to recombination, to the present. Phys. Rev. **D 60**, 045005 (1999)
20. Berry, A.: A short history of astronomy. Murray, London (1898)
21. Bertolami, O., et al.: Extra force in f(R) modified theories of gravity. Phys. Rev. **D 75**, 104016 (2007)
22. Birrell, N.D., Davies, P.C.W.: Quantum fields in curved space. Cambridge University Press, Cambridge (1982)
23. Bize, S., et al.: Testing the stability of fundamental constants with the $^{199}Hg^{+}$ single-ion optical clock. Phys. Rev. Lett. **90**, 150802 (2003)
24. Bondi, H.: Cosmology. Cambridge University Press, Cambridge (1961); Translation, Lampugnani Nigri (ed.), Milano, p. 150 (1970)
25. Bournaud, F., et al.: Missing mass in collisional debris from galaxies. Science **316**, 1093 (2007)
26. Buchbinder, I.L., Odintsov, S.D., Shapiro, I.L.: Effective action in quantum gravity. IOP Publishing, Bristol (1992)
27. Buchert, T.: Dark energy from structure: A status report. arXiv:0707.2153
28. Burkert, A.: The structure of dark matter halos in dwarf galaxies. Astroph. J. **447**, L25 (1995)
29. Caldwell, R.R., et al.: Early quintessence in light of WMAP. Astrophys. J. **591**, L75–L78 (2003)
30. Camerota, M.: Galileo Galilei e la cultura scientifica nell'età della Controriforma. Roma, Salerno editrice (2004)
31. Campbell, B.A., Olive, K.A.: Nucleosynthesis and the time dependence of fundamental couplings. Phys. Lett. **B 345**, 429 (1995)
32. Capozziello, S.: Curvature quintessence. Int. J. Mod. Phys. **D 11**, 483 (2002)
33. Capozziello, S., Lambiase, G.: Neutrino oscillations in Brans-Dicke theory of gravity. Mod. Phys. Lett. **A 14**, 2193 (1999)
34. Capozziello, S., Troisi, A.: Parametrized post-Newtonian limit of fourth order gravity inspired by scalar-tensor gravity. Phys. Rev. **D 72**, 044022 (2005)
35. Capozziello, S., Francaviglia, M.: Extended theories of gravity and their cosmological and astrophysical applications. Gen. Rel. Grav. Special issue on dark energy **40**, 357 (2008) [astro-ph/0706.1146]
36. Capozziello, S., et al.: Curvature quintessence matched with observational data. Int. J. Mod. Phys. **D 12**, 1969 (2003)
37. Capozziello, S., Cardone, V.F., Troisi, A.: Reconciling dark energy models with f(R) theories. Phys. Rev. **D 71**, 043503 (2005)
38. Capozziello, S., et al.: Cosmological viability of f(R)-gravity as an ideal fluid and its compatibility with a matter dominated phase. Phys. Lett. **B 639**, 135 (2006)
39. Capozziello, S., Cardone, V.F., Troisi, A.: Low surface brightness galaxy rotation curves in the low energy limit of Rn gravity: No need for dark matter? Mon. Not. R. Astron. Soc. **375**, 1423 (2007)
40. Capozziello, S., Stabile, A., Troisi, A.: Newtonian limit of f(R) gravity. Phys. Rev. **D 76**, 104019 (2007) [arXiv:0708.0723]
41. Capozziello, S., Stabile, A., Troisi, A.: Spherically symmetric solutions in f(R) gravity via the noether symmetry approach. Class. Quant. Grav. **24**, 2153 (2007)
42. Capozziello, S., De Laurentis, M., Francaviglia, M.: Higher-order gravity and the cosmological background of gravitational waves. Astrop. Phys. **29**, 125 (2008)
43. Carroll, S.M., et al.: Is cosmic speed-up due to new gravitational physics? Phys. Rev. **D 70**, 043528 (2004)

44. Carter, B.: Large number coincidences and the anthropic principle in cosmology. In: Confrontation of cosmological theories with observation data. Longair, M., Reidel, S., (eds.), Dordrecht, pp. 291–294 (1974)
45. Chand, H., et al.: Probing the cosmological variation of the fine-structure constant: Results based on VLT-UVES sample. Astron. Astrophys. **417**, 853 (2004)
46. Clavius, C.: Romani Calendari a Gregorio XIII P. M. Restituti Explicatio S. D. N. Clementis VIII P. M. iussu edita, Romae, Aloysium Zanettum, pp. 39, 96–97 (1603)
47. Clowe, D., et al.: A direct empirical proof of the existence of dark matter. Astrophys. J. Lett. **648**, L109 (2006)
48. Coc, A., et al.: Coupled variations of fundamental couplings and primordial nucleosynthesis. Phys. Rev. **D 76**, 023511 (2007)
49. Cole, S., et al.: The 2dF galaxy redshift survey: Power-spectrum analysis of the final data set and cosmological implications. Mon. Not. R. Astron. Soc. **362**, 505 (2005)
50. Coles, P., Lucchin, F.: Cosmology, the origin and evolution of cosmic structure. Wiley, New York (1995)
51. Combes, F., et al.: Galaxies and cosmology. Springer, Berlin, Heidelberg (1995)
52. Copeland, E.J., Sami, M., Tsujikawa, S.: Dynamics of dark energy. Int. J. Mod. Phys. **D 15**, 1753 (2006)
53. Cozzi, G.: Paolo Sarpi tra Venezia e l'Europa. Torino, Einaudi (1979)
54. Cyburt, R.H., et al.: New BBN limits on physics beyond the standard model from He-4. Astropart. Phys. **23**, 313 (2005)
55. Dallaporta, N.: Scienza e Metafisica, uno pseudo contrasto fra due domini complementari, Cedam (1997)
56. Dallaporta, N., Secco, L.: Il Principio Antropico in fisica ed in cosmologia, Giornale di Fisica XXXIV, N.3, 163 (1993)
57. Daly, R.A., Djorgovsky, S.G.: Direct determination of the kinematics of the universe and properties of the dark energy as functions of redshift. Astrophys. J. **612**, 652 (2004)
58. Davies, R.D., et al.: Sensitive measurement of fluctuations in the cosmic microwave background. Nature **326**, 462–465 (1987)
59. de Bernardis, P., et al.: A flat universe from high-resolution maps of the cosmic microwave background radiation. Nature **404**, 955–959 (2000)
60. de Blok, W.J.G., Bosma, A.: High-resolution rotation curves of low surface brightness galaxies. Astron. Astrophys. **385**, 816 (2002)
61. Dent, T., Stern, S., Wetterich, C.: Primordial nucleosynthesis as a probe of fundamental physics parameters. Phys. Rev. **D 76**, 063513 (2007)
62. Dickinson, C., et al.: High-sensitivity measurements of the cosmic microwave background power spectrum with the extended very small array. Mon. Not. R. Astron. Soc. **353**, 732–746 (2004)
63. Dirac, P.A.M.: The cosmological constants. Nature **139**, 323 (1937)
64. Dmitriev, V.F., Flambaum, V.V., Webb, J.K.: Cosmological variation of the deuteron binding energy, strong interaction, and quark masses from big bang nucleosynthesis. Phys. Rev. D **69**, 063506 (2004)
65. Dobrzycki, J.: Astronomical aspects of the calendar Reform. In: Gregorian reform of the calendar. Proceedings of the Vatican Conference to commemorate its 400th anniversary, pp. 1582–1982
66. Dodelson, S., Liguori, M.: Can cosmic structure form without dark matter? Phys. Rev. Lett. **97**, 231301 (2006)
67. Dolgov, A.D.: An attempt to get rid of the cosmological constant. In: The very early universe. Gibbons, G.W. et al. (eds.) Cambridge University Press, p. 449 (1983)
68. Doran, M., Schwindt, J.M., Wetterich, C.: Structure formation and the time dependence of quintessence. Phys. Rev. **D 64**, 123520 (2001)
69. Drake, S.: Galileo. Una biografia scientifica. Bologna, Il Mulino (1988)
70. Duff, M.J., Okun, L.B., Veneziano, G.: Trialogue on the number of fundamental constants. JHEP **0203**, 023 (2002)

71. Dvali, G., Gabadadze, G., Porrati, M.: 4D gravity on a brane in 5D minkowski space. Phys. Lett. **B 485**, 208 (2000)
72. Dyson, F.J.: Aspects of quantum theory. Salam, A., Wigner, E.P. (eds.), Cambridge University Press, Cambridge, p. 213 (1972)
73. Eisenstein, D.J., et al.: Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies. Astrophys. J. **633**, 560 (2005)
74. Fabrici d'Acquapendente, G.: De Visione, Voce, Auditu, Venetiis, per Franciscum Bolzettam, p. 93 (1600)
75. Favaro, A.: Galileo Galilei e lo Studio di Padova, 2 vols., Firenze, Successori Le Monnier (1883) and reprint, Padova, editrice Antenore (1966)
76. Favaro, A.: Le Opere di Galileo Galilei, 20 vols., Firenze, Barbera (1890–1907)
77. Favaro, A.: Amici e corrispondenti di Galileo. XL. Giuseppe Moletti. Atti del reale Istituto di SS, LL. e AA., 77, pp. 48–118 (1917–1918)
78. Favaro, A.: Galileo Galilei a Padova, Padova, Antenore (1968)
79. Favaro, A.: Scampoli galileiani. Rossetti, L., Soppelsa, M.L. (eds.), 2 vols., Trieste, ed. Lint (1992)
80. Favaro, A.: Adversaria Galilaeiana. Rossetti, L., Soppelsa, M.L. (eds.), 2 vols., Trieste, ed. Lint (1992)
81. Ferland, G., Baldwin, J.: Quasars and cosmology. ASP Conf. Ser. **162** (1999)
82. Ferris, T.: The red limit. The search for the edge of the universe. William Morrow, New York (1977); Quill, New York (1983) 2nd ed. revised and updated
83. Fischer, M., et al.: New limits on the drift of fundamental constants from laboratory measurements. Phys. Rev. Lett. **92**, 230802 (2004)
84. Fujii, Y., et al.: The nuclear interaction at Oklo 2 billion years ago. Nucl. Phys. **B 573**, 377 (2000)
85. Efstathiou, G., Sutherland, W.J., Maddox, S.J.: The cosmological constant and cold dark matter. Nature **348**, 705 (1990)
86. Einstein, A.: Ideas and opinions. Laurel Edition, 285 (1978)
87. Galilei, G.: Due lezioni all'Accademia fiorentina circa la figura, sito e grandezza dell'inferno di Dante Alighieri. In: Opere, ix. Firenze (1587–1588)
88. Galilei, G.: Sidereus Nuncius, In: Opere, iii, p. 6. Tomaso Baglioni ed. Venezia (1610)
89. Galilei, G.: Il Saggiatore, In: Opere, vi. Giacomo Mascardi, Roma (1623)
90. Galilei, G.: Dialogo. In: Opere, vii, pp. 150–151. Giovan Batista Landini ed. Firenze (1632)
91. Galilei, G.: Discorsi. In: Opere, viii, p. 7. Leida Elzeviri (1638)
92. Génova-Santos, R,.et al.: A very small array search for the extended Sunyaev-Zel'dovich effect in the Corona Borealis supercluster. Mon. Not. R. Astron. Soc. **363**, 79–92 (2005)
93. Gentile, G., et al.: Tidal dwarf galaxies as a test of fundamental physics. Astron. Astrophys. **472**, L25 (2007)
94. Geymonat, L.: Galileo Galilei. Torino, Einaudi (1980)
95. Gilbert, W.: De magnete magneticisque corporibus et de magno magnete tellure physiologia nova, plurimis et argumentis et experimentis demonstrata. Londini, excudebat Petrus Short, anno (1600)
96. Gingerich, O.: God's universe. Belknap Pressof Harvard University Press, Cambridge, MA (2006); Transl: Cercando Dio nell'universo. Lindau ed. (2007)
97. Grainge, K., et al.: The cosmic microwave background power spectrum out to $l = 1400$ measured by the very small array. Mon. Not. R. Astron. Soc. **341**, L23–L28 (2003)
98. Granett, B.R., Neyrinck, M., Szapudi, I.: An imprint of Super-Structures on the microwave background due to the integrated Sachs–Wolfe effect (2008) [arXiv:0805.3695].
99. Griselini, F.: Del genio di F. Paolo Sarpi. Venezia, appresso Leonardo Bassagli, ii, pp. 70–71 (1785)
100. Guth, A.H.: Inflationary universe: A possible solution to the horizon and flatness problems. Phys. Rev. **D 23**, 347–356 (1981)
101. Hancock, S., et al.: Direct observation of structure in the cosmic microwave background. Nature **367**, 333 (1994)

102. Harrison, D.L., et al.: A measurement at the first acoustic peak of the cosmic microwave background with the 33-GHz interferometer. Mon. Not. R. Astron. Soc. **316**, L24–L28 (2000)
103. Hellings, R.W., et al.: Experimental test of the variability of G using Viking lander ranging data. Phys. Rev. Lett. **51**, 1609 (1983)
104. Hildebrandt, S.R., et al.: COSMOSOMAS observations of the cosmic microwave background and galactic foregrounds at 11 GHz: Evidence for anomalous microwave emission at high galactic latitude. Mon. Not. R. Astron. Soc. **382**, 594–608 (2007)
105. Hoyle, F.: A new model for the expanding universe. Mon. Not. R. Astron. Soc. **108**, 372 (1948)
106. Hoyle, F.: Religions and the scientists. SCM press, London (1959)
107. Hoyle, F.: Il principio antropico e il principio cosmologico perfetto: Somiglianze e differenze. In: Il principio antropico. A cura di B. Giacomini, (ed). Spazio Libri Editori, p. 37 (1991)
108. Hubble, E.: A relation between distance and radial velocity among extra-galactic nebulae. Proc. Natl. Acad. Sci. **15**, 168 (1929)
109. Ichikawa, K., Kawasaki, M.: Constraining the variation of the coupling constants with big bang nucleosynthesis. Phys. Rev. **D 65**, 123511 (2002)
110. Ichikawa, K., Kawasaki, M.: Big bang nucleosynthesis with a varying fine structure constant and nonstandard expansion rate. Phys. Rev. **D 69**, 123506 (2004)
111. Jassal, H.K., Bagla, J.S., Padmanabhan, T.: WMAP constraints on low redshift evolution of dark energy. Mon. Not. R. Astron. Soc. **356**, L11–L16 (2005) [astro-ph/0404378]
112. Jha, S., Riess, A.G., Kirshner, R.P.: Improved distances to type Ia supernovae with multicolor light-curve shapes: MLCS2k2. Astrophys. J. **659**, 122 (2007)
113. Jones, M.H., Lambourne, R.J.: An introduction to galaxies and cosmology. The open University, Cambridge University Press, Cambridge (2003)
114. Kalberla, P.M.W., et al.: Dark matter in the Milky Way. II. The HI gas distribution as a tracer of the gravitational potential. Astron. Astrophys. **469**, 511 (2007)
115. Kapteyn, J.C.: First attempt at a theory of the arrangement and motion of the sidereal system. Astrophys. J. **55**, 302 (1922)
116. Kashikawa N. et al.: SUBARU deep Survey. III. Evolution of rest-frame luminosity functions based on the photometric redshift for a K-band selected galaxy sample. Astron. J. **125**, 53 (2003)
117. Kolb, E.W., Perry, M.J., Walker, T.P.: Time variation of fundamental constants, primordial nucleosynthesis, and the size of extra dimensions. Phys. Rev. **D 33**, 869 (1986)
118. Kolb, E.W., Matarrese, S., Riotto, A.: On cosmic acceleration without Dark Energy. New J. Phys. **8**, 322 (2006)
119. Kothawala, D., et.al.: Einstein's equations as a thermodynamic identity: The cases of stationary axisymmetric horizons and evolving spherically symmetric horizons. Phys. Letts, **B 652**, 338 (2007) [gr-qc/0701002].
120. Lahav, O., et al.: The 2dF galaxy redshift survey: The amplitudes of fluctuations in the 2dFGRS and the CMB, and implications for galaxy biasing. Mon. Not. R. Astron. Soc. **333**, 961 (2002)
121. Layzer, D.: The arrow of time. Astrophys. J. **206**, 559 (1976)
122. Liddle, A.R., Scherrer, R.J.: Classification of scalar field potentials with cosmological scaling solutions. Phys. Rev. **D 59**, 023509 (1999)
123. Linder, E.V.: Exploring the expansion history of the universe. Phy. Rev. Lett. **90**, 091301 (2003)
124. Linder, E.V., Jenkins, A.: Cosmic structure and dark energy. Mon. Not. Roy. Astron. Soc. **346**, 573 (2003)
125. Longair, M.S.: Galaxy formation, 2nd ed., Springer, Berlin (2008)
126. Lugmair, G.W., Galer, S.J.G.: Age and isotopic relationships among the angrites Lewis Cliff 86010 and Angra DOS Reis. Geochim. Cosmochim. Acta **56**, 1673 (1992)
127. Magnano, G., Ferraris, M., Francaviglia, M.: Nonlinear gravitational Lagrangians. Gen. Rel. Grav. **19**, 465 (1987)
128. Mannheim, P.D.: Conformal cosmology with no cosmological constant. Gen. Rel. Gravit. **22**, 289–298 (1990)

129. Mannheim, P.D.: Conformal gravity and the flatness problem. Astrophys. J. **391**, 429–432 (1992)
130. Mannheim, P.D.: Alternatives to dark matter and dark energy. Prog. Part. Nucl. Phys. **56**, 340–445 (2006)
131. Mannheim, P.D., Kazanas, D.: Exact vacuum solution to conformal weyl gravity and galactic rotation curves. Astrophys. J. **342**, 635–638 (1989)
132. Marion, H., et al.: Search for variations of fundamental constants using atomic fountain clocks. Phys. Rev. Lett. **90**, 150801 (2003)
133. McGaugh, S.S.: The baryonic Tully-Fisher relation of galaxies with extended rotation curves and the stellar mass of rotating galaxies. Astroph. J. **632**, 859 (2005)
134. McGaugh, S.S.: Observational constraints on the acceleration discrepancy problem (2006) [astro-ph/0606351]
135. Micanzio, F.: Vita di Paolo Sarpi, Teologo e Consultore della Serenissima Repubblica di Venezia, frate dell'Ordine dei Servi. Milano, per Giovanni Silvestri, (1824), 58; the on-line edition Fulgenzio Micanzio, *Vita del Padre Paolo*, ed. Corrado Vivanti, is available at www.liberliber.it, Manuzio Project
136. Mihalas, D., Binney, J.: Galactic astronomy. W.H. Freeman , San Francisco (1968)
137. Milgrom, M.: A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. Astrophys. J. **270**, 365–370 (1983)
138. Milgrom, M.: A modification of the Newtonian dynamics – implications for galaxies. Astrophys. J. **270**, 371–383 (1983)
139. Milgrom, M.: A modification of the Newtonian dynamics – implications for galaxy systems. Astrophys. J. **270**, 384–389 (1983)
140. Milgrom, M.: Dynamics with a nonstandard inertia-acceleration relation: An alternative to dark matter in galactic systems. Ann. Phys. **229**, 384 (1994)
141. Milgrom, M.: The modified dynamics as a vacuum effect. Phys. Lett. **A 253**, 273 (1999)
142. Milgrom, M.: The shape of dark matter haloes of disc galaxies according to MOND. Mon. Not. R. Astron. Soc. **326**, 1261 (2001)
143. Milgrom, M.: MOND and the mass discrepancies in tidal dwarf galaxies. Astrophys. J. Lett. **667**, L45 (2007)
144. Milgrom, M.: The MOND paradigm. In: Proceedings XIX Rencontres de Blois (2008) [arXiv:0801.3133M]
145. Moffat, J.W.: A modified gravity and its consequences for the solar system, astrophysics and cosmology. Int. J. Mod. Phys. **D J. Exp. Theor. Phys. 16**, 2075–2090 (2007)
146. Mukhopadhyay, A., Padmanabhan, T.: Holography of gravitational action functionals. Phys. Rev. **D 74**, 124023 (2006) [hep-th/0608120]
147. Muller, C.M., Schafer, G., Wetterich, C.: Nucleosynthesis and the variation of fundamental couplings. Phys. Rev. **D 70**, 083504 (2004)
148. Murphy, M.T., Webb, J.K., Flambaum, V.V.: Further evidence for a variable fine-structure constant from Keck/HIRES QSO absorption spectra. Mon. Not. R. Astron. Soc. **345**, 609 (2003)
149. Murphy, M.T., Webb, J.K., Flambaum, V.V.: Comment on limits on the time variation of the electromagnetic fine-structure constant in the low energy limit from absorption lines in the spectra of distant quasars. Phys. Rev. Lett. **99**, 239001 (2007)
150. Navarro, J.F., Frenk, C.S., White, S.D.M.: The structure of cold dark matter halos. Astrophys. J. **462**, 563–575 (1996)
151. Navarro, J.F., Frenk, C.S., White, S.D.M.: A universal density profile from hierarchical clustering. Astrophys. J. **490**, 493–508 (1997)
152. Nieto, M.M., Anderson, J.D.: Search for a solution of the pioneer anomaly. Contemp. Phys. **48**, 41 (2007)
153. Nollett, K.M., Lopez, R.E.: Primordial nucleosynthesis with a varying fine structure constant: An improved estimate. Phys. Rev. **D 66**, 063507 (2002)
154. Okun, L.B.: The fundamental constants of physics. Usp. Fiz. Nauk **161**, 177 (1991)
155. Olive, K.A., et al.: Constraints on the variations of the fundamental couplings. Phys. Rev. **D 66**, 045022 (2002)

156. Olive, K.A., et al.: Reexamination of 187Re bound. Phys. Rev. **D 69**, 027701 (2004)
157. Olmo, G.J.: The gravity lagrangian according to solar system experiments. Phys. Rev. Lett. **95**, 261102 (2005)
158. Ortolan, A., Secco, L.: Il Principio Antropico fra il dominio della scienza e quello della filosofia. Giornale di Fisica, XXXVII, N.4, 243 (1996)
159. Padmanabhan, T.: Gravitational field of the quantized electromagnetic plane wave. Gen. Rel. Grav. **19**, 927 (1987)
160. Padmanabhan, T.: Quantum structure of spacetime and entropy of schwarschild black holes. Phys. Rev. Letts. **81**, 4297 (1998) [hep-th/9801015]
161. Padmanabhan, T.: Classical and quantum thermodynamics of horizons in spherically symmetric spacetimes. Classical Quant. Grav. **19**, 5387, (2002) [gr-qc/0204019]
162. Padmanabhan, T.: Cosmological constant-the weight of the vacuum. Phys. Rept. **380**, 235 (2003) [hep-th/0212290]
163. Padmanabhan, T.: Vacuum fluctuations of energy density can lead to the observed cosmological constant. Classical Quant. Grav. **22**, L107 (2005) [hep-th/0406060]
164. Padmanabhan, T.: Gravity and the thermodynamics of horizons. Phys. Rept. **406**, 49 (2005) [gr-qc/0311036]
165. Padmanabhan, T.: Dark energy: Mystery of the millennium. AIP Conf. Proc. **861**, 179 (2006) [astro-ph/0603114]
166. Padmanabhan, T.: Gravity's immunity from vacuum: The holographic structure of semiclassical action. Third prize essay; Gravity Essay Contest 2006 Gen. Rel. Grav. **38**, 1547–1552 (2006)
167. Padmanabhan, T.: Gravity as an emergent phenomenon: A conceptual description. AIP Conf. Proc. **989**, 114 (2007) [arXiv: 0706.1654]
168. Padmanabhan, T.: Dark energy and gravity. Gen. Rel. Grav. **40**, 529 (2008) [arXiv: 0705.2533]
169. Padmanabhan, T., Choudhury, T.R.: The issue of choosing nothing. Mod. Phys. Lett. **A 15**, 1813 (2000) [gr-qc/0006018]
170. Padmanabhan, T., Choudhury, T.R.: A theoretician's analysis of the supernova data and the limitations in determining the nature of dark energy. Mon. Not. R. Astron. Soc. **344**, 823 (2003) [astro-ph/0212573]
171. Pais, A., Uhlenbeck, G.E.: On field theories with non-localized action. Phys. Rev. **79**, 145–165 (1950)
172. Paranjape, A., Sarkar, S., Padmanabhan, T.: Thermodynamic route to field equations in Lanczos-Lovelock gravity. Phys. Rev. D **74**, 104015, (2006) [hep-th/0607240]
173. Pastore-Stocchi, M.: Il periodo veneto di Galileo Galilei. In: Storia della Cultura veneta, iv, t. ii, pp. 37–66 (1984)
174. Peacock, J.A., Cole, S., Norberg, P.: A measurement of the cosmological mass density from clustering in the 2dF galaxy redshift survey. Nature **410**, 169 (2001)
175. Peacock, J.A.: Cosmological physics, Cambridge University Press, Cambridge (1999)
176. Pecchioli, R.: Introduzione. In: Paolo Sarpi, Istoria del Concilio Tridentino. Firenze, Sansoni editore, pp. xiii–xiv (1982)
177. Peebles, P.J.E., Dicke, R.H.: Cosmology and the radioactive decay ages of terrestrial rocks and meteorites. Phys. Rev. **128**, 2006 (1962)
178. Percival, W.J., et al.: Parameter constraints for flat cosmologies from CMB and 2dFGRS power spectra. Mon. Not. Roy. Astron. Soc. **37**, 1068 (2002)
179. Percival, W.J., et al.: Measuring the matter density using baryon oscillations in the SDSS. Astrophys. J. **657**, 51 (2007)
180. Percival W.J. et al.: Measuring the baryon acoustic oscillation scale using the SDSS and 2dFGRS. Mon. Not. Roy. Astron. Soc. **381**, 1053 (2007)
181. Perlmutter, S., et al.: Measurements of the cosmological parameters omega and lambda from the first seven supernovae at $z \geq 0.35$. Astrophys. J. **483**, 565 (1997)
182. Perlmutter, S. et al., Discovery of a supernova explosion at half the age of the universe. Nature **391**, 51 (1998)

183. Perlmutter, S.J., et al.: Measurements of Ω and Λ from 42 high-redshift supernovae. Astrophys. J. **517**, 565–586 (1999)
184. Quast, R., Reimers, D., Levshakov, S.A.: Probing the variability of the fine-structure constant with the VLT UVES. Astron. Astrophys. **415**, L7 (2004)
185. Ratra, B., Peebles, P.J.E.: Cosmological consequences of a rolling homogeneous scalar field. Phys. Rev. **D 37**, 3406 (1988)
186. Rebolo, R., et al.: Cosmological parameter estimation using very small array data out to $l = 1500$. Mon. Not. R. Astron. Soc. **353**, 747–759 (2004)
187. Rees, M.J.: Just six numbers. Basic Books, New York (2000). Transl: I sei numeri dell'universo. Rizzoli (2002)
188. Reeves, H.: La crescita della complessita' in un universo in espansione. p. 61. In: Il principio antropico. a cura di B. Giacomini, ed. Spazio Libri Editori (1991)
189. Renn, J., Valleriani, M.: Galileo and the challenge of the arsenal. In: Letture Galileiane. Florence, 21 March 2001 (online edition)
190. Riess, A.G., et al.: Observational evidence from supernovae for an accelerating universe and a cosmological constant. Astron. J. **116**, 1009–1038 (1998)
191. Riess, A.G., et al.: Type Ia supernova discoveries at $z > 1$ from the hubble space telescope: Evidence for past deceleration and constraints on dark energy evolution. Astrophys. J. **607**, 665–687 (2004)
192. Righini, G.: Contributo alla interpretazione scientifica dell'opera astronomica di Galileo. Suppl. Annali dell'Istituto e Museo di Storia della Scienza, fasc. 2, Firenze, 1978
193. Rocha, G., et al.: New constraints on varying alpha. New Astron. Rev. **47**, 863 (2003)
194. Roy, S., Kafatos, M., Datta, S.: Shift of spectral lines due to dynamic multiple scattering and screening effect: Implications for discordant redshifts. Astron. Astrophys. **353**, 1134 (2000)
195. Sakharov, A.D.: Vacuum quantum fluctuations in curved space and theory of gravitation. Sov. Phys. Dokl. **12**, 1040 (1968)
196. Sanders, R.H.: Modified gravity without dark matter. In: The invisible universe: Dark matter and dark energy. Third Aegean Summer School. (2006) [arXiv:astro-ph/0601431]
197. Sanders, R.H., McGaugh, S.S.: Modified Newtonian dynamics as an alternative to dark matter. Ann. Rev. Astron. Astrophys. **40**, 263 (2002)
198. Sánchez-Salcedo, F.J., Saha, K., Narayan, C.A.: The thickness of HI in galactic discs under MOdified Newtonian Dynamics: Theory and application to the Galaxy. Mon. Not. R. Astron. Soc. **385**, 1585 (2007) [arXiv:0712.0816]
199. Sarpi, P.: Lettere di fra Paolo Sarpi. ed. Filippo L. Polidori, preface of Filippo Perfetti, Firenze, G. Barbera ed. (1863)
200. Sarpi, P.: Lettere ai Gallicani. Boris Ulianich (ed.), F. Steiner, Wiesbaden (1961)
201. Sarpi, P.: Istoria dell'interdetto. Corrado Pin (ed.), introduction by William Shea, Conselve, Ed THINK ADV. (2006)
202. Scarpa, R.: Modified Newtonian dynamics, an introductory review. In: First crisis in cosmology. E.J. Lerner, J.B. Almeida (eds.) (2006) [astro-ph/0601478]
203. Scott, P. F., et al.: First results from the very small array. III. The cosmic microwave background power spectrum. Mon. Not. R. Astron. Soc. **341**, 1076–1083 (2003)
204. Scranton, R., et al.: Physical evidence for dark energy. [astro-ph/0307335v2] (2003)
205. Secco, L.: Oltre la quantità. In: "Un futuro per l'Uomo". Gabrielli Ed. (2005)
206. Seljak, U. et al.: Cosmological parameter analysis including SDSS Lyα forest and galaxy bias: Constraints on the primordial spectrum of fluctuations, neutrino mass, and dark energy. Phys. Rev. **D 71**, 103515 (2005)
207. Shea, W.R., Artigas, M.: Galileo in Rome. The rise and fall of a troublesome genius. Oxford University Press, Oxford, 2003
208. Shlyakhter, A.I.: Direct test of the constancy of fundamental nuclear constants. Nature **264**, 340 (1976)
209. Skordis, C., et al.: Large scale structure in Bekenstein's theory of relativistic modified Newtonian dynamics. Phys. Rev. Lett. **96**, 011301 (2006)
210. Smoliar, M.I., Walker, R.J., Morgan, J.W.: Re-Os Ages of group IIA, IIIA, IVA, and IVB iron meteorites. Science **271**, 1099 (1996)

211. Smolin, L.: Three roads to quantum gravity. Basic Books, New York (2001)
212. Smolin, L.: The trouble with physics. Houghton Mifflin Harcourt (2006)
213. Snyder, H.S.: Quantized space-time. Phys. Rev. **71**, 38 (1947)
214. Sotiriou, T.P., Liberati, S.: Metric-affine $f(R)$ theories of gravity. Ann. Phys. **322**, 935 (2007)
215. Spergel, D.N., et al.: Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. Astrophys. J. Suppl. **170**, 377 (2007)
216. Srianand, R., et al.: Limits on the time variation of the electromagnetic fine-structure constant in the low energy limit from absorption lines in the spectra of distant quasars. Phys. Rev. Lett. **92**, 121302 (2004)
217. Srianand, R., et al. Replay. Phys. Rev. Lett. **99**, 239002 (2007)
218. Starobinsky, A.A.: A new type of isotropic cosmological models without singularity. Phys. Lett. **B 91**, 99 (1980)
219. Steinhardt, P.J., Wang, L., Zlatev, I.: Cosmological tracking solutions. Phys. Rev. **D 59**, 123504 (1999)
220. 't Hooft, G.: The mathematical basis for deterministic quantum mechanics. In: Beyond the quantum, Th.M. Nieuwenhuizen et al. (eds.), World Scientific, Singapore, pp. 3–19 [quant-ph/0604008]
221. 't Hooft, G.: Emergent quantum mechanics and emergent symmetries, PASCOS 13, London, [hep-th/0707.4568]
222. Tegmark, M., et al.: Cosmological parameters from SDSS and WMAP. Phys. Rev. **D 69**, 103501 (2004)
223. Ulianich, B.: Saggio introduttivo. In: Paolo Sarpi, Lettere ai Gallicani. Boris Ulianich (ed.), Wiesbaden, F. Steiner, 1961
224. Uzan, J.P.: The fundamental constants and their variation: Observational status and theoretical motivations. Rev. Mod. Phys. **75**, 403 (2003)
225. Watson, R. A., et al.: Detection of anomalous microwave emission in the Perseus molecular cloud with the COSMOSOMAS experiment. Astrophys. J. **624**, L89-L92 (2005)
226. Webb, J.K., et al.: Evidence for time variation of the fine structure constant. Phys. Rev. Lett. **82**, 884 (1999)
227. Weinberg, S.: Anthropic bound on the cosmological constant. Phys. Rev. Lett. **9**, 2607–2610 (1987)
228. Weinberg, S.: The cosmological constant problem. Rev. Mod. Phys **61**, 1–23 (1989)
229. Wetterich, C.: Cosmology and the fate of dilatation symmetry. Nucl. Phys. **B 302**, 668 (1988)
230. White, S.D.M.: Fundamentalist physics: Why dark energy is bad for astronomy. Report Prog. Phys. **70**, 883 (2007)
231. Will, C.M.: The Confrontation between general relativity and experiment. Living reviews in relativity, **9**, http://www.livingreviews.org/lrr-2006-3
232. Zhang, P.: Testing gravity against the early time integrated Sachs–Wolfe effect. Phys. Rev. **D 73**, 123504 (2006)
233. Zhou, J., et al.: The generalized second law of thermodynamics in the accelerating universe. (2007) [arXiv:0705.1264]

Chapter 5
Next Challenges

Contributions by Matthias Bartelmann, Charles L. Bennett, Martin Bucher, Carlo Burigana, Massimo Capaccioli, Mauro D'Onofrio, Ruth Durrer, Isabella Gioia, Günther Hasinger, Charles Lawrence, Giuseppe Longo, Juan Francisco Macias-Perez, Piero Madau, John Mather, John Peacock, Lucia A. Popa, Rafael Rebolo, Wolfgang Reich, Matthias Steinmetz, Massimo Turatto, and Simon D.M. White

5.1 Outline of the Chapter

Previous chapters have offered a view of the status of several lines of research of interest for current cosmology, covering observational aspects and their theoretical interpretation(s) and focusing on the wealth of information they provide for our comprehension of the properties of the Universe. It is remarkable that, in spite of the huge and high quality developments achieved during the last decades, many important questions linked to both fundamental physics and cosmology are still open. As often occurs in science and as pointed out in the contributions of previous chapters, recent achievements and discoveries give satisfactory answers to some problems, but, at the same time, pose new crucial questions which have deep implications for our view of the cosmos.

This chapter is devoted to the current efforts of the astrophysical community aimed at improving our knowledge of the Universe. The emphasis is put on the forthcoming and future projects and on their scientific promises. In many of the next sections, a significant space will be left to technological aspects at the basis of next astronomical missions and facilities relevant for cosmology. We hope to offer to the reader a view of the huge and high quality work carried out by the teams dedicated to these challenges.

Certainly, the astrophysical community takes advantage of the general technological improvements. On the other hand, the new scientific goals, achievable only through the continuous improvement of astronomical instrumentation, stimulate, and in many cases, effectively lead to the technological development. In this sense, the "science driven" approach that will emerge from the following presentations reminds us, as a state of fact, how much the genuine desire of knowledge contributes to the human contemporary progress, or, more modestly, that it's been this way for years.

The various themes presented in the first and largest part of this chapter are organized almost according to the "classical scheme" in astrophysics, based on the subdivision of electromagnetic radiation spectrum according to frequency, starting from topics and projects at the longest wavelengths, the radio domain, to arrive at

M. D'Onofrio and C. Burigana (eds.), *Questions of Modern Cosmology: Galileo's Legacy*, DOI 10.1007/978-3-642-00792-7_5,

the perspectives in the X and γ rays, the so-called high energy astrophysics. We then include three sections dedicated to selected topics introduced in Chaps. 2 and 3.

The first interview with Wolfgang Reich, in Sect. 5.2, will highlight the main scientific promise of the new generation of large radio interferometers, of particular interest at least in two fields discussed in Chaps. 2 and 3: cosmological reionization and large scale magnetic fields.

We then move on to a broad section dedicated to the new perspectives in Cosmic Microwave Background (CMB) cosmology. The interview to John Mather (Sect. 5.3.1) attacks the important problem of the detection of the small CMB spectral distortions that are predicted to exist in standard, and less standard, models. We then discuss with Rafael Rebolo on the different roles of ground-based and space experiments and on their complementarity (Sect. 5.3.2).

In 2009, the European Space Agency (ESA) Planck *satellite will be launched. The whole cosmological and astrophysical community is waiting for its results, which should give definitive answers to some of the questions still open after the Wilkinson Microwave Anisotropy Probe (WMAP). Starting from the scientific motivations for* Planck*, the interview of Charles Lawrence in Sect. 5.3.3 will describe the main solutions adopted to carry out such accurate measurements and the cosmological impact of the* Planck *mission. The wide frequency coverage of* Planck *data will allow a significant improvement in the separation of the foreground from the cosmological signal. The interview to Juan Francisco Macias-Perez in Sect. 5.3.4 will point out how other projects at sub-millimeter, far-infrared, and centimeter wavelengths will complement* Planck *data in this respect. We then move in Sect. 5.3.5 to the next main benchmark of CMB cosmology: the search for primordial B-mode polarization anisotropy, or, in other words, for the background of primordial gravitational waves that is expected to be generated during the early stages of the Universe. The discovery of such signal is crucial to definitively test inflation. The interview with Martin Bucher will highlight the strong motivation, the difficulty, and, at the same time, the feasibility of this research.*

Section 5.4 is devoted to the important role for cosmology of galaxy surveys, mainly in the optical and infrared. The first two interviews with John Peacock and Charles Bennett are focused on the implications of future surveys for understanding the fundamental problem of present day cosmology, that is, the cosmological constant or dark energy, using, for example, Supernovae (SNe), lensing, Baryonic Acoustic Oscillations (BAO), also in view of next space projects.

In Sect. 5.4.1, the interviews to Massimo Capaccioli and Giuseppe Longo address the renaissance of ground-based optical surveys. They point out the "general purpose" of these surveys that might play a role in various astronomical fields, from Solar system objects to distant galaxies. The technological aspects are discussed, from hardware to data analysis issues, from the management of huge data sets to the exploration of the so-called "time domain".

Our discussion continues, in a more multi-frequency approach, with the interview with Piero Madau in Sect. 5.5 on the perspectives of new observations for understanding the first structures and the transition from dark to dawn ages.

Looking at the role of future surveys, and in continuation of the discussion with Peter Coles in Chap. 3, it is natural to ask what could be the contribution of numerical simulations aimed at modeling the structure formation, an aspect closely linked with the understanding of Dark Matter (DM) and feedback processes. We will discuss these aspects in the interviews with Matthias Steinmetz and Simon White (Sect. 5.6).

Future surveys will offer, in principle, the great opportunity to discover and study a large number of SNe up to redshifts significantly higher than those achieved up to now. In Chap. 2, we discussed the crucial role of these astronomical candles in mapping the expansion of the Universe and in the recent paradigm change from Cold Dark Matter (CDM) to ΛCDM. The possibilities opened by the next generation of infrared and optical surveys will be addressed in the interview to Massimo Turatto (Sect. 5.7).

Finally, the interviews to Günther Hasinger and Isabella Gioia in Sect. 5.8 will highlight the contribution to cosmology of future X and γ ray projects, the possible role of Gamma-ray bursts (GRB) as potential cosmological candles and high redshift tracers, and, in a more multi-frequency approach, the impact of new observations for our comprehension of galaxy cluster physics and their use in cosmology.

In the interview with Matthias Bartelmann on the expectations coming from future lensing studies (Sect. 5.9), we will come back to questions more linked to fundamental physics, from nonstandard metric theories and alternative dynamics and gravity theories to the problem of the detection of cosmic strings and Dark Energy (DE), through next generation optical surveys.

Then, Ruth Durrer will comment further on the role of lensing in the detection of topological effects, cosmic strings, in particular (Sect. 5.10). She will also address the future prospects for strings from the detection of gravitational waves and will compare the information coming from these kinds of observables with that based on CMB.

We conclude this chapter with an interview with Lucia Popa about the contributions provided in the future by different kinds of astronomical observations for a better understanding of neutrino physics (Sect. 5.11).

Let us start with Wolfgang Reich, who will illustrate the new generation of large radio interferometers and their main perspectives for cosmology.

5.2 New Perspectives from Radio Astronomy

Dear Wolfgang *(Reich)***, a crucial step in future radio-astronomy is expected by the development of a new generation of large interferometer projects, such as the Low Frequency Array (LOFAR) and the Square Kilometer Array (SKA). Can you describe the main concepts of such challenges? Can you discuss their main implications from our understanding of cosmological reionization and of the formation and evolution of magnetic fields on galactic and cosmological scales? Could HI tomography map the 21 cm spectral features that are predicted to be present prior of the full reionization?**

5.2.1 New Radio Telescopes Trace the Epoch of Reionization

Classical fully movable single-dish radio telescopes exist with diameters of up to 100 m like the Effelsberg and the Green Bank dishes. Telescopes operating at sub-millimeter wavelengths are much smaller. Very large telescopes like the 300 m Arecibo dish or the planned 500 m FAST telescope in China have a fixed main dish in a landscape depression and use moveable feeds for a limited sky access. Synthesis radio telescopes consist of numerous small telescopes, which are separated by larger distances to increase the intrinsically low angular resolution of radio telescopes due to the large wavelengths where they observe. These days new radio telescopes are under development, which entirely rely on digital electronics and have possibilities beyond those of classical telescopes. LOFAR is such a new telescope for long wavelengths (15 and 1.25 m), which is developed and build by ASTRonomisch Onderzoek in Nederland (ASTRON) in the Netherlands and will finally extend across Europe with distances between the individual antenna stations of about 1,000 km or even more. The basic antenna element of LOFAR is a simple dipole, which has the ability to permanently receive emission from nearly all directions of the visible sky over a wide frequency range. This opens new possibilities like observing all kinds of transient phenomena in the sky, which are not accessible with classical telescopes pointing always at a defined direction. LOFAR consists of antenna stations, where up to 96 fixed dipoles are located in an antenna field of about 65 m in diameter. The signals from the dipoles can be electronically combined, so that the waves from a certain sky direction are added "in phase", which is equivalent to point a classical telescope in a certain direction. LOFAR has no moving parts. Its digital electronics allow to observe in up to eight different directions at the same time. About 50 LOFAR stations will be distributed across Europe with a high concentration of stations in the central core located in the northern part of the Netherlands. The challenge of digital telescopes is the enormous data stream created by each antenna station, which amounts to 2 Gbit s^{-1} per LOFAR station. This data stream needs fiber connections being directed to a central processing unit, a supercomputer, where the data processing is performed in real time. No storage of the huge amount of raw data is planned. The digital technique realized with LOFAR is limited to low frequencies in the moment and is expected to become available for higher frequencies up to 1 or 2 GHz, which is important for planning the future SKA, the Square Kilometer Array, which is the major international project planned for 2020 and beyond by all leading radio observatories in the world. Its collecting area is aimed to be one square-kilometer and will be spread over distances of about 3,000 km. Its frequency range will likely cover the range from about 100 MHz up to 10 GHz or higher. The experience gained with LOFAR will be essential for planning the low-frequency part of the SKA. For higher frequencies, a classical synthesis telescope design with cheap small dishes is actually planned.

LOFAR exceeds all previous low-frequency radio telescopes in sensitivity by two orders of magnitudes and is thus suited for an attempt to detect signals from the Epoch of Reionization (EOR). During this phase, first stars and galaxies are formed from density fluctuations of neutral hydrogen, which is the state of matter during

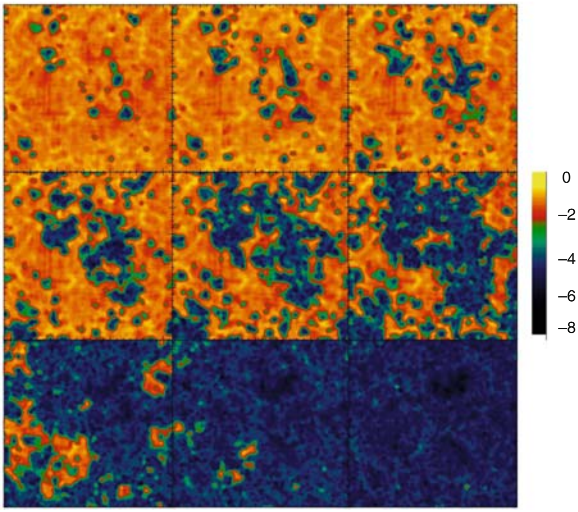

Fig. 5.1 EOR simulations. Differential antenna temperature in logarithmic scale are shown for redshifts between $z = 13.5$ (*upper left panel*) and $z = 8.1$ (*lower right panel*). The maps refer to the emission from a co-moving 220 kpc slice of gas. From [28]

the *Dark Ages* lasting for several hundred million years after the *Big Bang*. The UV-flux of stars and galaxies ionizes the neutral matter around them. The ionized bubbles are embedded in a neutral hydrogen environment. Neutral hydrogen emits a narrow spectral line at $\lambda 21$ cm or 1.420 MHz, which is the most intense radio spectral line observed from the milky way. However, the EOR happens quite early in the Universe. Present estimates expect this phase to set in around $z \sim 15 \div 20$ and being finished at around $z = 6$. These redshifts correspond to a shift of the $\lambda 21$ cm line to about 70–95 MHz (start of EOR) and 235 MHz (end of EOR). Figure 5.1 shows an example of EOR simulations, where ionized bubbles form in the neutral hydrogen as a function of z. LOFAR will be able to observe the frequency range from 110 to 240 MHz, which means that the very early phase might be missed. The reason that LOFAR cannot be observed between 80 and 110 MHz is due to FM radio stations distributed across Europe. These stations are so strong emitters that they mask any weak signal from the Universe. To observe this frequency range, a low-frequency array like LOFAR has to be built in Australia or Siberia, where the FM range is not used for broadcasting.

The major problem in detecting EOR signals is the removal of the complex foreground emission, which dominates the expected signal by about three orders of magnitudes. The foreground consists of Galactic and extragalactic continuum and recombination line emission and in particular of emission from weak compact radio sources, which need to be resolved and to be subtracted. This requires very long baselines to achieve arcsecond angular resolutions at these low frequencies. Unresolved compact sources confuse and are the limiting noise term (confusion limit). Longer exposure times will then not help to lower the noise. Presently, there are

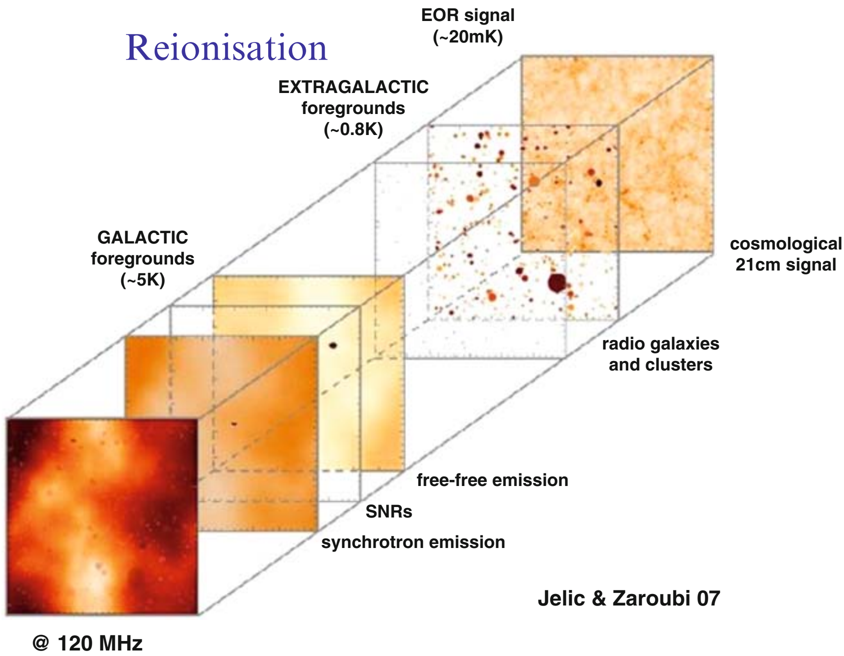

Fig. 5.2 Foreground component fluctuations, which need to be removed to uncover the weak embedded EOR signals. This figure was produced according to foreground simulations for the LOFAR–EOR experiment by Vibor Jelić and Saleem Zaroubi (private communication) and is available on the LOFAR–EOR website (see web page list)

so many unknowns in the foreground estimates at low frequencies that it is not surprising that various authors argue for or against the possibility of a successful foreground removal to detect EOR signals (see Fig. 5.2 for illustration). In any case, LOFAR will give an answer on the foreground contamination issue and also has a chance to trace signals from the early Universe.

The SKA and LOFAR are expected to make significant contributions to our understanding of synchrotron emission and its polarized component. Their sensitivity and angular resolution exceed that of any existing facility largely. LOFAR will provide new information on the magnetic fields in distant galactic halos and cluster halos. Cosmic ray electrons, which were accelerated in galactic disks to high energies, have already lost much of their initial energy when diffusing out into the halo. They illuminate the weak magnetic fields there, which become detectable by low-frequency synchrotron emission, where LOFAR will be observed. The question of magnetic field evolution from very weak seed fields in the early Universe to structured magnetic fields with a strength several orders of magnitude larger is of high interest, but not been answered yet. Pushing for highest sensitivity, LOFAR has a certain chance to detect magnetic fields in the Inter Galactic Medium (IGM) for the first time [12] and study their relation to the structure formation in the early Universe. This is expected in case the intergalactic magnetic field is not much weaker than its presently quoted upper limits of 10^{-8}–10^{-9} Gauss. In case the intergalactic magnetic field turns out to be weaker, the SKA will have another chance for its first detection and several scenarios have already been proposed how to extract its signature from SKA observations.

Thank you very much Wolfgang.
We also expect that the extreme quality of these new radio data will provide an important contribution for the subtraction of the foreground signals in the next generation of CMB experiments, discussed in the next section. We will start with the interview with John Mather on future CMB spectrum projects.

5.3 New Perspectives in CMB Cosmology

5.3.1 Ideas for New Spectrum Experiments

Dear John *(Mather)*, **Cosmic Background Explorer/Far Infrared Absolute Spectrophotometer (COBE/FIRAS) established the Planckian shape of CMB spectrum. On the other hand, very small distortions are predicted as results of primeval stages of structure formation or in some particle physics models beyond the standard one involving particle decays, annihilations, etc. What are the possibilities for the next decades to discover them? What is your opinion about next challenges in this direction?**

The COBE mission set the standard for CMB spectrum measurements, with millikelvin precision on the measured temperature, and 50 parts per million comparisons of the spectrum with a Black Body (BB) form.

An experiment in progress[1] is the ARCADE mission led by Alan Kogut at Goddard. It is a balloon payload with microwave receivers and antennas and a full-beam external calibrator body, all cooled by a flow of cold helium gas to the temperature of the CMB. The instrument concept is shown in Fig. 5.3. This is capable of millikelvin precision at centimeter wavelengths, where emission from hot electrons by the free-free process might be observable, and the equipment has already survived several flights. A successor mission could be developed for space flight, and could achieve much greater instrument precision. Whether this is required for progress will depend on the full analysis of the ARCADE data. If foreground radiation from the Milky Way galaxy already limits the CMB measurements, then a satellite mission might not be helpful. I expect that the opposite will be true, and the ARCADE data will show that the foregrounds are well enough understood to justify a space mission.

[1] *After the first submission of the manuscript of this book to Springer, some preprints presenting the results of Absolute Radiometer for Cosmology, Astrophysics, and Diffuse Emission (ARCADE) appeared on the arXiv archive [57, 68, 78, 123].*

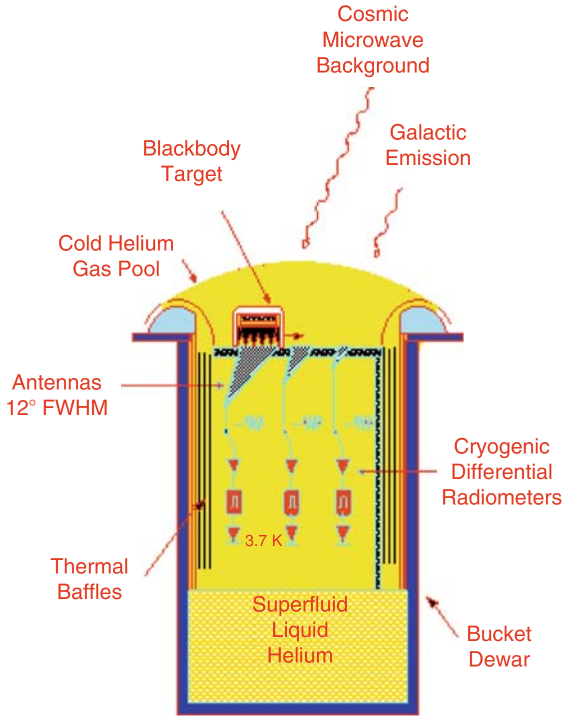

Fig. 5.3 ARCADE concept. Cold helium gas rushes out to cool the instrument apertures and the full-beam calibrator. From the ARCADE web site (see web page list)

Beyond FIRAS: Measuring CMB Spectrum Distortions

Although the FIRAS results effectively ended the debate about the validity of the Hot Big Bang theory (at least compared with the Steady State theory), continued analysis has shown the possibility of distortions of the spectrum that could be detectable. There is no current plan to make the required measurements, but they may be feasible. In my opinion, it ought to be possible to make measurements with about two orders of magnitude more sensitivity and accuracy for broad spectral distortions like the Bose–Einstein or Compton forms [55], and even stronger limits could be set for narrow-band distortions arising from disequilibrium processes at the decoupling.

Rashid Sunyaev and his colleagues have calculated that there are slight nonequilibrium effects during the cosmic decoupling, owing to the immense optical depth of the hydrogen recombination lines. As a result, there should be a small imprint, at the level of parts per billion, on the spectrum of the CMB. The pattern would be recognizable as the redshifted hydrogen spectrum, blurred in frequency by several percent by the extended time of recombination. There might also be some sign of the helium lines. Other disequilibrium phenomena have also been discussed due to the formation of the first molecules, LiH, for instance. The amplitude of the signal in all cases is small, as no particular cause for large disequilibrium has been recognized.

Instrument Concept

If an improved instrument were proposed, it would probably be totally symmetrical (the FIRAS was not, as the internal BB was not the same as the external one); totally isothermal (the FIRAS had many different temperatures in the chamber); far more sensitive (current detectors are at least 1,000 times as sensitive as those on FIRAS); and, far from Earth, to avoid issues about stray light from the Earth or Sun. It would also use improved thermometers, far more stable than those used on FIRAS, and might use a large parabolic reflector to reduce the field of view and enable the instrument to observe in the few small regions where Galactic dust is least bright. It might also use some kind of isolators between the detectors and the instrument, in the event that detectors must still be operated at temperatures well below the CMB temperature.

The ultimate limits to these measurements are probably set by our local astrophysical environment. Dust, atoms, and molecules in our Galaxy and others are quite bright relative to the proposed spectral distortions of the CMB, but they are localized in particular directions and hence (probably) recognizable. None of them would have the same frequency distribution as the hydrogen recombination lines.

At the moment, there is no particular reason to think that the theoretical calculations are incorrect, or that there is an important cosmological surprise lurking that could be discovered with a new instrument. But that could change: cosmology has been full of surprises, including both DM and DE. Now that they are widely accepted, they no longer seem as strange as they did before they were discovered.

Although there are easily enough photons available to measure the time dependence of the CMB temperature due to the expansion of the Universe, there is not yet a sufficiently stable temperature reference to enable such a measurement. Over the course of a 13.7 year mission, the temperature change would be only one part per billion.

To Summarize

The cosmic microwave background radiation has been observed with extraordinary precision and accuracy to reveal traces of the (presumably quantum mechanical) processes of the Big Bang itself. The COBE satellite mission began the era of precision cosmology, with the first major results in 1990 (the CMB spectrum) and in 1992 (the CMB anisotropy).

There is yet much to learn from the CMB, with better angular resolution on the fluctuations, with measurement of the polarization of the fluctuations, and potentially, with better measurements of the spectrum of the radiation. The spectrum can be measured significantly better at wavelengths from many centimeter down to less than 1 mm. Smooth distortions of the spectrum, mediated by hot electrons, are expected due to the eventual warming of the IGM. These smooth distortions could be measured one or two orders of magnitude better over the whole spectrum range. There may also be absorption or emission lines in the spectrum, due to atomic (hydrogen and helium) or molecular (LiH) processes. Even without major energy releases, these systems can be slightly out of equilibrium as the Universe cools through the cosmic decoupling at the age of 3.89×10^5 years. These narrow systems of atomic or molecular lines might be observable down to the parts per billion level if foreground emissions do not interfere.

Thanks a lot John. Indeed, current data set only upper limits on spectral distortions, in spite of the fact that they are predicted to exist. Their detection through a substantial progress in CMB spectrum measurements will represent an important step in fundamental cosmology.

An important problem in astronomy, and even more crucial in CMB experiments, is that of the choice of the observational sites, or, in the extreme case, the competition between ground and space observations. Although historical, it is continuously renewed by technological improvements. We discuss this point with Rafael Rebolo in the next interview.

5.3.2 The Future of CMB Experiments: Ground vs. Space?

Dear Rafael *(Rebolo)*, **what is your point of view about the different pros and cons of ground-based and of space experiments dedicated to CMB anisotropies in temperature and polarization?**

While space missions have obtained major achievements in the study of the CMB (COBE and WMAP), thanks to full sky coverage with high sensitivity and good control of systematics, ground-based and balloon-born experiments have also made extremely interesting contributions to the field. In particular, all the information we have on the CMB at angular scales below 12 arcmin comes from this kind of experiments. In the near future, *Planck* will be a major breakthrough, improving the present resolution, sensitivity, and frequency coverage of current full-sky maps in the microwave range. Mapping the microwave sky with a resolution better than 5 arcmin will then be a niche for future ground and balloon experiments.

The greatest challenge in CMB observations will be the measurement of polarization with sufficient sensitivity to detect the imprint of primordial gravitational waves. Unfortunately, theory cannot inform on the amplitude of the B-modes and only experiments can tell us the value of this crucial parameter. While a number of inflationary scenarios appear to predict values of the tensor to scalar ratio in the range 0.01–0.1, many other give values orders of magnitude below.

The ground-based experiments will explore the sky in the windows allowed by the atmosphere, mostly in the 10–30 GHz and in the range 90–100 GHz. Balloons with bolometer-based experiments will mainly cover the higher frequency domain from 100 to 300 GHz. Both types of experiments will be affected by different systematics and therefore detail comparison of data obtained for the same sky region will be essential. Given the limited sky area that ground- and balloon-based experiments can cover (typically less than 10^4 square degree) and the expected sensitivities, these experiments may be able in the next decade to tell whether the tensor to scalar ratio is higher than 0.01, but will face serious difficulties to measure a value for this parameter if significantly lower. The typical cost of these experiments will be in the range few to ten million euros. More than 10 such experiments are likely to be conducted during the next decade. They will be crucial as technology demonstrators for future instruments on board satellites, which will be designed to produce a full-sky coverage with high resolution and sensitivity. A space mission will cost several hundred million euros, but may be mandatory to push the detection limits below $r = 0.001$. The technology developments required for such an exciting enterprise are within reach, there is no obvious showstopper. We need to prove the new devices in the most advanced instruments we can build for ground and balloon experiments and then qualify them for use in space. It is likely that major breakthroughs will come again from space missions, but this will not be possible if we do not generate first a new suite of experiments on ground that demonstrate valid solutions that have been found improving technology limits. This new generation of experiments is being developed now and will see soon the microwave sky. We are fortunate to live in an epoch that has allowed a dramatic improvement in our possibilities to explore the origin of the Universe.

Thank you very much Rafael.
Among the various CMB projects of the next few years, the ESA Planck satellite is certainly the most ambitious. It can be considered as the definitive all-sky mission for total intensity anisotropies up to a resolution of about 10 arcmin and will provide a spectacular improvement also for polarization anisotropies, as Charles Lawrence will explain in the next interview.

5.3.3 *Planck*, A Forthcoming Space Mission

Dear Charles *(Lawrence)*, **after COBE and various excellent balloon-borne and ground-based experiments, and the spectacular results by WMAP, the** *Planck* **mission is the forthcoming most promising project dedicated to CMB anisotropy. Could you describe the fundamental experimental guidelines and technical solutions originally adopted for** *Planck*? **What are the most relevant improvements of** *Planck* **with respect to WMAP? Can you describe the most critical problems encountered during the development of this project and the solutions adopted to solve them?**

Planck is the third-generation space mission to measure the anisotropies of the CMB. Designed to extract essentially all the information contained in the temperature anisotropies, it will also improve dramatically our knowledge of the polarization anisotropies. To achieve this performance, *Planck* incorporates numerous technological innovations, including new detectors, coolers, and thermal design. *Planck* will refine and possibly change our understanding of the Universe, and will address a broad range of science from the solar system to the edge of the Universe.

The Origins of *Planck*

The CMB, first identified in 1964, is our most important source of information about the geometry and contents of the Universe. Two distinct kinds of measurements must be made: (1) the frequency spectrum of the radiation, that is, how its intensity varies with frequency; and (2) the variation in intensity of the radiation as a function of position on the sky, that is, the *anisotropy* of the CMB. The frequency spectrum of the CMB was measured definitively from 30 to 600 GHz by the COBE FIRAS instrument [56]; it is the most perfect BB spectrum ever observed in nature (The *Planck* mission is named after Max Planck, who first derived this spectrum of matter and radiation in equilibrium).

The first measurements of anisotropy were also made by COBE, with the Differential Microwave Radiometers (DMR) instrument [124]. COBE had the sensitivity and angular resolution to detect anisotropy only on large angular scales ($>10°$), but it provided powerful confirmation of the basic picture that structure in the Universe grew by gravitational instability from tiny fluctuations in the very early Universe. It also showed the angular resolution and sensitivity that would be required to mine all of the information from the CMB anisotropies, and confirmed the overwhelming scientific importance of making those measurements. The design of follow-on experiments began immediately. In Europe, this design activity led to *Planck*. In the US, it led to WMAP, which was launched in 2001.

Planck and WMAP are complementary. WMAP, with sensitivity and angular resolution intermediate between COBE and *Planck*, was launched first, and has measured roughly 10% of the information in the temperature anisotropies and 1% of the information in the polarization. *Planck* is much more ambitious, with the goal of

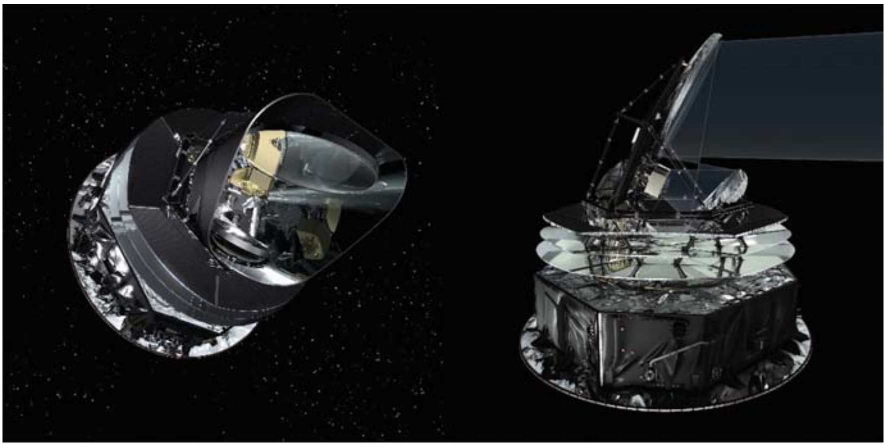

Fig. 5.4 Two views of a computer model of the *Planck* spacecraft. The spacecraft spins at 1 rounds per minute (rpm) on an axis normal to the solar panel on the circular end. The spin axis is always within a few degrees of the sun, maintaining full illumination of the solar panels, at the same time keeping everything else in the dark. This is the first key to the thermal design of *Planck*

measuring essentially all of the information in the temperature anisotropies; it will also measure about 10% of the information in the polarization. To do this, significant technical innovations were required.

Technical Innovations and Solutions Required for *Planck*

To extract essentially all the information in the temperature anisotropies of the CMB, *Planck* requires greater sensitivity and angular resolution than WMAP. To see why this is the case, consider first the current state of measurements. Figure 5.5 shows the angular power spectrum (C_ℓ vs. ℓ) of the temperature anisotropies as measured by WMAP. A best-fit cosmological model is superimposed. At about $\ell = 750$, the uncertainties start to increase rapidly, due to a combination of angular resolution (the WMAP beam at 94 GHz is about 0.25° Full Width Half Maximum (FWHM)) and noise. To go to higher multipoles (finer structure) requires both greater angular resolution and sensitivity. Figure 5.6 shows that a great deal of information is expected at the higher multipoles, and that *Planck* should be able to measure it. Figure 5.7 shows a similar comparison for polarization.

Wider frequency coverage is required to deal with confusing foreground radiation. When we measure the sky, we measure not only the CMB, but also every other source that radiates in the relevant frequency range. If we measure the sky

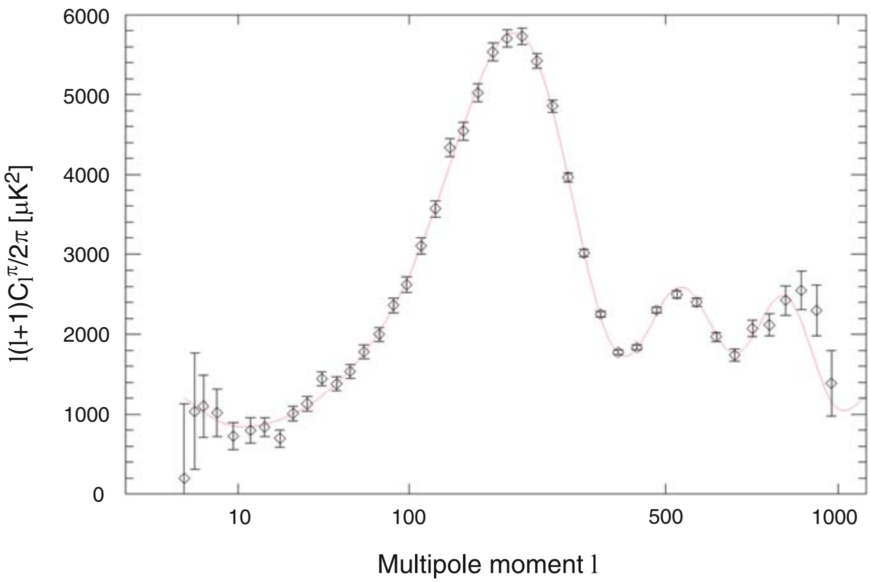

Fig. 5.5 Current data from WMAP [100]. The superimposed solid line is a best-fit model from [47] (Table 2) with ($\Omega_b h^2$, $\Omega_m h^2$, Δ^2_R, n_s, τ, H_0) = (0.0227, 0.131, 2.41, 0.961, 0.089, 72.4). WMAP can measure the first two and a half peaks in the angular power spectrum (APS) $\ell = 90$ corresponds to an angular scale of 2°

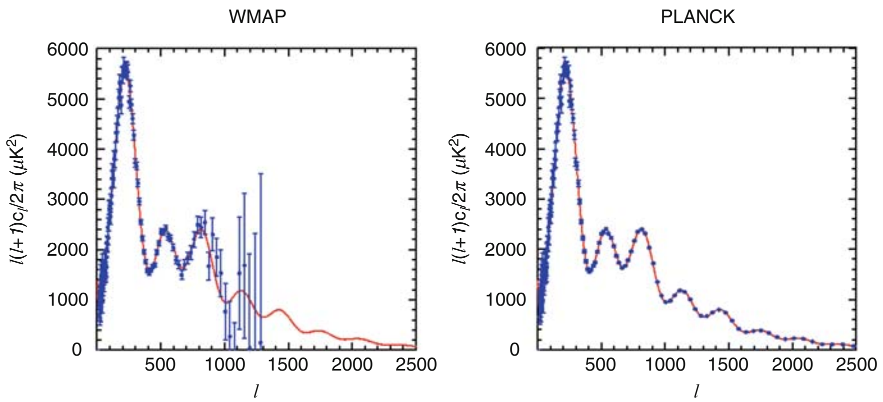

Fig. 5.6 Comparison of expected performance of WMAP and *Planck* for temperature anisotropies. As in Fig. 5.5, WMAP can measure the first two and a half peaks. *Planck* can measure seven or eight, well into the so-called "damping tail" of the CMB, where the signal becomes weak due to well-understood physical processes. There is little more information to be gained from these primary anisotropies, and that is the sense in which we can say that *Planck* will measure essentially all the information in the temperature anisotropies. Each multipole measured with signal-to-noise ratio greater than unity is an independent piece of information. The total number of multipoles goes as $\ell(\ell+1)$. WMAP measures $\ell \lesssim 750$. *Planck* should be able to measure $\ell \lesssim 2{,}500$. WMAP therefore measures only about 10% of the information contained in the anisotropies, whereas *Planck* measures it all. From [107]

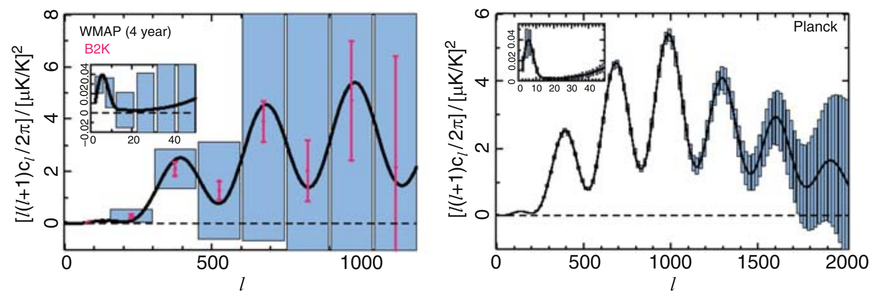

Fig. 5.7 Comparison of expected performance of WMAP and *Planck* for polarization anisotropies. From [107]

at enough different frequencies, we can separate the various foregrounds from the cosmic background. Figure 2.18 in Chap. 2 shows the spectrum of the CMB fluctuations compared to the spectra of Galactic synchrotron, free-free, and dust emission. The foreground minimum occurs at about 70 GHz. To separate the CMB from foregrounds, *Planck* must cover enough frequencies to get a good determination of both the synchrotron and free-free emission and the dust. WMAP's highest frequency was 94 GHz, not high enough to measure the dust.

To measure all the structure shown in the angular power spectrum in Fig. 5.6, *Planck* must have more than an order of magnitude greater sensitivity than WMAP, along with angular resolution three times greater, and frequency coverage that extends high enough to make a good measurement of the Galactic dust foreground. Achieving this sensitivity required significant innovation in detectors, coolers, and overall thermal design. The Low Frequency Instrument (LFI) (30, 44, and 70 GHz) uses amplifiers cooled to 20 K, while the High Frequency Instrument (HFI) (100, 143, 217, 353, 545, and 857 GHz) uses bolometers cooled to 0.1 K. The need for cryogenic detectors makes thermal design one of the most basic and important aspects of *Planck*.

There is one more fundamental design aspect of *Planck* that deserves mention. To reach the low overall noise levels in measuring the CMB required of *Planck*, multiple detectors must observe the sky simultaneously, with their output averaged together appropriately. The sky signals add up coherently; the noise, which is largely independent from detector to detector, averages down as the square root of the number of detectors, and so signal-to-noise ratio increases as the square root of the number of detectors. COBE and WMAP used a small number of detectors, which could be hooked up to measure differences in temperature between two separated spots on the sky directly. WMAP fed these detectors from two identical telescopes pointed about 140° away from each other, with a complicated system of waveguides to bring signals together for differencing. This has been the time-tested method of measuring the CMB. It is, however, impossible to scale such an arrangement to the large number of detectors required for *Planck*. As a result, *Planck* has a single telescope that scans the sky at 1 rpm, providing differencing in a different way than COBE or WMAP. Simulations show that this method will work very well.

Thermal Design

Planck will be the first astrophysics mission in space to achieve cryogenic temperatures without the use of stored cryogens (e.g., liquid helium). Four stages of cooling are used:

- Passive radiative cooling from multiple surfaces to cold space
- 20 K H_2 sorption cooler
- 4 K ^{4}He compressor and Joule–Thomson (JT) system
- 0.1 K ^{3}He^{4}He dilution cooler

Two of the coolers, the 20 K sorption cooler and the 0.1 K dilution cooler, represent the first of their kind to cool an instrument in space. The 20 K sorption cooler was developed and built at Jet Propulsion Laboratory (JPL) [150]. The 0.1 K open cycle dilution/JT cooler was been developed at the Centre de Recherches des Très Basses Temperatures (CRTBT) in Grenoble [14].

The key to the cryogenic system is the overall thermal design, in particular, careful isolation of cold from warm parts of the flight system and aggressive use of radiative cooling to do most of the job. The basic scheme is that one end of the flight system is warm, and the other end cold. Figure 5.4 shows two views of the flight system. The solar array covers the circular panel 4.2 m in diameter facing away in the images, which in flight is always nearly normal to the Sun, and operates at roughly the temperature of boiling water. The octagonal Service Vehicle Module (SVM) is near room temperature. Moving from solar panel towards telescope, the next section of the flight system comprises three large "V-grooves" thermally connected to low-conductivity struts that connect the telescope structure with the SVM. The top surface of the SVM and both surfaces of the bottom two V-grooves have low emissivity (i.e., they are mirrors). The top surface of the top V-groove has high emissivity (i.e., it is black). The telescope baffle, seen in the left view of Fig. 5.4 but removed from the right for visibility, is mirror-like on the inside and black on the outside.

The V-grooves provide excellent thermal isolation between the warm SVM and the cold telescope and focal plane unit, and at the same time, excellent coupling of radiation to cold space. They serve a dual role intercepting heat and radiating it to space. The temperature on top of the SVM is roughly 300 K, while the third V-groove will run below 50 K. At first glance, it may seem paradoxical that they can do these two things simultaneously and well. The reason is that the combination of mirror-like surfaces and the 7° or so angle between adjacent surfaces means that thermal radiation from the SVM has a very difficult time propagating to the telescope and its surroundings, but a very easy time being sent into cold space. They are so good at this that V-grooves will be nearly ubiquitous in future cryogenic missions such as James Webb Space Telescope (JWST). Although V-grooves have been used before in space, their use on *Planck* is the first on this scale, and at these temperatures.

The telescope baffle blocks far sidelobes of the telescope/feed system and provides a large, high emissivity surface for further radiative cooling. The telescope is expected to be below 40 K in space, colder than required for *Planck*, but providing added excellent margin against thermal emission from the telescope.

Detectors

Planck must cover a wide frequency range to deal with foreground emission. When *Planck* was designed, bolometers offered good sensitivity at high frequencies, but had not been built below 100 GHz, while amplifiers provided good sensitivity at low frequencies, but were too noisy at high frequencies. These facts dominated the choice of detectors and the technology developments needed.

For bolometers, significant issues that had to be addressed included sensitivity to ionizing radiation, sensitivity to microphonics, and polarization. The first two problems were solved by absorbing incoming radiation in a grid that was opaque to the relevant frequencies but essentially transparent to cosmic rays [18] (see Fig. 5.8). The grid is extremely stiff, with resonant frequencies in the tens of kilohertz, eliminating microphonic susceptibility. Polarization sensitivity was achieved by

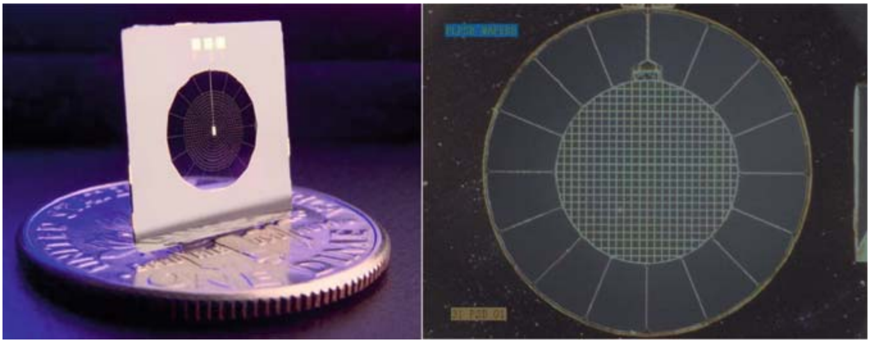

Fig. 5.8 Spider-web (*left*) and polarization sensitive (*right*) bolometers. The absorbing grid of the spider-web is opaque to the relevant frequencies, but transparent to rays. The resonant frequency of the structure is so high that microphonic susceptibility is extremely low. The PSB consists of two separate parallel meshes, oriented at right angles to each other, with a slight separation. Each one is sensitive to only one linear polarization

combining two parallel-mesh grids each with its own temperature sensor, mounted at right angles to each other, into one unit, the Polarization Sensitive Bolometer (PSB) [72]. *Planck* has spiderweb bolometers at 143, 217, 353, 545, and 857 GHz, and PSBs at 100, 143, 217, and 353 GHz. To achieve the necessary sensitivity, all bolometers must be at a physical temperature of 0.1 K.

For amplifiers, significant issues that had to be addressed were inherent gain fluctuations with an approximately $1/f$ spectrum [121], and power dissipation in the focal plane. The gain fluctuations were addressed by a pseudo-correlation radiometer design, in which signals from the sky and a BB reference load are combined by a hybrid coupler, amplified in two independent amplifier chains, and separated out by a second hybrid (Fig. 5.9). The sky and the reference load power are measured and differenced. The reference signal is subject to the same gain variations in the two amplifier chains as the sky signal, so when the difference is taken the fluctuations subtract out and the true sky power is recovered. To achieve the best elimination of gain fluctuations, the reference loads must be as close to the temperature of the sky as possible. This is achieved for *Planck* by mounting the loads themselves on the HFI, whose outer shell is maintained at a temperature under 5 K by the ^{4}He J-T cooler.

The *Planck* amplifiers are based on low-power InP transistors, but power dissipations are still 2–4 mW per transistor. The required heat lift determines the size of the sorption cooler, which has a big effect on the overall flight system. To minimize the dissipation in the focal plane, the amplification needed before the detector diodes was divided into two parts, one in the cryogenic focal plane assembly and the other on the room-temperature SVM. The two are connected with phase-preserving waveguides, which were a challenge to design and build, and must be carefully structured and mounted to avoid parasitic heat input to the focal plane assembly.

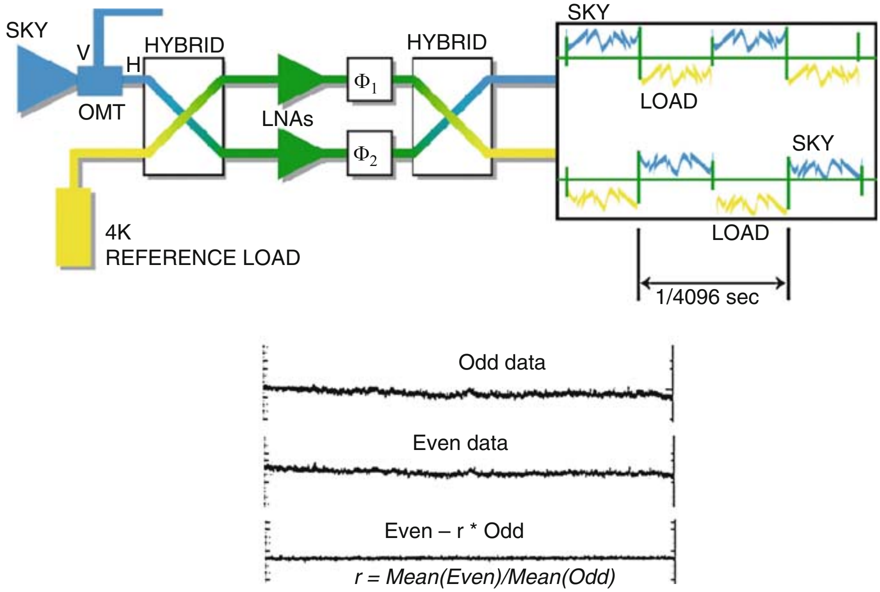

Fig. 5.9 (*Top*) Schematic of a *Planck* pseudo-correlation radiometer. (*Bottom*) Measured radiometer output. The signals from the two detector diodes ("odd" and "even" samples) correspond to the sky and reference load, in which the noise is dominated by a non-white, $1/f$-type component. The $1/f$ component is highly correlated in the two diodes, however, and strongly suppressed in the difference. From [107]

Coolers

To achieve their full low-noise potential, the *Planck* amplifiers must be cooled to $\leq$20 K, and the bolometers to 0.1 K. Such temperatures cannot be achieved with passive cooling alone. *Planck* has three mechanical coolers, two of which were developed specifically for *Planck*. A schematic of the cooling system is shown in Fig. 5.10. Figure 5.11 shows the fully redundant hydrogen sorption cooler system [150] being integrated into the SVM. It is easy to see from the photograph that the sorption cooler system is a major component of the flight system. It is harder to see all of the attachment points, thermal and mechanical, between various pieces of the sorption cooler system and the other parts of the flight system, but there are many. The cooler will produce a two-phase liquid–gas mixture that precools the HFI to 18 K, and cools the LFI to between 19 and 20 K.

The 4-K cooler uses a standard Stirling compressor developed originally by ESA in the 1990s for potential use in *Herschel*, which compresses ^{4}He mechanically and then sends it through a JT valve to produce $\sim$4.7 K for the HFI.

The 0.1 K ^{3}He–^{4}He dilution cooler [14] is the first such cooler designed for operation in space. One of the key differences between the space and ground versions is that in space it is not feasible to capture the ^{3}He and ^{4}He after they are mixed together and to separate and recirculate them. Instead, the gases are vented to space. The ^{3}He and ^{4}He gas supply is stored in high-pressure tanks in the SVM, precooled

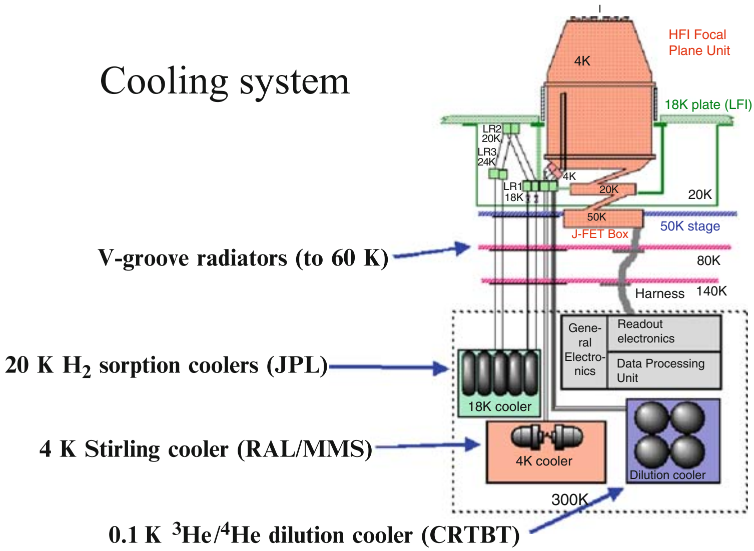

Fig. 5.10 The HFI cooling chain comprises the hydrogen sorption cooler providing 18 K, the closed-loop ^{4}He Joule–Thomson refrigerator providing 4 K, and a dilution refrigerator providing 0.1 K to the bolometers. From [107]

by the V-grooves, sorption cooler, 4 K cooler, and final JT expansions from 4.7 to 1.6 K (not shown in Fig. 5.10). This last precool stage gives the 1.6 K temperature required for the feeds and other passive optical components of the HFI. After that, the gases enter the dilution cooler system properly, producing 0.1 K to cool the bolometers, before being vented to space. The gas supply is expected to last for four complete surveys of the sky, or roughly 30 months including cruise to L_2 and checkout.

Two types of detector operating at different temperatures in the same focal plane, extensive passive cooling, plus three mechanical coolers all add up to quite a complicated thermal system. Mixing amplifiers and bolometers in the same focal plan is complicated; one might ask if it was necessary. It was true at the time of the design of *Planck*, and is still true today, that neither bolometer technology nor amplifier technology covers the entire frequency range required (Fig. 2.18 in Chap. 2). Thus mixing the detector technologies provides the best possible performance for *Planck*.

Cosmology and Astrophysics with *Planck*

The higher precision in the knowledge of CMB anisotropy achievable with *Planck* will allow to greatly refine the accuracy on a wide set of cosmological parameters. On the other hand, it will allow also to test more complex cosmological models, providing in this respect not only a quantitative but also a qualitative improvement in our comprehension of the Universe with respect to previous projects. Could you please discuss this aspect?

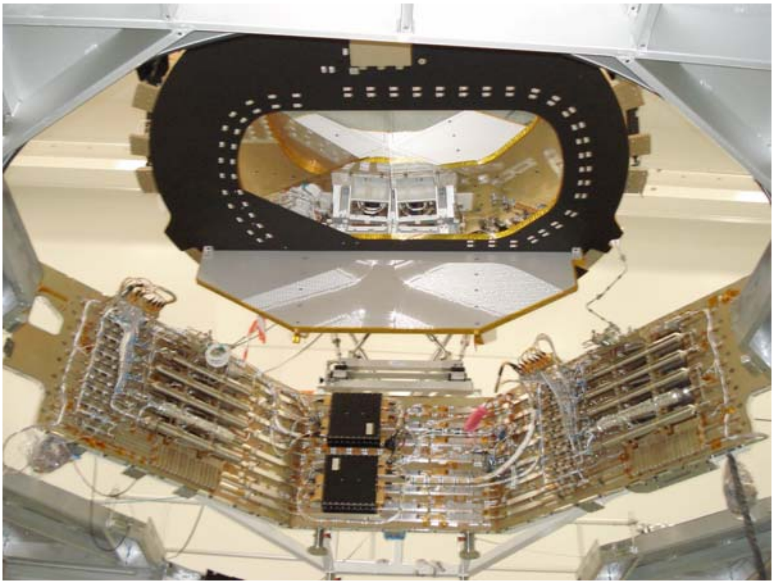

Fig. 5.11 20 K H_2 sorption cooler system, about to be integrated into the SVM. The cooler cools by a standard constant-enthalpy Joule–Thompson expansion of hydrogen through a capillary tube. The high and low pressures are provided by controlling the temperature of six beds containing a LaNiSn compound. At ∼260 K the pressure of H_2 is ∼1/3 atmosphere; at 465 K the pressure is ∼50 atmospheres. The system is fully redundant. The *left and right-hand panels* (which form two sides of the octagonal SVM) hold the compressor beds for the two independent coolers. The back side of all three panels shown above is covered with a radiator to dump heat from the compressor beds to space when they are cooled from 465 to 260 K. Heat pipes, which can be seen crossing into the *middle panel* in the photograph, distribute the thermal load. The cooler provides almost 1 W of heat lift for an input power of about 300 W at beginning of life. The high pressure supply gas and the low pressure return gas travel through a 1/8 arcs-inside-1/4 arcs tube-in-tube heat exchanger about 10 m long, which is heat sunk to all three of the V-grooves before reaching the focal plane unit. The tubing for the right-hand cooler can be seen adjacent to the top middle of the right-hand panel above

What is the chance of *Planck* to detect primordial gravitational waves through the discovery of the so-called B-mode polarization?

As shown in Figs. 5.6 and 5.7, *Planck* will improve dramatically our measurement of the temperature and polarization anisotropies of the CMB. Will it change our understanding of the Universe? In one way, this is an easy question to answer affirmatively. Observations made by CMB experiments, including WMAP, galaxy surveys such as 2-Degree Field (2dF) and Sloan Digital Sky Survey (SDSS), etc., are reproduced well by a relatively small number of parameters in an overall model including CDM and a nonzero value of Λ, the so-called concordance ΛCDM model (the reader could refer to a dedicated web page[2] for a contemporary compilation).

[2] See Legacy Archive for Microwave Background Data Analysis (LAMBDA) in web page list.

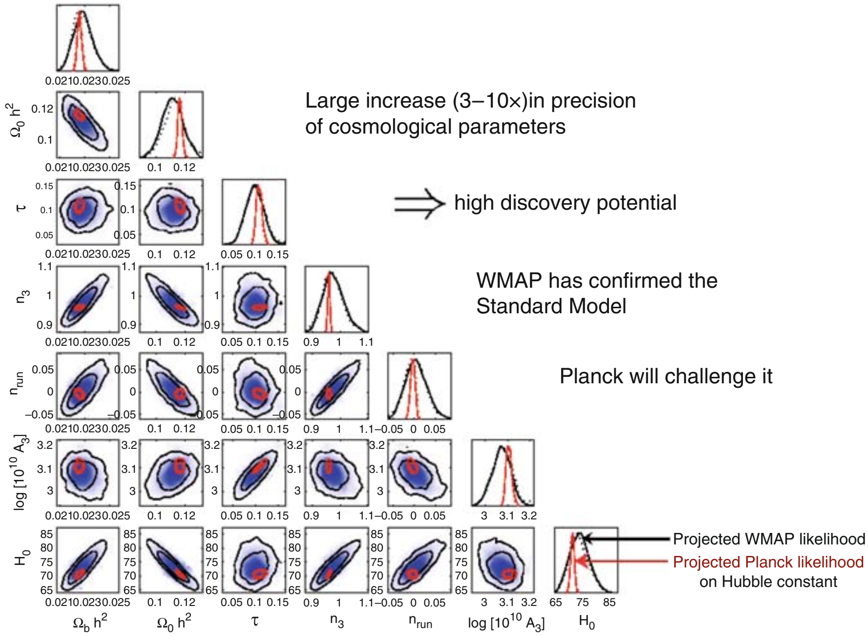

Fig. 5.12 Expected improvement in the uncertainties of cosmological parameters in a ΛCDM model due to *Planck*. With such a great improvement comes not only the certainty of refining our understanding of the Universe, but also significant discovery potential. It is not possible to predict the direction of such discoveries (that is the nature of discovery!), but it would be surprising if there were none. From [107]

In this overall scheme, we can predict with confidence that *Planck* will improve the uncertainties in these parameters by factors of roughly between 3 and 10 (Fig. 5.12). For example, our current knowledge about the distance to the so-called surface of last scattering (i.e., the shell that appears as the source of the CMB, and which sets the size of the observable Universe at $z = 1{,}100$) is limited by our knowledge of the average density of matter in the Universe, which has an uncertainty of about 6%. *Planck* will reduce this uncertainty in distance from about 3% to about 0.2%. It is amazing to think that we will know something as vast as the size of the observable Universe (in a very well-defined sense) to one part in 500, far better than we know most measured quantities in astronomy! In fact, after *Planck* we will know the properties of the Universe at a redshift of 1,100 better than we know the properties of the Universe in our immediate vicinity (at $z = 0$).

Planck's frequency bands were chosen partly to optimize detection of galaxy clusters through the Sunyaev–Zel'dovich effect, and we expect *Planck* to find thousands of clusters over the whole sky. Unlike almost any other signal in astrophysics, the strength of the Sunyaev–Zel'dovich Effect (SZE) is independent of redshift. Therefore, the *Planck* cluster sample will be rich in high redshift clusters, which are particularly important for cosmological studies.

One of the ways that *Planck* will influence fundamental physics is by measuring the masses of neutrinos. We know that neutrinos have mass from measurements of Solar neutrinos and cosmic rays. *Planck*'s precise measurement of the matter power spectrum on angular scales where neutrino masses make a difference will give new and much tighter constraints on their mass.

An exciting possibility is that *Planck* could detect evidence of a primordial gravitational wave background. This would show up as a specific pattern of polarization on large angular scales. The existence of such a background is a strong prediction of inflationary theories, although its amplitude could vary over an enormous range depending on the details, and so there is no guarantee that *Planck* will detect it. Detection of such a background, however, would provide a direct confirmation of inflation, and a measurement of its energy scale. Understanding what happened 10^{-35} s after the Big Bang by measuring photons that have traveled for 13 billion years would be a dramatic achievement indeed.

Another exciting possibility arises from *Planck*'s ability to measure non-Gaussianity in the primordial fluctuations. Non-Gaussianity is a higher-order statistical property of the fluctuations that depends on the fundamental physical processes of the very early Universe, and may have great power to distinguish between various classes of models. Recent work may have detected non-Gaussianity in the WMAP 3 year data [152]. *Planck* will be an order of magnitude more sensitive to non-Gaussianity than WMAP.

So clearly, *Planck* will refine our understanding of the Universe in many ways. Will it change it dramatically? Of course, I don't know – predicting the future is an uncertain business! Nevertheless, the dramatic increase in measurement capability of *Planck* has the potential to reveal new things, challenging rather than merely confirming the standard model. And despite the phenomenal improvement in our cosmological knowledge over the last few decades, we do not have the faintest idea why the Universe is the way it is! Theoretical estimates of the value of Λ are wrong by 120 orders of magnitude. We know by its gravitational effect that one quarter of the mass-energy in the Universe is in some form on non-electromagnetically interacting DM, but we do not know what it is. Inflation has been an enormously useful concept for which there is no equally satisfactory substitute, but we have never observed a scalar field, the underlying physical mechanism for inflation. There is certainly reason to hope that *Planck* will provide powerful new clues.

***Planck* will realize an all-sky survey of astrophysical sources at millimeter and sub-millimeter wavelengths. What kind of astrophysical information provided by *Planck* will be more relevant for cosmology, complementing the information contained in the CMB?**

Planck will produce the most sensitive all-sky surveys ever made at its nine frequencies. At frequencies above 100 GHz, it will make the *first* all-sky surveys sensitive enough to detect discrete sources. *Planck* will detect many types of objects, including asteroids, star forming regions in the Milky Way, dusty galaxies, radio galaxies, and quasars. At the lower frequencies, hundreds to thousands of sources will be detected. At the higher frequencies, tens of thousands of sources will be detected.

A broad range of science investigations will result. Among these will be studies of transient radio sources, cold cores, the cosmic infrared background, and the 3D structure of the Galactic magnetic field. *Planck*, the ultimate cosmology mission, will thus play an important role in Solar system, Galactic, and extragalactic science as well.

Thanks a lot Charles. Of course, these astrophysical signals, from the microwaves to far-infrared, that Planck *will allow us to accurately study are sources of contamination for CMB observations and need to be separated through dedicated algorithms, as discussed in Chap. 3. In the next interview, Juan Francisco Macias-Perez will report on other data recently accumulated or expected to be taken in the near future that will contribute to the separation of the various microwave emission components, with attention to the dusty side of foreground emission.*

5.3.4 Surveys to Map Dust Foreground Emission

Dear Juan *(Francisco Macias-Perez)***, what kind of new surveys are needed and foreseen for the future to precisely map the sub-millimeter and far-IR emission and extinction from dust grains in the IGM of our Galaxy, a significant source of contamination for cosmological observations?**

All the models of dust extinction and thermal foreground emission are based on the FIRAS and Infrared Astronomical Satellite (IRAS) data. To improve them, we need new observations at infrared wavelengths with better signal-to-noise ratio and higher resolution. The AKARI satellite (see [135] for details), which just finished observations, has been designed to fulfill these objectives. AKARI has observed at four frequencies bands between 50 and 185 μm covering the peak of the dust thermal emission. The analysis of the data is in process although the first AKARI full sky maps can be observed in [135]. With no doubt, the AKARI data will improve the local estimates of dust extinction because of its greater resolution. Furthermore, the better signal-to-noise will probably allows us to get a more physical picture of the dust thermal emission by refining the two component models of [54]. AKARI will be complemented by the *Herschel* satellite, (see for details [136]), to be launched in 2009. *Herschel* will observe in the range 60–675 μm permitting both photometry and spectroscopy. *Herschel* is a high resolution instrument but will not cover the full sky. From the foreground point of view, *Herschel* will be useful to study on particular regions the variation of the dust spectral index with frequency.

To significantly improve on the estimation of the dust foreground emission, we really need dedicated surveys at millimeter and radio wavelengths. This is the case of the *Planck* satellite experiment (see [107] for details), which will be launched in 2009. *Planck* is designed to measure the CMB temperature and polarization anisotropies with unprecedented accuracy. For this purpose, foreground emission removal is a very important issue and therefore *Planck* consists of seven polarization frequency bands from 30 to 353 GHz and two unpolarized ones at 545 and

857 GHz. The *Planck* resolution varies from 5 arcmin at high frequency to 33 arcmin at low frequency. The total intensity dust emission will be accurately monitored by the high frequency channels and accurate subtraction should be possible even on the lower dust emission regions. In particular, a detailed study of the dust spectral index will be possible at relative high resolution and therefore we might be able to discriminate between different populations of dust grains. However, as *Planck* does not have very low frequency channels (from 10 to 20 GHz), the physical modeling and subtraction of spinning dust remain a challenge. To help on this issue, the Q, U, I Joint Tenerife (QUIJOTE) CMB experiment [138] is currently designed and constructed and is expected to be operational for the *Planck* flight. QUIJOTE will have four polarized frequency bands from 11 to 30 GHz and 1° resolution. From the polarization point of view, the *Planck* data should significantly improve our current knowledge on vibrational dust polarization emission at low and intermediate Galactic latitudes. At very high Galactic latitudes and for the cleanest areas of the sky, if the results from the Balloon Observations Of Millimetric Extragalactic Radiation and Geophysics (BOOMERanG) (see [92] for details) are confirmed, a detailed study will be very challenging even at 353 GHz. Furthermore, in the analysis of the CMB B polarization modes, associated to primordial gravitational waves, a very accurate polarization foreground removal will be needed. The residuals need to be at the noise level as the signal may well be at that level or below. Notice that a better foreground cleaning increases the area of the sky available for CMB analysis and therefore improves the signal-to-noise ratio. The removal of the dust polarization emission will be crucial for this analysis as the high frequency data are much more sensitive.

Thank you very much Juan.
Increasing efforts from the CMB community have been dedicated to the study of polarization anisotropies and of their correlation with those of total intensity. The accurate study of the weak B mode polarization anisotropy, of extreme relevance for cosmology and fundamental physics, requires capabilities well beyond Planck in terms of sensitivity and control of systematics and foregrounds. The next interview with Martin Bucher will clarify these aspects.

5.3.5 Beyond Planck

Dear Martin *(Bucher)*, **recent CMB experiments have successfully mapped the CMB temperature anisotropy and detected the CMB E-mode polarization anisotropy.** *Planck* **will study the T and E anisotropies with even greater accuracy and also has a chance of detecting a primordial B-mode polarization anisotropy if we are lucky and the relative amplitude of primordial tensor perturbation lies only slightly below the present observational constraints. However, for many models of inflation, this signal is predicted to be too low for** *Planck*, **calling for a new generation of dedicated experiments in the context of next calls of the ESA Cosmic Vision program or the Beyond Einstein**

program of National Aeronautics and Space Administration (NASA). Could you discuss the cosmological relevance of such projects? Why are they so important also for particle physics? Is the physical information contained in such kind of observation really unique?

The search for a primordial B-mode polarization anisotropy in the CMB constitutes an extremely important and unique opportunity to probe inflation, and by extension fundamental physics near the Planck scale. Inflationary cosmology predicts the generation of two kinds of primordial perturbations. First there are the scalar perturbations, corresponding to quantum fluctuations of the inflaton field initially present in the incoming vacuum that become stretched and frozen in during the postulated epoch of inflationary expansion [11, 62, 66, 85, 98, 116, 129]. These are the perturbations manifested in the beautiful CMB maps provided by COBE [124], BOOMERanG [38], WMAP [13] and their progeny. As far as we know now, there is no primordial signal in these maps attributable to nonscalar perturbations, although the admissible fraction of tensor perturbations consistent with the present data is embarrassingly large ($r < 0.22$ at 95% confidence according to WMAP 5 year analysis [79]).

Inflation also predicts a spectrum of primordial gravity waves, generated in much the same way as the scalar perturbations, except that it is the vacuum quantum fluctuations of a tensor field, the linearized fluctuations of the spacetime metric, rather than those of a scalar field, that become stretched to cosmological scales and re-enter the horizon at late times [2, 128]. Without inflation there is no reason why there should be any a priori connection between the scalar and tensor perturbations. Nor is there any reason to believe that the tensor perturbations should have a very red primordial spectrum (characterized by an inordinate amount of power on the largest accessible scales) that in the absence of inflation would suggest an "acausal" mechanism for their generation. Such gravity waves have not yet been observed, and their discovery would constitute a satisfying confirmation of a truly unique prediction of inflation. The extreme red nature of their spectrum would make it extremely difficult to conjure up alternative explanations using more compact astrophysical gravitational wave sources, which would because of the absence of large-scale spatial correlations yield a white noise spectrum on cosmological scales probed by B-modes of the CMB. Whenever there are no long-range correlations in a spatial stochastic process, its power spectrum falls off as $P(k) \to$ (constant) as $k \to 0$, if not faster.

The observation of the B-mode polarization of the CMB anisotropy directly measures gravity waves predicted to have been generated from inflation. In the linear theory, it is impossible to generate such a polarization pattern by means of a scalar degree of freedom. The argument for the absence of B-modes at linear order from any model describable solely by means of a scalar degree of freedom is very robust, because it relies on symmetry properties and not on trusting complicated simulations for the predicted value of a nonzero quantity. In practice, things are not quite so simple. At higher order, a gravitational lensing contaminant signal arises with an essentially white noise spectrum on large angular scales having an amplitude of approximately 5 μ Karcmin. The signal, however, is well understood and can be calculated precisely. There is also the formidable problem of foregrounds. Until now, foregrounds have not posed a serious problem for CMB analysis. People have worried a lot about foregrounds and rightly so. But Nature has been kind and the central frequency bands have proved to be relatively free of foregrounds. For searching for B modes, this all is likely to change. More aggressive foreground subtraction will be necessary. We do not know a lot about polarized foregrounds, particularly about polarized dust emission, so it is difficult to say how kind Nature will be for B-mode hunters. Another challenge associated with searching for B modes is the shape of the spectrum. While the CMB scalar anisotropies have a red spectrum at low ℓ, the B-mode anisotropy has a spectrum quite close to that of white noise, roughly up to the scale of the horizon at last scattering. Consequently, from an instrumentalist's point of view, the greatest challenges lie at low-ℓ, where exquisite control of systematics is necessary. This situation is completely unlike that for the temperature anisotropy, where the low-ℓ signal in the sky is large and most of the difficulties lie at large ℓ.

Figure 5.13 illustrates the various anisotropies predicted in the standard inflationary model, assuming the cosmological parameters taken from the WMAP best-fit model. The green curves indicate the various power spectra for the scalar mode. On

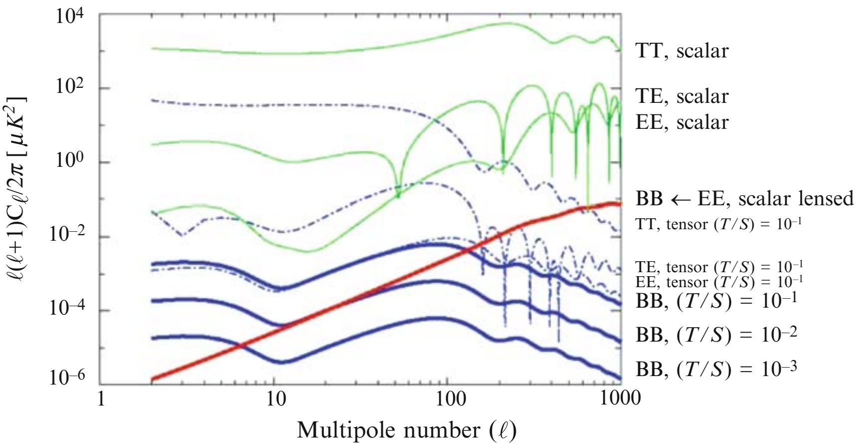

Fig. 5.13 Inflationary prediction for the CMB temperature and polarization anisotropies for the scalar and tensor modes. The *horizontal axis* indicates the multipole number ℓ and the *vertical axis* indicates $\ell(\ell+1)C_\ell^{AB}/(2\pi)$ in units of $(\mu K)^2$, which is roughly equivalent to the derivative of the power spectrum with respect to $\ln \ell$. The *green curves* indicate the TT, TE, and EE power spectra (*from top to bottom*) generated by the *scalar* mode assuming the parameters from the best-fit model from WMAP 3 year data ($H_0 = 70.4$, $\Omega_b = 0.044$, $\Omega_{cdm} = 0.22$, $\Omega_\Lambda = 0.73$, $\tau = 0.073$, and $n_s = 0.95$). The BB scalar component (indicated by the heavy *red curve*) results from the gravitational lensing of the E-polarized CMB anisotropy at the last scattering surface $z \approx 1{,}100$ by structures situated mainly around redshift $z \approx 2$. The *top* four *blue curves* (*from top to bottom on the left*) indicate the TT, TE, BB, and EE spectra (BB is the heavy *solid curve*) resulting from the *tensor* mode assuming a scale-invariant ($n_T = 0$) primordial spectrum and a tensor-to-scalar ratio (T/S) of 0.1. This value is situated a factor two below the upper limit established by WMAP. The *bottom* two *blue curves* indicate the tensor BB spectrum for (T/S) equal to 0.01 and 0.001, respectively. For the TE cross-correlations we have plotted the log of the absolute value, hence the downward spikes corresponding to sign changes. From [39]

the angular scales shown, WMAP provided primarily cosmic variance limited measurement of the TT anisotropy, a fairly good characterization of the TE anisotropy, and a very noisy characterization of the EE anisotropy. On these scales, *Planck* will provide a characterization of the E polarization limited principally by cosmic variance. The blue curves show the CMB anisotropy predictions for the tensor mode. The upper dashed blue curves show the TT, TE, and EE tensor anisotropies for the tensor-to-scalar ratio set to $(T/S) = 0.1$. For lower values of (T/S) these curves would slide downward by a corresponding amount. If (T/S) is low, the TT, TE, and EE anisotropies are not particularly useful for detecting B-modes, because the scalar and tensor power spectra add in quadrature to give the total power spectrum. Consequently, B-modes can be detected only by analyzing the shape, which is likewise affected by many other parameters. The BB power spectrum, however, is very promising because, as previously mentioned, the scalar prediction is zero at linear order. On the plot, we show the predictions for $(T/S) = 0.1$, 0.01, and 0.001. The red curve shows the predicted contamination from gravitational lensing of the E-mode into B-mode by clustered matter situated between us and the last scattering surface.

There are two strategies for detecting B-modes. First one may choose to live with the lensing contamination, treating it as a well-characterized noise in much the same way that one deals with detector noise. Alternatively, one may endeavor to subtract the lensing contaminant. For the former strategy, the ideal experiment would have a detector noise level (limited primarily by photon counting statistics) roughly equal to[3] $5\,\mu$K arcmin, the level of the lensing noise. Achieving a lower instrument noise level would be superfluous, because the instrument noise adds in quadrature to the lensing noise. A rather coarse angular resolution suffices, because almost all the useful signal lies at $\ell \lesssim 100$, as illustrated in Fig. 5.14. Consequently, for a space mission, where the bulky optics contribute substantially to the cost, no very large mirror or lens is required.

A more extravagant strategy would be to "clean" the map of its lensing contaminant. Strategies for doing this have been proposed by Okamoto and Hu [101] and by Hirata and Seljak [67]. Such cleaning effectively relies on reconstructing the lensing potential and making a map of the E-mode extending out to very small angular scales. Once the harmonic expansion coefficients $a^{E}_{\ell m}$ and $a^{\Phi}_{\ell m}$ (where Φ is the projected gravitational lensing potential) are known, the B-mode coefficient due to lensing $a^{\mathrm{B,lensing}}_{\ell m}$ can be calculated as a sort of convolution and subtracted at the map level rather than at the power spectrum level. This procedure in principle removes the limitation illustrated by the red curve in Fig. 5.13. Unfortunately, carrying out this strategy requires about an order of magnitude better angular resolution and almost as much additional sensitivity compared to a survey that simply accepts the presence of the B-mode lensing as a noise component. A mission that would provide the data necessary for cleaning lies beyond our present capabilities. Moreover, for such cleaning to be successful, exquisite foreground removal is re-

[3] This means that an arcminute square pixel would have a $5\,\mu$K rms polarized temperature variation. For other pixel sizes, the rms temperature varies inversely with the square root of the area, hence this kind of units for characterizing the strength of the white noise background.

quired. Consequently, it seems more prudent to design an experiment accepting the lensing as noise. The $c_\ell^{\mathrm{BB,lensing}}$ spectrum is both calculable theoretically and measurable directly.

There are two windows for detecting B-modes, one located at the reionization bump and another at intermediate angular scales centered about $\ell \approx 50$. On very large angular scales, the B-mode signal is greatly enhanced relative to its approximately white noise spectrum on intermediate angular scales. This bump (at $\ell \lesssim 10$) is known as the "reionization bump". The physical origin of the reionization bump is easy to understand by means of the following heuristic argument. The order of magnitude of the CMB polarization anisotropy may be estimated as $d^2(\partial^2 T/\partial x^2)$, where d is the co-moving distance between last and next-to-last scattering and $(\partial^2 T/\partial x^2)$ is the second-derivative of the temperature anisotropy. In reality, this second-derivative is a smooth kernel of a quadrupole shape whose support has a nonzero width d. For those photons emanating from the last scattering surface, one has a contribution approximately equal to $(1-\tau)d_{\mathrm{ls}}^2(\partial^2 T/\partial x^2)$, where τ is the reionization optical depth. For photons scattered off electrons arising from reionization, the contribution is approximately $\tau d_{\mathrm{reion}}^2(\partial^2 T/\partial x^2)$. Despite the fact that τ is small, $\tau d_{\mathrm{reion}}^2 \gg (1-\tau)d_{\mathrm{ls}}^2$. Consequently, for angular separations larger than or comparable to the angle subtended by d_{reion} on the celestial sphere, the contribution from reionization is dominant by approximately two orders of magnitude. However, on smaller scales, the contribution from reionization is washed out by the finite width of the kernel. The signal at the reionization bump is probably only accessible from space, because its measurement requires full-sky coverage, with exquisite control of systematic errors affecting points far apart on the celestial sphere.

The second window for detecting B-modes is centered around $\ell \approx 50$. Figure 5.14 shows on which angular scales the signal is situated. Ninety percent of the signal is situated at $\ell \lesssim 100$. However, to make a convincing detection that does not rely on an absolute calibration to measure c_ℓ^{BB}, one must also observe in the same survey the downturn in the tensor B mode spectrum, by passing into the regime dominated by the lensing signal, and this requires extending the angular coverage to about $\ell \approx 100$.

There are currently several ground-based efforts underway endeavoring to detect B modes, and space missions with the same goal have been proposed both in Europe[4] and in the United States. The necessary increase in sensitivity is attained by massively increasing the number of detectors. This is because the fundamental limitation on the sensitivity arises from photon counting statistics. Therefore, better sensitivity can be attained only by collecting more photons. If one uses single-mode detectors, which have the advantage of forming a clean beam with well-defined profile and polarization properties, the sensitivity is directly proportional to the number of detectors. Consequently, there is great interest in developing bolometer arrays, allowing one to envisage $O(10^4)$ or more detectors, rather than a less compact technology with one horn per bolometer. Nevertheless, horns present several advantages, because they form a beam with rapidly falling off side-lobes and good cross polar-

[4] See B-Polarization satellite mission (B-Pol) in web page list.

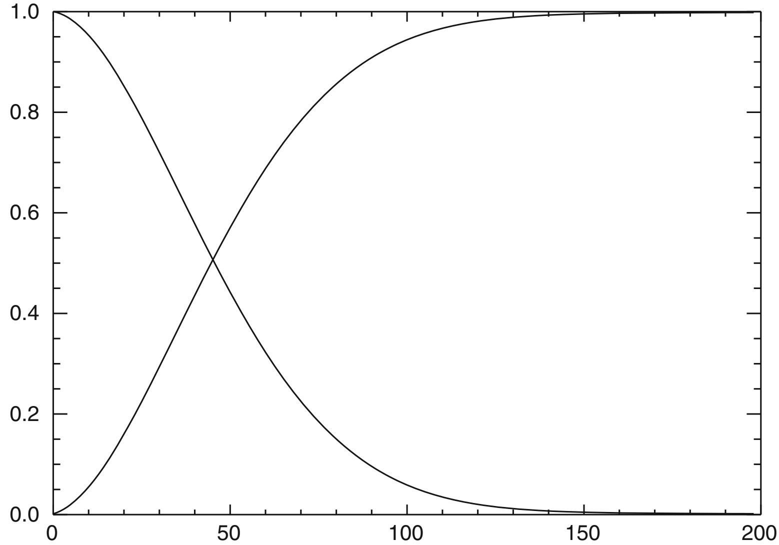

Fig. 5.14 Where does the information lie? We show the shape of the cumulative χ^2 for a B-mode detection using the intermediate window ($\ell \approx 50$) taking into account only noise from gravitational lensing of the primordial E mode into the B mode. In this plot, instrument noise has been neglected and the reionization bump has been artificially removed. The cumulative normalized χ^2 per multipole number, which is proportional to the integral of $\frac{1}{2}(2\ell + 1)(c_\ell^{\mathrm{B,tensor}}/c_\ell^{\mathrm{B,lensing}})^2$, is shown integrated for both descending and ascending ℓ

ization properties. There is also great interest in developing modulation techniques for measuring polarization more directly and rejecting parasite polarization generated within the instrument.

To my mind, the search for B-modes should be given a very high priority in view of the stakes involved. This observation offers a unique new probe into the presumed epoch of inflationary expansion. A measurement of the primordial B-mode would allow us to measure the Hubble parameter during inflation and thus directly establish the energy scale of inflation. It would, moreover, allow us to establish that inflation actually happened. Such a discovery would also have profound implications for particle theory because it would tell us something direct about physics near the Planck scale, about which there is thriving speculation but no data, except of the most indirect sort. Critics rightfully point out possible obstacles from foregrounds and the fact that models of inflation can be concocted having low levels of gravity waves. They are right that success cannot be guaranteed, but I think the consequences of such a discovery are so profound that the risk is well worth taking. When the *Planck* data comes in, we will have a much better understanding of the foreground situation. In the meantime, ground-based experiments will improve and demonstrate the technology and may make a detection if we are lucky. In any case, we will want to go to space, either to confirm a ground-based detection or to push further the constraints that have been established from the ground. In deciding on a strategy, it is important to consider technology development. A space mission planned now would necessarily be based on technology proved to work today. There are many opportunities to improve on the present technology, for example, by developing and demonstrating modulation strategies to measure the polarization directly rather than as a difference and by developing bolometer arrays to beat photon counting statistics. When evaluating experiments, one should look beyond promises of what values of the tensor-to-scalar ratio T/S can be reached and consider the uncertainties of each experiment and how technology development would be furthered.

Thanks a lot Martin.
While the mapping of CMB polarization anisotropies at the extreme accuracy level necessary to achieve the crucial detection of primordial B mode could require a significant technological improvement, substantial progress in cosmology could come on shorter timescales from future optical and infrared surveys. John Peacock will open this discussion in the next interview.

5.4 Perspectives from New Galaxy Surveys

Dear John *(Peacock)***, which kind of surveys should be planned for future significant improvement of present day observational cosmology?**

What are we trying to achieve when we survey the Universe? Probably we should distinguish fundamental aims from astrophysical ones. The former means attempting to measure some simple physical issue that impacts on the global nature of the Universe (the overall matter density, the neutrino mass, the vacuum density, etc.); the latter means in effect understanding the whole process of galaxy formation. In the past, serendipity would also have been a strong motivation: discoveries such as voids in the galaxy distribution were largely unexpected. Probably we have lower expectations these days: the standard ΛCDM framework works so well that we are less disposed to look outside its framework. This could be a big mistake, and it certainly makes the subject less exciting. Fortunately, the datasets we need for the big fundamental questions are sufficiently large and general that there will be plenty of scope for finding the unexpected.

The big driver for future surveys is geometrical: we want to take forward the great successes of the ΛCDM model in defining physical standard rods based on the horizon size at key epochs such as matter–radiation equality. Observing these rods in the CMB (WMAP etc.) and galaxy distribution (2-Degree Field Galaxy Redshift Survey (2dFGRS), SDSS) has told us profound facts: the total density of the Universe is within 1% of critical, and 75% of this is in the form of vacuum energy. Is it worth pushing this further? Who cares if, for example, the Hubble constant is 72 or 73? That in itself seems of no great importance, but (at least) two big fundamental questions remain: (1) do part of the CMB anisotropies originate from a gravity-wave background? (2) is the vacuum energy a cosmological constant, or does it vary with time? Both these questions can be addressed by extending the geometrical approach that has served us so well, mapping the CMB and the galaxy distribution to the limits imposed by cosmic variance.

Over the past few years, the cosmological community worldwide has debated at length the way in which this programme can be carried forward, and the answers are pretty clear. Comprehensive summaries of the issues can be found in the reports by the Dark Energy Task Force [3] and the ESA–ESO (European Southern Observatory) Working Group on Fundamental Cosmology [103]. The key to progress in primordial gravity waves is CMB polarization, trying to measure the curl-like "B-mode" polarization patterns produced by tensor modes. Tensor modes tend to be important only on large scales, but cosmic variance can prevent us attaining the necessary accuracy; this leads to a preferred multipole number $\ell \simeq 100$, or scales of a few degrees. Detection of this signal would be a historic event, and there are certainly some grounds for optimism. Probably the greatest recent advance in cosmology has been the rejection of the simplest model of pure scalar fluctuations with a scale-invariant $n = 1$ spectrum. The data are a significantly better match to the simplest possible inflation model: a single scalar field with a mass-like potential, which predicts $n < 1$ but also a rather healthy tensor fraction. This model should be tested within the next 5 years or so, assuming the foreground emission from the Milky Way can be understood.

The CMB data will be complemented by information from large-scale imaging and spectroscopic surveys. Here, the main techniques under discussion are gravitational lensing (correlated image distortions from light rays traveling through foreground potential perturbations) and further exploitation of the galaxy power spectrum, especially through the BAO imprinted by the photon–baryon component. These geometrical probes give us the extra information needed to measure the properties of DE more accurately. Doing this will not be easy, as the next steps require very high accuracy. The initial target is to measure the equation of state of the DE, $w \equiv P/\rho c^2$, and see if it differs from the $w = -1$ of a cosmological constant. Following any such event, the big debate will be whether we are witnessing a breakdown of GR, or something like scalar-field dynamics. But first we have to disprove the simple Λ model. Current data yield roughly $w = -1 \pm 0.06$, so we should be aiming for percent-level precision. The trouble is, however, that the central quantity involved (the distance–redshift relation) responds only weakly to changes in w, so we need to aim for a fractional precision in geometrical quantities of 0.001 or better. With massive surveys, there is no issue with achieving such precision statistically, and the real question is how hard we have to work to beat systematics; projects under discussion plan to image gravitational lensing over essentially the full extragalactic sky (2×10^4 square degree), and to measure 10^8 redshifts. It is remarkable that the current goals in cosmology require us to survey most of the visible Universe: 20 years ago, the idea of doing this would have seemed beyond fantasy.

Thank you very much John. Let us continue with a comment by Charles Bennett on future surveys from space.

Dear Charles *(Bennett)***, what is your opinion on new ideas and projects of surveys dedicated to DE and BAO in the context of next cosmic visions of space agencies?**

The WMAP data confirm the existence of DE and aid in placing new constraints on its nature. One possibility for the DE, Albert Einstein's "cosmological constant", is fully consistent with the WMAP and other cosmological data. In this scenario, the Universe will expand forever. However, if the DE is due to some other physical effect, the conclusion about the fate of the Universe could change. The BAO, calibrated by the CMB, can be measured as a function of redshift. This provides the expansion rate as a function of redshift $H(z)$, which will be of great importance in tracing the development of DE. The physics of this approach is fully understood and the systematic errors are small.

The luminosity distance derived from type Ia SNe is another approach for the measurement of $H(z)$. This approach assumes that all type Ia SNe are scalable to a standard light curve, which is not known. The physics of the SNe explosions themselves are not yet understood.

Finally, weak gravitational lensing provides a technique for determining the three-dimensional mass distribution of the Universe, and also bears on the nature of the DE.

I lead the NASA/DOE Advanced Dark Energy Physics Telescope (ADEPT) mission study. ADEPT primarily uses BAO to probe the effects of the DE. ADEPT also uses this information to check its own large sample of high-redshift SNe.

Among the many potential missions in ESA's Cosmic Visions plan is a mission to probe the dark Universe. The Dark UNiverse Explorer (DUNE) candidate mission, studied for many years, observes SNe and weak lensing, while the new Spectroscopic All-sky Cosmic Explorer (SPACE) mission candidate probes the BAO. It is unclear, as I write, what direction the ESA plan will take and what relationship it will have to the NASA and DOE plans.

It is advantageous to maximize the sky area for all DE experiments and missions. As ground-based efforts can probe up to almost a redshift of $z = 1$, a space mission should cover where the ground leaves off, at $z > 1$. That is, a DE space mission should be a wide-field infrared space mission.

It is remarkable that humankind has been able to discern the large-scale properties of the Universe as well as has been accomplished to date. I eagerly look forward to a new generation of exciting new cosmological results.

Thanks a lot Charles.
As shown by experience, advanced space satellites typically require about 10 years (or more) to be completed and arrive to the launch and data acquisition phase. Then, in a few years, they provide data of extreme quality and of decisive impact. On the contrary, ground-based programs, although marked by well defined steps in their progress, typically evolve more continuously and allow to accumulate data over very different timescales. In the next interview, Massimo Capaccioli will start the presentation of promising projects and results achievable in this field in about 5 years.

5.4.1 *Ground-based Optical Surveys and Related Technological Aspects*

Dear Massimo *(Capaccioli)*, **the VLT (Very Large Telescope) Survey Telescope (VST) is an instrument that will perform the deepest optical sky surveys with a single frame field of view of almost a square degree. VST will follow other innovative and successful surveys, such as the SDSS, the Classifying Objects by Medium-Band Observations – a spectrophotometric 17-filter survey (COMBO17), the NOAO Fundamental Plane Survey (NFPS), the Wide-field Nearby Galaxy-clusters Survey (WINGS), etc. Can you summarize the most important cosmological results of some of these surveys? Why you decided to build a telescope completely dedicated to surveys? What is the main advantage of VST with respect to previous surveys? What forthcoming programs will have an impact for the current cosmological ideas?**

After a rather uninteresting period marked by the transition from the photographic to the digital paradigm, the advent of a new generation of wide field digital detectors and dedicated telescopes has started what can be considered the "renaissance" of optical surveys. Panchromatic, multi-epoch survey data federated within the International Virtual Observatory Alliance (IVOA) infrastructure will surely lead to significant advances in all fields of astronomy, from planetary sciences to cosmology. The exploitation of such data sets, of unprecedented size and complexity, poses, however, problems that are still far from being resolved.

Survey data covering in an homogeneous way large patches of the sky have been and still are what we can define our "memory of the sky" and since the beginning of observational astronomy, they have been crucial to provide those large and homogeneous samples of objects that are needed to most fields of research: from planetary studies (e.g., searches for transneptunian objects and Near Earth Asteroids) to studies of Galactic structure, to extragalactic astronomy and cosmology (for an overview of the early developments refer to the papers in [10, 24, 31, 88, 93]). Furthermore, at the contrary of what happens to most astronomical data, which are usually acquired to pursue specific scientific goals and after being used by their owners are in most cases forgotten, survey data are by their nature general purpose and remain valuable over long periods of time.

The period from the late seventies to the late nineties of the last century marked a sort of "grey age" for survey work in the optical bands. During that period, in fact, the lack of suitably large digital detectors and the low quantum efficiency and photometric accuracy of photographic plates (usually placed at the focus of relatively small Schmidt telescopes) conspired to relegate optical surveys in the rather ancillary role of providing samples of (not too faint) objects to be afterward targeted with larger and more powerful instruments (significant examples being the Second Palomar Observatory Sky Survey or POSS-II [110] and the ESO/SERC Survey). New impetus came from the extensive digitalization [43] and accurate Charge Coupled Device (CCD) photometric calibration [34] of plate material and their distribution through public archives and data centers.

Photographic surveys, however, were not deep and accurate enough to match the demand of the new 8 m class telescopes and the "grey-age" ended abruptly in a sort of "survey renaissance" only when, at the turn of the century, the advent of a new generation of wide field digital detectors and dedicated survey telescopes boosted the quality (in terms of sampling and accuracy) of survey data, increasing their scientific relevance. There is no doubt that in this new era, the role of Michelangelo was played by the SDSS team [154] who pioneered the field, paved the way to all subsequent survey projects, and thought the community many important facts.

First of all, SDSS has shown the necessity of a strict and collaborative interaction between astronomers and computer scientists. In fact, a common aspect of all digital survey projects is the huge sizes of the resulting data sets that are in the range of many tens or hundreds of Terabytes of digital information, with detections of many millions or even billions of sources, and several hundreds of parameters measured for each detected source. The distribution, exploitation, and understanding of such vast amount of data presents a great technological challenge, which requires to exploit to their limits the new possibilities offered by fast computer networks, database management, high performance and GRID computing.

It would therefore not be too far from truth to say that the Digital Palomar Observatory Sky Survey (DPOSS) and SDSS experiences have been among the main triggers that at the end of the last century [6, 23] started the world-wide effort to build the Astronomical Virtual Organizations, which are nowadays being grouped under the flag of the IVOA [141].

Second, the SDSS experience marked a sort of methodological change in a large part of the astronomical community, which had to learn to trust data reduced by other teams in an almost fully automatic way (no other approach would have been possible with data of such size and complexity), thus breaking the consolidated tradition, which makes astronomers skeptical against any software tool that does not leave them full control on each single operation performed on their pixels.

Third, the SDSS showed that if survey data are properly reduced, calibrated, and documented, their scope and utility goes well beyond the original scopes of the survey itself. The extraordinary success of the SDSS science being certified by more than a thousand of scientific papers appeared in less than 10 years in the main scientific journal.

Before proceeding, it is necessary to make one additional comment: nowadays, the word "survey" has become highly fashionable and a quick search in the specialized literature leads to discover hundreds of different projects which, more or less rightly, are labeled as surveys. To restrict the subject of our discussion, we shall therefore focus on cosmological surveys mainly, and adopt the definition introduced in the 1990s by the Russian astronomer Valentin Lipovetsky [87] who considered to "survey data" any set of homogeneous data covering an area larger than the largest cosmic structure (a few hundreds Mpc), which in the local (i.e., $z < 0.05$) Universe corresponds to $\sim$1,000 square degree. The latter figure needs to be taken with some caution, as the new digital data probe a much larger depth than the photographic ones, and a rescaling with redshift needs to be introduced.

In what follows, we shall focus on three aspects of this survey renaissance, namely: the technological breakthrough that led to the present generation of survey dedicated instruments, the computational problems posed by the management, processing and distribution of survey data, and on some new types of science that will be made possible by the planned digital surveys.

The Hardware Revolution

The first telescope entirely devoted to digital survey work was the SDSS Apache Point 2.5 m telescope which, even though equipped with a 120 Mpxl CCD camera covering 1.5 square degree, was operated in drift scan mode. In this sense, therefore, the prototype of the modern survey telescopes is the MPI/ESO 2.2 m telescope, which in 1999 was refurbished and equipped with the Wide Field Imager [7]: a mosaic of eight $2k \times 4k$ CCD chips. WFI was used to perform the first European public survey (ESO-Pilot Imaging Survey; [37]). At the time, other CCD mosaics were already in use (more notably the French MegaCAM, 36 chips $2k \times 4.6k$ each) but they were installed at the focus of larger nondedicated telescopes and were used either for specific projects or for small surveys aimed at feeding the new generation of 8 m class telescopes (EIS [111]). More recently, the refurbishing of the Samuel Oshkin Schmidt telescope at Mount Palomar with the addition of a mosaic of 132 CCD detectors [9] has opened the era of multi-epoch surveys. In fact, since 2003, this instrument has been used in drift scan mode (∼150 s per scan) to perform the Palomar Quest Digital Synoptic Survey (PQ; [45]), which aims at covering 15,000 square degree in four bands with multiple (5–10) observations with time baselines ranging from hours to years. The PQ, in spite of the problems encountered in the cleaning and calibration of the data, must be considered as a pioneeristic effort to explore the time domain, which will be discussed in some detail below.

The real breakthrough in the field of survey telescopes will take place, however, at the end of 2008, when both the VLT Survey Telescope or VST, and the Visible and Infrared Survey Telescope for Astronomy (VISTA) will become operational (Fig. 5.15).

VST [5] is a 2.6 m telescope with a corrected field of view of 1 square degree and has been built by the Astronomical Observatory of Capodimonte in Italy to perform surveys in the optical bands. VST is completed by a $16k \times 16k$ CCD mosaic camera built by the European consortium OmegaCam [82].

The VST+Omegacam system will provide a pixel scale of 0.21 arcsec, and will operate in pointed mode, thus ensuring survey data of unprecedented depth and accuracy. This obviously implies a huge data throughput, estimated in ∼ 15 TB per year (∼ 5 TB of calibration data and 10 TB of raw science data).

VISTA [49] is instead a 4 m telescope with a corrected Field of View (FOV) of 1.65 degree (diameter) built by a consortium of UK institutes to perform surveys in the near infrared (from 0.85 to 2.3 μm). VISTA is equipped with Wide Field Camera (WFCAM): a CCD mosaic of 16 2048×2048 pixels. Both VST and VISTA will be operated by ESO and will dedicate ∼75% of the time to perform public surveys and the remaining time will be used for smaller proprietary projects.

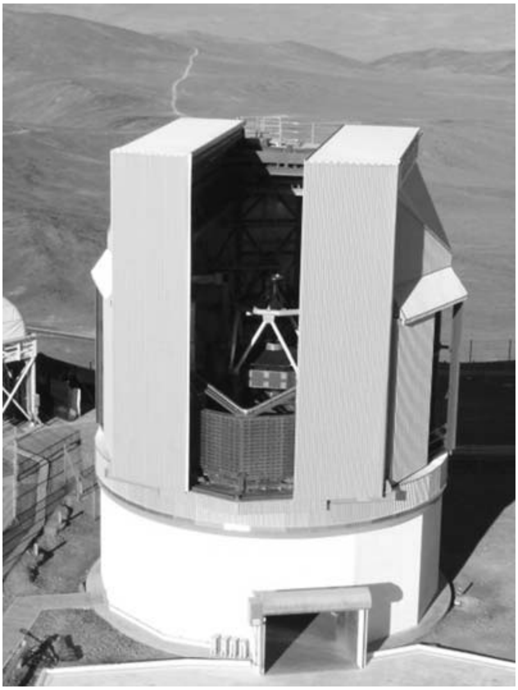

Fig. 5.15 The VST dome and the VST telescope in January 2008 during the installation of the telescope

It needs to be stressed that the final performances of a survey system are strongly dependent on the stability of the system itself and on the optimization of the acquisition and calibration strategies. In the case of VST+Omegacam, the data acquisition, calibrations, and pipeline reductions were strictly procedurized with the aim of maintaining the instrument and not the individual data sets calibrated at all times. ESO will operate the instrument in service mode, optimizing the observing programme to ambient conditions, and routinely taking calibration data.

Thus, each night the instruments overall responsivity and also the transmission of the atmosphere will be monitored in the u, g, r, and i bands irrespective of the schedule of science observations.

The ultimate survey telescope, however, will be the Large Synoptic Survey Telescope (LSST), expected to start operation in 2013: a 8.4 m telescope with a corrected FOV of 10 square degree and a 1 Gpxl camera. LSST will cover the entire sky every three nights and the expected data flow is an impressive 17 TB per observing night.

Thank you Massimo.
The last point made by Massimo underlines the crucial role of computer technology in the management of such enormous amounts of data. In the next interview, Giuseppe Longo will present the various aspects related to this problem.

Dear Giuseppe *(Longo)*, **the future big surveys will be a big problem for the present day computer technology of data storage and reduction. Could you please comment on that?**

Table 5.1 Data throughput of some ongoing and planned future survey facilities and experiments

Survey	Year(s)	Raw data	Catalogue data
SDSS	2001–2008	40 TB	3.5 TB
PQ	2003–2008	30 TB	2.5 TB
VST	2008 on	15 TB/year	2.0 TB/year
VISTA	2008 on	100 TB/year	$\sim$5 TB/year
PanSTARRS	2009	400 TB/year	$\sim$10 TB/year
LSST	2013 on	$\sim$5 PB/year	$\gg$10 TB/year

As it was mentioned by Massimo Capaccioli, the new generation of survey telescopes will produce a data avalanche, which in same cases needs to be processed in almost real-time. To deal with such data flows is not an easy task and, without entering in much detail, we shall just recall that, broadly speaking, the process leading from raw data to their scientific exploitation requires three steps: (1) from raw data to calibrated and stacked images (pre-reduction); (2) from stacked images to catalogues (extraction); (3) from catalogues to their interpretation (analysis and visualization).

For what the first step – pre-reduction – is concerned, it needs to be stressed that even though most of the operations that need to be performed on the individual pixels are the usual ones (i.e., bias correction, cosmetics, flatfielding, astrometric and photometric calibration, etc.), survey pipelines cannot be obtained by refurbishing old software, but require a new generation of software methods and tools. Sky surveying modes require in fact a larger data throughput, a very stable and reliable operation over long periods of time, and greater data flows than it is the case for most "old-style" astronomical observing. Furthermore, because of the long amounts of time required to complete a survey (often many years), a great deal of care must be exercised to monitor the overall performance of the survey to ensure uniform data quality and to include all relevant metadata and ancillary information in the same database. Common sense and the DPOSS and SDSS experience have in fact shown that to make the data truly useful, the raw data as well as the processed ones need to be stored into specifically designed databases together with all those metadata that are needed for further re-calibration and/or analysis. Furthermore, a crucial need for the processed data is to provide each single record with an history of all operations performed to derive it. Most (if not all) of these operations need to be performed in fully automatic way and many international teams have produced automatic data acquisition and data reduction pipelines, which successfully addressed the pre-reduction steps (i.e., from raw data to calibrated and stacked images) by introducing automatic quality controls and labeling processed data with quality flags [132] or by introducing new levels of abstraction in the data modeling. These pipelines can be broadly grouped in three main types, which we shall call traditional, forward chained, and backward chained. The traditional one (adopted, for instance, by SDSS, VISTA [64], and TERAPIX [139] teams) is centered on the public releases of complete sets of data with well described methods, qualifications, and calibrations. While it is suitable for homogeneous survey projects, such approach cannot match the needs of planned surveys both because methods are not scalable to the expected huge data streams, and, second, because when different users or communities want to make a different use of the data using new computational methods, calibration strategies, new insights or improved code, the re-processing of the raw data and the storing of the results turns out to be almost impossible.

A second solution, forward chaining, was pioneered at the Hubble Space Telescope Science Institute (HST-ScI), and allows the user to ask for a certain data product that will then run the calibration pipeline on-demand using the best available calibrations [131]. This approach, however, does not allow the user to go backwards (i.e., from the results to the raw data) and, for instance, prevents him from inspecting the robustness of the results by checking the dependencies on the uncertainties of calibration parameters and applied methods.

The third and more innovative approach, namely the backward chaining, has been developed within the Astronomical Wide-field Imaging System for Europe (AstroWISE) pipeline and seems to be the most promising one. The AstroWISE system [143] was designed enforcing a global data acquisition and processing model, while retaining flexibility. In AstroWISE, the data model has been translated into an object model, with full registration of all dependencies, and all I/O of the project are stored in a single, distributed database, containing all metadata describing the bulk data (e.g., images) and the derived results in catalogue form (e.g., lists of celestial sources). The database is connected to a federated file server that stores the bulk data, and to a computing GRID that sends jobs (including clients) to single nodes or parallel clusters (Fig. 5.16).

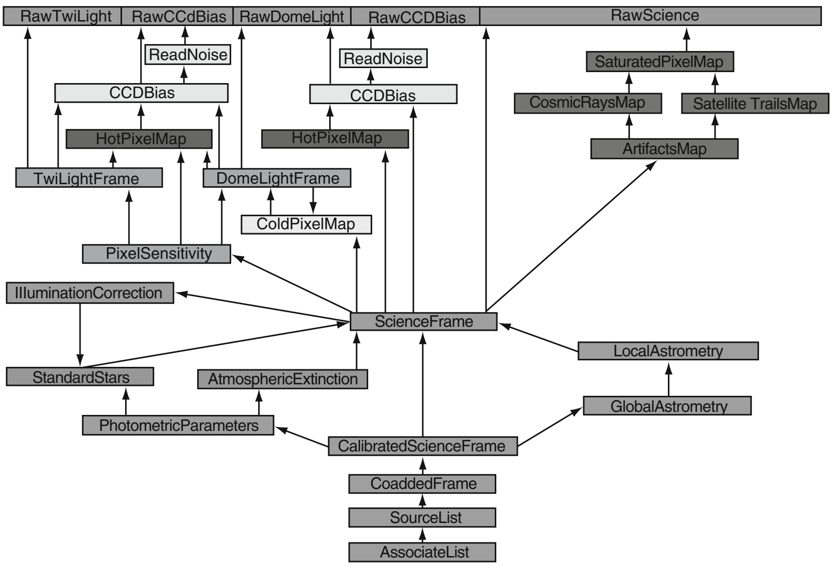

Fig. 5.16 A target diagram of AstroWISE, i.e., a slightly simplified view of the dependencies of targets to the raw observational data. *Arrows* indicate the backward chaining to the raw data. From [143]

For what the second step, namely the automatic detection and extraction of objects from the images and the measurement of their parameters is concerned, many issues remain still to be solved. In fact, the choice of the optimal algorithm for source extraction is very much depending on both the nature of the data and on the type of science that needs to be performed. First of all, the extraction process requires a good understanding of the noise properties (which vary from night to night and even from frame to frame), which determine the detection significance threshold and therefore the completeness (the fraction of real sources detected) and the contamination (due to noise peaks mistaken for real sources or to extended objects erroneously split in several parts) of the resulting catalogues. Most source detection algorithms require a certain minimum number of adjacent or connected pixels above some signal-to-noise thresholds for detection (either in the real or in some transformed space); the optimal choice of these thresholds depending on the sampling and on the power spectrum of the noise. In many cases, the detection process involves some type of smoothing or optimal filtering, for example, with a Gaussian whose width approximates that of an unresolved point source. Unfortunately, this also introduces a preferred scale for source detection which, accordingly to the specific needs, is usually optimized either for the unresolved sources (e.g., stars) or for the resolved ones (e.g., faint galaxies). In other words, the source detection and extraction algorithms introduce selection biases not only on the source flux but also on their shape or contrast (i.e., in the limiting surface brightness averaged over some specific angular scale).

A typical example of how such biases can affect the final catalogues is the fact that most methods (e.g., S-Extractor [17]) miss the detection of low surface brightness galaxies, which can be recovered only by re-processing the data with specifically tailored (e.g., multi-scale methods) algorithms.

Once individual sources are detected, structural parameters need to be measured: fluxes in a range of apertures or isophotal thresholds, diameters, radial moments of the light distribution, etc. from which a suitably defined, intensity-weighted centroid is computed. To avoid useless re-processing of the data, the set of parameters measured for each object needs to be kept as large as possible (the SDSS lists almost a thousand of measured parameters and associated errors for each object), thus causing an additional inflation in the dimensionality of the parameter space.

The size, complexity, and high dimensionality of the parameter space are behind the third, more intriguing and still largely unsolved problem: the visualization, manipulation, and extraction of the information contained in the final survey catalogues. The vastity of the problem can be easily summarized by reminding that such volumes of data can hardly be explored using traditional, interactive approaches, and that also automatic methods that are common in other disciplines (such as bioinformatics, market trend-analysis) cannot be easily applied. From a mathematical point of view, most of the operations astronomers perform on their data can be reconduced to common data mining tasks such as classification, unsupervised clustering, regression, and more complex forms of pattern recognition.

All these methods scale badly with the number of records (N) and even worse with the number of features (parameters) (D) in the parameter space. As a rule of thumb, the computational cost of the more common DM tasks scale as

- Clustering as: $\sim N \times \log N \times N^2$, and as $\sim D^2$,
- Correlation as: $\sim N \log N \times N^2$, and as $\sim D^k \quad (k \geq 1)$,
- Likelihood or Bayesian as: $\sim N^m \quad (m \geq 3)$, and as $\sim D^k \quad (k \geq 1)$,

which means that even taking into account the current trends in the growth of both disk I/O speed and in Central Processing Unit (CPU) performances, the archives produced by modern astronomical surveys cannot be effectively dealt with, unless innovative data mining methods are adopted. To add further complexity to the problem, we must mention that astronomical parameters are often highly correlated and that their significance is conditioned by flags that are often under the form of logical or discrete variables, which cannot be effectively approached with traditional algorithms. Furthermore, we must take into account the fact that, as it has been stressed by many authors, one of the main science product of properly conducted surveys is that they will allow to create a unified, multi-wavelength, multi-epoch view of the Universe, from X-rays through UV, optical, near-IR to radio.

The federation of homogeneous and well conducted surveys, while on the one end is highly desired, on the other end it poses further challenges which, up to now, are still poorly matched by existing algorithms and methods, that is, the presence in the data sets of a large fraction of nondetections (missing data or upper limits) in the intersection of different archives [132].

Finally, it is apparent that with the current networking technologies, it is not even thinkable to transfer over the net any large survey archive and it is instead necessary to move the programs where the data are. It is therefore not an exaggeration to say that the exploitation of survey data is one of the main reasons why the astronomical community became interested to the use of distributed computing environments.

The perception of the relevance of this problem to the exploitation of existing and planned surveys was behind the suggestion made in 2001 by the US National Science Foundation which, in its decadal survey [6], recommended as top priority among the so-called *small-sized* projects, the implementation of a National Virtual Observatory (NVO): that is, a Cyber-infrastructure aimed at the standardization, federation, and interoperability of all astronomical archives produced by ground-based and space-borne telescopes. Even though it is impossible to summarize in a few lines the outcome of almost 10 years of world-wide effort [141], we just wish to point out that while the federation of different survey archives has already become a reality (most survey data archives are already interoperable within the VOb infrastructure), the implementation of user friendly and effective query languages and tools for visualization and data mining is still much less advanced.

Special problems are posed by data mining of survey data in the time domain. In this respect, we wish to stress that any survey is bound to produce time flagged data both due to the dithering acquisition procedure intrinsic to the use of CCD mosaics (short sampling), to the fact that in order to ensure proper calibration, frames of adjacent regions need to overlap by a nonnegligible amount and, finally, to the fact that in most cases multiple images are frequently obtained at the same position of the sky for follow up co-addition as it helps one go deeper and look for fainter objects. The targets of this time domain studies are mainly of two types: intrinsically variable objects such as variable stars or AGN, SNe, etc. or astrometric transients such as asteroids or transneptunian objects, which occupy different positions in different epochs and pose a real challenge to data mining algorithms as they appear as missing data (or upper limits) in most epochs.

Can you formulate an example of how these new data mining methodologies can affect the extraction of information from the present and future massive astronomical data sets?

A typical example is what George Djorgovski calls the "discovery of the rare and of the unknown" [44], namely the well known fact that every time a new technology opens a new window in the astronomical parameter space, or a new way to explore it allows a better sampling of its characteristics, new phenomena are bound to be discovered. Rare objects are objects that are either intrinsically rare (e.g., one out of a billion) or more simply hard to find in a given data sets. The traditional approach to their search consists in using their known or expected properties (e.g., typical broadband spectra, or variability properties) convolved with the survey selection functions (e.g., bandpass curves, flux limits, angular resolution, etc.) to define regions in the parameter space where they are more likely to be found or even to design experiments where such objects can be distinguished from the "uninteresting" majority of other objects. Typical examples being the searches for high redshift quasars (usually performed by using their known properties to define specific regions in the colors space [51, 77, 112–114]); high redshift galaxies [40]; brown (L and T) dwarfs [83], etc.

As an exemplification, in Fig. 5.17 we show how the criteria for the search of quasars must evolve with the increasing complexity and size of the available survey data. If we take into account that at intermediate Galactic latitudes there is about one quasar per million stars (down to $r \simeq 19.5$), it becomes evident that a good color discrimination and a good star-galaxy separation are essential to avoid an excessive contamination of the spectroscopic follow-up samples by mismeasured stars or misclassified galaxies. Panel (a) is a diagram from [77], which shows how high redshift ($3.8 < z < 4.3$) quasars were selected in the DPOSS catalogue.

Normal stars form a well defined temperature sequence (the banana-shaped locus in panels a and b) and the quasars appear as more or less well defined outliers. The cuts to be applied to disentangle the two types of objects can be derived by folding the spectra of template quasars through the survey filter curves. Photometric errors, peculiar stellar spectra, and varying quasar redshift produce a region of ambiguity where the contamination from stars misclassified as candidate quasars is very high. By adding an additional dimension to the diagram, that is, by taking into consideration one more color, the degeneracy can be solved and the two groups (stars and quasars) are better separated (panel b). Adding more dimensions helps to obtain a more accurate discrimination and in panel (c), taken from [36], we present the results of a new method applied to SDSS data using a color space of higher

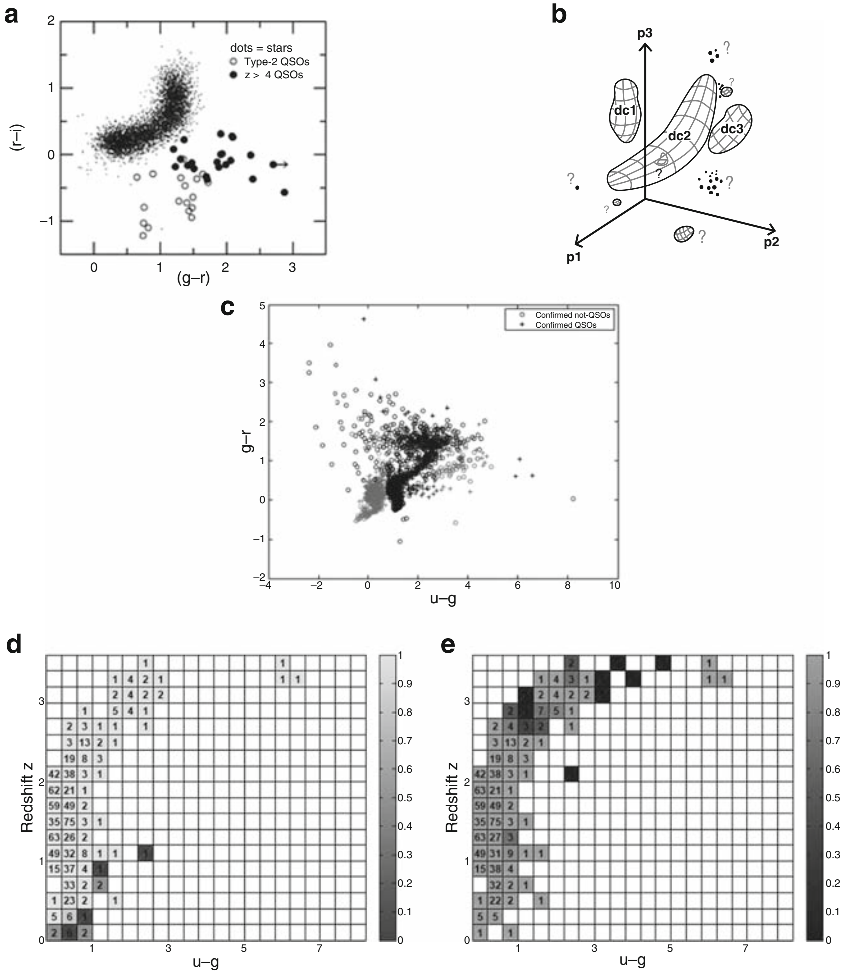

Fig. 5.17 (**a**) Selection of high redshift quasar candidates in the DPOSS data. (**b**) Scheme showing how the addition of more features (dimension of the parameter space) may help in disentangling objects that are projected in the same region of a 2D diagram. (**c**) Results of the unsupervised clustering performed on SDSS data by [36]. In (**d**) and (**e**) are shown, respectively, the efficiency and completeness evaluated from the base of knowledge as a function of the redshift (see the text and the original paper for more details)

dimensionality. The method is based on the unsupervised clustering of data (using the Probabilistic Principal Surfaces algorithm [127]) followed by a labeling phase performed using the knowledge extracted from the SDSS spectroscopic database. Without entering into the details of the method, which can be found in [36], we shall just mention that, by exploiting a larger amount of information, the method is capable to identify candidate quasars that would otherwise be lost in the stellar locus. The high accuracy (efficiency ∼95.4% and completeness ∼94.7%; panels d and e in Fig. 5.17) is obtained at the price of 11 days of CPU per experiment.

Even more intriguing is the prospect that survey data will unveil a wealth of previously unknown types of astronomical objects and phenomena in a sort of planned serendipity. Possible examples of new kinds of objects (or at least extremely rare or peculiar sub-species of known types of objects) have already been found in the SDSS [52, 133], DPOSS [45], and UK Infrared Telescope (UKIRT) Infrared Deep Sky Survey (UKIDSS) [144] data sets (to quote only a few cases).

It is clear, however, that in the future, the most effective way to pursue this type of researches will be the use of statistical methods to partition the parameter space in N distinct classes (or clusters) of similar objects; these can then be labeled using the information contained in a suitable base of knowledge (i.e., an exhaustive sets of well studied templates) and it has to be expected that the majority of the objects will fall into a relatively small number of (easy to recognize) classes, such as stars, early type or late type galaxies, etc. with only a few objects grouped in a small number of less populated clusters. These poor and/or unlabeled clusters, as well as individual outliers, will form the hunting ground where to look for either rare or unknown objects.

It needs to be stressed, however, that at the moment the use of unsupervised clustering in the astronomical context has just started due to several factors. First of all, the computational cost of most unsupervised methods, when applied to realistically large data sets, imposes the use of parallel and/or distributed computing and so far no such code has yet been made available.

Second, astronomical data sets are usually plagued by a large fraction of missing data (i.e., upper limits or non detection), which are not effectively dealt with by most machine learning methods, which interpretate them as singularities in the parameter space. Finally, there is not yet a suitable base of knowledge, that is, an exhaustive set of templates sampling in a more or less uniform way the parameter space. All these three items, however, are being addressed by several teams working under the IVOA umbrella and preliminary applications are beginning to appear [25, 36].

An alternative, less effective but more common, approach sees the use of supervised classification methods such as decision trees, artificial neural networks, or support vector machines, trained on extensive set of templates. Peculiar or new types of objects can in this case be identified by anomalously large values of some metric distance from the known groups of objects. Supervised methods belong to the normal analysis practice of modern survey data as they are used for a variety of tasks: from the evaluation of photometric redshifts [35], to star-galaxy separation in the optical or NIR bands (see [46] and references therein); for the identification of variable objects (see next section), to derive galaxy morphologies [8].

Thank you very much Giuseppe. It is remarkable how much the astronomical research still continues to be one of the main motors of the continuous progress in computer science and data storage, reduction, and analysis, other than in instrument technology. In the next interview, Massimo Capaccioli will comment on the opportunities opened by new astronomical facilities for the study of various time-dependent astrophysical phenomena.

Dear Massimo *(Capaccioli)*, can you comment on the role that future surveys will play in the exploration of the time domain?

Time variability of astronomical objects is one of the most intriguing and less explored aspects of modern astrophysics.

Over the past decade, many projects have been handling moderate (MB–GB per night) data rates searching for transient phenomena (see for instance, the Deep Lens Survey [151], the High-z SN Search [120], the Nearby Galaxies Supernova Search [130] to quote just a few) or for microlensing effects (Massive Compact Halo Objects (MACHO) [4] and the Optical Gravitational Lensing Experiment (OGLE) [140]), leading to outstanding discoveries: from the accelerated expansion of the Universe [104], to the discovery of planetary transits [26], to the discovery of optical flashes from GRBs (e.g., GRB990123).

One of the main science drivers of LSST and other planned surveys is the exploration of the time domain in search of astrometric (objects which change their position) and photometric transients (variable objects). Astrometric transients are Solar system objects such as Kuiper Belt objects, Near Earth Objects or distant comets. They appear either as elongated trails in long exposures (if fast moving) or, more likely as missing data or upper limits in all but a few temporal slices of the time domain. At least two epochs are needed to identify previously unknown asteroids in any synoptic survey, and their baseline will define the effective time resolution of any transient search (we note that at least three properly spaced epochs are needed to compute even a rough orbit). It needs to be stressed that even when the scope of the survey is to search for intrinsically variable objects only, the detection and recognition of astrometric transient is crucial, as slow moving objects are the principal source of contaminants. In the PQ, for instance, there are >100 asteroids for each astrophysical transient (down to mag 21 and depending on the Ecliptic latitude) and, therefore, to optimize the scientific return of future surveys, a crucial requirement is to improve the existing catalogs of asteroids [45].

Once astrometric transient are removed from the survey data, the remaining transients are the photometric ones and, by extrapolating the number of photometric transients discovered in the PQ data and in the SDSS [122] surveys, it is possible to estimate that the Panoramic Survey Telescope and Rapid Response System (PanSTARRS) and LSST will observe $\sim 10^4/night$ and 10^5–10^6 variable objects per night, respectively. Extrapolating what has been found by the SDSS team, $\sim 2\%$ of unresolved optical sources will be variable at the 0.05 mag level (rms) in both the g and r bands and the majority (2/3) of these are low-redshift ($z < 2$) quasars. LSST should therefore be able to obtain well-sampled 2% accurate, multi-color lightcurves for ~ 2 million low-redshift quasars, and discover at least 50 million variable stars. Among the other objects many will be known, highly variable types of objects, where the "low state" is below the detection of the baseline data, with variable stars of different kinds dominating on the short time scales (from minutes to months), and AGN (mainly Blazars) dominating on the longer time scales (years and longer). Some other objects will be of unknown type and real-time spectroscopic and other follow-ups will become necessary to understand them.

Achieving high completeness and low contamination (a few false alarms) is therefore both a must and a yet unsolved challenge. In fact, in large data sets, the most unlikely things will happen, and to minimize the number of false detections, a robust and reliable data cleaning is required. Furthermore, in most cases, to effectively study many of these transients, especially those that are variable over only one cycle such as SNe, one has to detect them not off-line but as they occur in order to allow for immediate follow-up and analysis. Once more, to prioritize the foreseen spectroscopic and photometric follow-ups of the most interesting events, it will be necessary to recur to automated classification methods largely based on machine learning.

Thanks a lot Massimo.
Having discussed the observational expectations over various frequency bands, it is natural to ask if they could be decisive for our understanding of the formation of first structures in the Universe, a crucial problem in cosmology. Piero Madau will address this aspect in the next interview.

5.5 New Key Observations Dedicated to the First Structures

Dear Piero *(Madau)***, in Chap. 2 you discussed the most important aspects connected to the first structure formation and the dawn age of the Universe. Which observations may help solving the actual controversies?**

The dialog in Chap. 2 should make it clear that, despite much recent progress in our understanding of the formation of early cosmic structure and the high-redshift Universe, the astrophysics of first light remains one of the missing links in galaxy formation and evolution studies. We are left very uncertain about the whole era from 10^8–10^9 year – the epoch of the first galaxies, stars, SNe, and massive Black Holes (BHs). Some of the issues discussed earlier are likely to remain a topic of lively controversy until the launch of the JWST, ideally suited to image the earliest generation of stars in the Universe. If the first massive BHs form in pregalactic systems at very high redshifts, they will be incorporated through a series of mergers into larger and larger halos, sink to the center owing to dynamical friction, accrete a fraction of the gas in the merger remnant to become supermassive, and form binary systems [149]. Their coalescence would be signaled by the emission of low-frequency gravitational waves detectable by the planned Laser Interferometer Space Antenna (LISA).

An alternative way to probe the end of the dark age and discriminate between different reionization histories is through 21 cm tomography [90]. Prior to the epoch of full reionization, 21 cm spectral features will display angular structure as well as structure in redshift space due to inhomogeneities in the gas density field, hydrogen ionized fraction, and spin temperature. Radio maps will show a patchwork (both in angle and in frequency) of emission signals from H I zones modulated by H II regions where no signal is detectable against the CMB [29]. The search at 21 cm for the epoch of first light, while remaining an extremely challenging project, remains a tantalizing possibility within range of the next generation of radio arrays.

While the above cosmological puzzles can be tackled directly by studying distant structures, it has recently become clear that many of today's "observables" within the Milky Way and nearby galaxies relate to events occurring at very high redshifts, during and soon after the epoch of reionization (see, e.g., [96] and references therein). In this sense, galaxies in the Local Group can provide a crucial diagnostic link to the physical processes that govern structure formation and evolution in the early Universe, an approach termed "near-field cosmology". It is now well established that the hierarchical mergers that form the halos surrounding galaxies are rather inefficient, leaving substantial amounts of stripped halo cores or "subhalos" orbiting within these systems. Small halos collapse at high redshift when the Universe is very dense, so their central densities are correspondingly high. When these merge into larger hosts, their high densities allow them to resist the strong tidal forces that acts to destroy them. Gravitational interactions appear to unbind most of the mass associated with the merged progenitors, but a significant fraction of these small halos survives as distinct substructure. The "Via Lactea Project", a suite of the most detailed N-body simulations of Milky Way CDM substructure to date [41, 42, 81, 331], has shown that, in the standard CDM paradigm, galaxy halos should be filled with tens of thousands subhalos that appear to have no optically luminous counterpart (see Fig. 5.18).

As shown in Fig. 5.19, such finding appears to exacerbate the so-called missing satellite problem", the large mismatch between the twenty or so dwarf satellite galaxies observed around the Milky Way and the predicted large number of massive CDM subhalos. Even if most DM satellites have no optically luminous counterparts, the substructure population may be detectable via flux ratio anomalies in strong gravitational lenses [94], or possibly via γ-rays from DM annihilation in their cores (e.g., [16, 30]). We are coming into a new era of galaxy formation and evolution studies, in which fossil signatures accessible today within nearby galaxy halos will allow us to probe back to early epochs, and in which the basic building blocks of galaxies will become recognizable in the near-field.

Thank you very much Piero.
As discussed in previous chapters and underlined also in the above section, an accurate modeling of DM through detailed N-body simulations is of fundamental importance for the understanding of the structure and substructure formation process in the DM paradigm. We will continue this discussion in the next two interviews with Matthias Steinmetz and Simon White.

5.6 N-Body Simulations

Dear Matthias *(Steinmetz)*, **what are the future possibilities and the limits of numerical simulations in the DM research?**

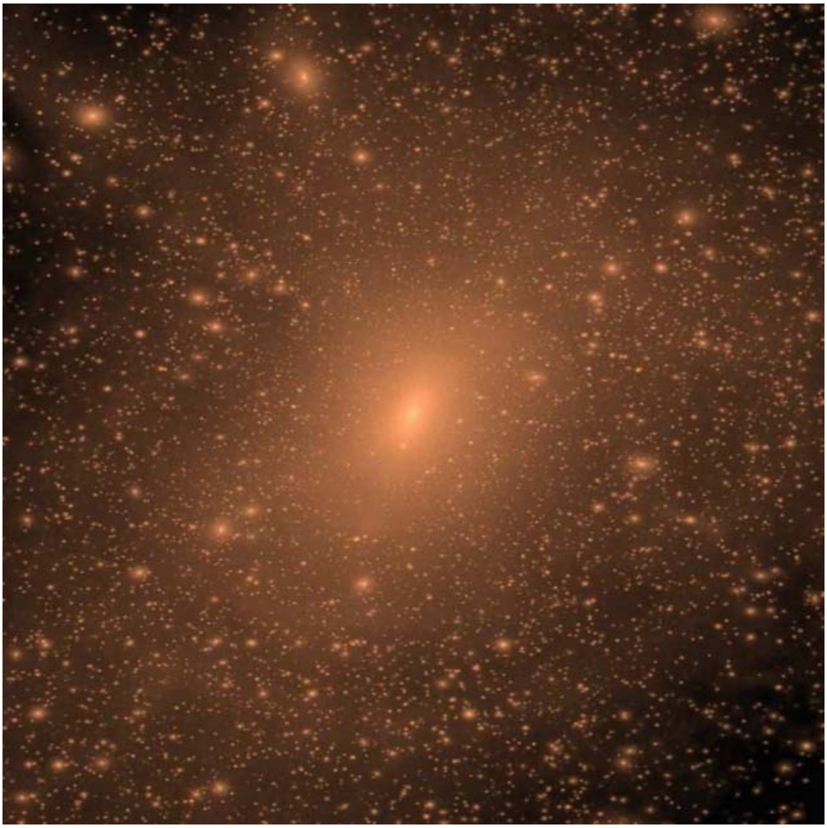

Fig. 5.18 Projected DM density-squared map of the Via Lactea halo at the present epoch. The simulation follows the growth of a milky way-size halo in a ΛCDM universe from redshift 50 to the present. It was performed with the parallel treecode PKDGRAV [126] and samples the galaxy-forming region with 234 million particles of mass 2.1×10^4 $M_\odot$. The image covers an area of 800×600 kpc, and the projection goes through a 600 kpc-deep cuboid containing a total of 110 million particles. The logarithmic color scale covers 20 decades in density-square. Note the wealth of resolved substructure

Numerical simulations have been an integral part in the detailed analysis of the virtues of the ΛCDM scenario. Only numerical techniques can account for the highly irregular structure formation process and for at least some of the complicated interaction between gravity and other relevant physical processes such as gas dynamical shocks, star formation, and feedback processes. Simulations also provide the required interface to compare theoretical models with observational data and are able to link together different epochs.

N-body simulations that follow the evolution of a collision-less component (DM, stars) are nowadays an extremely robust and reliable tool for theoretical studies, the capabilities and limits are fairly well understood and under control. Simulations with hundred of millions, if not a few billion particles, have provided considerable insight into the formation of the large scale structure and the dynamics of DM halos.

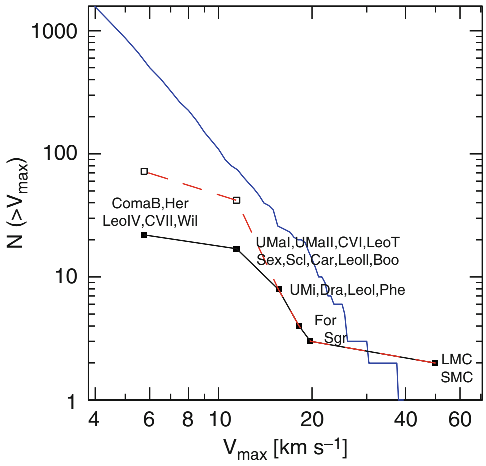

Fig. 5.19 Cumulative number of Via Lactea subhalos (*solid curve*) as well as all milky way satellite galaxies within 420 kpc (*filled squares*), as a function of circular velocity. The data points assume a maximum circular velocity of $V_{\max} = \sqrt{3}\sigma$, where σ is the measured stellar velocity dispersion. The *short-dashed curve* connecting the empty squares shows the expected abundance of luminous satellites after correcting for the sky coverage of the SDSS. From [331]

They also have been able to highlight some of the key issues of the ΛCDM model as discussed in this articles: the core vs. cusps problem and the substructure issue. The challenge in N-body simulation is, however, nowadays to a lesser amount the actual performance of large simulations, but rather to master the large amount of data provided by some of the largest simulations performed so far. For many applications of N-body methods, we are indeed no longer limited by compute power but by data mining techniques and by man power to analyze the data.

The situation is quite different for gas-dynamical simulations. With exception of the low density IGM, there is hardly any application of pure gas dynamics in extragalactic astrophysics, and indeed for that particular application (IGM), much of what has been said for N-body simulations is valid as well, the field is rather mature. The situation changes, however, dramatically if other applications are considered, like cooling flows in galaxy cluster or in particular the intricate details of the galaxy formation process. The evolution of a galaxy system is dominated by the so-called feedback processes, that is, for the time being physical processes far below the resolution scale of a simulation. Consequently, considerable discrepancies between the applied methods and/or at least uncertainties on their validity exist. First steps towards a consistent treatment of multi-phase models (see, e.g., [125, 153]) are, however, quite promising. To a large extent, progress in this field will depend on the development of reliable multi-phase models.

N-body simulations and gas dynamical simulations of the IGM will continue to be routine tools for extragalactic astrophysics. The frontier is where small scale processes like star formation or accretion onto BHs dominate. Progress in this area will not solely depend on the increased availability of compute power or of more advanced simulations methods, but rather by our ability to design and test realistic sub-resolution models of the interaction of gas dynamics, cooling, star formation, and various feedback processes.

Thanks a lot Matthias.

Dear Simon *(White)*, in which way N-body simulations may help to better understand the properties and the nature of DM? How is the situation improved, thanks to the recent supercomputer facilities?

Historically, N-body simulations have played a major role in suggesting likely properties for the DM. A priori, the most plausible elementary particle candidate for the DM (and indeed the only one that is currently *known* to exist) is a massive neutrino, for example, the τ-neutrino. Nevertheless, neutrinos were discarded as viable candidates soon after they were first suggested, because N-body experiments showed conclusively that nonlinear growth from Hot Dark Matter (HDM) initial conditions produces low-redshift structure, which is incompatible with observation. Similar experiments started from CDM initial conditions produced a much better match, thereby substantially enhancing the case for CDM.

Simulations are most useful for predicting structure in the DM distribution on nonlinear scales. The fundamental unit of cosmic structure appears to be the DM halo, and simulations have been essential for predicting the abundance, clustering, and internal structure of dark halos. Although there are analytic methods that can give insight into the first two of these issues, precise predictions still require simulations. The simulation-based demonstration that dark halos should have "universal" density profiles and "universal" distributions of shape, spin, and internal substructure is at the basis of many proposed low-redshift tests of the standard paradigm, for example, those based on gravitational lensing reconstruction of the mean mass density surrounding galaxies or galaxy clusters.

In principle, the structure of the innermost regions of dark halos and the abundance of very small halos are expected to differ in Cold and Warm Dark Matter universes. Differences are also expected if the DM particles have a large enough elastic collision cross-section that their mean free path becomes comparable to the size of a halo. These effects can be explored by simulations, and considerable effort has been put into such work because some properties of dwarf galaxies appear inconsistent with simple models based on CDM simulations. Unfortunately, the regions where discrepancies occur are strongly influenced by processes in the observed stellar and gaseous components, so that it is unclear whether they reflect a need to modify the properties of the particles, or just an incomplete understanding of the evolution and structure of the baryonic component.

Computing capabilities continue to grow exponentially and so the scale and complexity of the calculations we can complete is increasing rapidly. Nevertheless, it is important to realize that major results have typically been obtained with quite small simulations. The HDM model was excluded by cosmological simulations that represented the entire cosmic matter distribution using only 32,768 particles. The universal density structure of dark halos was established using simulations with only about 10^4 particles per halo. The universal shape and spin distributions of dark halos were established with even fewer particles per halo. The understanding of the relation between the Lyα forest in quasar spectra and the diffuse IGM was also achieved with quite modest simulations. Although later work at higher resolution refined all these results, it was not necessary to achieve the original insights. In my view, our current understanding of galaxy formation is limited primarily by our ability to describe the relevant "small-scale" physics correctly, rather than by the computer power available to us.

Increased computer power *has* recently led to a qualitative step forward in our ability to simulate large regions of the Universe with sufficient resolution to produce realistic mock galaxy surveys for comparison with recent studies of the nearby and distant Universe. This is proving critical both for calibrating the extent to which cosmologically precise information can be obtained from such surveys, and for isolating how they can illuminate the process of galaxy formation. For example, I believe substantial further work in this direction will be necessary to understand how precisely cosmological parameters can be constrained through measurements of the baryon oscillation signal in the *galaxy* distribution.

Thank you very much Simon.
Another important opportunity for cosmology opened by the next generation of optical and infrared surveys is the high redshift mapping of the expansion of the Universe through the observation of a very large sample of SNe. Massimo Turatto will discuss these perspectives in the next interview.

5.7 Future Perspectives from SNe

Dear Massimo *(Turatto)***, on the SN side, what are the most ambitious research initiatives now, and what we will see in the nearby future?**

Because of their central role in stellar evolution and in cosmology, the SNe are at the core of a number of initiatives, both space and ground-based, some already in progress, others envisaged for the near future. Indeed, SNe will maintain the privileged role as the most effective probe of DE and, in combination with weak gravitational lensing, BAO and galaxy clusters. The goal is to constrain the parameters of the DE equation of state w_0 and w_a to a few and 10%, respectively.

The ongoing SN DE projects like the Supernova Legacy Survey (SNLS)[5] and Equation of State SuperNovae trace Cosmic Expansion (ESSENCE)[6] are expected to deliver their final results very soon. At that point also combining the data sets (e.g., [293]), the uncertainty on w_0 will be of the order of 5% statistical and 5% systematic. These surveys will find natural extensions in future wider-field surveys on larger telescopes, which will attach the DE issue with the multiple, cross-checking techniques mentioned earlier. For the SN surveys, the goal is to have larger samples, on wider range of redshifts and to use self-consistent treatment of the systematics. It will be also crucial to have higher throughput at the redder wavelengths and very accurate knowledge of the response of the photometric systems. At the beginning of the next decade, the Dark Energy Survey (DES)[7] will use a several square degree im-

[5] See SNLS in web page list.

[6] See ESSENCE in web page list.

[7] See DES in web page list.

ager on the prime focus of the Blanco 4 m telescope at CTIO in Chile to carry out an overall 5,000 square degree survey in g, r, i, and z passbands. For SNe, the goal is to discover over 2,000 SNe-Ia with $0.3 < z < 0.8$ over the course of a 5 year program. A dramatic increase in the SN sample might come in 2015 from the LSST, which will provide multi-color light curves for millions of SNe (with mean $z \sim 0.45$), and deeper and better sampled light curves for tens of thousands events to a limiting $z > 1$ [69]. The new generation wide-field spectrographs (Wide-Field Multi-Object Spectrograph and Smart Fast Camera) that are envisaged for the 8–10 m class telescopes will have the fundamental role of spectroscopically confirming as $true$ type Ia SNe at least a fraction of the discovered SNe, to validate the classification criteria and to study the systematics.

Three missions have been selected by NASA and the US Department of Energy (DOE) as candidates for a Joint Dark Energy Mission (JDEM): ADEPT, DESTINY[8] and SuperNova Acceleration Probe (SNAP)[9]. All of them make a combined use of SNe, weak lensing, and BAO to characterize the DE, though with different emphasis. In particular, SNAP will make accurate optical and near-IR imaging and spectroscopy of 2,000 SNe up to $z \sim 1.7$, thanks to a tight control of the systematic uncertainties by mean of careful calibration and repeated observations. ESA has started the study of a parallel mission, named *Euclid*, which combines the approaches of DUNE (focused on weak lensing) and SPACE (focused on BAO) and has obvious interest also for SN science. Discussion for possible cooperation between ESA and NASA are in progress.

The improvement in our knowledge of the DE does not come only from pushing at high-redshifts. The distant SNe in the Hubble diagram need a reliable sample of low-redshift SNe, free from uncertainties and biases, to be anchored to. These SNe should be located in the smooth, nearby Hubble-flow. It has been shown [86] that between 300 and 900 type Ia SNe around $z \sim 0.05$ are required. To this aim, dedicated surveys have been carried out in the past and others are in progress (e.g., Nearby Supernova Factory (SNFactory)[10]), but still we are far from reaching this goal.

Even in the future, the physics of the SN explosions will be best studied in the nearby objects for which plenty of photons are available. But discovering local and Hubble-flow SNe is, in some respect, more difficult than picking up high$-z$ SNe, which can be discovered on-demand with the wide-field imagers on large telescopes while vast areas of sky need to be searched to sample sizable volumes of local Universe. Presently, the largest contribution to the discovery of the local SNe comes from automated surveys on targeted galaxies (e.g., Lick Observatory SN Search (LOSS)[11]) or from a vast array of amateur astronomers. The situation will change when a new generation of dedicated, wide-field survey telescopes will enter into operation in both hemispheres. The PanSTARRS[12] will soon begin to

[8] See DESTINY in web page list.

[9] See SNAP in web page list.

[10] See SNFactory in web page list.

[11] See LOSS in web page list.

[12] See PanSTARRS in web page list.

survey the full available sky from Hawaii with its prototype PT1 1.8 m telescope. One day PanSTARRS will be a system of four telescopes surveying the sky once per week, producing thousands of SNe per year, for hundreds of which, a detailed monitoring with other mid-size telescopes will be possible. In the southern hemisphere SkyMapper, a 1.3 m telescope featuring a 5.7 square degree camera at Siding Spring Observatory will provide soon about 100 nearby SNe-Ia per year just using the leftover time from other programs [76], and then again one day there will be the whole-sky LSST survey.

Thanks to frequent visits, all sky surveys will provide not only a wealth of bright targets discovered soon after the explosion but also multicolor light curves. The systematic spectral classification, the selection of the most promising targets, and the intensive follow-up will be impossible with current projects on conventional telescopes. There is therefore the need for 2 m-sized robotic telescopes, possibly equipped with integral-field spectrographs, operating in coordination. The Las Cumbres Observatory Global Telescope (LCOGT) network[13] is an interesting initiative in this direction.

But new epochal advances in our comprehension of the *phenomenon SN* are expected to come with the next generation of Extremely Large Telescopes (ELT), for example, European-ELT[14], which will combine larger collecting area to high spatial resolution. At that point, we will have direct information on the precursor stars and the local environment, together with enough photons to play with spectral modeling, to make spectrophotometry, high resolution, and late-time studies to an extraordinary precision.

Thanks a lot Massimo.
Although mainly devoted to the exploration of high energy processes in astrophysics, future X-ray missions are expected to play an important role also for cosmology. The next interview with Günther Hasinger will explain why.

5.8 New Perspectives in High Energy Astrophysics and Galaxy Clusters

Dear Günther *(Hasinger)***, what is your point of view concerning the relevance of next generation of X-ray space missions for cosmology?**

The future is indeed bright for X-ray astronomy. Several new projects have been approved or have made significant steps towards approval. On the European side, these are in particular the extended Röntgen Survey with an Imaging Telescope Array (eROSITA) aboard the Russian Spektr-Rntgen-Gamma mission, *Simbol*-X and XEUS.

[13] See LCOGT in web page list.
[14] See E-ELT in web page list.

eROSITA

The nature of the mysterious DE that is driving the Universe apart is one of the most exciting questions facing astronomy and physics today. It may be the vacuum energy providing the Cosmological Constant in Einstein's theory of GR, or it may be a time-varying energy field. The solution could require a fundamental revolution in physics. As discussed earlier, the discovery of DE has come from three complementary techniques, among which the study of large scale structure and in particular of clusters of galaxies plays an important role. The amount and nature of DE can be tightly constrained by measuring the spatial correlation features and evolution of a sample of about 100.000 galaxy clusters over the redshift range $0 < z < 1.5$. Such an X-ray survey will discover all collapsed structures with mass above $3.5 \times 10^{14} h^{-1} M_{\odot}$ at redshifts $z < 2$. Above this mass threshold, the tight correlations between X-ray observables and mass allow direct interpretation of the data. DE affects both the abundance and the spatial distribution of galaxy clusters. Measurements of the number density $\mathrm{d}^2 N/\mathrm{d}M\,\mathrm{d}z$ and the three-dimensional power spectrum P(k) of clusters are complementary and have different parameter degeneracies with respect to other DE probes, such as type Ia SNe or CMB anisotropies, and precisely constrain cosmological parameters [63]. In particular, a survey of 10^5 clusters of galaxies will allow to measure the "baryonic acoustic wiggles" imprinted on the power spectrum of primordial fluctuations, which gives an independent measurement rod for precision cosmology. As these quasi-periodic acoustic fluctuations are present in any baryonic component of the Universe, there may also be a chance to detect them at higher redshifts using large numbers of AGN sampled across the whole sky.

The X-ray telescope eROSITA aboard the Russian Spektr-Röntgen-Gamma mission will perform the first imaging all-sky survey in the medium energy X-ray range up to 10 keV with an unprecedented spectral and angular resolution [109]. In the energy band 0.5–2 keV, the eROSITA survey will have about a factor of 50 higher sensitivity then the ROentgen SATellite (ROSAT) All-Sky-Survey, and in the 2–10 keV band, it will be two orders of magnitude deeper than the High Energy Astrophysical Observatory (HEAO)–1 All-Sky-Survey. eROSITA will detect more than 3 Million AGN and in particular isolate systematically all obscured accreting BHs in nearby galaxies and many (>170,000) new, distant AGN in the hard (2–10 keV) band. eROSITA will detect the hot IGM of 50–100 thousand galaxy clusters and groups and hot gas in filaments between clusters to map out the large scale structure in the Universe and to find in particular the rare massive distant clusters of galaxies for the study of DE. But such a sensitive X-ray survey will also provide essential new data for a large variety of astrophysical Galactic topics to study in detail the physics of Galactic X-ray source populations, like pre-main sequence stars, SN remnants, and X-ray binaries. The effective area at 1 keV of seven eROSITA telescopes is about twice the effective area of one X-ray Multi Mirror Satellite XMM-*Newton* telescope, and its solid-angle-area-product ("grasp") is about three times that of all three XMM-*Newton* telescopes.

Simbol-X

The use of grazing incidence X-ray optics, first below 2–3 keV with the Einstein and ROSAT satellites, later up to 10 keV with the Advanced Satellite for Cosmology and Astrophysics (ASCA), Chandra, XMM-*Newton* and finally Suzaku, has led to a dramatic increase in sensitivity of X-ray telescopes and opened a huge new discovery space. This technique has been so far limited to energies below ∼10 keV. Hard X-ray and gamma-ray imaging instruments are so far utilizing coded mask imaging techniques, such as those aboard the International Gamma Ray Astrophysics Laboratory (INTEGRAL) or SWIFT missions. At energies above 10 keV, however, new emission mechanisms start to appear, which are associated with nonthermal components in the X-ray sources, both in AGN and in the hot plasma of SN remnants and possibly clusters of galaxies. Also, the most heavily obscured Compton-thick sources appear only above 10 keV. A clear requirement for future high energy astrophysics missions is therefore to bridge this gap of sensitivity by offering an instrumentation in the hard X-ray range with a sensitivity and angular resolution similar to that of the current imaging X-ray telescopes. To do this, a hard X-ray focusing optics is needed. Such an optics can readily be implemented by a simple extension of the current X-ray mirror technology to small grazing angles and thus long focal lengths. *Simbol*-X is a hard X-ray mission, operating in the ∼0.5–80 keV range, proposed as a collaboration between the French and Italian space agencies with participation of German laboratories for a launch in 2014 [53]. Relying on two spacecraft in a formation flying configuration, *Simbol*-X uses for the first time a 20 m focal length X-ray mirror to focus X-rays with energy above 10 keV, resulting in over two orders of magnitude improvement in angular resolution and sensitivity in the hard X-ray range with respect to nonfocusing techniques. *Simbol*-X will serve as a first demonstrator for the technique of formation flying, adding impetus to technology development that will also benefit XEUS. Because of its higher angular resolution, *Simbol*-X will also be significantly more sensitive than the Nuclear Spectroscopic Telescope Array (NuStar) (by NASA) and New X-ray Telescope (NeXT) mission (by Japan Aerospace Exploration Agency/Institute of Space and Astronautical Science; JAXA/ISAS), which are planned in a similar time frame.

XEUS

XEUS, the X-ray Evolving Universe Spectroscopy mission, is one of the three large missions selected for study by ESA within the ESA Cosmic Vision program. It represents ESA's next generation X-ray observatory and will provide a facility for high-energy astrophysics fully complementary to other major future observatories operating across the electromagnetic spectrum such as SKA, Atacama Large Millimeter/submillimeter Array (ALMA), JWST, ELT, and Atacama Cosmology Telescope (ACT), but also to planned future particle and gravitational wave detectors (km^3 NEutrino Telescope (KM3NET) and LISA, respectively). The XEUS concept envisages a pair of spacecrafts in a formation flying configuration in which an X-ray telescope of novel design and unprecedented collecting area feeds a suite of state-of-the art instruments. Formation flying will be at the heart of many large astronomical observatory planned for the future and thus will receive a technological push in the near future. The huge improvement in sensitivity compared to current X-ray telescopes, coupled with a high spatial and spectral imaging capability, will make XEUS a unique facility for studying high-energy phenomena and processes over the full span of the observable Universe. Of particular importance are its contribution to the Evolution of Large Scale Structure and Nucleosynthesis, where XEUS will enable studies of the genesis of groups and clusters of galaxies and the Cosmic Web at up to $z \sim 2$, and the evolution of the physical state and chemical abundances of the IGM, as well as the coeval evolution of galaxies and their supermassive BHs, where XEUS will be able to study the birth and growth of supermassive BHs at $z \sim 10$ and their accretion and spin history thereafter.

Thank you Günther.
As discussed earlier, future X-ray observations of galaxy clusters will provide rich cosmological information. Physical processes in galaxy clusters are successfully studied in a multifrequency approach and we expect that this will be even more fruitful in the future. Isabella Gioia will comment on this point in the next interview.

Dear Isabella *(Gioia)*, **in Chap. 2 you discussed the astrophysical properties and the cosmological information of galaxy clusters. To your opinion, what kind of new observations will allow a significant progress in this field?**

The SZE will soon be used as a new band for detecting clusters at high redshift. SZE surveys will be a tremendous source of new information in the near future. In particular, surveys like the South Pole Telescope (SPT)[15] [117] or the Atacama Cosmology Telescope (ACT)[16] [80]) will produce catalogs of clusters unbiased in redshift. Some of the planned SZE instrumentation is now reality. I am thinking of the Sunyaev-Zel'dovich Array (SZA), an eight-element interferometer that enables one to achieve high sensitivity with respect to single dish observations even for extended low-surface brightness emission. During the commissioning period, the SZA demonstrated that it can be used to study distant ($z \geq 1$) clusters [97].

The soon to be flown *Planck* satellite[17] will extend our knowledge of the CMB beyond the limits set by past and present experiments (for instance WMAP). *Planck* will survey the whole sky and will provide a large dataset of clusters expected to be at $z > 1$. Blind SZE surveys in the near future will discover thousands of clusters. As the SZE signal is independent of redshift, the limit of such surveys will be a mass limit. Such cluster surveys can be used to determine cosmological parameters with

[15] See SPT in web page list.

[16] See ACT in web page list.

[17] See PLANCK in web page list.

high precision. The *Planck* mission will likely lead us to a full comprehension of the CMB temperature anisotropies and it will be crucial as a test of the robustness of the CDM Concordance Model.

The nonthermal components of clusters of galaxies will be revealed by the future radio telescopes. When arrays like the SKA[18], the Long Wavelength Array (LWA)[19], or the LOFAR[20] will become operational, they will reveal new radio halos especially in distant clusters. One can then be able to compare the statistics between the observational data and the expectations from models of cluster and structure formation. The combination of radio and hard X-ray data will be crucial to measure the energy content in the form of relativistic electrons and magnetic field in the Intra Cluster Medium (ICM). The proposed new generation hard X-ray telescope *Simbol*-X[21] (a jointly supported Italian–French mission with German participation), which will operate in the 0.5–80 keV, is expected to reveal and map the nonthermal emission in clusters of galaxies.

I believe the time is certainly mature to have a new medium-depth X-ray all-sky survey of clusters carried out with a dedicated satellite with a good point spread function (similar or better than XMM), optimized optics for wide-field X-ray imaging and low background. An all-sky survey, and its associated large sample of clusters, would be crucial to investigate the relationship between X-ray observables and masses. In addition, many new clusters at high redshift will be discovered. We need more objects to observe and study. Several ideas for such a survey have been proposed by the scientific community to the various space agencies, but none has been approved so far. In the meantime, we have to make do with the invaluable archives of both *Chandra* and XMM-*Newton*, which are providing interesting new results (see among others [21, 22, 50, 99, 145, 148]) and with ongoing X-ray cluster surveys like the XMM Cluster Survey (XCS) [115] or the XMM-LSS survey [106]. The first will produce a catalog of several thousand serendipitously detected clusters in over 500 square degree to beyond $z = 1$. See [118] for a recent paper forecasting the constraints on the values of Ω_m, σ_8, and cluster scaling relations parameters expected from the XCS survey. The second survey, the XMM-LSS, has recently produced a combined analysis of weak lensing and X-ray blind surveys [15]. Meanwhile, the continuing program of *Chandra* and XMM observations will contribute to increase the cluster statistics. The *Planck* satellite will provide new large datasets of clusters identified through the SZE. These will be new targets for the future X-ray observatories like the NASA mission Constellation-X[22] and the ESA mission XEUS[23] that will allow us to carry out more precise studies on the nature and content of the DM and DE of the Universe.

[18] See SKA in web page list.

[19] See LWA in web page list.

[20] See LOFAR in web page list.

[21] See *Simbol*-X in web page list.

[22] See Constellation-X in web page list.

[23] See XEUS in web page list.

Thank you very much Isabella.
Let us consider now observations in the highest frequency band. Astronomy in γ-rays allows to explore the most violent processes in astrophysics and extremely bright sources, as GRB, that can be detected up to very high redshifts. Why not use them to trace cosmic evolution? We will hear the opinion of Günther Hasinger in this respect.

Dear Günther *(Hasinger)*, **GRB may open a new window for cosmological studies. Since the discovery of GRB050904 at $z = 6.3$, measured by VLT and** *Subaru*, **these phenomena may offer the opportunity to investigate the environment in which they develop, the properties of the interstellar medium, and the cosmic abundances up the reionization epoch. Would you comment on this possibility? Do you believe that GRB can be used as standard distance candles to probe the expansion and properties of the very distant Universe?**

Because of their high intrinsic luminosity, gamma-ray bursts are indeed among the populations with the highest redshifts, in competition only with Quasi Stellar Objects (QSOs) and more recently protogalaxies. As such, they provide the chance to detect some of the earliest objects in the Universe, and to pinpoint and study their environment. Indeed, almost all of the future large astrophysical observatories aim at studying the Universe at redshifts $z \sim 10$ and GRBs are one of the few possibilities to localize such distant objects. Already in the current generation of GRBs from SWIFT, a substantial fraction of bursts at redshifts higher than six are expected ($\sim$5%). The fact that so far we have detected only one of these (GRB050904) is mainly due to the extreme difficulty of identifying and localizing them. At such high redshifts, the optical light of the GRB is completely wiped out by the Hydrogen Lyman absorption, so that fast and sensitive observations in the near-infrared range of the electromagnetic spectrum are required. Indeed, almost half of the SWIFT bursts are "optically dark" bursts in the sense that their afterglows are not detectable with the on-board optical monitor telescope or other fast reaction telescopes on the ground.

The 7-channel imaging camera Gamma-ray burst Optical/Near-infrared Detector (GROND) [61] was specifically designed for rapid follow-up observations of gamma-ray burst afterglows. It allows simultaneous imaging in the Sloan $g-r-i-z$ and near-infrared $J - H - K$ bands. In this way, simultaneous spectral energy distributions can be measured and fast photometric redshifts derived. GROND was commissioned at the MPI/ESO 2.2 m telescope at La Silla (Chile) in April 2007, and since then performs rapid follow-up observations of GRB and other fast transients. Interesting high-redshift candidates are used to trigger immediate follow-up observations at the VLT or other large telescopes. It is only a matter of time until the redshift distribution of gamma bursts above $z = 3$ will be measured well enough to constrain the fraction of very high redshift bursts.

As the question poses, GRB have also been proposed as standard candles with which the geometry of the Universe can be charted at very high redshifts, similar to what has been achieved using type Ia SNe at lower redshifts. Personally, I am very skeptical about this possibility, because the detailed properties of GRB are so varied and by far not fully understood. However, as discussed earlier, GRB are potentially the best tracers of matter and protogalaxies at the highest redshifts and thus may act as pathfinders for high redshift cosmological studies.

Thanks a lot Günther.
We have seen above how the global properties and evolution of the Universe and the cosmic development of the structures it contains will be studied through dedicated observations of stars, galaxies, and clusters of galaxies at different frequency domains, from radio wavelengths to γ-rays. When cosmic structures are formed, they also offer the opportunity to test cosmological models through the lensing effect they produce. In the next interview with Matthias Bartelmann, we will continue the discussion started in Chap. 3, but focusing on the future prospects in this field.

5.9 Cosmological Expectations from Lensing

Dear Matthias *(Bartelmann)***, could you comment if lensing can be used to test standard and nonstandard metric theories? Could accurate observations of lensing provide new fundamental insights about theories of modified gravity?**

Light deflection in gravitational fields follows already from the equivalence principle and is therefore more fundamental than general relativity (GR). The mere fact that light deflection occurs is therefore no proof of GR or any other metric theory of gravity. The specific dynamics of a particular metric theory, expressed by its field equations relating the metric to the matter–energy content of the space-time, must then determine by how much a certain distribution of matter or energy can deflect an actual light path. In the weak-field limit, GR relates the amount of deflection to the Newtonian gravitational potential, by which both space and space–time are curved.

Any metric theory of gravity must repeat the impressive successes of Newtonian gravity in the Solar system and its neighborhood. The field equation of Newtonian gravity is the Poisson equation to which the field equations of any metric theory must return when restricted to weak gravitational fields. Also, light deflection by the Sun has been measured to an impressive accuracy, in particular by using radio interferometry to measure how the positions of distant radio sources relative to the Sun change as the Sun moves past them in the course of a year. Thus, light deflection within the Solar system is not only qualitatively, but also quantitatively determined.

Apart from lensing phenomena, alternative metric theories must allow the construction of cosmological models that must be required to be similarly successful as the standard Friedmann models in GR. In particular, they must explain the existence and appearance of the CMB, primordial nucleosynthesis, the expansion behavior of the Universe determined from type Ia SNe, the evolution of linear and nonlinear cosmic structures, patterns in the distribution of galaxies, and so forth.

Therefore, we must expect of any nonstandard metric theory that it has the right weak-field limit and that it reproduces the appearance of the Universe at large. Theories obeying these limits are quite hard to test in terms of gravitational lensing.

Let us consider one specific example. The Tensor–Vector–Scalar (TeVeS) theory was proposed by Bekenstein as an alternative metric theory satisfying the same construction principles as GR, but avoiding the need for DM. It is a generalization of Modified Newtonian Dynamics, which asserts that Newton's law of gravity must be modified at gravitational acceleration levels below some small threshold. This threshold was introduced ad hoc to allow an explanation of stellar dynamics in galaxies without the need for DM.

Compared to Modified Newtonian Dynamics, TeVeS has the advantage that it enables us to construct cosmological models as well as to model light deflection by localized masses in these same cosmologies. It has the substantial disadvantage of leading to a nonlinear Poisson equation describing the lensing effects, which is hard to solve and severely hampers the construction of lens models. Currently, it seems that lenses in TeVeS either required implausibly massive neutrinos, or DM on top of its tensor, vector, and scalar fields. Should lens models for specific observed lens systems indeed be able to prove that TeVeS cannot avoid DM despite having been constructed for this purpose, it would lose much if not all of its credit. We do not seem to be as far yet.

Summarizing, it seems fair to say that lensing can indeed test metric theories of gravity, provided the lensing mass and its spatial distribution are known well enough. In this way, it can for instance constrain Jordan–Brans–Dicke theories in the Solar system to a degree that makes them indistinguishable from Einstein's GR. On larger, cosmological scales, however, it seems that lensing effects necessarily need to be combined with additional information on the amount and the distribution of lensing masses before alternative metric theories can be constrained.

Do you believe in the possible detection of cosmic strings by means of weak lensing?

Cosmic strings can be seen as one-dimensional defects in space–time, giving rise to very peculiar features. Light rays coming from objects that we see behind a cosmic string can arrive at our position going around either side of the string. Such objects would thus appear doubly imaged, with the string passing between the images and one image being mirror-reversed compared to the other.

Any object behind a string would appear in this way. As the string would run across the sky like a curved line, it would be accompanied by pairs of mirror images straddling it. Off and on, close pairs of galaxies were interpreted as arising from lensing by a string, but none of these cases could so far be considered as a convincing example. In particular, neighboring galaxy pairs in the continuation of the claimed strings could not be demonstrated. Most likely, these cases were similar, but physically different, individual galaxies.

Could you tell us about the proposed future missions dedicated to a better understanding of the lensing phenomenon relevant for the development of cosmology?

Without any doubt, weak gravitational lensing by large-scale structures has significantly been detected, and relevant constraints of cosmological parameters have been derived from the results. But these measurements are very difficult. First of all, very many small and faint distant galaxies need to be detected and their shapes be measured because the cosmic lensing effect changes the radii of images by at most a few percent only, and because galaxies are intrinsically irregularly shaped. Second, imperfections of the optical systems used for weak-lensing surveys cause additional distortions. Tiny astigmatism of the telescope optics, detectors somewhat out of focus, slight tilt of detector frames out of the focal plane and similar effects cause secondary image distortions that have of course nothing to do with the cosmological shear signal. Third, turbulence in the Earth's atmosphere, familiar from the twinkling of the stars on the night sky, blurs images on angular scales comparable to the sizes of the small galaxy images.

Even if it was possible to precisely measure and correct all these effects, two main uncertainties remain. First, cosmological parameters can be derived only if the distance distribution of the faint galaxies is known well enough. Second, nonlinear evolution of the cosmic structures needs to be modeled at least as precisely as the intended measurement because it affects the weak-lensing signal on angular scales below approximately 10 arcmin.

Accordingly, there are currently four major challenges: surveys need to be extended to measure the shear from substantially larger data sets than now, optical imperfections need to be determined and corrected at the best possible level, the distance distribution of the lensed galaxies needs to be measured more precisely, and our understanding and quantitative description of nonlinear structure formation needs to be improved.

Why should all this be done? The main motivation is the sensitivity of lensing to the evolution of cosmic structures over time, which is in turn determined by the possibly dynamical behavior of the DE. Weak cosmological lensing is one of the few indicators promising to allow precise determinations of the cosmic dynamics, provided the measurement accuracy as well as the theoretical understanding can be improved substantially beyond the current level.

This will be possible only if the change of the shear signal with the distance to the sources can be measured, because this will reflect how the lensing structures themselves grow over cosmic time. So many distant galaxies need to be observed that they can be grouped into distance bins without compromising the measurement accuracy of the cosmic shear.

Three requirements for future surveys can immediately be derived from here. First, the area covered by the surveys need to be substantially increased. While surveys covering of order 100 square degrees are currently at the forefront, surveys of 1,000 or more square degrees are being planned. Within a few years, ground-based surveys will be able to cover of order 10^4 square degrees with the required depth and accuracy. The whole sky has $\sim 4 \times 10^4$ square degrees, and a good fraction of it will be covered with suitable observations within years.

The second challenge concerns the precision of the imaging optics, or at least the accuracy to which any imperfections can be discovered and removed. Much progress has already been achieved in this regard, but there is still a long way to go on the way towards precise and mathematically sound, rather than merely empirically calibrated, techniques for correcting imaging imperfections. This needs to go hand in hand with the development of more reliable ways of measuring the shapes of faint and irregular images with the goal of precise shear determinations.

The third challenge concerns the measurement of source distances. While it is simply impossible to take spectra and measure redshifts for all or many of the images used for shear measurements, photometric redshift information can be collected. Photometric redshift determinations can be compared to spectroscopy at low resolution. Image fluxes are measured in several frequency bands and compared with the fluxes expected from certain galaxy types moved through redshift space. The redshift where the expected collection of fluxes matches the observed one, the best is assigned to the object as a photometric redshift. Accuracies can be sufficiently high, provided one of the bands is in the near infrared because otherwise ambiguities occur.

A satellite mission has been proposed to the ESA under the provisional name of DUNE, for Dark Universe Explorer. This satellite is designed specifically for obtaining optical images deep enough to detect faint distant galaxies well on all the sky not affected by the emission and absorption of the Milky Way galaxy. Its several filter bands will allow the measurement of sufficiently precise photometric redshifts of the background galaxies to improve our knowledge of their distance distribution and to group them reliably into different redshift bins so as to separately analyze how the weak-lensing signal increases with the distance to the sources.

This satellite mission, if approved, combined with the much improved ground-based surveys upcoming in the next years, is almost guaranteed to improve our knowledge of weak cosmological lensing to a degree sufficient to use it for precise constraints on the dynamics of structure formation, and therefore on the enigmatic DE.

Thank you very much Matthias.
As seen earlier, lensing could be used to probe cosmic strings. In the next interview, Ruth Durrer will go more deeply into the controversial problem of cosmic string detection through non-CMB observations, and will summarize her opinion on future perspectives.

5.10 Future Tests for Topological Defects

Dear Ruth *(Durrer)*, **what are in your opinion the most promising cosmological observations, different from the CMB, to test topological defect models?**

In Chap. 3, we have analyzed what we may learn about cosmological topological defects from observations of the CMB. Here we discuss some other cosmological observations that may lead to the detection of topological defects, especially cosmic strings.

We concentrate on cosmic strings also as they are much better motivated than global defects. It actually has been shown that the majority on Grand Unified Theory (GUT) inflationary scenarios do have a symmetry breaking phase transition at the end of inflation, which leads to the formation of cosmic strings [70]. In addition, the possibility of cosmic superstrings that could be a relevant consequence of string theory and might lead to a network with quite different properties than ordinary cosmic string has revived this subject of research.

Furthermore, lensing and gravitational radiation, the two most promising effects of cosmic strings other than the CMB, are unimportant for global defects. Lensing by global monopoles or textures is highly improbable and global defects do not emit significant amounts of gravitational radiation. They decay predominantly via the emission of massless Goldstone bosons.

5.10.1 What are the Characteristics of Lensing by Cosmic Strings?

The metric of a straight cosmic string is flat, but has a conical singularity at the position of the cosmic string, that is, the azimuthal angle around the cosmic string goes only from zero to $2\pi(1 - 4\mathrm{GM}^2)$. The geometry of the planes normal to the string is conical with a deficit angle $\Delta = 8\pi\mathrm{GM}^2 \simeq (\mathrm{GM}^2/10^{-6})5$ arcsec. A source lying behind the string (sufficiently close to the line of sight between us and the cosmic string) is therefore seen twice; the images are separated by an angle Δ as shown in Fig. 5.20.

The lensing probability depends on GM^2, as the source is lensed if it lies within an angle Δ of the line, which is normal to the cosmic string and connects it with the observer. In addition, of course, the number density of long strings is relevant, which is typically of the order of a few, say five per horizon volume, but independent of GM^2. Furthermore, strings are not entirely straight but do have wiggles and a finite curvature radius. Recent analyses of the problem [60, 89] have shown that by

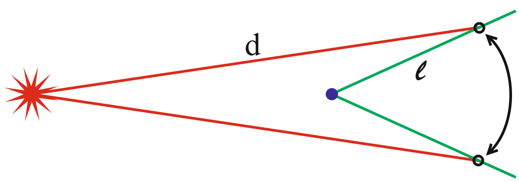

Fig. 5.20 The plane normal to the cosmic string (•) is shown. A source (*star*) behind a cosmic string seen by an observer (*circle*). In this representation, the observer is at two positions, which have to be identified (*double arrow*). The observer sees the source in the two direction, the two lines connecting the observer positions and the source. These directions are separated by an angle $\Delta = 8\pi\mathrm{GM}^2$

scanning an area of 1 square degree, one should see typically $8(\mathrm{GM}^2/10^{-6})$ events if one has an angular resolution better than about $(\mathrm{GM}^2/10^{-6})3$ arcsec. Therefore, if $\mathrm{GM}^2 > 10^{-7}$, one should see a lensing event by a long cosmic string already in the Hubble Space Telescope (HST) Advanced Camera for Survey (ACS) data (HST-ACS archive). In [60], it is argued that planned radio surveys with angular resolution of 0.028 arcsec can achieve the limit $\mathrm{GM}^2 < 10^{-8}$ if they see no lensing event within a few square degrees of the sky. On the other hand, if no lensing events from loops are seen in future radio surveys such as SKA, one might be able to reduce this limit even to $\mathrm{GM}^2 < 10^{-9}$ [89]. This limit, coming from loops, however, depends strongly on the assumptions made on the population of cosmic string loops, which are quite uncertain [33, 108] as we shall discuss later.

It is interesting to mention that a few years ago, a pair of galaxies has been interpreted as a lensing event by a cosmic string. In the mean time, however, high resolution observations with the HST have revealed that the object in question is a physical galaxy pair [119].

Also global monopoles and texture have interesting lensing signatures, but their frequency is small as these are point-like or even event-like objects and it is very improbable that we detect them via lensing.

5.10.2 Would Cosmic Strings Lead to an Observable Gravitational Wave Background?

Already very early on, it has been argued that cosmic string loops, which oscillate with relativistic speeds emit gravitational waves [48, 142]. The typical gravitational wave power from a loop has been estimated to be $\Gamma \mathrm{GM}^4$, with $\Gamma \sim 50$, a parameter that depends on the loop configuration. The spectrum and amplitude of this stochastic gravitational wave background has been estimated and it has been found that, depending on the value of GM^2, this background might be seen as random fluctuations in the timing of binary pulsars. These are binary systems containing a rapidly rotating magnetized neutron star that emits beamed synchrotron radiation in the radio band, which may pass the earth with a frequency of a few milliseconds.

A somewhat uncertain theoretical analysis of the stochastic gravitational wave emission from cosmic string loops gives [108, 146]

$$\Omega_{\mathrm{g}} h^2 \sim 3 \times 10^{-3} (\alpha \mathrm{GM}^2)^{1/2} \tag{5.1}$$

for frequencies in the interval $10^{-16}\,\mathrm{Hz}/(\mathrm{GM}^2) < \omega < 3 \times 10^{-11}\,\mathrm{Hz}/(\mathrm{GM}^2)$. Here $\Omega_{\mathrm{g}} = \rho_{\mathrm{g}}/\rho_{\mathrm{c}}$ is the energy density in gravitational waves divided by the critical energy density of the Universe, $\rho_{\mathrm{c}} = 3H_0^2/8\pi G \simeq 10^4\,\mathrm{eV\,cm^{-3}\,h^2}$ and h is the reduced Hubble constant. The measured value is $h = 0.72 \pm 0.1$. The parameter α is loop size divided by the horizon size at formation. This number is very uncertain and its most optimistic value is $\alpha \simeq 0.1$. One expects the value or α to depend on the details of the cosmic string network, for example, it will be different for cosmic superstrings and for the usual cosmic strings from symmetry breaking in the Abelian Higgs model.

The first thorough analysis of the available data [74] yields a limit for the density parameter in gravitational waves of $\Omega_g h^2 < 6 \times 10^{-8}$ at angular frequency $\omega \simeq 1/\text{year} \simeq 3 \times 10^{-8}$Hz. More recent results give about $\Omega_g h^2 < 4 \times 10^{-8}$ at the same frequency [71]. Using the above prediction with $\alpha = 0.1$, this results in the most stringent upper limit for the string energy scale, $\text{GM}^2 \lesssim 10^{-9}$.

The problem here is that the typical loop size parametrized by α is very uncertain. It ranges from $\alpha = 0.1$ down to the prediction that the loops reduce their size by inter-commutation until they reach a size that corresponds roughly to the string thickness, after which they decay by the emission of gauge bosons [147]. If this is correct, the gravitational wave production from cosmic strings is irrelevant. But the gauge bosons emitted from the collapsing loops lead to a background of cosmic rays that constrains the symmetry breaking parameter, leading to a similar value, $\text{GM}^2 < 10^{-9}$. The most recent limits come from the AUGER experiment [137]. However, more recent simulations and analytical [102, 108] arguments indicate two populations, one of large loops containing about 10–20% of the string length and one that consists of very small loops which are chopped off near the "cusps", the typical points which cosmic string loops exhibit. If this is the correct interpretation, gravitational waves from cosmic strings still lead to the presently best upper limit of $\text{GM}^2 \lesssim 10^{-8}$. The evolution of a network of superstring is still quite unknown.

To Summarize

We have seen in Sect. 3.3 of Chap. 3 that CMB anisotropies and polarization constrain topological defects, leading to a limit of $\text{GM}^2 \lesssim$ a few $\times 10^{-7}$. As discussed here, also gravitational lensing presently gives similar limits. These limits can be improved by about a factor of 10 in the foreseeable future, or they lead to the detection of cosmic strings if $\text{GM}^2 \gtrsim 10^{-8}$.

Let us compare CMB constraints with other present and future observations that constrain topological defects. The constraints from gravitational radiation are uncertain as long as the question about the cosmic string loop population is not resolved. Are there sizable cosmic string loops that are significantly larger than $L_f \sim \alpha t_f$ with $\alpha > \Gamma \text{GM}^2$, where L_f and t_f indicate the length and time at formation? If this is satisfied, cosmic strings emit a considerable gravitational wave background, which can be measured by advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) even if $\text{GM}^2 \simeq 10^{-9}$ [108]. However, if there are no loops of appreciable size and, especially, if they become of order $L \sim 1/M$ and then decay by emission of gauge bosons, cosmic strings do not emit a substantial gravitational wave background. This is in my opinion the most important theoretical problem left to solve: what is the characteristic length distribution of the cosmic string loop population formed by the Kibble mechanism during a symmetry breaking phase transition? And especially what are the characteristics, the typical loop sizes, of a network of superstrings?

Fortunately, the CMB results do not strongly depend on the loops as has been shown in [59] and we expect them to be reliable.

It is also interesting to note that there is no way either with present or with planned future experiments to observe cosmic topological defects if their energy scale is too low, $M \lesssim 10^{14}$ GeV. The reason for this is the weakness of gravity, or what amounts to the same thing, the huge Planck mass. Therefore, topological defects can be detected in the future only if the energy scale of the symmetry breaking phase transition that has generated them lies in the relatively narrow window of 4×10^{15} GeV$> M > 10^{14}$ GeV.

This is the main reason why, when asked to bet whether we shall ever see cosmic strings in the sky, I would bet NO. What a shame ...

Thanks a lot Ruth for your passionate discussion, in spite of this pessimistic but frank conclusion.
We conclude our interviews with a discussion with Lucia Popa on a transversal topic, linked to fundamental physics, that could benefit from various kinds of cosmological observations.

5.11 New Perspectives for Neutrino from Astrophysical Cosmology

Dear Lucia *(Popa)***, to your opinion, what are the projects in cosmology most promising to answer to the current questions on neutrino physics?**

The quality of WMAP 5 year data [79] has shown the importance of CMB data as a probe of the neutrino mass. There are several other projects in operational or design stage, aimed mainly at improving the sensitivity towards small scales and polarization anisotropies. To measure with precision the neutrino masses, it will be useful to combine some large scale structure observations with the best possible CMB dataset to reduce as much as possible parameter degeneracies. We can expect some very interesting results regarding the neutrino mass from a number of ground-based experiments mapping the CMB anisotropies only in a small regions with excellent sensitivity and resolution: ATC [117], Background Imaging of Cosmic Extragalactic Polarization (BICEP) [75], designed for large angular scale, and Q and U Extra-galactic Sub-mm Telescope (QUEST) at Degree Angular Scale Interferometer (DASI) (QUaD) [27], designed for small angular scales. These experiments are optimized for measuring E-polarization or even B-polarization anisotropies, which will still be poorly constrained after the completion of WMAP and even after *Planck* in the case of B-polarization.

A second set of experiments is scheduled in Antarctica at the French–Italian Concordia station and in the Atacama plateau in Chile, for unprecedented precision measurements of the B-mode for $l < 1{,}000$ (which is particularly useful for probing the primordial gravitational waves from inflation): the B-modes Radiation measurement from Antarctica with a bolometric Interferometer (BRAIN) [105] instrument for measuring large scales, and the CℓOVER [91] instrument for intermediate scales. The *Planck* satellite [134] has already been built and will be launched in 2009 by the ESA. The temperature sensitivity of *Planck* will be so good that it can be thought as the ultimate observation in the sector of temperature anisotropies. With *Planck* it will be possible to improve the measurement of E-polarization on small angular scales but B-polarization will be poorly constrained. On intermediate scales, the ground-based experiments should be quite efficient, but progress will still be needed on very large scales (requiring full-sky coverage) and very small scales (requiring both high resolution and excellent sensitivity). Beyond *Planck*, at least two space projects are under investigation. B-Pol [20] is targeted for large scales aiming to improve the Cℓ ObserVER (CℓOVER) measurement of the B-mode for $l < 1{,}000$, thanks to its full-sky coverage and slightly better sensitivity. Inflation Probe [19] is a NASA satellite project aiming to make the ultimate measurement of E-polarization, like *Planck* for temperature, and a very good measurement of B-polarization on all scales. The satellites experiments as *Planck* and B-Pol or the combination of them will reach a 1σ error on the neutrino mass in the range 0.3–0.4 eV, the same order of magnitude as the current neutrino mass bounds, while the Inflation Probe mission will provide a sensitivity on neutrino mass ranging from 0.15 to 0.25 eV.

The Galaxy Redshift Surveys allow for the reconstruction of the matter power spectrum. The models for nonlinear structure formation show that the relation between the reconstructed power spectrum and the total matter linear power spectrum should be scale-independent at least up to a wavenumber $k_{\rm max} \sim 0.15 h\,{\rm Mpc}^{-1}$, where nonlinear corrections start to induce scale-dependent biasing.

Beyond SDSS, there are various projects for larger surveys, like for instance the SDSS-II LEGACY project and the Advanced Large Homogeneous Area Medium Band Redshift Astronomical (ALHAMBRA) Survey [95]. On October 2007, two candidate missions addressing the study of DM and DE have been selected by ESA for a further assessment and consideration for launch in 2018: DUNE and SPACE[24]. While they propose to use different techniques, DUNE is proposed as a wide-field imager and SPACE is proposed as a near-infrared all-sky surveyor, they address the same basic science goal: measure of the variations of the linear growth factor to get some good handle on the DE variables. In the follow-up study phase, a trade-off will be performed leading to the definition of a proposal for a European DE mission. These survey will go to such high redshift that it will be possible to compare the observed power spectrum with the linear one up to wave-numbers significantly larger than the usual $k_{\rm max} \sim 0.15\,h\,{\rm Mpc}^{-1}$ and to reconstruct the matter power spectrum in various redshift bins, corresponding to different times in the evolution of the Universe. These measurements will also be ideal for neutrino mass extraction, as massive neutrinos induce a very peculiar redshift-dependence on the matter power

[24] The ESA space project coming from the merging of DUNE and SPACE has been subsequently called *Euclid*.

spectrum. The sensitivity of large redshift surveys alone to the neutrino mass gives essentially no information on m_ν because of the various parameter degeneracies. For the SDSS+WMAP combination, the 1σ error on the neutrino mass is of the order 0.3 eV, while for SDSS+*Planck* is of order 0.2 eV.

The issue of degeneracies between the neutrino mass and other parameters, such as those describing the DE evolution, needs further investigation, the data being affected by a significant degeneracy between m_ν and w. In breaking this degeneracy, it is useful to be taken into account the fact that the CMB and Large Scale Structure (LSS) data are not statistically independent. Because of the late Integrated Sachs–Wolfe (ISW) effect, the power spectrum of temperature anisotropies encodes some information on neighboring structures, like galaxy clusters. So, there is a nonzero cross-correlation between temperature and galaxy maps, which has actually been already measured [58] with a rather low level of significance.

Limits on neutrino mass can also be obtained by using future CMB experiments only, assuming that the lensing power spectrum could be extracted. In this way, it is possible to combine information on CMB acoustic oscillations and on the surrounding LSS of the Universe at redshifts of order 3, avoiding many of the complicated features of galaxy redshift surveys related to the mass-to-light bias and to the strongly nonlinear evolution of small-scale perturbations at small redshift. For example, by using the CMB lensing extraction, the *Planck* satellite will achieve a sensitivity of 0.15 eV (1σ) on the neutrino mass [73], which shows that this technique should improve the global sensitivity in a rather spectacular way at least for this parameter. In the case of the CMB Polarization (CMBPol) mission project, an impressive 1σ error of 0.044 eV is predicted [84].

The galaxy weak lensing, or the cosmic shear, changes the apparent shape of galaxies: galaxies that would be apparently spherical without weak lensing can look in fact elliptical, stretched in one direction, and squeezed in the orthogonal direction. This effect, which is coherent over the angular size of the lensing gravitational, is called cosmic shear and can be detected by using very dense sample of galaxies obtained with enough resolution to measure each individual shape. To reconstruct the lensing gravitational field, it is also needed to know either the redshift of each source galaxy or at least their number density in the redshift space and in the direction. There are various ongoing and planned cosmic shear surveys, like for instance the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) 20, the SNAP, the Pan-STARRS, the LSST. The sensitivity of this type of observations to neutrino masses has been addressed [1, 32, 65], showing that the most ambitious hypothetical experiment of this type combined with the *Planck* data can reach a sensitivity of 0.045 eV, while combined with CMBPol can reach a sensitivity of 0.027 eV, which means that a 2σ detection of the neutrino mass would occur even for the smallest m_ν of order 0.05 eV.

Thank you very much Lucia.

References

1. Abazajian, K., Dodelson, S.: Neutrino mass and dark energy from weak lensing. Phys. Rev. Lett. **91**, 041301 (2003)
2. Abbott, L.F., Wise, M.B.: Constraints on generalized inflationary cosmologies. Nucl. Phys. **B 244**, 541–548 (1984)
3. Albrecht, A., et al.: Causality, randomness, and the microwave background. Phys. Rev. Lett. **76**, 1413 (1996)
4. Alcock, C., et al.: Possible gravitational microlensing of a star in the large magellanic cloud. Nature **365**, 621 (1993)
5. Arnaboldi, M., et al.: VST: VLT survey telescope. The Messenger **93**, 30 (1998)
6. Astronomy and astrophysics survey committee, commission on physical sciences, mathematics, and applications, national research council http://www.nap.edu/catalog.php?record_id=9839
7. Baade, D., et al.: The wide field imager at the 2.2-m MPG/ESO telescope: First views with a 67-million-facette eye. The Messenger **95**, 14 (1999)
8. Ball, N.M., et al.: Galaxy types in the sloan digital sky survey using supervised artificial neural networks. Mon. Not. R. Astron. Soc. **348**, 1038 (2004)
9. Baltay, C., et al.: The QUEST large area CCD camera. Pub. Astron. Soc. Pacific **119**, 1278 (2007)
10. Banday, A.J., Zaroubi, S., Bartelmann, M.: Mining the sky. Springer, Berlin (2001)
11. Bardeen, J.M., Steinhardt, P.J., Turner, M.S.: Spontaneous creation of almost scale-free density perturbations in an inflationary universe. Phys. Rev. **D 28**, 679–693 (1983)
12. Beck, R.: Measurements of cosmic magnetism with LOFAR and SKA. Adv. Radio Sci. **5**, 399 (2007)
13. Bennett, C.L., et al.: First-year Wilkinson microwave anisotropy probe (WMAP) observations: Preliminary maps and basic results. Astrophys. J. Suppl. **148**, 1–27 (2003)
14. Benoit, A., Caussignac, M., Pujol, S.: New types of dilution refrigerator and space applications. Physica **B 197**, 48 (1994)
15. Berg, J., et al.: Combined analysis of weak lensing and X-ray blind surveys. Mon. Not. R. Astron. Soc. (in press) [arXiv:0712.3293]
16. Bergstrom, L., et al.: Clumpy neutralino dark matter. Phys. Rev. **D 59**, 043506 (1999)
17. Bertin, E., Arnout, S.: SExtractor: Software for source extraction. Astron. Astrophys. Suppl. Ser. **117**, 393 (1996)
18. Bock, J.J., et al.: A novel bolometer for infrared and millimeter-wave astrophysics. Space Sci. Rev. **74**, 229–235 (1995)
19. Bock, J., et al.: Task force on cosmic microwave background research (2006) [astro-ph/0604101v1]
20. BPol collaboration (2007); http://www.b-pol.org/index.php 727–737 (2007)
21. Branchesi, M., et al.: High redshift X-ray galaxy clusters. II. The $L_X - T$ relationship revisited. Astron. Astrophys. **472**, 739–748 (2007)
22. Branchesi, M., et al.: High redshift X-ray galaxy clusters. I. The impact of point sources on the cluster properties. Astron. Astrophys. **472** (2007)
23. Brunner, R.J., Djorgovski, S.G., Szalay, A.S.: Virtual observatories of the future. ASP Conf. Ser. **225** (2001)
24. Capaccioli, M.: Astronomy with Schmidt-type telescopes. LXXVIII IAU Colloquium, Reidel, Dordrecht (1983)
25. Cavuoti, S.: Classification of AGN using support vector machines. PhD thesis, University of Napoli Federico II (2007)
26. Charbonneau, D., et al.: Detection of planetary transits across a sun-like star. Astrophys. J. **529**, L45 (2000)
27. Church, S., et al.: QUEST on DASI: A south pole CMB polarization experiment. New Astron. Rev. **47**, 1083 (2003)

28. Ciardi, B., Madau, P.: Probing beyond the Epoch of Hydrogen Reionization with 21 Centimeter Radiation. Astrophys. J. **596**, 1 (2003) [arXiv:0303249]
29. Ciardi, B., Ferrara, A., White, S.D.M.: Early reionization by the first galaxies. Mon. Not. R. Astron. Soc. **344**, L7 (2003)
30. Colafrancesco, S., Profumo, S., Ullio, P.: Multi-frequency analysis of neutralino dark matter annihilations in the coma cluster. Astron. Astrophys. **455**, 21 (2006)
31. Colombi, S., Mellier, Y., Raban, B.: Wide field surveys in cosmology. Editions Frontieres, Gif sur Ivette (1998)
32. Cooray, A.R.: Weighing neutrinos: Weak lensing approach. Astron. Astrophys. **348**, 31 (1999)
33. Copeland, E.J., Kibble, T., Steer, D.: Evolution of a network of cosmic string loops. Phys. Rev. **D58**, 043508 (1998) [hep-ph/9803414]
34. Cunow, B., et al.: Photometric CCD sequences for calibration of the ESO/SERC atlas. Astron. Astrophys. Suppl. Ser. **125**, 71 (1997)
35. D'Abrusco, R., et al.: Mining the SDSS archive. I. Photometric redshifts in the nearby universe. Astrophys. J. **663**, 752 (2007)
36. D'Abrusco, R., Longo, G., Walton, N.A.: Quasar candidates selection in the virtual observatory era. MNRAS, submitted [astro-ph/08050156]
37. Da Costa, L., et al.: ESO imaging survey: Past activities and future prospects. The Messenger **98**, 36 (1999) [astro-ph/9912224]
38. de Bernardis, P., et al.: A flat universe from high-resolution maps of the cosmic microwave background radiation. Nature **404**, 955–959 (2000)
39. de Bernardis, P., et al.: B-Pol: Detecting primordial gravitational waves generated during inflation. Exp. Astron. **23**(1), 5 (2009) [arXiv:0808.1881]
40. Dickinson, M., et al.: The unusual infrared object HDF-N J123656.3+621322. Astrophys. J. **531**, 624 (2000)
41. Diemand, J., Kuhlen, M., Madau, P.: Dark matter substructure and gamma-ray annihilation in the milky way halo. Astrophys. J. **657**, 262 (2007)
42. Diemand, J., Kuhlen, M., Madau, P.: Formation and evolution of galaxy dark matter halos and their substructure. Astrophys. J. **667**, 859 (2007)
43. Djorgovski, S., et al.: The palomar digital sky survey (DPOSS). In: Wide field surveys in cosmology. Colombi S., Mellier Y., Raban B. (eds.), p. 89 (1998)
44. Djorgovski, S., et al.: Searches for rare and new types of objects. In: Virtual observatories of the future. Brunner, R.J., Djorgovski, S., Szalay, A.S. (eds.), ASP Conf. Ser. **225**, p. 52 (2001)
45. Djorgovski, S., et al.: The palomar-quest digital synoptic sky survey. Astron. Nachr. **329**, 263 (2008)
46. Donalek, C.: Neural methods for the Star/galaxy separation. PhD thesis. University of Napoli Federico II (2006)
47. Dunkley, J. et al.: Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from WMAP Data. Astrophys. J. Suppl. **180**, 306 (2009)
48. Durrer, R.: Gravitational angular momentum radiation of cosmic strings. Nucl. Phys. **B 328**, 238 (1989)
49. Emerson, J., et al.: Visible and infrared survey telescope for astronomy: Prog. Report. The Messenger **126**, 41 (2006); **117**, 24 (2004)
50. Ettori, S., et al.: Scaling laws in X-ray galaxy clusters at redshift between 0.4 and 1.3. Astron. Astrophys. **417**, 13–27 (2004b)
51. Fan, X., et al.: High-redshift quasars found in sloan digital sky survey commissioning data. Astron. J. **118**, 1 (1999)
52. Fan, X., et al.: The discovery of a high-redshift quasar without emission lines from sloan digital sky survey commissioning data. Astrophys. J. **526**, L57 (1999)
53. Ferrando, P., et al.: Simbol-X: Mission overview. SPIE **6266E**, 11 (2006)
54. Finkbeiner, D.P., Davis, M., Schlegel, D.J.: Extrapolation of galactic dust emission at 100 microns to cosmic microwave background radiation frequencies using FIRAS. Astrophys. J. **524**, 867–886 (1999)

55. Fixsen, D.J., Mather, J.C.: The spectral results of the far-infrared absolute spectrophotometer instrument on COBE. Astrophys. J. **581**, 817–822 (2002)
56. Fixsen, D.J., et al.: The cosmic microwave background spectrum from the full COBE FIRAS data set. Astrophys. J. **473**, 576–587 (1996)
57. Fixsen, D.J., et al.: ARCADE 2 measurements of the extra-galactic sky temperature at 3–90 GHz. Submitted to Astrophys. J. (2009) [arXiv:0901.0555]
58. Fosalba, P., Gaztanaga, E., Castander, F.J.: Detection of the integrated Sachs-Wolfe and Sunyaev-Zel'dovich effects from the cosmic microwave background-galaxy correlation. Astrophys. J. Lett. **597**, L89–L92 (2003)
59. Fraisse, A., et al.: Small-sngle CMB temperature anisotropies induced by cosmic strings. Phys. Rev. **D 78**, 043535 (2008) [arXiv:0708.1162]
60. Gasperini, M.A., et al.: Direct observation of cosmic strings via their strong gravitational lensing effect: I. Predictions for high resolution imaging surveys. Accepted for publication in MNRAS (2008) [arXiv:0710.5544v2]
61. Greiner, J., et al.: GROND – a 7-Channel Imager. Pub. Astron. Soc. Pacific **120**, 405–424 (2008) [arXiv:0801.4801]
62. Guth, A.H., Pi, S.Y.: Fluctuations in the new inflationary universe. Phys. Rev. Lett. **49**, 1110–1113 (1982)
63. Haiman, Z.: An X-ray galaxy cluster survey for investigations of dark energy (2005) [arXiv:0507013]
64. Hambly, N.A., et al.: VISTA data flow system survey access and curation: The WFCAM science archive. In: Optimizing scientific return from astronomy through information technologies. Quinn, P.J., Bridger, A. (eds.). Proc SPIE 5493, 423 (2004)
65. Hannestad, S., Tu, H., Wong, Y.Y.: Measuring neutrino masses and dark energy with weak lensing tomography. J. Cosmol. Astropart. Phys., JCAP06(2006)025
66. Hawking, S.W.: The development of irregularities in a single bubble inflationary universe. Phys. Lett. **B 115**, 295–297 (1982)
67. Hirata, C.M., Seljak, U.: Reconstruction of lensing from the cosmic microwave background polarization. Phys. Rev. **D 68**, 083002 (2003)
68. Interpretation of the extragalactic radio background. (2009) [arXiv:0901.0559]
69. Ivezic, Z., et al.: American Astronomical Society, AAS Meeting #213, #460.03; Bulletin of the American Astronomical Society, Vol. 41, 366 (2009)
70. Jeannerot, R., Rocher, J., Sakellariadou., M.: How generic is cosmic string formation in SUSY GUTs. Phys. Rev. **D 68**, 103514 (2003) [hep-ph/0308134]
71. Jenet, F., et al.: Upper bounds on the low-frequency stochastic gravitational wave background from pulsar timing observations: Current limits and future prospects. Astrophys. J. **653**, 1571 (2006) [astro-ph/0609013]
72. Jones, N., Stoica, H., Tye, S.H.: The production, spectrum and evolution of cosmic strings in brane inflation. Phys. Lett. **B 563**, 6 (2003) [hep-th/0303269]
73. Kaplinghat, M., Knox, L., Song, Y.S.: Determining neutrino mass from the cosmic microwave background alone. Phys. Rev. Lett. **91**, 241301 (2003)
74. Kaspi, V., Taylor, J., Ryba, M.: High-precision timing of millisecond pulsars. 3: Long-term monitoring of PSRs B1855+09 and B1937+21. Astrophys. J. **428**, 713 (1994)
75. Keating, B.G., et al.: BICEP: A large angular scale CMB polarimeter. SPIE **4843**, 284 (2003)
76. Keller, S.C., et al.: The SkyMapper telescope and the southern sky survey. Pub. of the Astron. Soc. of Australia **24**, 1 (2007)
77. Kennefick, J.D., et al.: The discovery of five quasars at $z > 4$ using the second palomar sky survey. Astron. J. **110**, 78 (1995)
78. Kogut, A., et al.: ARCADE 2 observations of galactic radio emission. Submitted to Astrophys. J. (2009) [arXiv: 0901.0562]
79. Komatsu, E., et al.: Five-year Wilkinson microwave anisotropy probe (WMAP) observations: Cosmological interpretation. Astrophys. J. Suppl. **180**, 330–376 (2009) [arXiv:0803.0547v1]
80. Kosowsky, A.: The atacama cosmology telescope. New Astron. Rev. **47**, 939–943 (2003)
81. Kuhlen, M., Diemand, J., Madau, P.: The shapes, orientation, and alignment of galactic dark matter subhalos. Astrophys. J. **671**, 1135 (2007)

82. Kuijken, K., et al.: OmegaCAM: The 16x16k CCD camera for the VLT survey telescope. The Messenger **110**, 15 (2002)
83. Leggett, S., et al.: The missing link: Early methane ("T") dwarfs in the sloan digital sky survey. Astrophys. J. **536**, L35 (2000)
84. Lesgourgues, J., Pastor, S., Perotto, L.: Probing neutrino masses with future galaxy redshift surveys. Phys. Rev. **D 70**, 045016 (2004)
85. Linde, A.: Chaotic inflation. Phys. Lett. **B 129**, 177–181 (1983)
86. Linder, E.V.: Importance of supernovae at $z < 0.1$ for probing dark energy. Phys. Rev. **D 74**, 103518 (2006)
87. Lipovetsky, V.A.: The importance of wide-field imaging. In: Astronomy from wide field imaging. IAU Symp. 161, p.3. Kluwer, Dordrecht (1994)
88. MacGilliwray, H.T., et al.: Astronomy from wide-field imaging. Kluwer, Dordrecht (1994); Space Sci. Rev. **76/3**, 363 (1996)
89. Mack, K.J., Wesley, D.H., King, L.J.: Observing cosmic string loops with gravitational lensing surveys. Phys. Rev. **D 76** 123515 (2008) [astro-ph/0702648]
90. Madau, P., Meiksin, A., Rees, M.J.: 21 centimeter tomography of the intergalactic medium at high redshift. Astrophys. J. **475**, 429 (1997)
91. Maffei, B., et al.: CLOVER: The CMB polarization observer. EAS Pub. Ser. **14**, 251 (2005)
92. Masi, S., et al.: Instrument, method, brightness, and polarization maps from the 2003 flight of BOOMERanG. Astron. Astrophys. **458**, 687 (2006)
93. McLean, B.J., et al.: New horizons from multi-wavelength sky surveys. Kluwer, Dordrecht (1998)
94. Metcalf, R.B., Madau, P.: Compound gravitational lensing as a probe of dark matter substructure within galaxy halos. Astrophys. J. **563**, 9 (2001)
95. Moles, M. et al.: The ALHAMBRA Survey: For a systematic study of cosmic evolution (2005) [astro-ph/0504545]
96. Moore, B., et al.: Globular clusters, satellite galaxies and stellar haloes from early dark matter peaks. Mon. Not. R. Astron. Soc. **368**, 563 (2006)
97. Muchovej, S., et al.: Observations of High-Redshift X-Ray Selected Clusters with the Sunyaev-Zel'dovich Array. Astrophys. J. **663**, 708 (2007)
98. Mukhanov, V.F., Chibisov, G.V.: Quantum fluctuations and a nonsingular universe. J. Exp. Theor. Phys. Lett. **33**, 532–535 (1981)
99. Mullis, C.R., et al.: Discovery of an X-ray-luminous galaxy cluster at $z = 1.4$. Astrophys. J. Lett. **623**, L85–L88 (2005)
100. Nolta, M. et al.: Five-year Wilkinson microwave anisotropy probe (WMAP) observations: Angular power spectra. Astrophys. J., submitted
101. Okamoto, T., Hu, W.: CMB lensing reconstruction on the full sky. Phys. Rev. **D 67**, 083002 (2003)
102. Olum, K., Vanchurin, V.: Cosmic string loops in the expanding universe. Phys. Rev. **D 75** 063521 (2007) [astro-ph/0610419v3]
103. Peacock, J.A., et al.: Report by the ESA-ESO Working Group on Fundamental Cosmology (2006) [astro-ph/0610906]
104. Perlmutter, S. et al., Discovery of a supernova explosion at half the age of the universe. Nature **391**, 51 (1998)
105. Piat, M., et al.: Precise measurement of CMB polarisation from Dome-C: the BRAIN and CLOVER experiments. In: Semaine de l'Astrophysique Francaise. Combes, F., Barret, D., Contini, T., Meynadier, F., Pagani, L. (eds.) EdP-Sciences, Conf. Ser., p. 707 (2004)
106. Pierre, M., et al.: The XMM-large scale structure catalogue: X-ray sources and associated optical data. Mon. Not. R. Astron. Soc. **382**, 279–290 (2007)
107. *Planck* Collaboration: The Scientific Programme of *Planck*. ESA-SCI(2005)1 [astro-ph/0604069]
108. Polchinski, J.: Cosmic string loops and gravitational radiation. In: Proceedings of the 11th Marcel Grossmann Meeting on General Relativity (2007) [arXiv:0707.0888v2]
109. Predehl, P., et al.: eROSITA. SPIE **6266E**, 19 (2006)

110. Reid, I.N., et al.: The second palomar sky survey. Pub. Astron. Soc. Pacific **103**, 661 (1991)
111. Renzini, A., Da Costa, L.: The ESO imaging survey. The Messenger **87**, 23 (1997)
112. Richards, G.T., et al.: Colors of 2625 quasars at $0 < z < 5$ measured in the sloan digital sky survey photometric system astron. J. **121**, 2308 (2001)
113. Richards, G.T., et al.: Spectroscopic target selection in the sloan digital sky survey: The quasar sample. Astrophys. J. **123**, 2945 (2002)
114. Richards, G.T., et al.: Efficient photometric selection of quasars from the sloan digital sky survey: 100000 $z < 3$ quasars from data release one. Astrophys. J. Suppl. Ser. **155**, 257 (2004)
115. Romer, A.K., et al.: Cluster Survey with XMM: Expected Catalog Properties and Scientific Applications. Astrophys. J. **547**, 594–608 (2001)
116. Rubakov, V.A., Sazhin, M.V., Veryaskin, A.V.: Graviton creation in an inflationary universe and the grand unification scale. Phys. Lett. **B 115**, 189–192 (1982)
117. Ruhl, J., et al.: The South Pole telescope. Proc. of SPIE **5498**, 11 (2004)
118. Sahlén, M., et al.: The XMM cluster survey: Forecasting cosmological and cluster scaling-relation parameter constraints. [arXiv:0802.4462v1] (2008)
119. Sazhin, M., et al.: Gravitational lensing by cosmic strings: What we learn from the CSL-1 case. Mon. Not. R. Astron. Soc. **376**, 1731 (2007) [astro-ph/0611744]
120. Schmidt, B., et al.: The high-Z supernova search: Measuring cosmic deceleration and global curvature of the Universe using type IA supernovae. Astrophys. J. **507**, 46 (1998)
121. Seiffert, M., et al.: 1/f noise and other systematic effects in the Planck-LFI radiometers. Astron. Astrophys. **391**, 1185–1197 (2002)
122. Sesar, B., et al.: Exploring the variable sky with the sloan digital sky survey. Astron. J. **134**, 2236 (2007)
123. Singal, J., et al.: The ARCADE 2 Instrument. (2009) [arXiv:0901.0546]
124. Smoot, G.F., et al.: Structure in the COBE differential microwave radiometer first-year maps. Astrophys. J. Lett. **396**, L1–L5 (1992)
125. Springel, V., Hernquist, L.: The history of star formation in a Λ cold dark matter universe. Mon. Not. R. Astron. Soc. **339**, 312 (2003)
126. Stadel, J.: Cosmological N-body simulations and their analysis. PhD thesis, University of Washington (2001)
127. Staiano, A.: Probabilistic principal surfaces and PPS ensembles. PhD thesis, University of Salerno (2005)
128. Starobinsky, A.A.: Spectrum of relict gravitational radiation and the early state of the universe. J. Exp. Theor. Phys. Lett. **30**, 682–685 (1979)
129. Starobinsky, A.A.: Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. Phys. Lett. **B 117**, 175–178 (1982)
130. Strolger, L.: The rate of supernovae in the local universe. PhD thesis, University of Michigan (2003)
131. Swade, D.A., Hopkins, E., Swam, M.S.: HST data flow with on-the-Fly Reprocessing. In: ASP Conf. Ser. ADASS X. **238**, 295 (2001)
132. Szalay, A.A., et al.: Microsoft Research Technical Report MS-TR-99-30 (1999) [astro-ph/9907009]
133. Szkody, P., et al.: Two rare magnetic cataclysmic variables with extreme cyclotron features identified in the sloan digital sky survey. Astrophys. J. **583**, 902 (2003)
134. Tauber, J.A.: The Planck mission. Adv. Space Res. **34**, 491 (2004)
135. The AKARI satellite web site, http://www.ir.isas.jaxa.jp/AKARI/Outreach/index_e.html
136. The Herschel satellite web site, http://herschel.esac.esa.int/overview.shtml
137. The Pierre Auger Collaboration: Upper limit on the cosmic-ray photon flux above 10^{19} eV using the surface detector of the Pierre Auger observatory. Astropart. Phys. **29**, 243 (2008)
138. The QUIJOTE experiment web site, http://www.iac.es/project/cmb/quijote/index_esp.php
139. Traitement Élémentaire, Réduction et Analyse des PIXels de Megacam. http://terapix.iap.fr/
140. Udalski, A., et al.: The optical gravitational lensing experiment. Discovery of the first candidate microlensing event in the direction of the Galactic Bulge. Acta Astron. **43**, 289 (1993)

141. URL: http://www.ivoa.org/
142. Vachaspati, T., Vilenkin, A.: Formation and evolution of cosmic strings. Phys. Rev. **D 30**, 2036 (1984)
143. Valentijn, E.A., et al.: Astro-WISE: Chaining to the Universe. In: ASP Conf. Ser., ADASS XVI **376**, 491 (2007)
144. Venemans, B.P., et al.: The discovery of the first luminous $z \sim 6$ quasar in the UKIDSS large area survey. Mon. Not. R. Astron. Soc. **376**, L76 (2007)
145. Vikhlinin, A., et al.: Chandra sample of nearby relaxed galaxy clusters: Mass, gas fraction, and mass-temperature relation. Astrophys. J. **640**, 691–709 (2006)
146. Vilenkin, A., Shellard, E.P.S.: Cosmic Strings and Other Topological Defects. Cambridge University Press, Cambridge (1994)
147. Vincent, G., Hindmarsh, M., Sakellariadou, M.: Scaling and small scale structure in cosmic string networks. Phys. Rev. **D 56**, 637 (1997) [astro-ph/9612135]
148. Voevodkin, A., Vikhlinin, A.: Constraining amplitude and slope of the mass fluctuation spectrum using a cluster baryon mass function. Astrophys. J. **601**, 610–620 (2004)
149. Volonteri, M., Haardt, F., Madau, P.: The assembly and merging history of supermassive black holes in hierarchical models of galaxy formation. Astrophys. J. **582**, 559 (2003)
150. Wade, L.A., et al.: The Planck mission 9. Advances in Cryogenic Eng. **45**, 499–506 (2000)
151. Wittman, D., et al.: Deep lens survey. Tyson, J.A., Wolff, S. (eds.). In: Astronomy with large telescopes: Survey and other telescope technologies and discoveries. Proc. SPIE **4836**, 73 (2002)
152. Yadav, A.P.S., Wandelt, B.D.: Detection of primordial non-Gaussianity (fNL) in the WMAP 3-year data at above 99.5% confidence, Phys. Rev. Lett. submitted
153. Yepes, G., et al.: Hydrodynamical simulations of galaxy formation: effects of supernova feedback. Mon. Not. R. Astron. Soc. **284**, 235 (1997)
154. York, D.G., et al.: The sloan digital sky survey: Technical summary. Astron. J. **120**, 1579 (2000)

Chapter 6
Concluding Remarks

Mauro D'Onofrio and C. Burigana

After so many pages of discussion on the present situation in cosmology, we still have, as editors, a crucial question to answer: can we conclude that the goals motivating the realization of this book have been achieved?

The first of our remarks concerns the structure of the book as a whole and the character and style of each interview. It is apparent that the presentation of some arguments is addressed in more than one interview, although in different perspectives, and the style of the book is far from homogeneous, some themes being more suitable for expert readers than others. The reasons of these peculiarities are at least two: (1) the questions we sent to each contributor are quite general, covering various astrophysical themes linked to the present cosmology and permitting a lot of freedom in the possible answers; (2) each author has personally interpreted the "Galilean spirit", as requested by the editors at the beginning of this enterprise, adopting a personal point of view to answer our proposed questions. The resulting lack of homogeneity may appear as a fault of the book, but we believe that it also has the positive aspect that readers of different scientific backgrounds may find it of interest.

In general, we are confident that, from the point of view of a reader, the book offers a panoramic view of the great efforts carried out in cosmology over the last decades. Although incomplete, both in selected arguments and references, the interviews included here provide a global description of the enormous research activity both in terms of experiments/observations and in the production of theoretical models. It is also evident from the various contributions that the linking of cosmology to a puzzle, proposed in our introduction, works appropriately. Remarkably, the current standard Concordance Model, in which the Universe is going through an accelerated expansion dominated today by the cosmological constant or by the energy density of the dark energy (DE), represents the emerging picture of a puzzle coming from different branches of astrophysics and observational cosmology. Indeed, Big Bang Nucleosynthesis (BBN) tests, Cosmic Microwave Background (CMB) studies, Large Scale Structure (LSS) observations, galaxy cluster properties, and Supernova (SN) searches, together with constraints from the Hubble Space Telescope (HST) cosmological distance ladder, Lyα forest, and luminous red galaxies, globally agree on a common set of cosmological parameter values, in spite of the particular sensitivity of each data set to a specific subset of them.

M. D'Onofrio and C. Burigana (eds.), *Questions of Modern Cosmology: Galileo's Legacy*, DOI 10.1007/978-3-642-00792-7_6,

The most uncomfortable shortcoming of the ΛCDM model is represented by the enormous fraction of the energy density of the Universe that is unknown (up to 96%, with 20% dark matter (DM) and 76% DE), while only about 4% of it is in the form of baryonic matter. Although the cosmological parameters of the model have been better and better determined and restricted in small intervals of possible values, the understanding of what the Universe is made of is actually escaping every theory. It is also very difficult to explain how, from the initial conditions, we arrive today at a Universe with energy density parameters, Ω_m and Ω_Λ, of the same order of magnitude, jointly providing an almost flat Universe. It may be reminiscent, in some way, of the flatness problem of the Universe "solved" by the inflationary scenario, shifted now towards the nature of the scalar field driving inflation. In this context, we can understand the increasing activity towards alternative models of gravity aimed at explaining the cosmological constant problem, or, in a more standard approach, towards tracking solutions for (quintessence) DE models, both of them trying to circumvent the fine-tuning argument.

At present, we have no more a clear idea of the evolution and fate of the Universe, this being eventually determined by the equation of state of DE, whatever it is, in each scenario.

Are the problems recalled above symptoms of an upcoming revolution in physics? In some way, certainly yes, whatever the indications coming from the future large set of experiments/observations/surveys.

We observe at first that this question has two faces: one properly refers to the scientific shift of paradigm, the other concerns the methodological aspects of working in cosmology. Historically, science develops thanks to two different and complementary approaches: one inductive, the other deductive. These two methodologies are particularly mixed in practice in cosmology and their complex interplay is likely the only possible way of working in astrophysics and cosmology today, in a sort of loop. This does not mean that they are completely equivalent and always had the same weight during the evolution of astronomy and cosmology. We could then argue that future answers and discoveries will not only impact purely on scientific aspects but will also have consequences for our methodological choices. This is not a mere abstract problem, but is relevant for a society that intends to invest in the future of astrophysics.

If the Concordance Model, with its detailed predictions, were to pass future high accuracy tests, we would have not only a fully exhaustive verification of the scientific paradigm change of the last decade (from the old CDM scenario to the ΛCDM model), but also an excellent example of the "scientific method", introduced by Galileo and Bacon, in which empirical evidence drives theoretical hypotheses. We cannot forget how much an unbiased, empirical approach to science, free from prejudices, contributed to the development of cosmology. In this case, although some observations/experiments could suggest a possible failure of the standard model of physics, it was the trust in it that motivated the original introduction of DM and DE. Their successful discovery in the near future (e.g., through the Large Hadron Collider (LHC) experiments which may detect the weak interacting massive particles (WIMP) DM candidates, or by astronomical determination of the equation of state *of DE with* $w \neq -1$*) will enforce the relevance of the "inverse approach" to the "scientific method", in which theoretical deductions largely anticipate the experimental confirmation. Without neglecting the crucial role in cosmology of advanced experiments and observations, this will be a strong argument against a purely empirical approach in science. This is the old question of the inductivist turkey made by Russell and settled by Popper with the statement that theory always precedes observations.*

If a direct identification of DM and DE and a deeper comprehension of their nature continues to escape our tests, the crisis of the standard model of physics is largely unavoidable, leading to a radical shift of scientific paradigm. In particular, the current interpretation of gravity should be greatly revisited. The way in which this will occur is unclear as yet. Nevertheless, in this case we will assist to a deeper trust in the "genuine version" of the "scientific method" as the best way to make significant progress in cosmology. It will be the revenge of those who strongly believe in a strict empirical approach in scientific activity, but, at the same time, it will require a much more profound revision of current theoretical physics. New models and ideas, however, should not only explain the data that prompted their formulation, but also have new predictive power towards subsequent generations of experiments and observations.

In principle, we should expect that, in both cases, the prevailing vision will be full of consequences in terms of the distribution of economical and human resources. In the former case, scientific development will be mainly driven by projects devoted to the exploitation of theoretical models in line with the most accredited scenarios that will be largely supported. Otherwise, a larger fraction of resources will be dedicated to research programs promoting views alternative to those that have become standard.

We believe that it emerges clearly from the book that progresses in the comprehension of the Universe come from a joint development of various branches of astrophysics and technology connected in various ways to cosmology. The best example of recent times is offered by type Ia SNe: despite the fact that the DE requirement was already present at the time of the crisis of the CDM scenario, as remarked by John Peacock and others in this book, the discovery that prompted the shift towards the Λ*CDM paradigm came, in large part, from the observations of these objects. This is, by the way, a warning message for all those who believe that only a few selected researches merit economical support from our society. The history of physics has demonstrated that the key to progress often comes from unpredictable fields of research.*

From a scientific perspective, the book has impartially given space both to experiments and theories supporting the standard cosmological scenario and alternative, even radical, ideas. We have seen, however, only a few examples of theories that claim an alternative explanation of gravity, accounting for the DM and DE phenomenology. A quick look at the Smithsonian Astrophysical Observatory/National Aeronautics and Space Administration (SAO/NASA) Astrophysics Data System (ADS) shows that the research activity in this field is extraordinary: in the first 10 months of 2008, more than 250 articles in journals with referees have in their title *the words "Dark Energy". A lot of work is still necessary to assess the global validity of the proposed models, despite their partial successes in accounting for certain astrophysical and cosmological data.*

As far as the scientific content of the included interviews is concerned, we decided to avoid as much as possible any comment on the single interviews, with some circumstantial exceptions. We make below some general remarks on specific points addressed in many of the interviews, with a certain attention to the implications for the future.

The first comment is that we are really impressed by the pulsating fervor shown by all the scientific projects carried out in almost every field of astrophysics, in the near past and foreseen for the near future, aimed at testing the various pieces of the puzzle of the standard cosmological model. Although the ΛCDM scenario has passed up to now several tests based on precise experiments and observations, the overall impression is that we are drawing a lot of cosmological conclusions from still limited and sometimes uncertain datasets. We believe that an exhaustive solution of the DM and DE problem will probably come from the cooperative efforts in different branches of astrophysics, rather than from specific projects that will provide more precise estimates of the cosmological parameters.

A substantial improvement will come from a better knowledge of the role of evolution of astrophysical systems during the history of the Universe, in particular by the comprehension of the complex nonlinear phases. Another way of saying the same thing is that we should better understand the role of transfer functions, which provide the detected astrophysical signal after the modifications introduced by evolution, environment, detectors, etc.

In this context, we have to stress that many efforts should still be made to bypass the "nature vs. nurture" question. This is true in every field of astrophysics, from SNe as distance indicators, to galaxy and cluster morphology and properties, up to the metallicity content and luminosity of stars, galaxies, and QSOs. What we actually miss is a clear identification of the progenitors at high redshifts of the present day astrophysical sources. We cannot be sure, for example, that type Ia SNe at $z \sim 1.5$, belonging to younger galaxies with more massive and rapidly evolving stars living in a low metallicity environment, can be compared with those found at low redshifts. The same can be said for galaxies and clusters, whose properties evolve (passively?) with time.

The nonlinear epoch of galaxy formation, the so-called dark and dawn ages, are extremely important for present day cosmology. If the new generation of large ground-based and space telescopes (e.g., the Extremely Large Telescope (ELT), the James Webb Space Telescope (JWST)) are able to detect the light of the first galaxies and, at the same time, if it is possible to constrain the properties of the high-redshift Intergalactic Medium (IGM) through the upcoming radio surveys, we will have made a crucial step forward in understanding the details of cosmological reionization.

It is also clear that the hierarchical DM scenario is too simplistic, and a lot of things still await for a common explanation: the dwarf/giant galaxy number ratio and distribution, the downsizing effect, the universal halo properties, and the Hubble *morphological sequence are some examples of the problems unsolved by the present framework of baryon collapse within DM halos.*

The combination of lensing and stellar dynamics seems promising in the definition of the properties of DM on small scales. The few data already available suggest the presence of inner cores within DM halos that could imply the existence of warm DM, for example, consisting of neutrinos in the keV mass range. It appears, however, that the difficulties associated with the mass measurements by means of strong and weak lensing are quite large and possibly biased by the unknown distances of the sources, which can be estimated only via redshifts assuming a cosmological model. The large number of projects in this direction will reveal if lensing is indeed the powerful dynamical probe that, coupled with the distribution of luminous mass, may provide the mass profile of galaxies and clusters at increasing redshifts.

We believe that the systematic studies of astrophysical sources at increasing steps in redshifts, although long and laborious, can give the key to understanding the complex aspects of evolution. Computer simulations are certainly important in this game, but their predictive power is still unsure. In some cases they have been tuned to explain some particular observation with ad hoc assumptions rather than to derive it from basic physical principles. Also, as recently recognized, they should be able to jointly include both gravitational, thermodynamical, and feedback processes, often difficult to implement in a single simulation at various physical scales, particularly in the nonlinear phase of the Universe's evolution. We recall that even the star formation process, from a primordial cloud, is a problem that has not been solved yet and the star formation history in galaxies is still poorly understood, although phenomenological approaches have provided some insight on this question. Major progress, however, is expected in the field of numerical simulations, thanks to the continuous development of both software and hardware. This progress, coupled with the better understanding and treatment of small scale physics, will greatly enhance the role of numerical simulations and the quality of their comparison with observations.

The development of our knowledge of the early Universe is another crucial point for cosmology. This book touches only the problem of the missing link of gravity with the other fundamental interactions in the context of quantum mechanics and the corresponding implications for cosmology. This question is so relevant and difficult as to merit by itself a dedicated book. The opinion of Gerard t'Hooft leaves a certain optimism about the possibility of solving this problem in the future within the framework offered by current particle physics theory. On the other hand, the general impression we get from the interviews about gravity on cosmological scales is that the phenomenon is currently not firmly understood. The various authors discussed the tests of General Relativity (GR) and the possible alternatives or generalization of it, presenting some of the ideas proposed up to now, particularly in connection with the problem of the cosmological constant and/or DE phenomena. It is likely that the foreseen revolution in physics will include a new way of looking at gravity, in which the quantum properties at the Planck epoch can be reconciled with the better known properties at larger length scales, within a unified scheme that will link gravity with the other forces.

The current emergence of the Anthropic Landscape, prompted by String Theory, could be a symptom of our ignorance about gravity and fundamental physics. Substantial improvements in this field will come slowly, until theoretical and observational progress reveals some further piece of the puzzle. The Anthropic Principle offers today a way to escape the Landscape labyrinth, but it could simply reflect our inability to link gravity with the other forces, rather than being a working physical hypothesis to be tested in the view of its predictive power. In the history of physics, the Anthropic view has always demonstrated such limits.

It is likely that the CMB will continue to provide the widest and deepest cosmological window to investigate the early phases of the Universe. In fact, although its properties carry integrated information on the whole history of the Universe up to current times, these are not so influenced by nonlinear processes affecting the formation and evolution of cosmic structures and astrophysical objects, as other cosmological observables.

Certainly, a big step forward will be the detection of the primordial stochastic background of gravitational waves or of their signature on the CMB polarization anisotropy through the discovery of primordial B-modes at large and intermediate angular scales. If ground-based balloon or space CMB experiments succeed in detecting them, leading to a firm determination of the tensor to scalar ratio of primordial perturbations, then the energy scale of inflation will be fixed and the set of possible inflationary models will be significantly restricted. Also, in the case of a lack of such detection, we could still constrain many models through more stringent upper limit analyses.

The control of systematic effects of instrumental origin, astrophysical foreground contamination, and their interplay is much more crucial for B-modes than for total intensity (or temperature), E-modes, and temperature–polarization cross-correlation anisotropies. In particular, the degree of polarization of the CMB is typically smaller than that of diffuse foregrounds, and the B-modes are subdominant with respect to E-modes in the case of the CMB. An extremely accurate knowledge of the astrophysical sources masking the CMB signal is then demanded. This further underlines the synergy between cosmological and astrophysical studies as a necessary step for progress, together with technological and data analysis improvements. Anyway, it is important to note that, even in the case of a missing of B-mode identification, the accurate measure of E-mode polarization anisotropy is of extreme relevance for breaking some degeneracies in the determination of cosmological parameters and to better determine them. It will contribute to a deeper understanding of several specific topics in cosmology, such as, for example, the cosmological reionization process and its link to the emerging of primeval objects and their subsequent evolution. This has been already indicated by the excellent results obtained by the Wilkinson Microwave Anisotropy Probe (WMAP) team.

The great focus on CMB polarization should not mask the relevance of continuously improving the maps of temperature anisotropies at different angular scales (or multipole regimes). Again, only the extremely accurate removal of systematics and foregrounds will permit the extremely promising analysis of the Gaussianity of the CMB pattern and to possibly find the small deviations from Gaussianity predicted *in many inflationary models, or to set more stringent constraints on them, as well as to better assess the problem of the large scale anomalies and of the geometry of the Universe. The intriguing, but extremely difficult, possibility of firmly identifying signatures from topological defects in the CMB, as claimed by some recent works, still has to catch the above kinds of problems.*

The forthcoming Planck *mission could significantly improve the current status of many of the topics mentioned earlier. At the same time, the Low Frequency Instrument (LFI) onboard* Planck *will also perform, to a certain extent, an absolute measurement of the CMB temperature at wavelengths between 1 and 0.43 cm, allowing an independent consistency test of the Cosmic Background Explorer/Far Infrared Absolute Spectrophotometer (COBE/FIRAS) results. We point out that while some observational details of the CMB still need to be better assessed and fully understood, the globality of its properties, the amplitude of the Sunyaev–Zel'dovich effects observed towards galaxy clusters, and the agreement with the ratio of baryon to photon number densities and light element abundances as derived in the BBN theory cannot be explained at present far from the fundamental recipes of the current cosmological model.*

Coming back to polarization projects dedicated to B-modes, we note that, independent of their success in terms of detection, they should also be able to carry out geometrical and Gaussianity analyses for the E-mode analogous to those currently carried out and planned for temperature anisotropies. We therefore expect that many of the problems still open after WMAP will find clear solutions.

We also have to remember the relevance of conducting CMB research covering, possibly with a combination of different experiments, both large, intermediate, and small angular scales. This is crucial, for example, to accurately separate the B-modes induced by lensing from the primordial modes. Some controversial questions about the existence and amplitude of primordial magnetic fields (or, more precisely, of the amplitude of their fluctuations) could be better assessed by looking at small scales, which are also of extreme relevance for studying the statistical signatures of the thermal and kinetic Sunyaev–Zel'dovich effect.

Another problem we see comes from the ever more complex procedures of data analysis in connection with data storage and handling. Verification and reproducibility of scientific results are an important step for scientific progress, besides a fundamental requirement of the scientific method. How it will be possible in the future to exhaustively verify, check, and reproduce, or possibly confute, the scientific results achieved through the efforts of a very large team by means of complex methods of data analysis is a nontrivial problem. This could potentially become a new big problem to solve. This may occur for two reasons: first the technical difficulty of collecting both infrastructure and human resources necessary to carry out such a huge work of data analysis; and second, the work needed to verify results already obtained by another group may sometimes be perceived as unattractive in the context of making new discoveries. Clearly, both these aspects are mitigated by the fact that often the analysis of the same set of data can be performed by various groups through different and complementary methods, possibly not on the whole chain of the data analysis procedure but on different specific pieces of it, more *difficult or more strongly dependent on underlying assumptions. Note also that this in part happens within large teams, as a step necessary to cross-check the results and their generality and to arrive at agreed and robust conclusions. In general, the feasibility of this important verification work relies on the availability of the various crucial intermediate products of the data analysis procedure, which should then be organized limiting as much as possible a "black box" approach to the production and delivery of the scientific results.*

Another goal of this book is the discussion on the relationship between science and society and the self-organization of the research activity in astronomical communities. We have tried to open here such a discussion. Although these themes are far from the everyday work of scientists, several colleagues have taken advantage of this possibility and have expressed their opinions. We are happy that the discussion has seen different points of view. A recognition common to different interviews is that even in our times scientific research can be limited in its progress by several influences and constraints operating in society. Readers living today the experience of scientific research are often conscious that it is not easy to escape social and economic constraints. We should not forget that research activity today is a cooperative effort and that we are quite far from the epoch of Galileo when the kind of collaboration and links among scientists were certainly not comparable with those enjoyed by contemporary scientists. Certainly, the possibility to exchange data and ideas is today faster and more efficient and represents a great opportunity for researchers. In this context, it seems a paradox that in parallel to this freedom and spirit of collaboration, we assist sometimes to the creation of self-referential research groups, almost univocally oriented towards common frameworks, leaving in practice little room for alternative ideas. This could be particularly dangerous for young researchers who may suffer not only the conditioning of the scientific context in which they work but also the difficulty in finding support far from well established projects, in an age potentially rich in new creativity. This could limit the development of science. It may sound strange, but a "Galileo case" may happen again, although in a different way. We believe that, particularly today and independently of the size of the project/collaboration, a fruitful management of human resources should properly balance the necessity of achieving the goals of a given program with those of effectively allowing and stimulating a certain freedom and creativity of young researchers.

The astronomical and physical communities are therefore asked to look at their inner self-organization not only to share in a democratic and intelligent way the economical resources, but also to avoid blind preclusions and hostile attitudes to alternative ideas.

We hope that the fruitful collaboration between particle physicists and astrophysicists, who have already reached several important scientific results as pointed out by Simon White, will continue, respecting the peculiarities of both fields of research. An effort should be made in this direction to get the required synergies that may permit several astrophysical questions to be addressed. This is obviously particularly important for some topics, as, for example, those connected with neutrino physics and their role in astrophysics and cosmology, jointly addressed by very *different kinds of experiments and observations. Also, this could become more and more important in the case of the discovery of the particle(s) responsible for DM. We expect in fact that it will call for a more motivated and precise treatment of the physical properties of DM particles in revised, more accurate astrophysical models.*

In general, the most critical problem in all scientific communities is that related to the financial support of research activities, both in absolute quantitative value and in continuity of available resources. In spite of the promises of governments and formal agreements among them, in too many countries the economical support for research and development is still only $\simeq$1% of the Gross Domestic Product. We can add to this, that, sometimes, astrophysical and cosmological studies, as well as other branches of fundamental research, are perceived by several people as too far from more urgent needs of our society, although the appeal of our science is widely recognized. May be, we need to dedicate more efforts to enforce the perception not only of the beauty of our researches, but also of their middle-long term relevance for the technological and cultural development of society. This means that it is important to remember continuously the numerous applications of our science already present in many widely used technological devices (e.g., Global Positioning System (GPS), imaging techniques for diagnostic and Earth mapping, new kinds of receivers, new informatic and numerical methods, etc.), and its role in the global perception of mankind in the Universe.

Looking at the variety of current and future projects, we can identify several broad categories: those devoted to a wide class of astrophysical/cosmological problems and those aimed at the solution of a single or of an extremely restricted set of questions; those mainly focussed on the detection and/or analysis of phenomena predicted by the current model and/or already preliminarily observed, and those mainly devoted to fundamental tests. Note also that in each region of the electromagnetic spectrum, a given project may belong to one or more of these categories and that the relative relevance of the various themes addressed by a given multi-purpose program can change from the time of its proposal to that of its realization.

For example, in the context of CMB projects, the Planck *mission will provide data of extreme interest for a variety of cosmological and astrophysical questions, some of them left open or even raised by WMAP and by recent observations and almost ignored at the time of Cosmic Background Radiation Anisotropy Satellite – Satellite for Measurement of Background Anisotropies (COBRAS–SAMBA, the previous name of the* Planck *mission) proposal. In contrast, the next generation of polarization anisotropy experiments focussed on the discovery of primordial B-modes (like, e.g., B-Polarization Satellite Mission (B-Pol), CMB Polarization (CMBPol) Mission, etc.) are strongly motivated by a single, crucial task and their full scientific return is a bet (being unknown the tensor-to-scalar ratio of primordial perturbations), at least in the "positive sense" of a detection that will be rich of consequences for cosmology and fundamental physics. From the other side, a significantly better measure of the shape of the CMB spectrum at long wavelengths (see, e.g., the Absolute Radiometer for Cosmology, Astrophysics, and Diffuse Emission (ARCADE)) will represent a fundamental consistency test for the cosmological model, as well as the detection at frequencies already covered by FIRAS of the very* *small deviations from a Planckian shape that are expected to be generated during the dark and dawn ages (the main goal of some of the ideas presented by John Mather).*

Similar considerations hold for infrared and optical surveys. Ground-based projects, achievable, for example, with the VLT (Very Large Telescope) Survey Telescope (VST) at optical wavelengths and the Visible and Infrared Survey Telescope for Astronomy (VISTA) in the near infrared, are clearly multipurpose in nature. They will provide crucial data for dedicated astronomical studies, at the same time improving our comprehension of a large variety of astrophysical problems, many of them closely linked to fundamental cosmological questions and to the evolution of classes of astrophysical objects. To a certain extent, we can also say that their capabilities of solving controversial astrophysical questions might result in a necessary step for constraining cosmological models using the same data.

In the same context, several space projects are and will be suitable to the investigation of different kinds of observables, as, for example, the Advanced Dark Energy Physics Telescope (ADEPT), designed to study Baryonic Acoustic Oscillations (BAO) and high-redshift SNe, or Dark Universe Explorer (DUNE), aimed at SNe searches and weak lensing analyses. In some cases, a single mission may merge original ideas targeted to different specific aims, as, for example, Euclid*, combining DUNE with the Spectroscopic All-sky Cosmic Explorer (SPACE), the latter focussed to BAO and various evolutionary aspects of galaxies and clustering. In contrast, other projects, like, for example, the SuperNova Acceleration Probe (SNAP) dedicated to the DE equation of state through the discovery of high-redshift SNe, have been designed to provide a clear answer to a specific, crucial, and well defined cosmological question, providing in parallel a significantly better understanding of the corresponding astrophysical framework.*

Typically, low and high energy projects (like, e.g., the Low Frequency Array (LOFAR), the Square Kilometer Array (SKA), the extended Röntgen Survey with an Imaging Telescope Array (eROSITA), Symbol-X*, and the X-ray Evolving Universe Spectroscopy (XEUS) mission) are more targeted to several crucial astrophysical questions, like, for example, reionization, cosmic magnetism, primeval structures, thermal and nonthermal processes in galaxy clusters, feedback between galaxies and their environment, Black Holes (BH) and Active Galactic Nuclei (AGN) physics, etc. These kinds of research are directly linked to the comprehension of cosmic structure formation and evolution, and indirectly contribute to the cosmological models and their parameters. A further common characteristic of many of these future projects is high spatial/angular resolution that assumes an increasing relevance. This is also true in almost all wavelength bands: we recall projects such as the Sunyaev-Zel'dovich Array (SZA) and the Atacama Large Millimeter/submillimeter Array (ALMA) in the microwave domain,* Herschel *and JWST in the infrared, and ELT in the optical band. High resolution laser measurements are also at the basis of the direct detection, up to now missing, of gravitational waves in interferometric projects like the Laser Interferometer Gravitational-Wave Observatory (LIGO) and the Laser Interferometer Space Antenna (LISA).*

On longer timescales, space agencies of several countries and state communities could decide to significantly invest in a wide, long-term challenge (a dream?), *hopefully in the form of an international collaboration, with many technological, space exploration, economical, and scientific implications. The two main alternatives discussed so far considered the possibility of a renaissance of Moon exploration and/or a new generation champaign to Mars, ultimately aimed at the realization of bases for a human presence. Clearly, the former is potentially much more appealing for cosmological studies. Currently, a mission for an experiment on the Moon costs significantly more than an analogous space mission that could be implemented on a free-flyer, thus the possible advantage of the Moon relies more on a global vision than on an analysis of pros and cons for a single project. The Moon has the really unique, attractive opportunity of the absence of atmospheric contamination as well as to the possibility of constructing massive instrumentations. Many astronomical and fundamental physics ideas have proposed the use of the Moon (the reader could refer to the web pages of a couple of dedicated conferences*[1] *for a more complete view), some of them particularly related to the cosmological topics discussed in this book, such as, for example, a new generation of Lunar Laser Ranging (LLR) experiments able to improve by some orders of magnitude the accuracy of current LLR tests on GR; radio arrays at extremely long wavelengths (*$\lesssim$ *some tens of megaHertz, filtered out by the terrestrial atmosphere) with baselines of* $\sim 0.1 \div 10\,km$ *that could put new light on synchrotron emission by relativistic electrons and positrons in distant galaxies and clusters with implications for our understanding of the cosmic magnetism and, more generally, of the formation of early structures in the dark age of the Universe; CMB experiments of large size, dedicated to observe the polarization anisotropy with a wide frequency and multipole coverage and to determine the CMB radiation spectral shape, in particular at centimeter and decimeter wavelengths, where the large size of the necessary experimental equipment makes difficult or impossible the use of free-flyers.*

We believe that the existence of a variety of projects in the different branches represents a richness for the future of cosmology, particularly in view of its puzzle-like character, and to a certain extent, could greatly contribute to keeping "astrophysical cosmology" close to its original spirit.

In summary we do not believe that the present phenomenological formulation of the Λ*CDM Concordance Model will be the last word in cosmology. If we will be able to maintain the extraordinary wealth of astrophysical information contributing to the whole cosmological puzzle, we will for sure go through unexpected discoveries, whose revolutionary implications will strongly impact on fundamental physics.*

[1] See moon conferences in web page list.

Web Pages

ACBAR	Arcminute Cosmology Bolometer Array Receiver. http://cosmology.berkeley.edu/group/swlh/acbar/
ACT	Atacama Cosmology Telescope. http://www.physics.princeton.edu/act/
ADEPT	Advanced Dark Energy Physics Telescope. http://universe.nasa.gov/program/probes/adept.html
AKARI	Previously known as ASTRO-F or IRIS - InfraRed Imaging Surveyor. http://www.ir.isas.jaxa.jp/ASTRO-F/Outreach/index_e.html
ALMA	Atacama Large Millimeter/submillimeter Array. http://www.alma.nrao.edu/
ARCADE	Absolute Radiometer for Cosmology, Astrophysics, and Diffuse Emission. http://arcade.gsfc.nasa.gov/instruments.html
ASCA	Advanced Satellite for Cosmology and Astrophysics. http://www.astro.isas.ac.jp/asca/
AstroWISE	Astronomical Wide-field Imaging System for Europe. http://www.astro-wise.org/
BICEP	Background Imaging of Cosmic Extragalactic Polarization. http://bicep0.caltech.edu/ bicep/bicep_front.htm
BOOMERanG	Balloon Observations Of Millimetric Extragalactic Radiation and Geophysics. http://oberon.roma1.infn.it/boomerang/
B-Pol	B-Polarization Satellite Mission. http://www.b-pol.org/index.php
CERN	Conseil Européen pour la Recherche Nucléaire. http://public.web.cern.ch/public/
CFHT	Canada-France-Hawaii Telescope. http://www.cfht.hawaii.edu/
CFHTLS	Canada-France-Hawaii Telescope Legacy Survey. http://www.cfht.hawaii.edu/Science/CFHLS/
CHANDRA-Photo	Chandra Photo Album. http://chandra.harvard.edu/photo/

COBE	Cosmic Background Explorer. http://lambda.gsfc.nasa.gov/product/cobe/
COMBO17	Classifying Objects by Medium-Band Observations – a spectrophotometric 17-filter survey. http://www.mpia.de/COMBO/combo_index.html
Constellation-X	NASA mission. http://constellation.gsfc.nasa.gov/
COSMOGRAIL	COSmological MOnitoring of GRAvItational Lenses. http://www.cosmograil.org/
COSMOSOMAS	COSMOlogical Structures On Medium Angular Scales. http://www.iac.es/project/cmb/cosmosomas/
CTA	Cherenkov Telescope Array. http://www.cta-observatory.org/
CUORE	Cryogenic Underground Observatory for Rare Events. http://crio.mib.infn.it/wigmi/pages/cuore.php
CXC	Chandra X-ray Center. http://cxc.harvard.edu/
DASI	Degree Angular Scale Interferometer. http://astro.uchicago.edu/dasi/
DES	Dark Energy Survey. http://www.darkenergysurvey.org/
DESTINY	Dark Energy Space Telescope. http://www.noao.edu/noao/staff/lauer/destiny.htm
DIRBE	Diffuse Infrared Background Experiment. http://lambda.gsfc.nasa.gov/product/cobe/dirbe_overview.cfm
DPOSS	Digital Palomar Observatory Sky Survey. http://www.astro.caltech.edu/ george/dposs/
DUNE	Dark UNiverse Explorer. Mission merged into *Euclid*. http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=42266
E-ELT	European Extremely Large Telescopes. http://www.eso.org/public/astronomy/projects/e-elt.html
eROSITA	extended Röntgen Survey with an Imaging Telescope Array. http://www.mpe.mpg.de/projects.html#erosita
ESO	European Southern Observatory. http://www.eso.org/
ESSENCE	Equation of State SupErNovae trace Cosmic Expansion. http://www.ctio.noao.edu/essence/
FIRAS	Far Infrared Absolute Spectrophotometer. http://lambda.gsfc.nasa.gov/product/cobe/firas_overview.cfm
GROND	Gamma-Ray burst Optical/Near-infrared Detector. http://www.mpe.mpg.de/ jcg/GROND/
HEAO	High Energy Astrophysical Observatory. http://heasarc.gsfc.nasa.gov/docs/heao3/heao3.html

HEGRA	High Energy Gamma Rays Astronomy. http://www.mpi-hd.mpg.de/hfm/CT/CT.html
HST-ScI	Hubble Space Telescope Science Institute. http://www.stsci.edu/resources/
IAC	Instituto de Astrofísica de Canarias. http://www.iac.es/
IGEX	International Germanium Experiment. http://www.unizar.es/lfnae/ipaginas/ip0400.html#migex
INTEGRAL	International Gamma Ray Astrophysics Laboratory. http://sci.esa.int/science-e/www/area/index.cfm?fareaid=21
IRAS	Infrared Astronomical Satellite. http://irsa.ipac.caltech.edu/IRASdocs/iras.html
IVOA	International Virtual Observatory Alliance. http://www.ivoa.net/
JDEM	JointDarkEnergyMission. http://nasascience.nasa.gov/missions/jdem
JPL	Jet Propulsion Laboratory. http://www-b.jpl.nasa.gov/
JWST	James Webb Space Telescope. http://www.jwst.nasa.gov/
KATRIN	KArlsruhe TRItium Neutrino. http://www-ik.fzk.de/tritium/
KM3NET	km^3 NEutrino Telescope. http://www.km3net.org/home.php
LAMBDA	Legacy Archive for Microwave Background Data Analysis. http://lambda.gsfc.nasa.gov/
LCOGT	Las Cumbres Observatory Global Telescope. http://lcogt.net/
LHC	Large Hadron Collider. http://lhc.web.cern.ch/lhc/
LIGO	Laser Interferometer Gravitational-Wave Observatory. http://www.ligo.caltech.edu/
LISA	Laser Interferometer Space Antenna. http://lisa.nasa.gov/
LOFAR	Low Frequency Array. http://www.lofar.org/; LOFAR-EOR website http://www.astro.rug.nl/ vjelic
LOSS	Lick Observatory SN Search. http://astro.berkeley.edu/ bait/kait.html
LSST	Large Synoptic Survey Telescope. http://lsst.org
LWA	Long Wavelength Array. http://lwa.unm.edu/
MACHO	MAssive Compact Halo Objects. http://wwwmacho.anu.edu.au/

MAGIC	Major Atmospheric Gamma-ray Imaging Cherenkov. http://wwwmagic.mppmu.mpg.de/magic/index.html
Manuzio	Manuzio project. Online edition of *Le Opere*. http://www.liberliber.it/
MAX	Millimeter wavelength Anisotropy eXperiment. http://cosmology.berkeley.edu/group/cmb/MAX_experiment /max_page.html
MAXIMA	Millimeter Anisotropy eXperiment IMaging Array. http://cosmology.berkeley.edu/group/cmb/ Overview.html#max
Moon Conferences	9th ILEWG International Conference on Exploration and Utilization of the Moon (ICEUM9/ILC2007), October 22–26, 2007, Sorrento, Italy. http://sci.esa.it/science-e/www/object/index.cfm? fobjectid=40925 Observation of the Universe from the Moon, May 7, 2007, LNF INFN, Frascati, Italy. http://www.lnf.infn.it/conference/moon07/
NASA	National Aeronautics and Space Administration. http://www.nasa.gov/
NeXT	New X-ray Telescope (Mission). http://www.astro.isas.ac.jp/future/NeXT/
NFPS	NOAO Fundamental Plane Survey. http://astro.uwaterloo.ca/ mjhudson/fp200/
NuStar	Nuclear Spectroscopic Telescope Array. http://www.nustar.caltech.edu/
NVO	National Virtual Observatory. http://www.us-vo.org/
OGLE	Optical Gravitational Lensing Experiment. http://ogle.astrouw.edu.pl/
OSIRIS	sOptical System for Imaging and low/intermediate-Resolution Integrated Spectroscopy. http://www.iac.es/project/OSIRIS/
PanSTARRS	Panoramic Survey Telescope And Rapid Response System. http://pan-starrs.ifa.hawaii.edu/public/
Planck	*Planck* mission. http://www.rssd.esa.int/index.php?project=Planck
QUaD	QUEST at DASI. http://www.stanford.edu/ schurch/quad.html
QUIJOTE-CMB	Q, U, I JOint TEnerife CMB. http://www.iac.es/project/cmb/quijote/
RAVE	RAdial Velocity Experiment. http://www.rave-survey.aip.de/rave/
ROSAT	ROentgen SATellite. http://heasarc.gsfc.nasa.gov/docs/rosat/rosat.html

SAX	Satellite per Astronomia X. http://www.asdc.asi.it/bepposax/
SCUBA	Submillimetre Common-User Bolometer Array. http://www.jach.hawaii.edu/JCMT/continuum/s
SDSS	Sloan Digital Sky Survey. http://www.sdss.org/
SEGUE	Sloan Extension for Galactic Understanding and Exploration. http://segue.uchicago.edu/
Sidereus	Online English translation of the Sidereus Nuncius by Edward Stafford Carlos. http://www.chlt.org/sandbox/lhl/gal1880/page.1.a.php?size=480x640
Simbol-X	*Simbol*-X mission. http://www.asdc.asi.it/simbol-x/
SIM PQ	Space Interferometry Mission Planet Quest. http://planetquest.jpl.nasa.gov/index.cfm
SIRTF	Space InfraRed Telescope Facility. http://www.aerospaceguide.net/telescope/sirtf.html
SKA	Square Kilometer Array. http://www.skatelescope.org/
SNAP	SuperNova Acceleration Probe. http://snap.lbl.gov
SNFactory	Nearby Supernova Factory. http://snfactory.lbl.gov/
SNLS	Supernova Legacy Survey. http://cfht.hawaii.edu/SNLS/
SPT	South Pole Telescope. http://pole.uchicago.edu/
THEMIS	Télescope Héliographique pour l'Étude du Magnétisme et des Instabilités Solaires. http://webast.ast.obs-mip.fr/people/paletou/Themis/
UKIDSS	UKIRT Infrared Deep Sky Survey. http://www.ukidss.org/
VISTA	VISTA. Visible and Infrared Survey Telescope for Astronomy. http://www.vista.ac.uk
VLA	Very Large Array. http://www.vla.nrao.edu/
VLBI	Very Long Baseline Interferometry. http://www.vlba.nrao.edu/
VST	VLT Survey Telescope. http://vstportal.oacn.inaf.it/
WINGS	WIde-field Nearby Galaxy-clusters Survey. http://web.pd.astro.it/wings/

WMAP	Wilkinson Microwave Anisotropy Probe. http://map.gsfc.nasa.gov/
XCS	XMM Cluster Survey. http://xcs-home.org/
XEUS	X-ray Evolving Universe Spectroscopy (Mission). http://www.rssd.esa.int/index.php?project=XEUS
XMM-Newton	X-ray Multi Mirror Satellite. http://xmm.esac.esa.int/s
2dFGRS	2-Degree Field Galaxy Redshift Survey. http://www2.aao.gov.au/2dFGRS/

Further Reading

1. Akbar, M., Cai, R.G.: Friedmann equations of FRW universe in scalar tensor gravity, f(R) gravity and first law of thermodynamics. Phys. Lett. **B 635**, 7 (2006)
2. Akbar, M.: Thermodynamic interpretation of field equations at horizon of BTZ black hole. Chinese Phys. Lett. **24**(5), 1158–1161 (2007)
3. Albrecht, A., et al.: Report of the dark energy task force (2006) [astro-ph/0609591]
4. Alcaniz, J.S.: Dark energy and some alternatives: A brief overview. Braz. J. Phys. **36 A**, 1109–1117 (2006)
5. Allemandi, G., et al.: Charges and energy in Chern Simons theories and lovelock gravity. Classical Quant Grav **20**, 5103A (2003) [gr-qc/0308019]
6. Ashtekar, A., Rovelli, C., Smolin, L.: Weaving a classical metric with quantum threads. Phys. Rev. Lett. **69**, 237 (1992)
7. Bagla, J.S., Padmanabhan, T., Narlikar, J.V.L.: Crisis in cosmology: Observational constraints on ω and H_0. Comments on Astrophys. **18**, 275 (1996) [astro-ph/9 511102]
8. Barr, S.M.: Attempt at a classical cancellation of the cosmological constant. Phys. Rev. **D 36**, 1691 (1987)
9. Barris, B.J.: Twenty-three high-redshift supernovae from the institute for astronomy deep survey: Doubling the supernova sample at $z > 0.7$. Astrophys. J. **602**, 571 (2004)
10. Benson, A.J., et al.: The nature of galaxy bias and clustering. Mon. Not. R. Astron. Soc. **311**, 793 (2000)
11. Bullock, J.S., Kravtsov, A.V., Weinberg, D.H.: Hierarchical galaxy formation and substructure in the galaxy's stellar halo. Astrophys. J. **548**, 33 (2001)
12. Cai, R.G., Cao, L.M.: Unified first law and thermodynamics of apparent horizon in FRW universe (2006) [gr-qc/0611071]
13. Chaboyer, B.: The age of globular clusters in light of hipparcos: Resolving the age problem? Astrophys. J. **494**, 96 (1998)
14. Chiang, L.-Y., Coles, P., Naselsky, P.: Return mapping of phases and the analysis of the gravitational clustering hierarchy. Mon. Not. R. Astron. Soc. **337**, 488 (2002)
15. Choudhury, T.R., Padmanabhan, T.: Cosmological parameters from supernova observations: A critical comparison of three data sets. Astron. Astrophys. **429**, 807 (2005) [astro-ph/0311622]
16. Conley, A., et al.: Is there evidence for a hubble bubble? The nature of type Ia supernova colors and dust in external galaxies. Astrophys. J. **664**, 13 (2007) [arXiv:0705.0367]
17. Deliyannis, C.P., Demarque, P., Kawaler, S.D.: Lithium in halo stars from standard stellar evolution. Astrophys. J. Suppl. **73**, 21 (1990)
18. DeWitt, B.S.: Gravity: A universal regulator? Phys. Rev. Lett. **13**, 114 (1964)
19. Dirac, P.A.M.: A new basis for cosmology. Proc. Roy. Soc. of London **A 165**, 199 (1938)
20. Dolgov, A.D.: Cosmology and new physics (2006) [hep-ph/0606230]
21. Dressler, A.: The hubble sequence. In: From stars to galaxies: Building the pieces to build up the universe. ASP Conf. Ser. **374**, 415 (2007)
22. Dressler, A., et al.: Evolution since $z = 0.5$ of the morphology-density relation for clusters of galaxies. Astrophys. J. **490**, 577 (1997)

23. Durrer, R., Gangui, A., Sakellariadou, M.: Doppler peaks in the angular power spectrum of the cosmic microwave background: A fingerprint of topological defects. Phys. Rev. Lett. **76**, 579 (1996) [astro-ph/9507035]
24. Durrer, R., Kunz, M., Melchiorri, A.: Cosmic microwave background anisotropies from scaling seeds: Global defect models. Phys. Rev. **D 59**, 123005 (1999) [astro-ph/9811174]
25. Efstathiou, G., Bond, J.R., White, S.D.M.: COBE background radiation anisotropies and large-scale structure in the universe. Mon. Not. R. Astron. Soc. **258**, 1 (1992)
26. Fan, X., et al.: High-redshift quasars found in sloan digital sky survey commissioning data. II. The spring equatorial stripe. Astron. J. **119**, 1 (2000)
27. Fan, X., et al.: L dwarfs found in sloan digital sky survey commissioning imaging data. Astron. J. **119**, 928 (2000)
28. Fan, X., et al.: The discovery of a luminous $z = 5.80$ quasar from the sloan digital sky survey. Astron. J. **120**, 1167 (2000)
29. Fixsen, D.J., et al.: The cosmic microwave background spectrum from the full COBE FIRAS data set. Astrophys. J. **473**, 576–587 (1995)
30. Fontanini, M., Spallucci, E., Padmanabhan, T.: Zero-point length from string fluctuations. Phys. Lett. **B 633**, 627 (2006) [hep-th/0509090]
31. Ford, L.H.: Cosmological-constant damping by unstable scalar fields. Phys. Rev. **D 35**, 2339 (1987)
32. Gal, R.R., et al.: The digitized second palomar observatory sky survey (DPOSS). II. Photometric calibration. Astron. J. **128**, 3082 (2004)
33. Galilei, G.: Opere. Available at http://www.liberliber.it
34. Garay, L.J.: Quantum gravity and minimum length. Int. J. Mod. Phys. **A 10**, 145 (1995)
35. Gaskell, C.M.: A redshift difference between high and low ionization emission-line regions in QSOs - Evidence for radial motions. Astrophys. J. **263**, 79 (1982)
36. Gonzáles, J.J.: Line-strength gradients and kinematic profiles in elliptical galaxies. PhD Thesis. University of California, Santa Cruz (1993)
37. Hannestad, S.: Dark energy and dark matter from cosmological observations. Int. J. Mod. Phys. **A 21**, 1938 (2006)
38. Hasinger, G.: The new XEUS science case, SPIE **6266E**, 49 (2006)
39. Hebecker, A., Wetterich, C.: Quintessential adjustment of the cosmological constant. Phy. Rev. Lett. **85**, 3339 (2000)
40. Hebecker, A.: On dynamical adjustment mechanisms for the cosmological constant. In: Particles, strings, and cosmology. Proceedings of the Eighth International Conference of North Carolina, Chapel Hill. Frampton, P. and Jack Ng. (eds.). Rinton Press, p. 361 (2001) [hep-ph/0105315]
41. Hesser, J.E.: The galactic globular cluster system. Astrophys. Space Sci. **267**, 57 (1999)
42. Hirata, C.M., et al.: Correlation of CMB with large-scale structure: II. Weak lensing (2008) [arXiv:0801.0644v1]
43. Huang, C.G., Sun, J.R.: Thermodynamic properties of spherically-symmetric, uniformly-accelerated reference frames. (2007) [gr-qc/0701078]
44. Jacobson, T.: Thermodynamics of spacetime: The Einstein equation of state. Phys. Rev. Lett. **75**, 1260 (1995)
45. James, D.F.V.: The wolf effect and the redshift of quasars. Pure Appl. Opt. **7**, 959 (1998)
46. James, D., Savedoff, M. Wolf, E.: Shifts of spectral lines caused by scattering from fluctuating random media. Astrophys. J. **359**, 67 (1990)
47. Jassal, H.K., Bagla, J.S., Padmanabhan, T.: Observational constraints on low redshift evolution of dark energy: How consistent are different observations? Phys. Rev. **D 72**, 103503 (2005) [astro-ph/0506748]
48. Jassal, H.K., Bagla, J.S., Padmanabhan, T.: Understanding the origin of CMB constraints on Dark Energy (2006) [astro-ph/0601389]
49. Jones, W.C.: A polarization sensitive bolometric receiver for observations of the cosmic microwave background, Proc. SPIE **4855**, 227–238 (2003)
50. Kauffmann, G., et al.: Stellar masses and star formation histories for 10^5 galaxies from the sloan digital Sky Survey. Mon. Not. R. Astron. Soc. **341**, 33 (2003)

51. Kennicutt, R.C., Freedman, W.L., Mould, J.R.: Measuring the hubble constant with the hubble space telescope. Astron. J. **110**, 1476 (1995)
52. Kinney, A.L., Rivolo, A.R., Koratkar, A.P.: A study of the Baldwin effect in the IUE data set. Astrophys. J. **357**, 338–345 (1990)
53. Kogut, A., et al.: An instrument to measure the temperature of the cosmic microwave background radiation at centimeter wavelengths. Astrophys. J. Suppl. Ser. **154**, 493 (2004)
54. Lewis, A., Challinor, A.: Weak gravitational lensing of the CMB. Phys. Report **429**, 1–65 (2006)
55. Makela, J.: Quantum-mechanical model of spacetime. (2007) [gr-qc/0701128]
56. Maltoni, M., Schwets, T.: Sterile neutrino oscillations after first MiniBooNE results. Phys. Rev. **D 76**, 093005 (2007) [hep-ex/0705.0107]
57. Marziani, P., et al.: Comparative analysis of the high- and low-ionization lines in the broad-line region of active galactic nuclei. Astrophys. J. Suppl. **104**, 37 (1996)
58. Millennium Simulation: http://www.mpa-garching.mpg.de/millennium
59. Mishustin, I.N., Ruzmaikin, A.A.: Formation of priming magnetic fields during the formation of protogalaxies. J. Exp. Theor. Phys. **61**, 441–444 (1971)
60. Mishustin, I.N., Ruzmaikin, A.A.: Occurrence of priming magnetic fields during the formation of protogalaxies. J. Exp. Theor. Phys. **34**, 233–235 (1972)
61. Naselsky, P.D.: Hydrogen recombination kinetics in the presence of low-mass primordial black holes. Pis'ma Astronomicheskii Zhurnal **4** 387 (1978)
62. Nesseris, S., Perivolaropoulos, L.: Tension and systematics in the Gold06 SnIa data set. J. Cosmol. Astropart. Phys. **02**, 025 (2007)
63. Nobbenhuis, S.: The cosmological constant problem, an inspiration for new physics. (2006) [gr-qc/0609011]
64. Nojiri, S., Odintsov, S.D.: Unifying inflation with ΛCDM epoch in modified f(R) gravity consistent with Solar System tests. Phys. Lett. **B 657**, 238 (2007)
65. Ostriker, J.P., Steinhardt, P.J.: The observational case for a low density universe with a non-zero cosmological constant. Nature **377**, 600 (1995)
66. Padmanabhan, T.: Physical significance of planck length. Ann. Phys. **165**, 38 (1985)
67. Padmanabhan, T.: Limitations on the operational definition of spacetime events and quantum gravity. Classical Quant. Grav. **4**, L107 (1987)
68. Padmanabhan, T.: Duality and zero-point length of spacetime. Phys. Rev. Lett. **78**, 1854 (1997) [hep-th/9608182]
69. Padmanabhan, T.: Hypothesis of path integral duality. I. Quantum gravitational corrections to the propagator. Phys. Rev. **D 57**, 6206 (1998)
70. Padmanabhan, T.: Event horizon of a schwarzschild black hole: Magnifying glass for planck length physics. Phys. Rev. **D 59**, 124012 (1999) [hep-th/9801138]
71. Padmanabhan, T.: Is Gravity an intrinsically quantum phenomenon? Mod. Phys. Lett. **A 17**, 1147 (2002) [hep-th/0205278]
72. Padmanabhan, T.: Thermodynamics of horizons. Mod. Phys. Lett. **A 17**, 923, (2002) [gr-qc/0202078]
73. Padmanabhan, T.: Topological interpretation of the horizon temperature. Mod. Phys. Lett. **A 18**, 2903 (2003) [hep-th/0302068]
74. Padmanabhan, T.: Gravity as elasticity of spacetime. Int. J. Mod. Phys. **D 13**, 2293–2298 (2004) [gr-qc/0408051]
75. Padmanabhan, T.: Entropy of static spacetimes and microscopic density of states. Classical Quant. Grav. **21**, 4485 (2004) [gr-qc/0308070]
76. Padmanabhan, T.: A new perspective on gravity and dynamics of space-time. Int. J. Mod. Phys **D 14**, 2263–2270 (2005) [gr-qc/0510015]
77. Padmanabhan, T.: Dark energy: The cosmological challenge of the millennium. Curr. Sci. **88**, 1057 (2005) [astro-ph/0411044]
78. Padmanabhan, T.: Gravity: A new holographic perspective. Int. J. Mod. Phys. **D 15**, 1659 (2006) [gr-qc/0609012]
79. Padmanabhan, T., Paranjape, A.: Entropy of null surfaces and dynamics of spacetime. Phys. Rev. **D 75**, 06 4004 (2007)
80. Padmanabhan, T., Narasimha, D.: Constraints on the shape of the density spectrum from COBE and galaxy surveys. Mon. Not. R. Astron. Soc. **259**, 41P (1992)

81. Pecker, J.C., Narlikar, J.: Current issues in Cosmology. Cambridge University Press, Cambridge (2006)
82. Perivolaropoulos, L.: Accelerating universe: Observational status and theoretical implications. AIP Conf. Proc. **848**, 698 (2006)
83. Postman, M., et al.: The morphology-density relation in $z \sim 1$ clusters. Astrophys. J. **623**, 721 (2005)
84. Rees, M.J.: Origin of cosmic magnetic fields. Astron. Nachr. **327**, 395–398 (2006)
85. Rosenberg, A.: Photometric catalog of nearby globular clusters. II. A large homogeneous (V,I) color-magnitude diagram data-base. Astron. Astrophys. Suppl. **145**, 451 (2000)
86. Rowan-Robinson, M.: Cosmological parameters: Do we already know the final answer? Proc. of the 3rd International Workshop on the Identification of Dark Matter, N.J.C. Spooner, V. Kudryavtsev (eds). World Scientific, Singapore, p.20 (2001)
87. Saha, A.: Cepheid distances to SNe Ia host galaxies based on a revised photometric zero point of the HST WFPC2 and New PL relations and metallicity corrections. Astrophys. J. Suppl. **165**, 108 (2006)
88. Sami, M., Padmanabhan, T.: Viable cosmology with a scalar field coupled to the trace of the stress tensor. Phys. Rev. **D 67**, 083509 (2003) [hep-th/0212317]
89. Scannapieco, E., Ferrara, A., Madau, P.: Early enrichment of the intergalactic medium and its feedback on galaxy formation. Astrophys. J. **574**, 590 (2002)
90. Sheykhi, A., et al.: Thermodynamical properties of apparent horizon in warped DGP Braneworld. (2007) [hep-th/0701198]
91. Singh, T.P., Padmanabhan, T.: An attempt to explain the smallness of the cosmological constant. Int. J. Mod. Phys. **A 3**, 1593 (1988)
92. Spergel, D.N., Zaldarriaga, M.: CMB polarization as a direct test of Inflation. Phys. Rev. Lett. **79**, 2180 (1997) [astro-ph/9705182]
93. Srinivasan, K., Sriramkumar, L., Padmanabhan, T.: The hypothesis of path integral duality II: corrections to quantum field theoretic results. Phys. Rev. **D 58**, 044009 (1998) [gr-qc/9710104]
94. Sulentic, J.: Tests of the discordant redshift hypothesis. In: New ideas in astronomy. F. Bertola, J.W. Sulentic, B. Madore (eds.). Cambridge University Press, Cambridge, p. 123 (1988)
95. Szalay, A.S., Brunner, R.R.: Astronomical archives of the future: a virtual observatory. Future Gener. Comput. Syst. (Netherlands), **16/1**, p. 63–72 (1999) [astro-ph/9812335]
96. Taylor, A.C., et al.: First results from the very small array. II. Observations of the cosmic microwave background. Mon. Not. R. Astron. Soc. **341**, 1066-1075 (2003)
97. Tegmark, M., et al.: Cosmological constraints from the SDSS luminous red galaxies. (2006) [astro-ph/0608632]
98. Thakar, A.R., et al.: Data organization in the SDSS data release 1. In: ADASS XII. H.E. Payne, R.I. Jedrzejewski, R.N. Hook (eds.), ASP Conf. Ser. **295**, 217 (2003)
99. Tonry, J.L., et al.: Cosmological results from high-z supernovae. Astrophys. J. **594**, 1 (2003)
100. van der Wel, A., et al.: The evolution of the field and cluster morphology-density relation for mass-selected samples of galaxies. Astrophys. J. **670**, 206 (2007)
101. Vittorio, N., Silk, J.: Anisotropies of the cosmic microwave background in nonstandard cold dark matter models. Astrophys. J. Lett. **385**, L9 (1992)
102. Volovik, G.E.: Superfluid analogies of cosmological phenomena. Phys. Rept. **351**, 195 (2001)
103. Volovik, G.E.: Vacuum energy: Myths and reality. (2006) [gr-qc/0604062]
104. Wampler, J.: Workshop overview. Quasars and cosmology, G. Ferland, J. Baldwin (eds.) ASP Conference Series **162**, p. 1–11 (1999)
105. Wang, Y., Mukherjee, P.: Observational constraints on dark energy and cosmic curvature. Phys. Rev. **D 76**, 103533 (2007)
106. White, S.D.M., Frenk, C.S., Davis, M.: Clustering in a neutrino-dominated universe. Astrophys. J. **274**, L1 (1983)
107. Xian-Hui, G.: First law of thermodynamics and Friedmann-like equations in braneworld cosmology. Phys. Lett. B **651**, 49 (2007) [hep-th/0703253]
108. Yi, S.K., Demarque, P., Kim, Y.C.: The Y^2 isochrones. Astrophys. Space Sci. **291**, 261 (2004)

Index