A Science Vision for European Astronomy

ASTRONET SVWG

DRAFT December 19, 2006

Contents

Note

This is a draft document, written by the members of the Science Vision Working Group and its supporting panels (see Contributors). The document is intended as a starting point for community input via the ASTRONET website `www.astronet-eu.org` and the Science Vision discussion website `http://www.strw.leidenuniv.nl/sciencevision`, and during the ASTRONET Symposium in Poitiers, January 23–25, 2007. The aim is to improve the overview of the scientific priorities for European astronomy, and in particular to sharpen the recommendations which will be input for the development of a road map for the required infrastructure.

The final version will also have improved figure captions, and proper credits and references.

Chapter 1

Introduction

1.1 The role of science in society

Mankind has an innate curiosity, and a particular curiosity about the Universe. Perhaps this has arisen as an inevitable outcome of evolution: curiosity leads to exploration, to discovery, to learning, and in the most basic sense, to survival. At a higher level, societies develop and prosper through innovation, and through planning for the future. Many of the advances made by civilization in the last few centuries can be attributed to the advancement in our scientific understanding.

There is, perhaps fortunately, no obvious end to this process in sight. Further scientific advances should continue to contribute to the improved quality of life for future generations in countless ways. At the same time, societies also face momentous and troubling challenges, to which scientists will be expected to contribute solutions. To maintain this forward momentum, Europe must continue to contribute its share to the process. It can do so by inspiring, educating, and training new scientists, and by encouraging scientific advances through investment in all forms. Happily, the conditions seem reasonably favourable for this process to continue from strength to strength. Nevertheless, complacency would be ill-founded, and there are some worrying signs.

Most importantly, scientific progress relies on a continued influx of bright young scientists, but enrolment to the physical sciences has declined in recent years throughout the western world. To a large extent, young people are persuaded to embark on scientific careers because a field in which they have a potential interest is seen to be thriving. They see possibilities for exciting and rewarding careers. Major advances – whether as cures for diseases, options for cheaper energy, or technological innovations to make life more comfortable or enjoyable – all require a continuous injection of new ideas. And young people will only enter a field if it captures their imagination.

Scientific literacy: the foundation of democracy. But not only budding young scientists are fascinated by the Universe: even though not directly involved in the process of scientific advance, a large fraction of the public finds science intriguing and wants to learn more. Demonstrating advances and satisfying this curiosity are indispensable in acknowledging and repaying society's good will and financial support, elements so essential for its continued well being. And quite independently of their value for society, advances in scientific understanding should go hand-in-hand with other cultural advances, in literature, music, and the arts in general. It is reasonable to suppose that a field that is perceived to be stagnant, or in decline – whether in science or in any of these other areas – will have great problems in attracting young people.

However, the interest of lay people in the secrets of the Universe may turn towards superstition as well as science. Astrology, creationism, or conspiracy theories such as those surrounding the Moon landings can be considered as faintly amusing or harmless. Skepticism based on critical thinking is indeed healthy, but perversion of scientific understanding can make it hard for the public to distinguish between what is fact and what is fiction, and large-scale ignorance of science must be considered as dangerous. It is important that a society – and not just a select few – should be able to make informed decisions across a broad range of issues: from understanding the ways computers enter our lives, priorities for health care, judging the implications of global warming, the relative merits of fossil fuels or nuclear or wind-generated power, the case for genetically-modified food, and so on.

Astronomy in society. Astronomy, as a field of science, enters our daily lives directly. This includes everyday phenomena such as the influence of the Sun and the tilt of the Earth's axis on our seasons and climate, the influence of the Moon on the tides, and why the sky is blue but why sunsets are red. Yet ignorance of even these basic phenomena remains widespread. And while a Solar eclipse can be predicted with precision and viewed in awe as one of nature's great spectacles, a significant fraction of the population on Earth still views an eclipse with mysticism and fear.

Astronomy also plays a central role in a wide range of highly practical matters. Historically, the accurate measurement of star positions provided the basis for navigation at sea, an advance of enormous economic importance. The current state-of-the-art provides an accurate celestial reference framework at the foundation of modern-day satellite communications. Astronomical advances underpin today's global satellite navigation in which gravity, Earth rotation, and General Relativity are intimately involved – allowing airplanes to land, and the emergency services to act. Understanding the complex effects of the Earth's motion is important in predicting long-term climate change. Advances in observations and models of our Sun offer the prospects of predicting Solar storms, with the practical goal of averting potential damage to satellites and power lines.

In the last few years, astronomy has provided a deeper understanding of the role played by the impact of small asteroids in the biological evolution of the Earth, demonstrating that this is not only an ongoing process, but one that is amenable to prediction by astronomical observation, and possibly to the mitigation of an impending catastrophic collision.

The implications of astronomy on our lives are also deeply cultural. It addresses profound questions such as the origin of the Universe, of time and space, of our Galaxy and the stars within it, of our Solar System, of our Earth, and of life itself. In the last decade, astronomers have begun investigating the existence of other worlds beyond our Solar System – planetary systems around other stars, some with considerable similarity to our own. Exciting plans exist for the rapid development of this very young and dynamic field. It has already resulted in intense interdisciplinary research amongst astronomers, chemists, and biologists, with the aim of better understanding the conditions under which life might have developed elsewhere in the Universe, and how it might be detected – whether we are alone in the immensity of space, or immersed in a cosmos teeming with life. The general public has a deep desire to accompany scientists on this exciting journey of discovery, and is as vocal as many professionals in wanting to know how and why the Universe originated, and whether intelligent life exists elsewhere.

Contributions of astronomy to the next generation. In the last twenty years astronomy has made particularly impressive advances, both technically and theoretically. Despite a general downturn in science intake at university level, young scientists have continued to enter the field, intrigued by their innate curiosity, and motivated to contribute directly to advances in knowledge. Excited by astronomy, young and gifted minds are frequently attracted to related scientific disciplines, so that astronomy acts as a springboard and catalyst for wider scientific enquiry.

In addition to national initiatives, Europe now has a coherent space research programme orchestrated through the European Space Agency, and a vibrant ground-based astronomy programme in which a number of major facilities are being developed under the leadership of the European Southern Observatory. These have already contributed substantially to advances in science and technology, and have motivated gifted individuals to enter the field, to be trained, and to train others in their turn.

Breathtaking images, scientific discoveries, and renewed appreciation of the scale and diversity of nature provided by astronomy will continue to captivate people's imagination, inform and inspire teachers, and excite students and the public about science and exploration. Continued developments will guarantee advances in basic research and applied technology, helping to ensure that a healthy supply of scientists and engineers will be available to meet the broader needs of society in the twenty-first century.

Astronomy continues to be a highly dynamic field of research. New discoveries have led to the award of three Nobel Prizes for Physics in the domain of astronomy in the last 15 years. Yet hugely significant questions remain unanswered, such as the nature of dark matter and dark energy, with their discovery through astronomy now driving significant research in particle physics.

Happily, ideas abound for the future experimental and theoretical probing of the underlying fabric of the Universe, in the fields of gravitational wave research, stellar seismology, extra-solar planetary searches, Solar System exploration, and many others.

1.2 Astronomy

The dramatic progress of astronomical discoveries over the past decades is intimately connected to major advances in technology and instrumentation. Ever since the invention of the telescope in the early 17th century, telescopes have steadily increased in power. The opening of the radio domain in the middle of the past century provided the first new window on the Universe, and the ability to launch satellites into space provided observing facilities which now cover the entire electromagnetic spectrum. The challenging requirements of sensitivity and precision of astronomical measurements in turn have often driven the pace of technological capabilities, thus cross-fertilizing basic research and commercial applications.

Observing facilities on the ground and in space. The major astronomical space observatories that are currently active include the Hubble Space Telescope in the optical, the Spitzer Space Telescope in the infrared, and the Chandra, XMM and Integral telescopes at high-energies. Major missions under development include ESA's Herschel, Planck and Gaia satellites, and the NASA / ESA James Webb Space Telescope. Astronomy missions dedicated to specific topics include RXTE, SWIFT, Akari, CoRoT, and Kepler. Soho, Ulysses, and Cluster are studying the Sun and its surroundings. In the field of planetary exploration Cassini is active in the Saturn system, Rosetta is on its way to comet Chiriumov–Gerasimenko, orbiters are probing Venus and Mars, a.o. Venus Express and Mars Express , and the Mars rovers Spirit and Opportunity continue to provide stunning science and remarkable images.

On the ground the 8–10 m class optical / infrared telescopes (Gemini, Keck, Subaru, VLT) are being equipped with a full arsenal of instruments, including many that will take advantage of progress in adaptive optics. They are being linked interferometrically (e.g., VLTI) to obtain milli-arcsecond resolution. The GranTeCan is nearly finished. Numerous large-scale surveys are available, including 2MASS, GSC-II, USNO-B, the 2dF and Sloan Digital Sky Surveys. Many telescopes in the 2–4 m class now concentrate on challenging wide-field imaging surveys (MegaCam on the CFHT, Omegacam on the VLT Survey Telescope, and VISTA). The RAVE project will obtain millions of radial velocities for stars in the Galaxy using multi-fiber spectroscopy. Following the success of the micro-lensing surveys MACHO, EROS and OGLE, a new generation of synoptic facilities is being planned, with PanStarrs-1 already under construction. Various radio telescopes are under construction, including LOFAR which will provide a major advance in the study of objects which emit extremely long radio waves. The next decade will see the full power of the 8–10 m class optical / infrared telescopes exploited, with second-generation instruments and interferometric links, as well as the completion of the transformational (sub)millimetre telescope array ALMA.

The role of technology. Continuing improvements in semiconductor sensors, electronics, telescopes and computing have maintained an impressive doubling in detection sensitivity of radio

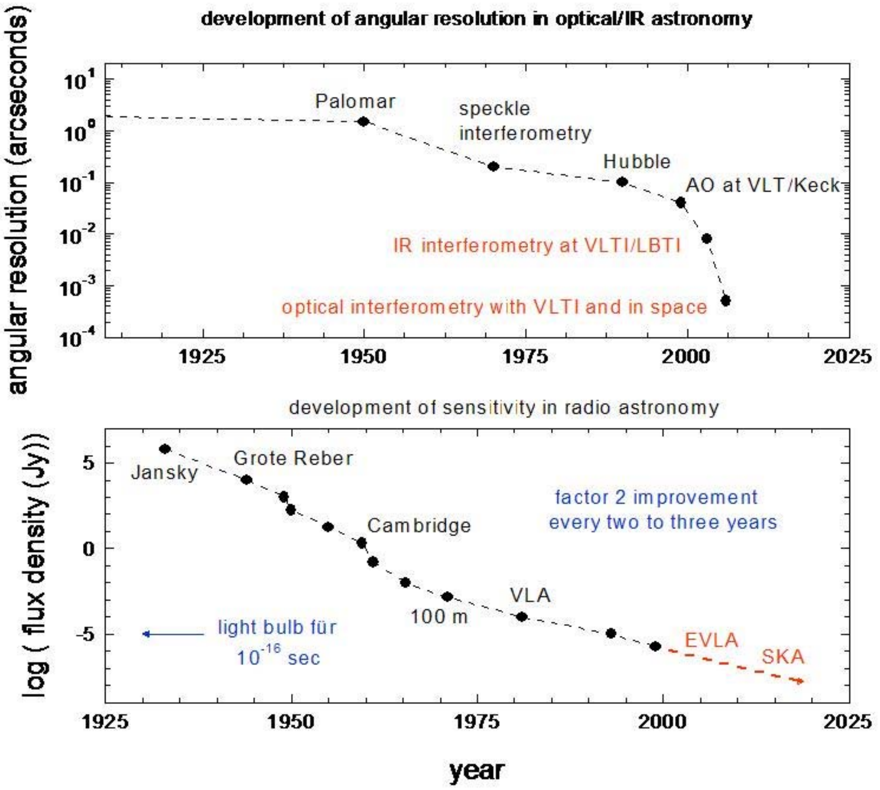

Figure 1.1: Top: improvement in angular resolution in optical / infrared astronomy. Bottom: improvement of sensitivity with time in radio astronomy.

telescopes every three years over the past seventy years (Figure 1.1, bottom). This has resulted, for example, in a total gain in sensitivity of radio astronomical measurements of twelve orders of magnitude since Karl Jansky's pioneering work in the 1930's. Further improvements in digital technology and computers, and the mass production of cheap, commercial radio dishes, are together expected to lead to another two orders of magnitude improvement over the next two decades. Similar dramatic advances have occurred at X-rays and optical / infrared wavelengths, especially in space. Continued progress in detector technology and telescope collecting area can be expected also in these wavebands in the foreseeable future, leading to correspondingly large gains in sensitivity in these fields.

Energy-resolving detectors represent another area of transformational technology. Such devices have been used very successfully in X-ray and gamma-ray astronomy. Current progress in superconducting devices will soon allow the development of energy-resolving, imaging detectors also for longer wavelengths, from the optical to the (near-)infrared, eliminating the need for dispersive devices in spectroscopic applications, and resulting in further gains of throughput and sensitivity.

A key area where substantial improvements in capabilities can be expected is in the angular resolution of astronomical measurements. Figure 1.1 (top) shows the development of angular resolution in optical / near-infrared astronomy over the past 70 years. While the adverse impact of the Earth's atmosphere prevented significant improvement of optical imaging until the middle of the twentieth century, dramatic advances have occurred since that time. They will continue almost certainly for the next one or two decades. The development of the Hubble Space Telescope was one key stepping stone toward much higher angular resolution by bringing an optical telescope above the Earth's atmosphere. Another was the development of techniques, such as speckle and adaptive-optics imaging, that correct for the blurring of the atmosphere from the ground. The combination of adaptive optics with large, lightweight optical mirrors has led to the dramatic improvement in ground-based angular resolution shown in Figure 1.1. Modern adaptive optics systems routinely allow diffraction limited imaging on 8–10 m class telescopes in the near-infrared. The next decade should see the application of this technique to 20–40 m class telescopes as well as to shorter wavelengths.

Spatial interferometry between several individual telescopes is another key development. Historically this technique was pioneered in radio astronomy. During the past decade wide-bandwidth interferometry has become feasible also at infrared and optical wavelengths, resulting in milli-arcsecond resolution. While infrared / optical interferometry is presently still a highly challenging and somewhat experimental technique, it is expected that further progress in single-mode optical fibers, integrated optics, lasers and fast control systems will make (sub) milli-arcsecond imaging interferometry widely applicable over the next ten years. A longer-term application will be precision interferometry from space, with enormous added benefits in wavelength range, stability and sensitivity. X-ray (and perhaps even gamma-ray) interferometry may also be feasible from space in the coming decades, resulting in micro-arcsecond angular resolution in these important fields.

Future progress in spectroscopic capabilities can be expected both in terms of multiplexing, and in spectral resolution and precision. The advent of large integral-field spectrometers and energy-resolving devices, in combination with ever larger imaging detectors, will likely allow very significant progress in the spatial and spectral multiplexing across all wavebands, from X-ray to radio wavelengths. The combination of very high resolution spectroscopy with ultra-stable laser clocks will make possible a new generation of fundamental time and frequency measurements.

Advances in remote sensing instruments, solar electric and micropropulsion technology, radiation-hardened electronic circuitry, digital instrumentation, high-bandwidth communications, on-board processing, advanced optical ceramics, and interplanetary navigation, have led to an equally remarkable surge in missions to explore the Solar System. Landing on Saturn's moon Titan, flybys and impacts of asteroids, and rovers operating on the surface of Mars, were essentially implausible even two decades ago. Now plans and capabilities exist to explore the inhospitable surfaces of Mercury and Venus at one extreme of temperature, and the icy crust of Europa at the other.

Finally, astronomy has always been pushing the boundary of technical possibilities in computing. Large astronomical simulations in cosmology, hydrodynamics and gravity have been among the key test cases for the fastest computers of each generation. The requirement of larger and more realistic simulations in astronomy, and the vastly larger data rates combined with much more complex data processing needs, has strongly motivated the development of yet faster and more capable devices. Physics and astronomy have been pushing and frequently leading the internet and Grid revolutions, and these developments are expected to continue for some time into the future.

1.3 Predicting the future

These technological developments now make it possible to observe planets around other stars, as well as peer deeper than ever into the Universe.promises tremendous progress in key astronomical questions such as the nature of dark matter and dark energy; physics under extreme conditions including black holes, supernovae and gamma-ray bursts; the formation and evolution of galaxies from first light to the present; and the formation of stars and planets, including the origins of our own Solar System and the beginning of life. These are amongst the most fundamental questions in all of science, and are of enormous interest to the general public. The world-wide astronomical community, together with national science foundations and space agencies, should be in a position to answer many of these questions using advanced telescopes on the ground and in space now in the planning or conceptual stage, supported by interpretative efforts and theoretical work.

Astronomers are preparing plans for a number of exciting observational facilities. These include extremely large telescopes for the optical and infrared, a giant radio telescope with a collecting area of a square kilometer, an 8 m class survey telescope which would provide deep imaging of the sky every four nights, a 4 m class advanced-technology Solar telescope, wide-field imagers in space, advanced and ambitious planetary and Solar missions, a mission to detect the gravitational wave signature of coalescing black holes, and next generation X-ray telescopes and space missions devoted to characterizing extra-solar planets. Detailed science cases are available for all these missions and facilities.

1.4 This document

The plans for astronomy are very ambitious, and are considered attainable. Within Europe, they would require a collective investment of several billion Euros for new instrumentation and associated hardware and operations. While some funding will be pursued through programs of the European Union, the bulk of the support will only be accessible from the national funding agencies.

While the national funding agencies have proven to be very supportive of astronomy and Solar System exploration over many years, funding of substantial facilities which are continuing to grow in complexity, and whose capital costs are growing commensurately, clearly becomes more problematic. In part, this is because the national and international funding agencies need to be presented with the collective priorities for ground-based and space-based astronomy at the pan-European level. These must be comprehensive including not only the capital and operational costs, but also the principle investments necessary for numerical simulations, laboratory experiments and theory, and data analysis, and the links with neighbouring disciplines. For these reasons, a number of national funding agencies established ASTRONET, an ERA-net with support by the European Union, to develop a strategic plan for European astronomy.

The first step of this process is the development of an integrated ‘science vision’, compiled by a working group and supporting panels appointed by the funding agencies. A subsequent step will be the preparation of a road map for the development of the required infrastructure.

The current document is a draft of this science vision. As a draft it will be neither balanced, nor fully representative of the wishes and scientific priorities of all sections of the relevant communities. Rather, it represents a starting point for wider consultation. Comment and further input from the astronomical community will be gathered in two main ways: via the Science Vision Discussion website, and via discussions at the forthcoming Symposium in Poitiers, in January 2007. Based on this, a final version will be produced, which will serve as input for the development of the road map for the required infrastructure. Together this will provide the long-term planning directed towards achieving a world-leading position for European astronomy.

In each of the four chapters that follow, the present situation is described, including the expected capabilities of facilities under construction or planned. In each case the key science questions are summarized, and the technique or facility needed to make substantial progress is identified. Much basic material is already available from national strategic plans, from science cases for specific proposed facilities, from the ESA Cosmic Vision document, and from the three ESA–ESO working group reports. Since the ERA-net ASPERA program concentrates on astro-particle physics in Europe, which has a clear link and overlap with the topics addressed in Chapter 2, care has been taken to harmonize the recommendations with ASPERA.

Chapter 6 gives a preliminary summary of the recommendations. A list of abbreviations provides background information for many of the facilities and acronyms used throughout this document.

Chapter 2

Do we understand the extremes of the Universe?

The challenge of astrophysics is to identify the physical concepts and tools that are relevant for understanding a given astronomical object. Sometimes this challenge is one of implementation of well-understood physics in the face of a complicated dataset, but there are often much greater issues of principle. Where the observational constraints are few, as with, e.g., gamma-ray bursts, the appropriate physics is highly uncertain. Astrophysics also offers unique possibilities for probing fundamental physics beyond the level that can be explored in the laboratory. The obvious examples are gravity in the strong-field regime, and particle physics at energies above the TeV scale. From this point of view, astrophysics becomes perhaps the prime arena where the frontiers of physics can be advanced. This is however not to say that the subject can be neatly divided into 'pure' and 'applied' subsets, since achieving any of the fundamental goals also requires detailed understanding and control of the normal-physics aspects of the objects under study.

These extreme applications of astrophysics deal with grand and general themes. At the greatest extreme of scale, we find questions of cosmology: how the Universe came to exist in its current form, and the nature of its contents. Here, astronomy has made what is indisputably its greatest contribution to physics by the detection of a non-zero vacuum density – the so-called dark energy. It is a common cliche to describe dark energy as the greatest unsolved problem in physics, but it is hard to disagree. There is a widespread feeling that the dark energy density should vary with time, possibly in a way that is detectable. It may also have been very much larger at very early times, causing a phase of inflation that started the current expansion. If so, fluctuations in density and radiation temperature are relics of this era, and have much to tell us about how it happened. The origin of the Universe is arguably the greatest challenge in strong-field gravity, but it remains to verify classical strong gravity in the Universe today. Many candidate black holes exist, but so far there is no direct evidence for an event horizon. Equally, gravitational waves are inferred only indirectly from binary systems containing pulsars. One probe of the very centre of a black hole may very well come from the phenomena of jets and outflows; another may come from a better understanding of the most powerful celestial explosions: supernovae and gamma-ray bursts. Finally, hyper-energetic particles may originate near black hole horizons, or from annihilation of dark matter particles. Therefore, the study of cosmic rays is included in the list of extreme areas of astronomy. With this motivation, Panel A focused on the following questions:

- How did the Universe begin?
- What is dark matter and dark energy?
- Can we observe strong gravity in action?
- How do supernovae and gamma-ray bursts work?
- How do black hole accretion, jets and outflows operate?
- What do we learn about the Universe from energetic radiation and particles?

We now discuss each in turn, concentrating on the most promising techniques by which each question may be attacked, followed by some details on the potential European experiments that are likely to be of most importance under each heading. On the cosmological front, deep optical and near-infrared imaging surveys of most of the sky can be envisaged, together with massive campaigns of galaxy spectroscopy. Much of this three-dimensional mapping of the Universe will also be enabled by large future radio telescopes. Completing this picture of the inhomogeneous Universe, detailed space-borne measurements of the microwave background polarisation will be needed. The latter may also probe primordial gravity waves, but their direct detection from space-borne interferometers such as LISA remains the ultimate goal. Direct imaging of black hole event horizons will require submillimetre and infrared interferometry. Future large X-ray and gamma-ray observatories will also probe these inner regions. The understanding of the nature of supernovae explosions and gamma-ray bursts will benefit from the same instruments, as well as from future extremely large optical/infrared telescopes. Cerenkov and radio detectors for high energy neutrinos employing huge volumes of water and ice will peer right into the central engines of these active objects. This is a highly active area, and there are many exciting European possibilities for the future under all these headings.

2.1 How did the Universe begin?

2.1.1 Background

The origin of the Universe is the ultimate question in physics, and it may seem presumptuous to claim that we have made any progress in answering it. Indeed, the description of the first instants of the Universe will require a quantum theory of gravity – which has yet to be fully formulated, despite encouraging progress in the area of string/M-theory. Nevertheless, cosmologists work with a ‘Standard Model’ that is extremely successful in explaining the evolution of the Universe from the first second to the present: the Big Bang theory. Its foundations are the so-called Friedmann-Robertson-Walker solutions of the general relativistic theory of gravity, plus confirmed observations of the expansion of the Universe; the abundances of primordially synthesised light elements, especially deuterium and helium; the thermal relic of the Big Bang in the form of an isotropic microwave background having a blackbody spectrum; and the matter structures such as galaxies, clusters and superclusters, formed by gravitational collapse of primordial fluctuations. What the theory leaves unanswered concern its initial conditions: the origin of the homogeneity and flatness of spatial sections; the origin of matter and radiation, in particular the origin of the matter-antimatter asymmetry; the origin of the primordial seeds for structure; the origin and nature of the dark matter and dark energy; and the origin of the Big Bang itself.

Today, these initial conditions are described within a broader paradigm – cosmological inflation. Inflation predicts a nearly uniform Universe, with negligibly small spatial curvature; and, more importantly, a Gaussian spectrum of small amplitude adiabatic density perturbations, the seeds for galaxies and anisotropies in the cosmic microwave background (CMB), plus a stochastic background of gravitational waves whose amplitude is directly related to the energy scale of inflation. Inflation is supposedly driven by the vacuum energy of a hypothetical scalar field; this necessarily involves physics beyond the well-understood Standard Model of particle physics, so there is no consensus on either the identity of the scalar field nor the initial conditions that lead to an inflationary phase. Nevertheless, the inflationary paradigm provides a useful framework for discussion of phenomenology – in particular the generation of the observed density perturbations.

At the end of inflation, the huge energy density driving the rapid expansion is presumed to convert into matter and radiation in a process known as reheating, which essentially marks the beginning of the classical ‘hot Big Bang’ phase. This process is not well understood since, as remarked above, the physical nature of the energy density responsible for inflation is as yet unknown. However it may have novel phenomenological signatures such as non-thermal production of the matter-antimatter asymmetry, supermassive dark matter particles, gravitational waves, and magnetic fields. In general, the early Universe has a rich potential for leaving relics in the present (see also § 3.1), and much can be inferred from their presence (or absence).

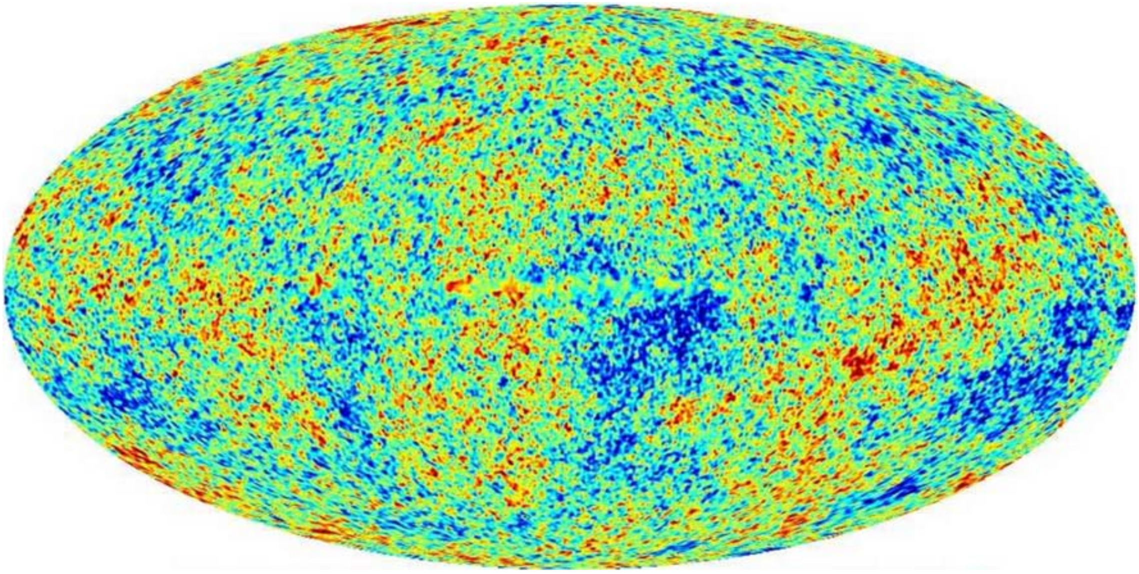

Figure 2.1: A map of the Universe showing the difference in the intensity of the microwave background. The maximum difference in this image is less than 10^{-5}. The map has been produced with the data of WMAP.

2.1.2 Key observables

The generic signatures of inflation are the 'tilt' of the spectrum of density perturbations (essentially, the extent to which the metric of space-time deviates from a pure fractal), together with a gravitational wave background whose amplitude would tell us the energy scale of inflation.

The data from the WMAP satellite (Fig. 2.1) indicate that the spectrum is indeed tilted below scale-invariance with a fitted power-law index of $n_s = 0.95 \pm 0.02$, but set only a weak upper limit of 45 per cent on the fraction of the large-scale angular power in the CMB contributed by gravitational waves (tensor modes). Future observations, particularly of so-called B-mode polarisation in the CMB, are expected to be sensitive to tensor fractions as small as a few per cent, as against a prediction of order 10 per cent from the simplest models. There are however many inflationary models that predict much lower levels. Other potential signals of the inflationary era could be features in the spectrum of density perturbations associated, e.g., with possible phase transitions occurring during inflation when the Universe supercools to a low temperature. Observations of large-scale structure provide a valuable complementary probe of such phenomena. It may also be possible to study inhomogeneities at the end of the so-called Dark Ages, when neutral hydrogen (H I) became reionized by the first objects (§ 3.1).

2.1.3 Future experiments

Eventually, one would hope to detect relic gravitational waves directly, using a post-LISA satellite (Fig. 2.2). However, the most immediate prospect is via the CMB and this has become the centrepiece of all future CMB experiments: the USA's SPT and ACT on small scales, and on large scales ESA's Planck mission followed by an all-sky polarisation mapper. One aim in this field is to increase the sensitivity to the nanoKelvin threshold with thousands of detectors at arcsecond angular resolution. Progress is likely to be limited by imperfect removal of foreground signals from our Galaxy, as well as from other galaxies, so an improved understanding of this will be essential in pursuing an understanding of the early Universe.

Although direct detection of high-frequency gravitational-wave relics from inflation is expected to be challenging, larger signatures are possible, in particular from so-called topological defects,

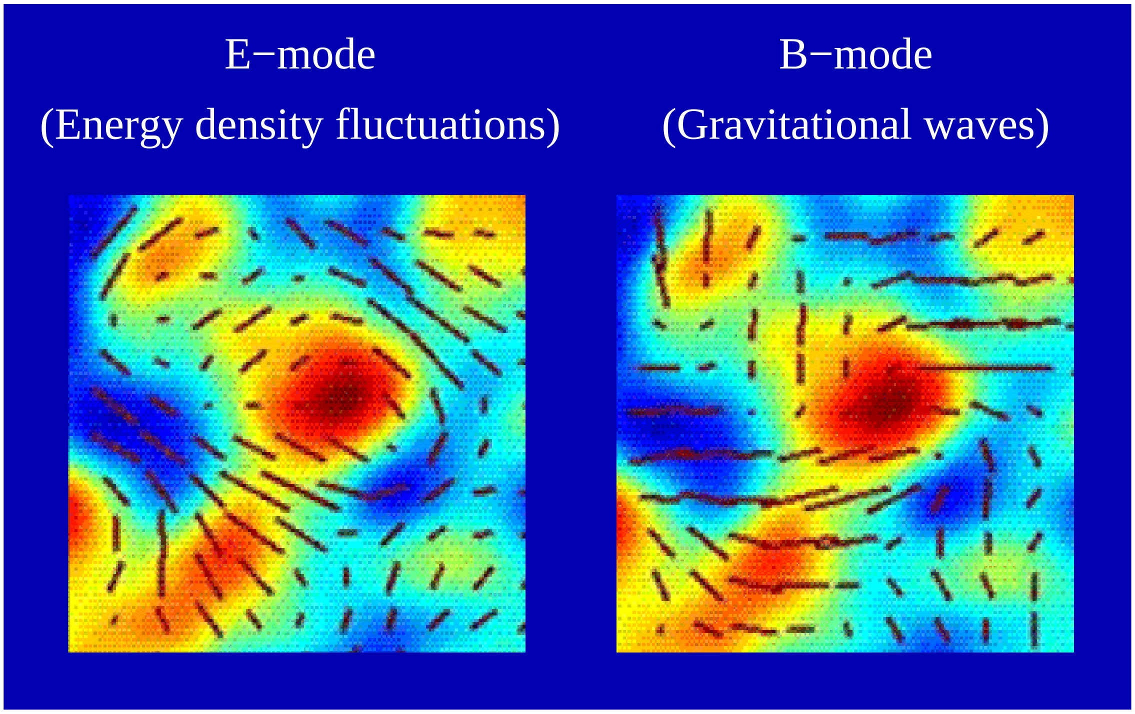

Figure 2.2: An illustration of possible signals in the polarization of the Cosmic Microwave Background. The color shows the CMB total intensity and the 'sticks' indicate the direction of polarization. The first panel shows scalar ('E-mode') signals, which arise from simple density fluctuations. The second panel shows tensor ('B-mode') signals, which arise only from primordial gravitational waves. Such waves would be generated in the very first instants of the expanding universe, and their detection is perhaps the single most important goal in fundamental cosmology.

which are structures in primordial scalar fields, analogous to, e.g., vortex rings in fluids. In addition to possible detections with ground-based detectors or with LISA, the high-precision timing of an array of radio millisecond pulsars complements the frequency coverage by being sensitive to gravitational waves in the nano-Hz to micro-Hz regime. The SKA will have the sensitivity and frequency coverage to discover and time about a thousand millisecond pulsars for this purpose.

These experiments, as well as numerical simulation of cosmological processes such as reheating after inflation, will require substantial supercomputing resources.

2.2 What is dark matter and dark energy?

2.2.1 Current status

All structures in the Universe bound by gravity, from individual dwarf, spiral and elliptical galaxies, to clusters and superclusters, appear to be dominated by unseen matter. Astronomical studies of galactic dynamics, gravitational lensing, formation of structure, X-ray emission of galaxy clusters, and CMB anisotropies provide good arguments that most of the dark matter is 'cold', i.e., consists of slowly moving non-relativistic particles that are non-interacting. However several observations on galactic scales suggest that the situation may be more complicated. It has even been proposed that dark matter may be an illusion due to modification of Newtonian dynamics (MOND) at very low accelerations. Establishing the identity of the dark matter or alternatively demonstrating that MOND is physically and cosmologically viable, is among the key challenges in cosmology (and fundamental physics).

A combination of recent cosmological observations has led to the amazing conclusion that the Universe is dominated by a uniformly distributed form of (dark) energy that behaves as Einstein's cosmological constant Λ. The chief evidence for this conclusion is that the expansion rate is accelerating (from distances to Type Ia supernovae, § 2.4); that the Universe is flat (from the CMB); and that dark matter cannot provide the critical density (from large-scale structure). Any two of these three observations establishes the reality of dark energy, so their unanimity is impressive. Dark energy could be dynamical, in the form of an ultralight scalar field that is slowly evolving down its potential to some asymptotic minimum which may be at zero – this has been named 'quintessence'. But the energy density of dark matter and dark energy today are comparable, even though they would have varied differently during the past evolution of the Universe. Some quintessence models with exponential scalar potentials can 'track' the energy density of dark matter, but fine-tuning is still necessary to obtain the negative pressure required to drive accelerating expansion. New possibilities emerge if the quintessence field can interact with dark matter. There may be violations of the equivalence principle due to the new long range force, and even time variations of coupling constants if quintessence is identified with 'moduli' of string / M theory.

Given that dark matter and dark energy have similar energy densities it is likely that they are somehow related in a fundamental theory. For example according to string theory we live in a world with ten dimensions but are able to experience only four of them. This picture can in principle explain why gravity is so much weaker than the other fundamental forces and suggests that dark energy may be mimicked by a modification of Einstein's General Relativity on very large scales (e.g., through the opening up of a new spatial dimension). Alternatively, the usual Friedman–Robertson–Walker cosmological model may be an oversimplified description of the real inhomogeneous Universe and dark energy may be an artifact of interpreting the data in the wrong model. It is essential in this regard to formulate new observational tests that can discriminate between these possibilities.

2.2.2 Experimental signatures

The simplest and most appealing route to the identification of cold dark matter particles lies in the laboratory, through direct detection in underground nuclear recoil experiments – this is expected to require a ton-class bolometric detector such as EURECA, 100-250 kg scintillation detectors such as ANAIS and the DAMA / LIBRA upgrade, or novel liquid noble gas detectors under development such as ArDM, XENON, WARP and ZEPLIN. Detection of plausible particle candidates such as supersymmetric neutralinos at accelerators such as the LHC will be a major advance, although it would not be possible to prove that such particles are sufficiently long-lived to be cosmologically relevant. A complementary approach is to search for the annihilation products (gamma-rays, neutrinos, antiprotons / positrons) of dark matter clustered gravitationally on various scales – from the centre of the Sun to the centre of the Galaxy. Such measurements will be pursued to interesting levels by the GLAST satellite and the planned upgrades of atmospheric Cerenkov detectors including HESS and MAGIC, as well as neutrino detectors such as ANTARES and IceCube. To exclude astrophysical sources of this gamma-ray signal requires detailed radio and X-ray surveys. On larger scales of clusters and beyond, gravitational lensing offers the main route to determining the dark matter distribution in conjunction with other probes of structure although these are affected by non-linear evolution effects. In addition, cosmological probes (CMB in combination with large-scale structure) are sensitive to the masses of neutrinos (which would constitute 'hot' dark matter), at a level below the reach of the best laboratory experiments.

Concerning dark energy, we want to know: Is it there? Is it Λ or quintessence? Is it a nonlinear effect of General Relativity, or something else? Does it interact with (dark) matter? This amounts to measuring the parameter w which defines the equation of state of the Universe ($P = w\rho c2$, where P is the pressure, ρ is the density, and c is the speed of light) as a function of redshift. This is possible because the history of dark energy affects both geometry (the relation between distance and redshift) and the rate at which density fluctuations develop under gravitational instability. Having these two independent probes is extremely valuable; any inconsistency between them could indicate a failure of General Relativity.

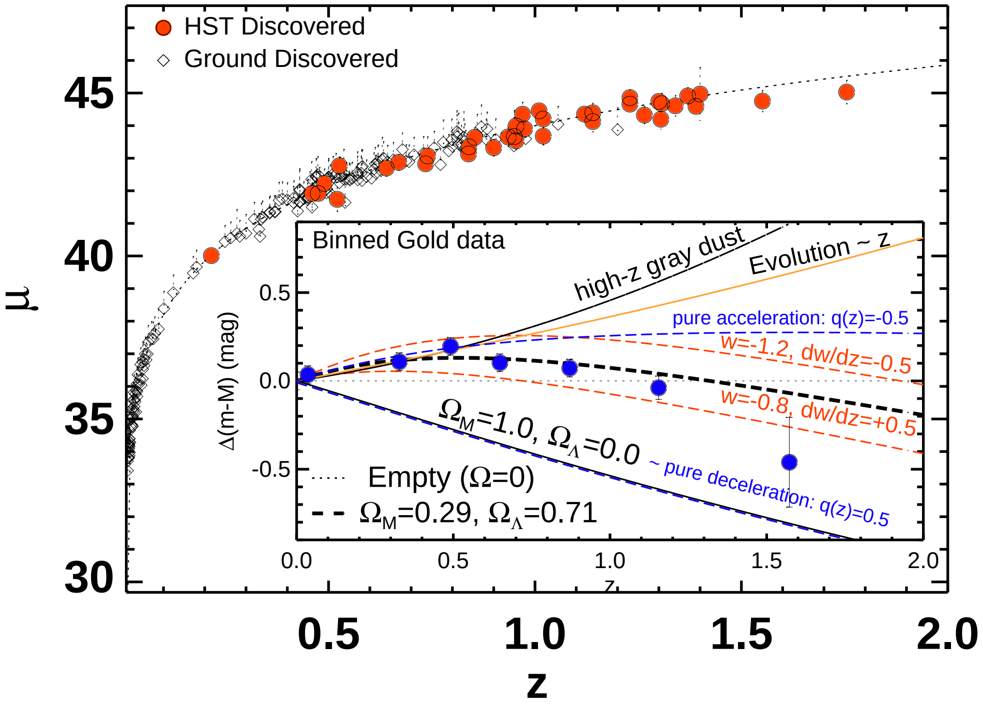

Figure 2.3: Hubble diagram for Type Ia supernovae observed with ground based telescopes and the Hubble Spece Telescope. The best fit line has $\Omega_M = 0.27$ and $\Omega_\Lambda = 0.73$ (Riess et al. 2006).

2.2.3 Future strategy

Dark energy has dominated the expansion of the Universe only since a redshift less than unity; thus the effects of dark energy are best observed at such relatively low redshifts. Higher redshift measurements are dominated by dark matter. However, variable dark energy could provide signatures at larger redshifts and hence these need to be explored as well. The main techniques for joint study of dark matter and dark energy are:

Cosmic Microwave Background. Ground-based experiments will probably measure the very small-scale anisotropies adequately, but a post-Planck satellite will be required (if foregrounds can be tamed: see §2.1.3), in particular to probe the scales larger than a degree, which contain most of the information on inflationary gravity waves. The CMB will be a common basis for many other tests, lifting the degeneracies inherent in these techniques. To remove foregrounds, deep radio surveys including polarisation need to be performed.

Gravitational Lensing. This technique can probe dark energy both via the distance-redshift relation, and its impact on the growth of density fluctuations. Lensing is thus in principle as direct a probe of cosmology as the CMB, although it does depend on an understanding of density fluctuations in the nonlinear regime. The key ingredient in the method is to obtain photometric redshifts, and a wide wavelength coverage is critical, extending to the near-infrared for redshifts beyond one. The way forward is to construct a very large CCD imager, similar to the proposed LSST – although the SKA could in principle carry out such work using radio measurements of gravitational image shear and redshifts of neutral hydrogen. Lensing studies would be greatly aided by data from space – both for improved image quality, and for the ability to probe the near-infrared, since the background levels in space are orders of magnitude below those for ground-based telescopes.

Baryonic Acoustic Oscillations. The baryon-oscillation signature of sound waves in the spectrum of large-scale structure can be measured using a dedicated survey of more than a million redshifts, which will have many other applications, and requires a dedicated ground-based facility. Either spectroscopic or photometric surveys are possible, although the latter needs to measure 30 times as many objects for the same precision. This work may also be possible using HI surveys with the SKA, at least out to redshifts of 1.5. In any case, a detailed understanding of the relation between galaxy tracers and dark matter will be required.

Supernovae. The current supernova surveys will provide limits on a constant equation of state parameter of around 7 per cent. The precision is limited by the intrinsic scatter of the peak luminosity of the supernovae, unknown extinction in the supernova host galaxies and possible secular evolution. Ways to address these limitations are a better understanding of the explosion and minimising extinction by observing at infrared wavelengths. Both of these issues can be addressed by systematic investigations of large samples of nearby supernovae, using robotic optical and near-infrared telescopes of about 2 m aperture. Evolution can only be controlled by detailed comparisons of spectroscopic data and spectral energy distributions. These require extremely large ground-based telescopes or a medium-size space telescope for the analysis. The numerical simulations of the explosions will need to continue.

As for more exotic possibilities, extra dimensions will appear through modifications of gravity, which can occur on any scale. The smallest-scale effects can be studied directly today with experiments to measure variations of the strength and nature of gravity on scales below a millimetre, plus improved traditional tests of General Relativity, such as violations of the equivalence principle. Important information will also come from Gaia astrometry and radio pulsar timing.

Large-scale effects manifest themselves through modifications of the Friedman equation and growth rates of structure. The CMB is again the best bet, but needs to be complemented by lower-redshift probes (which often extend the spatial range studied). The proposed CODEX experiment on an ELT could measure the accelerating expansion directly, rather than indirectly via properties of inhomogeneities. For varying fundamental constants, either precision measurement (timing with atomic clocks, e.g., ACES) to get derivatives at the present epoch, or detailed spectroscopy versus redshift, by virtue of very precise mapping of absorbers on the line of sight to quasars.

2.3 Can we observe strong gravity in action?

2.3.1 Background

Regions of strong gravity, where spacetime is significantly curved and orbital velocities are a significant fraction of the speed of light, occur around neutron stars and black holes. General Relativity is required to describe the behaviour of matter and radiation in these regions. Conditions become increasingly extreme as the event horizon of a black hole is approached, providing stringent tests of theory and of our understanding.

Strong gravity induces many observable effects, including large Doppler shifts caused by high velocities, large gravitational light bending near the neutron star or black hole, large deviations from simple Keplerian orbits, and large gravitational redshifts for us, the outside observers. Observations of neutron stars and black holes can therefore be used to test and refine our understanding of strong gravity and General Relativity, provided we can measure the motion of matter or photons at a few gravitational radii.[1] The radius of a neutron star is about seven gravitational radii

---

[1] The gravitational radius of an object with mass M equals GM/c^2, where G is the constant of gravitation and c is the

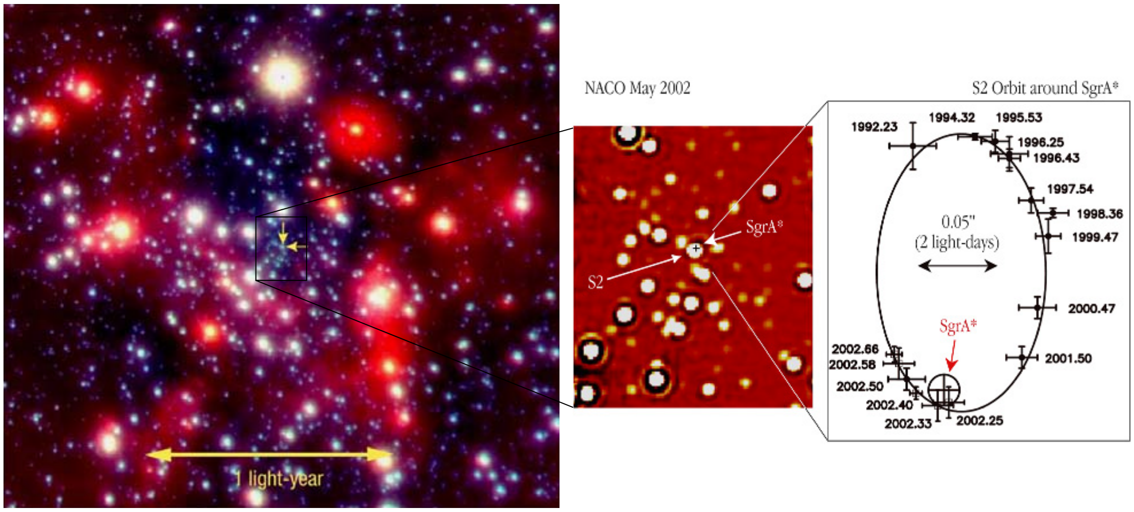

Figure 2.4: The centre of our Galaxy (Sgr A) and the star S2 which orbits it. The measured motions of this and other stars reveal the presence of a black hole of about three million Solar masses in the centre of the Galaxy. Based on work by the groups of R. Genzel and A. Ghez. Images courtesy of ESO.

and the radius of the event horizon is two gravitational radii for a non-spinning black hole, reducing to one as the spin approaches its maximum value. Observations can be made by detecting photons from matter orbiting around or accreting onto a neutron star or black hole, of matter on the surface of a neutron star, or of gravitational waves from neutron stars or black holes merging with each other. The mass of a black hole can be measured at large distances where gravity is weak. The determination of the spin requires probing the strong field regime, either by obtaining information from the immediate proximity of the black hole through imaging or by measuring relativistic spin-orbit coupling of pulsars in tight binary orbits about a black hole. The amount of spin depends on the growth history of the black hole, with many mergers of smaller black holes commonly leading to low spin and continuous accretion of material resulting in high values. The strong gravity regime generates the most luminous events in the Universe, but they make for challenging observations. Nevertheless the time is now ripe for rapid progress in the field.

Neutron Stars. Radiation from the surface of a neutron star tends to involve high energies, manifesting itself in the X-ray regime. Light bending and gravitational redshift both contribute to the observed profiles of the absorption lines seen during an X-ray burst. Quasi-periodic oscillations in the X-ray flux from accreting neutron stars with weak magnetic fields sometimes occur in the kHz range indicating they originate from radii close to or in the strong field regime.

Neutron stars are also observable as radio pulsars, acting as a precise clock and moving as a test mass in the gravitational field of a companion when found in a binary system. General relativistic effects are measurable to high precision in such systems, and the famous Hulse–Taylor binary pulsar provided the first observational evidence for gravitational waves. The loss of orbital energy due to gravitational wave emission in binary pulsars leads to a shortening of the orbital period, at an increasingly fast rate as the neutron stars approach and finally merge. The gravitational wave 'chirp' from the final few seconds is an excellent test of strong gravity. Pulsars orbiting stellar and (super-)massive black holes allow us to accurately measure their mass, spin and quadrupole moment. Such results would provide unprecedented tests of General Relativity and alternative theories of gravity, and would complement the results from methods described below.

Black Holes. The motion of the innermost orbiting stars around the three million Solar mass black hole in the Galactic Centre, coincident with the radio source Sgr A*, may probe relativistic gravitational effects down to hundreds of gravitational radii in the next ten years. The strong gravity

---

speed of light. This radius is 1.5 km for the Sun.

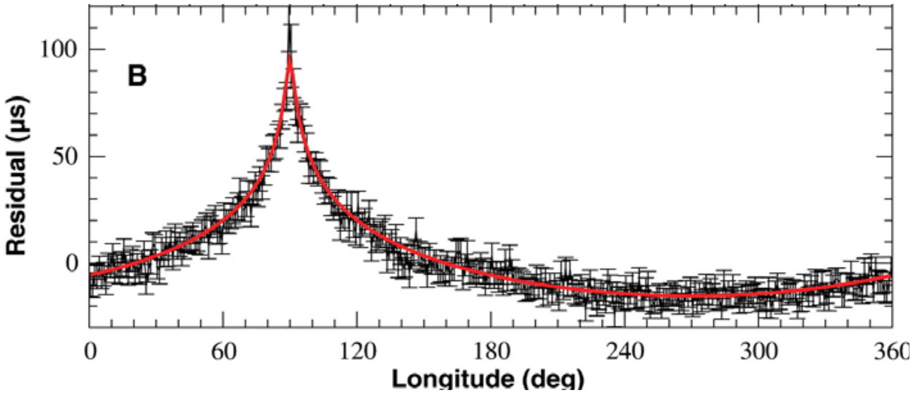

Figure 2.5: The Shapiro time delay measures the relativistic time delay experienced by the pulses from one of the pulsars in the the double pulsar system PSR J0737-3039A as they pass through the gravitational field of the other neutron star. The red line shows the predicted delay based on General Relativity, and the agreement is within 0.013% of the theory, providing one of the best tests in the strong-field limit. (From Kramer et al 2006)

regime can be imaged (appearing as a 'shadow') in the radio band if the emission region has an appropriate disposition. VLBI observations of Sgr A* already probe down to ten gravitational radii. Interstellar scattering and optical depth effects require observations in the submillimetre range, with baselines of $\simeq$ 4000 km. Optically thin emission from near the event horizon is also observed as near-infrared and X-ray emission. In particular, near-infrared interferometry can probe General Relativity through astrometric measurements of 'orbiting blobs' and flares.

The most luminous part of a radiatively-efficient accretion flow around a black hole is the innermost region. Comparison of the light emitted by accreting massive black holes (quasars and active galaxies) with the local mean black hole mass density shows that massive black holes attained most of their mass by radiatively efficient accretion. Indeed the rest-mass to energy conversion efficiency must be at least 10 per cent, which requires that the black hole is spinning reasonably rapidly. The inner luminous parts of the flow are then well within the strong gravity regime. The inner regions are strong X-ray sources, as emphasised by the rapid variability often seen. Iron is the most abundant element with strong emission lines to survive in the flow and provides a good diagnostic in terms of the 6.4-6.9 keV line in the rest frame, observed to be broadened by the strong Doppler and gravitational redshift effects. Observations of some Seyfert galaxies and also black hole binaries in our Galaxy show extreme broadening indicating that the accretion disc extends down to around two gravitational radii, implying that the black holes are spinning very rapidly. The broad iron line is induced in the accreting gas by reflection of X-ray radiation. The relative strengths of the apparent flux of the primary X-ray emission in the strong gravity regime and the secondary reflection component are strongly influenced by gravitational light bending. Observations suggest that such bending does take place. Future theoretical modelling coupled with sensitive observations should enable the effects to be separated clearly.

When galaxies merge their central black holes can approach and merge, generating enormously luminous events in gravitational waves. Space-based gravitational wave interferometers should detect such inspiral and merger events, which will give exquisite tests of General Relativity.

2.3.2 Experiments

Event horizon imaging can be achieved by combining individual submillimetre-telescopes – perhaps together with ALMA – in VLBI observations of the Galactic Centre. Near-infrared interferometry with the VLTI will probe general relativistic effects using astrometric techniques. Together these two experiments would not only provide direct evidence for the existence of an event horizon but also determine the spin of the black hole and test other effects predicted by General

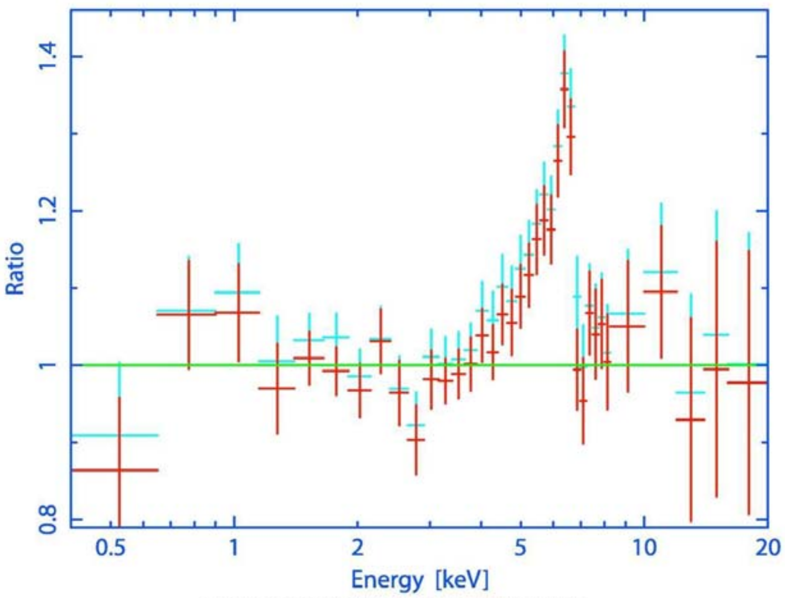

Figure 2.6: XMM-Newton spectrum showing a relativistic Fe K-line around active galactic nuclei. Image courtesy of MPE and ESA.

Relativity. The latter may require long-term observing of flares, possibly with dedicated facilities.

Gravitational waves are expected to be strong and distinctive during the final stage of the merger of neutron stars and black holes. For neutron stars the frequencies lie in the kHz band and the waves are detectable by ground based instruments such as LIGO & VIRGO. For massive black hole mergers the frequencies are milli-Hz and a space mission such as LISA is required. Radio pulsar timing can also detect gravitational waves in the nano-Hz band.

A Galactic census of pulsars is possible with the SKA if much of the collecting area is concentrated in the array core. It is expected that about 100 relativistic binary pulsars and pulsars orbiting stellar black holes and Sgr A* could be found, allowing much improved tests of General Relativity.

The broad X-ray iron lines from active galactic nuclei have low fluxes (see Fig. 2.6), and measuring the lines on timescales comparable to the light crossing time of the strong gravity region requires large collecting area, such as expected from the proposed missions XEUS and Constellation-X. The lines are brighter from Galactic black hole binaries, but cannot be studied on the appropriate short timescales. Rapid variability and burst studies of bright Galactic sources, such as neutron stars and black holes require large collecting areas and the ability to deal with high count rates.

2.4 How do supernovae and gamma-ray bursts work?

2.4.1 Current status

Supernova explosions represent the endpoint of stellar evolution and the most extreme conditions in the Universe with respect to temperature, matter and energy densities. They are responsible for the formation of neutron stars and stellar mass black holes, and plausibly for much of the energy input to the interstellar medium and the acceleration of galactic cosmic rays up to $\sim 10^{15}$ eV. The explosions create some of the most critical elements for life and are pivotal for the redistribution of higher elements throughout the Universe. Finally, supernovae represent the most important distance indicators for cosmology (§2.2.1). Two general physical mechanisms for supernova explosions are currently discussed: core collapse to a dense endpoint – neutron star or black hole – and the explosive thermonuclear incineration of carbon and oxygen in a white dwarf star. No single simulation has so far provided the required explosion energies from first principles for either class, and many questions regarding the explosion mechanisms remain open.

The most extreme explosions are the gamma-ray bursts, with two distinct types – short and long bursts. The recent excellent space-ground coordination has allowed a dramatic increase in our understanding of these objects. Many of the long duration bursts are closely related to very energetic supernovae. The exact relation of gamma-ray bursts and core-collapse supernovae is a topic of intense observational as well as theoretical interest. The origin of the short bursts remains unclear, although much information about their environment and energetics has been gained. The main theoretical contender for this class of gamma-ray bursts are merging neutron star–neutron star binaries or neutron star–black hole binaries although compelling evidence is lacking. For both cases theory predicts black hole formation and a relativistic jet with a high Lorentz factor.

While there are detailed numerical simulations for the evolution up to the explosion, the explosion physics, its aftermath, and the environmental effects, there remain severe observational and theoretical gaps. One reason is that the explosions themselves are hidden deep within the stars and, with the exception of neutrinos, become observable only when the shock or the thermonuclear flame breaks through the surface. The observed taxonomy of core collapse supernovae is suspected to be mainly the result of various degrees of mass loss before the explosion. The interaction of the supernova shock with circumstellar matter from radio to X-rays can shed light on the progenitor evolution. There is in this respect an important connection with supernova- and gamma-ray-burst afterglows, which can give crucial information about the progenitor and its local environment. The lack of any secure progenitor system for thermonuclear supernovae is one of the most critical impediments for their application as distance indicators and resolving this problem will consequently also have wide-ranging consequences for cosmology (§2.2.1).

The only direct observing window into the core collapse event is provided by neutrinos. However the detection of neutrinos is challenging and will probably be restricted to core collapses within the Galaxy and its nearest neighbours for many years to come.

Electromagnetic observations of supernovae are mostly concerned with the ashes of the explosion and their distribution. While optical, infrared- X- and gamma-rays probe the decay of newly synthesised elements and their distribution in the ejecta, the direct detection of the ^{56}Co gamma-lines would provide a unique calibration of thermonuclear supernovae. The optical and near-infrared emission provides a progressively deeper look into the ejecta and after about two to three months the ejecta become transparent. Observations of the longer-lived radioactive isotopes (^{57}Ni, ^{57}Co and ^{44}Ti) together with the most abundant, synthesized elements provide strong constraints on the explosion mechanism. Models of core collapse supernovae indicate that strong deviations from spherical symmetry are required to make the explosions work, mostly related to the neutrino transport regenerating the shock wave. Whether large asymmetries are expected in thermonuclear supernovae is still debated.

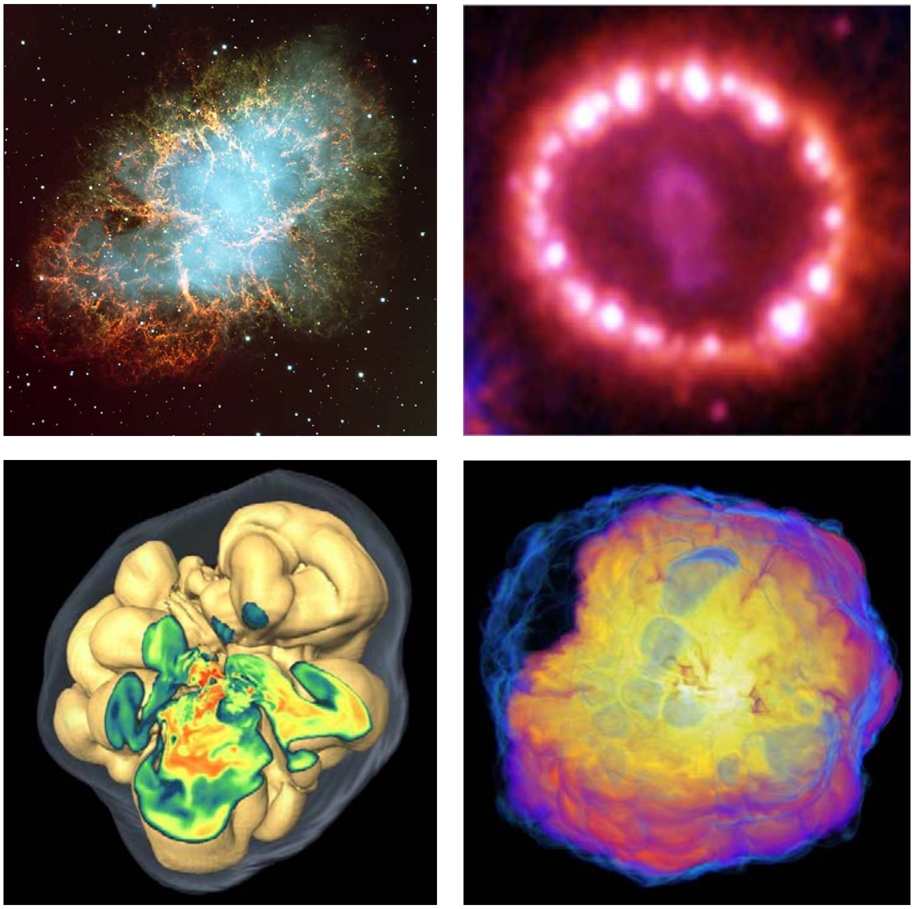

Figure 2.7: Top left: The Crab nebula, a supernova remnant resulting from a supernova in 1054 (image courtesy ESO). Top right: HST images of the supernova remnant resulting from the SN 1987A. Bottom: Two 3-D Simulation of the core collapse and resulting supernova for a SN Ia collapse. Pictures courtesy of Janka (left) and Hillebrandt (right).

2.4.2 Key questions

Which stars explode as supernovae? To date we have observed only a handful of progenitor objects and the picture is rather spotty. What are the nucleosynthetic yields of supernovae of different progenitor mass and metallicity? How do the first generation of massive stars end their lives?

Neutrinos are messengers from the core-collapse itself and have been observed only in SN 1987A in the Large Magellanic Cloud, a nearby companion of the Galaxy. Direct observation of the core collapse and its impact on the innermost regions of the supernova is needed to understand the conditions in the transition of matter to nuclear densities or even black holes.

The main unsolved problems for the core collapse in massive stars include understanding the properties of hot matter above nuclear density, the transport of neutrinos in the core, and the effects of rotation and magnetic fields. All these issues require sophisticated, multidimensional magneto-hydrodynamic simulations, including allowance for General Relativistic effects.

Another challenge for such studies is to explain the large variety of neutron star properties that is observed. What causes the large birth velocity of pulsars exceeding 1000 km per second? When and how is the superstrong magnetic field of magnetars created? What is the link of magnetars to high-magnetic field pulsars and gamma-ray bursts?

The gamma-ray bursts add new problems to this, including the formation of a possible accretion disc and the launch of the relativistic jet. A further unsolved problem is the acceleration of the relativistic particles and the formation of the magnetic field necessary to explain the non-thermal radiation in the burst itself and in the afterglow. The relation between the explosion engine creating the jets and the high-energy emission and the supernova itself remains an open question.

2.4.3 Future experiments

We have only a handful of pre-explosion observations of massive stars in nearby galaxies that were the progenitors of core-collapse supernovae. For future explosions high spatial resolution observations are required to securely identify the supernova progenitor relative to nearby, unrelated objects. This needs large telescopes with adaptive optics, as will be available with an ELT. Little is known about the Type Ia progenitors; a binary nature for thermonuclear supernovae has been postulated, but direct detection is out of the question, as the white dwarfs and the companions are too faint to be seen in external galaxies. Possible indirect detections include early X-ray, optical, or radio emission during the first few days of the explosion, when the shock may interact with circumstellar material. The statistics of merging white dwarf binaries, with future gravitational wave detectors, would improve the constraints on the progenitor models.

The next core-collapse supernova in the Galaxy will certainly be detected by a variety of neutrino observatories providing a unique tool for studying nuclear matter in the extreme state encountered in the supernova core, and the time evolution of the energy spectra of the various flavours are sensitive to (matter-enhanced) neutrino oscillations in the stellar envelope. Through detailed studies it will be possible to distinguish between the normal and inverted neutrino mass ordering, and perhaps even reconstruct the internal shock structure. A clear identification of subtle effects due to flavour oscillations can be made by a megaton-class water Cerenkov detector or a 50-kiloton-class liquid scintillation detector. Novel experimental set-ups such as a 100 kiloton-class liquid argon time projection chamber would be a valuable complement.

It is expected that high energy neutrinos are produced in the gamma-ray burst fireball by photonuclear interactions of protons accelerated by the internal shock(s) with the observed gamma-rays. The production of high-energy neutrinos in gamma-ray bursts will create a diffuse high energy neutrino background which should generate about 10 events per year in a kilometre-scale Cerenkov detector. It may be possible to detect individual nearby bursts in conjunction with accurate timing and positional information provided by observations of the electromagnetic burst.

A detailed observational analysis of the result of the explosion has only been possible for SN 1987A. The value of this object for supernova research has been enormous. Future facilities will allow investigation of more distant and fainter objects to the same level of detail. With an ELT, in combination with the infrared capabilities of JWST, the study of a representative sample of different types of supernovae, including supernovae related to gamma-ray bursts, will become possible. In particular, observations of the nucleosynthesis and deviations from spherical symmetry will provide important tests of hydrodynamical explosion models for these events.

The VLA can detect emission from supernova-circumstellar interaction only for the nearest supernovae. The rapid evolution of the gamma-ray bursts make observing them very difficult. The SKA would provide access to long radio wavelengths, which are important for using synchrotron self-absorption as a diagnostic. All-sky radio monitors, like LOFAR, will detect the supernova and gamma-ray-burst blast waves as early as possible in the radio and follow their evolution. Sensitive X-ray telescopes, such as XEUS, will provide a complementary view in the X-rays, offering diagnostics of both line and continuum emission. The large number of 20 000-30 000 pulsars expected from the SKA will provide excellent statistics of a direct birth product of supernovae and their kick velocities.

The invaluable information on supernovae and gamma-ray bursts provided by new observational facilities will have to be supplemented by complex computer simulations of the explosions, requiring the largest supercomputers. For the white dwarf explosions the computational requirements are daunting as the flames have to be resolved, while the effects on the complete star have to be calculated in time-dependent and fully three-dimensional simulations for thermonuclear supernovae. For the core-collapse simulations one has to add the neutrino physics and a full description of general relativistic effects. Accessibility to supercomputer facilities is therefore essential.

2.5 How do black hole accretion, jets and outflows operate?

2.5.1 Background

Accretion by black holes is the most efficient source of energy in the Universe. The energy output from accreting supermassive black holes in the nuclei of galaxies accounts for about five per cent of the total energy emitted in the Universe today. Feedback from these tiny, but massive, objects is believed to have a profound effect on the host galaxy, terminating star formation in the most massive galaxies (§§ 3.1, 3.4). More generally, energy from accreting compact objects is a significant fraction of the total input into the interstellar and intergalactic medium.

The principle involved is gravitational energy release from matter falling in the strong gravitational field of a black hole. The details, however, are unclear. In the case of luminous quasars, emission of radiation from a thin accretion disc appears to be responsible. In about ten per cent of objects, a powerful jet is found which can transport matter at high, sometimes relativistic, speed over intergalactic distances. The exact nature and origin of such jets is obscure. It is likely to involve energy transported via twisted magnetic fields from the innermost regions of a disc immediately around the black hole, but current models so far lack predictive power. What has been shown, however, is that there is a close coupling between the infalling matter in the accretion flow and the outflowing material in the relativistic jets in some kind of jet-disc symbiosis. This close coupling extends to black holes of all mass and mass accretion scales, suggesting a common and fundamental underlying mechanism. This allows one to start unifying stellar mass and supermassive black holes into one coherent picture where very similar physical processes are at work despite very different observational appearances.

It is argued that radiation pressure limits the fuelling rate and the luminosity of accreting black holes – the Eddington limit. Observations generally support this, although the possibility of super-Eddington accretion is much debated. At low Eddington rates the nature of the accretion flow appears to change, with jets becoming common. The accretion is often then radiatively inefficient and may involve different temperatures for the protons than for the electrons. Accretion often leads to highly time-variable emission, probably because the viscosity, which transports the angular momentum out while the mass flows in, is magnetic in origin. Many of the most spectacular variable objects in the sky are accreting black holes and neutron stars.

The big technical challenge in the study of these objects is the physics of multi-dimensional fluid flows, including turbulence, magnetic fields, and radiative transfer. These phenomena are not only worthy of understanding by themselves; all observed phenomena depend on the fluid dynamics and radiative transfer, hence if we are to use the phenomena to get at the fundamental properties of black holes (§ 2.3), we must understand the dynamics of the flows around them. At present, even such basic issues as whether jets are composed of only leptons or contain baryons, and whether relativistic jets differ fundamentally from non-relativistic ones, are unknown. Besides better observational studies, very large-scale numerical simulations are essential tools to increase our understanding of these flows and their sources. Many of these issues are also relevant to the gamma-ray bursts discussed in § 2.4.

Compact objects and their jets are also prime candidates for sources of non-electromagnetic messengers from space: gravity waves, neutrinos, and cosmic rays. A plethora of new experiments are being constructed or planned to observe these, and provide a truly multi-messenger approach to the study of these extreme and enigmatic objects. Very large-scale numerical simulations of collisionless shocks, tenuous plasmas with charge separation, and the particle acceleration in these

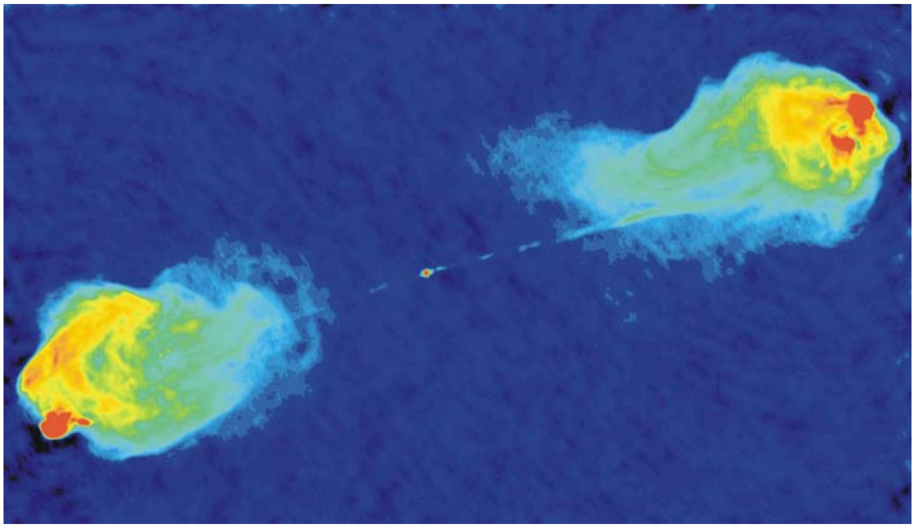

Figure 2.8: Radio image of the jet of the compact object Cyg A.

environments will be a key ingredient to understanding these phenomena.

The key questions to focus on for the foreseeable future are: How do we understand the inflows and outflows of compact objects? What messengers do we use to diagnose the phenomena? How does the energy output from these objects affect the Universe and its history?

2.5.2 Experiments

Accreting black holes and their jets produce radiation and phenomena throughout the entire electromagnetic spectrum, from radio galaxies at the lowest frequencies observed to the TeV emission from blazars. There are similarities, and differences, across the mass range from stellar mass black holes in X-ray binary systems, ultra-luminous X-ray sources, Seyfert galaxies up to the most powerful active galactic nuclei, the quasars, some of which harbour black holes with masses approaching ten billion times the mass of the Sun.

Individual studies are required of the brightest, nearest and most extreme objects, deep surveys to find the common factors and establish the probability for each mode of operation, and repeated all-sky surveys to find the transient sources. The latter can arise from sudden mass transfer from a binary companion, from accretion of a gas cloud or, for black holes of about ten million Solar masses, from the tidal disruption of a star that strays too close. There is a particular need for sensitive observations of the very innermost parts of both discs and jets. These can be identified either by high spatial resolution or through rapid variability. Some spectral features (such as X-ray lines) reveal the depth of the potential well and the proximity to the black hole (§ 2.3.1).

Because many of the most interesting objects in this class are transient or highly variable, with their stay in interesting bright states a tiny fraction of their life, all-sky monitors are key for studying black-hole physics, both in X- and gamma-rays from space and at radio wavelengths from the ground. Prospective facilities in this area are XEUS for X-rays, rapidly responding telescopes such as robotic optical telescopes or electronically-steered phased arrays in the radio regime (LOFAR, SKA). Many non-electromagnetic observatories require (or have their sensitivity greatly enhanced by) triggers from such monitoring instruments.

For most objects the horizon itself will remain outside our reach (except for sub-millimetre interferometry of the Galactic Centre). Since the phenomena from horizon to jet easily span twelve orders of magnitude in size, much progress is possible even if the horizon scale is not yet reached. Besides the spectroscopic observations already mentioned in § 2.3, a key tool will be high time-resolution observations with large-area detectors, to resolve and analyse the most rapid variability in the radiation from the inner regions of the flow. Present data on such phenomena as quasi-periodic oscillations already probe the very relativistic inner regions in this way, but larger detector areas are needed to bring more sources in reach of this technique, and to refine observations of the brighter ones.

2.6 What do we learn from energetic radiation and particles?

2.6.1 Background

The origin of cosmic rays has remained a mystery since they were first discovered nearly a century ago. There is an appealingly simple 'Fermi mechanism' for the acceleration of relativistic charged particles to high energies in astrophysical sources such as supernova remnants through stochastic processes involving energy exchange with the shock waves present in the turbulent magnetised plasma. The radio emission from such objects is understood as synchrotron emission by electrons being accelerated up to energies of at least 100 TeV, and the observed hard X-rays are consistent with inverse-Compton scattering of the photons present by the same electrons. The detailed acceleration mechanism and magnetic field generation is, however, not well understood. Neither is there any definitive observation of the distinctive (pion decay) gamma-rays which should result if protons and heavy nuclei are accelerated in such objects, although recent advances in TeV gamma-ray astronomy are bringing new data to bear on this issue. Such observations have also implicated new classes of objects as possible cosmic ray sources, including micro-quasars and binary pulsars.

An important consideration is that although the cosmic ray spectrum continues roughly as E^{-3} power-law up to energies E of at least 10^{11} GeV, sources in the Galaxy cannot accelerate particles beyond about 10^9 GeV – apart from the physical limitations for plausible sources (the 'Hillas criterion'), gross anisotropies should then be seen since the Galactic magnetic field is too weak to randomise the directions of such particles. The observed arrival directions are isotropic indicating that the sources at such high energies are very distant and therefore very energetic, perhaps gamma-ray bursts. In that case the interactions of the cosmic rays (if these are protons) with the CMB should cause a sharp 'Greisen-Zatsepin-Kuzmin' cutoff in the spectrum at about 5×10^{10} GeV. Presently there are contradictory observations concerning this and new definitive data from a large air shower array is eagerly awaited. If the spectrum is found to extend beyond this energy, it would suggest that the sources are in fact local – perhaps superheavy relic particles decaying or annihilating in the Galactic halo. Alternatively there may be something wrong with our understanding of cosmic ray propagation – it has even been proposed that Lorentz invariance may be violated at very high energies (perhaps through quantum gravity effects).

Ultra-high energy neutrinos should also be detectable from the extragalactic cosmic ray sources with forthcoming under-sea and under-ice neutrino telescopes, opening up a new window on the Universe as well as providing a unique laboratory for astronomically-long baseline studies of neutrino oscillations, sensitive to a variety of new physics. A large cosmic ray air shower array can detect ultrahigh energy neutrinos as quasi-horizontal showers and is sensitive to possible enhancements of their interaction cross-section due to new physics such as the opening up of new dimensions at the TeV scale as in some 'braneworld' models. For the propagation of cosmic rays in the Galaxy and throughout the Universe a detailed understanding of the interstellar and intergalactic magnetic field distribution is required.

2.6.2 Experiments

Present Cerenkov arrays, such as HESS and MAGIC, have demonstrated the power of these instruments for understanding the acceleration of particles in several classes of high-energy objects (Fig. 2.9). The next generation arrays, like the CTA, covering ~ 10 GeV to several TeV in energy,

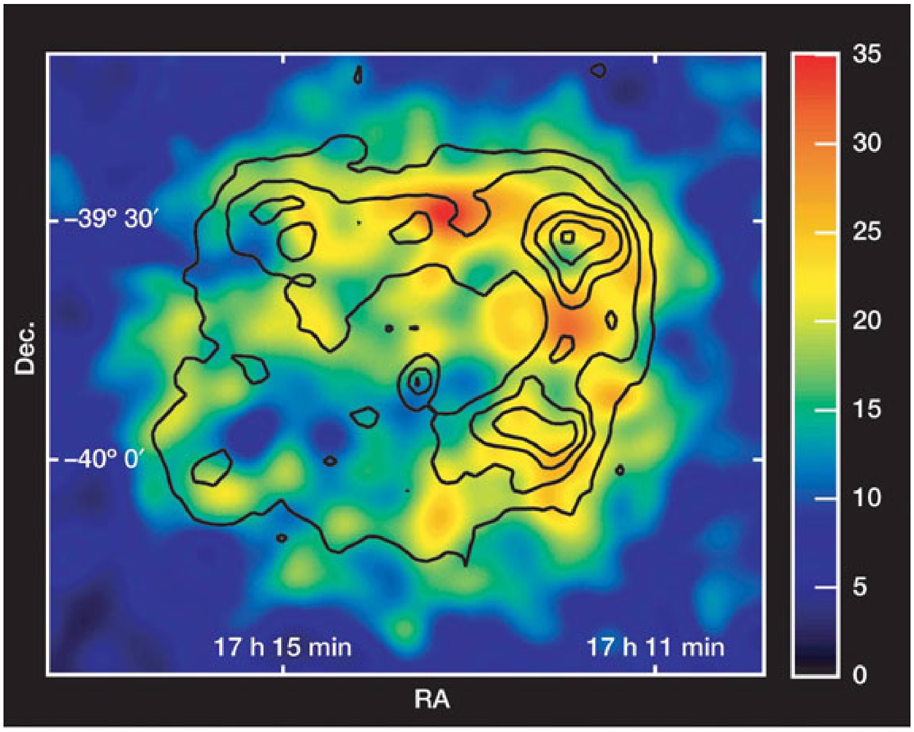

Figure 2.9: TeV gamma-ray image of the supernova remnant RX J1713.7-3946 obtained with the H.E.S.S. telescopes. The superimposed contours show the X-ray surface brightness seen by ASCA in the 1–3 keV range.

will greatly improve the understanding of the nature and acceleration mechanism of the particles.

Direct measurements are being made with balloon-borne experiments such as CREAM and TRACER which have had long duration flights in Antarctica, with more flights planned. The PAMELA magnetic spectrometer has just been launched on a Russian satellite and the bigger AMS magnetic spectrometer is awaiting launch on a space shuttle to the International Space Station. Meanwhile medium-scale air shower experiments such as KASKADE-GRANDE in Germany and TUNKA in Siberia are being expanded, and SPASE at the South Pole is being extended to IceTop. The HiRes experiment in Utah uses the air fluorescence technique to measure air showers and its successor the Telescope Array has just been approved. Meanwhile the Pierre Auger Observatory in Argentina, which employs both a giant ground array and fluorescence telescopes, is nearing completion and the site for its Northern hemisphere counterpart has been chosen in Colorado. Novel methods such as radio detection (LOPES, CODALEMA) are also being investigated. Looking into the future, a space observatory such as EUSO is being actively reconsidered by ESA while NASA has a similar project (OWL). At the very highest energies above 10^{21} eV, extremely large detector volumes are needed. Here experiments which use radio telescopes to look for radio signals from particles hitting the Moon hold promise.

Chapter 3

How do galaxies form and evolve?

We know very little about the nature of dark matter and dark energy, and we feel more familiar with 'ordinary matter', made of protons, neutrons, electrons, etc. Yet, much remains to be done to fully map the evolution of this 'baryonic' component of the Universe. The finite volume accessible to astronomical observations is bounded by the sphere at a redshift of about one thousand emitting the microwave background, beyond which the Universe is fully opaque to radiation. Within this volume, three quarters of all the baryons we could, in principle, detect lie between redshifts of seven and a thousand. These are the Dark Ages, between the epoch of recombination ($\sim$ 400 000 years after the Big Bang) and the most distant galaxy so far detected ($\sim$750 million years after the Big Bang). No direct evidence has yet been gathered for any kind of event in the Dark Ages, in spite of density fluctuations having grown by many orders of magnitude during this critical half billion years in the life of the Universe.

Nevertheless, our first glimpses at redshift seven reveal that the young Universe was by then almost completely re-ionized, while stars, galaxies, and quasars had begun to form and shine, many with the metal-rich signatures of even earlier generations of star formation. Understanding this rapid buildup of stars, metals, galaxies, and supermassive black holes, as well as the subsequent transformation of these young objects to the present-day Hubble sequence of galaxies is a major challenge. As the Universe was re-ionized, most baryons were heated by stellar radiation and mechanical energy input from exploding stars (supernovae) as well as from active galactic nuclei powered by supermassive black holes. As a result, over 90 per cent of the baryons were left in a diffuse intergalactic medium. Nearly half of even the local intergalactic baryons have yet to be detected. Understanding these cycles of formation and evolution of galaxies, and the intergalactic baryons and metals that link them, will require a clever combination of observational and theoretical approaches. The key questions addressed by Panel B are:

- How can we peer into the Dark Ages, and map the growth of matter density fluctuations from their tiny size at redshift one thousand to the formation of the first stars and galaxies?
- What are the dominant sources for re-ionization of the Universe: star light, black hole powered active galactic nuclei, or even decaying supersymmetric particles? How long did the process take?
- How did the structure of the cosmic web of galaxies and intergalactic gas evolve?
- What are the histories of the production and distribution of the metals in the universe, within and between the galaxies?
- How was the present-day Hubble sequence of galaxies assembled, as traced by the buildup of their mass, gas, stars, metals, and magnetic fields?
- What is the detailed history of the formation and evolution of our own Galaxy, and what lessons does it hold for the formation and evolution of galaxies generally?

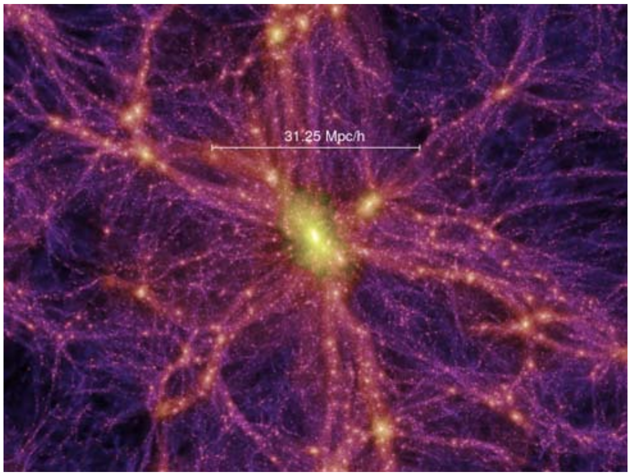

Figure 3.1: Image of a cluster of galaxies as calculated by the Millennium simulation, the largest cosmological N-body simulation carried out to date. It has been used to construct sophisticated semi-analytic models of galaxy formation that cover the complete galaxy population in a representative piece of the Universe down to stellar luminosities.

3.1 How did the Universe emerge from its Dark Ages?

3.1.1 Background

A number of important events occurred in the Dark Ages: the birth of the first stars, supernovae and black holes; the formation of mini-galaxies which, according to the hierarchical structure formation paradigm, constitute the building blocks of larger ones; emission of ionizing photons starting the process of cosmic reionization; finally, radiative, mechanical and chemical feedback processes shaping the evolution of the underlying cosmic structure. This fascinating cosmic period, and its equally interesting physics, have started only recently to be investigated, thanks to newly available and greatly improved computational and observational tools. The key problems that require a combination of theory and data include: Determine the Initial Mass Function of the first stars and its evolution; Constrain cosmic reionization; Understand the role of feedback from the first stars on subsequent galaxy formation; Determine the formation rate and fate of very massive black holes; Assess the direct observability of the first sources; Search for fossil imprints of the Dark Ages in the metallicity evolution / pattern of the intergalactic and intracluster medium, in low-redshift Ly-α emitters and in metal-poor halo stars in the Galaxy.

3.1.2 Key questions

The neutral Universe. The most promising technique to explore the evolution of the intergalactic medium during the Dark Ages is to use the 21 cm (1420 MHz) hyperfine transition of the ground state of neutral hydrogen, H I. This line, in principle, allows a superb tracing of the H I in the early Universe, and therefore a reconstruction of the reionization history as governed by the first luminous sources. H I 21 cm data from LOFAR, PAST, MWA, and, on a longer time-scale, SKA, will provide a breakthrough in the study of the Dark Ages, due to the astonishingly large number of independent patches in the sky which allow a study of those epochs with unprecedented large mass / size dynamical range. In order to exploit this large dataset properly, a substantial effort is required to understand the various contaminant radiation sources, including terrestrial ones.

An exciting possibility offered by a detection of the 21 cm signal from the Dark Ages is to determine the presence of annihilating / decaying dark matter radiation. At redshifts z between 50 and 100, the spin temperature of H I is locked to the kinetic temperature of the gas, as long as the collisional excitation rate of the hyperfine structure is sufficiently high. As the kinetic temperature is lower than the temperature of the Cosmic Microwave Background, due to the adiabatic cooling produced by the overall expansion of the Universe, this results in a 21 cm absorption signal at frequencies corresponding to 1420 MHz / $(1+z) \sim$ 20–30 MHz. The presence of such a feature is an undisputed signature of the linear evolution of the density field. Dark-matter-produced radiation could erase / modify this feature due to the corresponding increase of (i) the kinetic temperature, (ii) change in the HI ionization fraction, (iii) Ly-α pumping of the spin temperature. Although very challenging due to the opacity of the ionosphere below about 30 MHz, detection of this effect would provide one of the most clear indications on the nature of dark matter. HI 21 cm data could also constrain the size of ionized (HII) bubbles around sources located close to the end of reionization. The absence of neutral hydrogen within HII regions means that these are also sites where the sources can be studied with less interference from rest-frame Ly-α absorption. Some of these regions will contain luminous quasars, while most will host galaxies and groups of galaxies. The feasibility of this method requires careful assessment and calibration against radiative transfer cosmological simulations of the HII regions around bright high-redshift quasars in volumes as large as those investigated by the Millennium Run. Large volumes (to catch the large HII regions carved by luminous quasars and the density field on those scales) and reliable radiative transfer treatment (to account for clumping, shadowing and self-shielding effects of matter) are necessary.

First stars. Many arguments suggest that the first stars in the Universe were massive so that their Initial Mass Function was quite different from the one we observe today (§ 4.2). When and why the transition from this top-heavy mode to a normal one occurred is a matter of speculation. The existence of a critical metallicity above which the fragmentation properties of star-forming clouds might change drastically, appears to be the most intriguing possibility. The detailed physics of this fragmentation process must be put on firm and quantitative grounds and be confronted with observations. The precise value of the critical metallicity may depend on the presence of dust; if so, dust formation in the ejecta of the first supernovae must be understood. Other questions to be addressed are: is primordial star formation self-propagating through the collapse and fragmentation of the cold shells produced by supernova explosions? What are the environmental conditions leading to such events? What is the likely mass and metallicity range of this second generation of stars? Do these stars bear the signatures of the peculiar metal abundance patterns expected from the first massive stars? What is their connection with the metal-poor halo stars? Recent studies suggest that collapsing very massive stars en route to the formation of intermediate mass black holes could result in gamma-ray bursts. Understanding when the first stars formed and whether the very different primordial environment (no dust, no heavy elements, no dynamically important magnetic fields, low background radiation) produced noticeable effects concerning their properties and evolution is crucial. Stellar evolutionary models for a large range of masses, together with more precise metal yields, are required. Inherently three-dimensional effects as rotation, convection and instabilities of such objects should be included in the models.

Cosmic reionization. Observations of cosmic epochs closer to the present have established that the cosmic gas is in a (re-)ionized state. It is not yet known when the transition from the neutral state resulting from the recombination process (400 000 years after the Big Bang) to the ionized state started. This is partly due to uncertainty about the sources responsible for the production of ionizing photons: stars, quasars, dark matter particle annihilation / decays, virialized gas inside cosmic structures are the most plausible guesses. To complicate matters, a plethora of different possibilities exists for each of these sources: Metal-free or normal stars? Quasars or mini-quasars? What type of dark matter particles and hence radiation spectra? Gas in galaxy halos or groups?

A key problem is the determination of the history of reionization towards its completion. This evolution is governed by complex and interlinked physical processes ('feedback'): the ultraviolet radiation (and the associated heavy element production by stars) necessary to reionize the cosmic plasma affects galaxy and star formation, possibly depressing it and causing a delay in the progress of reionization. The post-reionization Universe is filled with precious information on how reionization proceeded and on the sources that caused it. Leftovers of reionization as

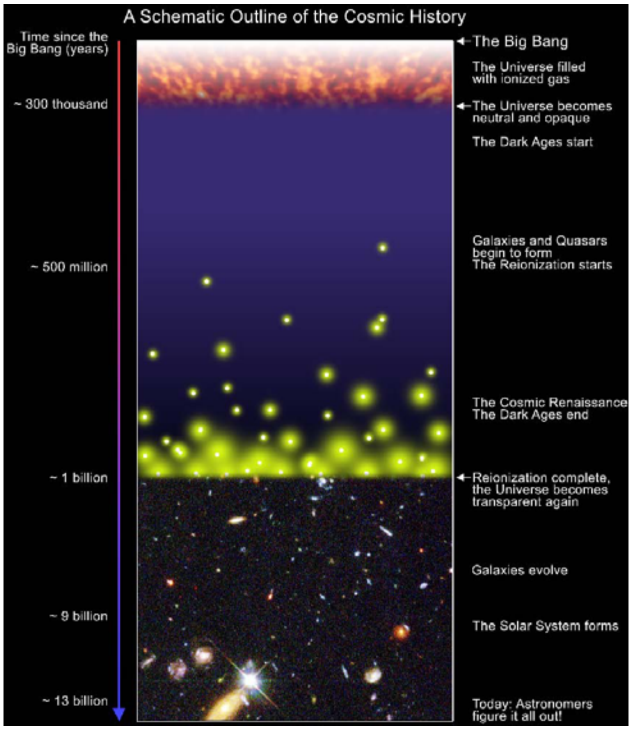

Figure 3.2: The reionization era placed within the evolution of our Universe. Image created by S.G. Djorgovski et al. & Digital Media Center, Caltech.

islands of H I that could not be completely turned into ionized patches are routinely detected by absorption-line experiments targeting distant quasars. Even the halo of our Galaxy may contain some of the most ancient (perhaps first) stars in the Universe, surviving as cosmic fossils.

The entire field will experience a strong burst of experimental activity in the near future. The Planck satellite will allow improved measurements of the CMB temperature and polarisation spectra. Sub-millimetre observatories including ACT and ALMA will explore for the first time the sub-arcminute features of the secondary anisotropies spectrum resulting, amongst other processes, from patchy reionization. High-redshift quasars and gamma-ray bursts can be used as targets for absorption-line experiments (in the near-infrared with the VLT and ELTs; in the radio with E-VLA and ALMA), to investigate the reionization process in exquisite detail.

The keys to improved cosmic reionization modelling are: Simulating large representative cosmic volumes; Improving radiative transfer treatments to include the effects of small-scale gas / radiation inhomogeneities and anisotropies; Including the effects of the different feedback types in detail. At least four types of feedback processes occur during the Dark Ages: radiative, chemical, mechanical and magnetic. The first one, through the emission of ultra-violet photons by the sources, governs the formation / destruction rate of the key gas-cooling species (molecular hydrogen H_2), and hence the collapse of the first (mini-)galaxies. Chemical feedback results in a metallicity dependence of the cloud fragmentation process driving the Initial Mass Function from a top-heavy to a normal one. Mechanical feedback due to energy deposition of supernova explosions heats and ionizes the gas, causing a possible delay of galaxy formation and affecting the properties of the intergalactic / intracluster medium. Primordial seed magnetic fields can be amplified to significant strengths by supernova-induced turbulence and / or by compression in collapsing protogalaxies.

Constraining the nature of the reionization sources requires both observations and theory, with the latter helping in predicting number counts, luminosity functions, supernova / gamma-ray burst rates associated with different (black hole-powered, stellar and other) source types. Finally, reionization could provide one of the most solid ways to constrain the nature of dark matter, if the latter consists of decaying / annihilating particles, thus complementing analogous more local studies that will be carried out by GLAST.

First black holes and their evolution. The massive black holes in the centres of nearby galaxies were likely born after the first generation of very massive and short-lived stars. These seed black holes had masses in the range of tens to hundreds of Solar masses, thus qualifying as Intermediate Mass Black Holes which could provide a link between stellar evolution and quasar activity. Whether such IMBHs indeed form is related to the ability of a protostar to accrete a sufficiently large gas mass before the end of its life (§ 4.1). Progress on understanding the accretion process onto protostellar cores requires specifically designed radiative-hydrodynamic codes. It is crucial to understand the emission properties of such objects as a function of their environment and assess their impact on cosmic reionization and feedback on star and galaxy formation.

Various physical processes might have contributed to the growth of these seed IMBHs to their current large sizes, ranging from millions to billions of Solar masses. Observations suggest that efficient gas accretion (producing copious X-rays) is most likely the dominant process. Tracing the history of massive black hole growth to its origins therefore requires sensitive X-ray observatories (like XEUS), capable of detecting the emission from IMBHs and spectral features (such as the Fe emission line) that can help their identification and study of their physical properties out to very high redshifts. Other processes also contribute to the massive black hole growth, such as black hole mergers in high-density environments and tidal capture of stars. These produce abundant gravitational waves, which can be potentially detected with LISA (§ 2.3).

Accreting massive black holes appear to reside in metal-rich (up to three times Solar) and dusty environments: How did metals and dust form so rapidly and in so large amounts in the early Universe? Does pristine dust share the same properties as the present one? Answering these questions is fundamental and will be made possible by radio / millimetre-instruments detecting dust and molecules at the highest redshifts (Herschel, ALMA and SKA).

3.1.3 Future strategy/opportunities

Pushing the border of the explored Universe to the epoch when the first stars formed is very challenging. Distant objects are small and faint; thus sensitivity is a key instrumental requirement which can be achieved only by large photon-collecting areas. Also, light from the early Universe is redshifted into the infrared / sub-millimetre bands. Direct detection of the first stars will probably require JWST or an ELT in order to achieve the sensitivity needed to reveal the faint clumps of mini-galaxies hosting them. Also, by monitoring large patches of the sky at suitable time intervals (months) JWST and JDEM will be able to identify the most distant supernova explosions, thus tracing uniquely the formation history of the first stars (an experiment that should be complemented by gamma-ray burst searches). The first stars are copious emitters of Ly-α photons, which are redshifted into the near-infrared bands. The excess light (with respect to known galaxies) and the amplitude of its angular fluctuations in the wavelength range of 0.8-24 micrometer would provide firm evidence of the earliest cosmic star formation. This would require wide-field imagers in space, to measure fluctuations on scales ranging from a few arcseconds to a degree, and at least a medium-resolution spectrometer to detect the Ly-α cutoff. Detecting the emission of high-redshift H_2 (and other molecules), i.e., the gas providing the fuel for the formation of the first stars, could be achieved with far-infrared / sub-millimetre interferometers in space.

The future of reionization studies relies heavily on the detection of secondary anisotropies on sub-arcmin scales, requiring ALMA. An almost completely unexplored window is the part of the radio spectrum below 300 MHz which will open a new channel to study the Universe before the birth of the first galaxies. The contribution of LOFAR and SKA could be ground-breaking, if the challenge of controlling radio-frequency interference of both cosmic and terrestrial nature, and the development of next-generation receivers is successful. The predicted signatures of feedback

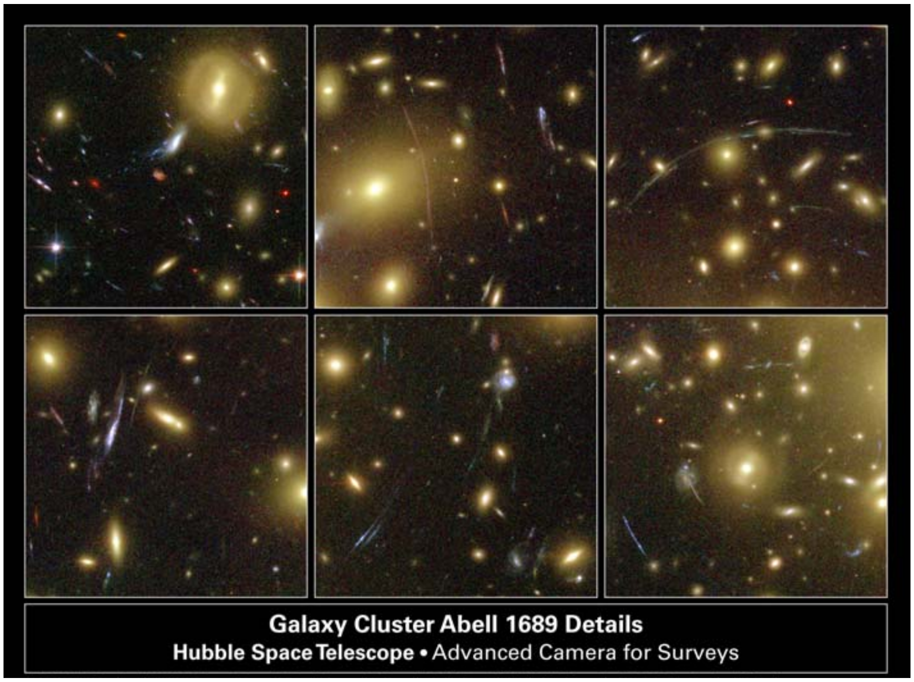

Figure 3.3: Detailed images of Abel 1689. The gravitational arcs caused by the bending of the light due to the massive cluster in front are clearly visible. Picture taken from the HST image archive.

left on the intergalactic medium (e.g., hot bubbles and ionization proximity effects around star forming galaxies) can be compared with quasar absorption lines (using the VLT and ELTs) and next-generation X-ray data. Exploiting this expected wealth of data will required enhanced numerical simulations of structure formation in large cosmic volumes, and of early stellar evolution and accretion process onto protostellar cores with inclusion of detailed physical processes.

3.2 How did the structure of the cosmic web evolve?

3.2.1 Background

The formation of the large-scale structure of matter can be understood as a competition of gravitational contraction of local overdensities and the overall expansion of the Universe. The predictions of numerical simulations of the dark matter evolution are consistent with observations of the distribution of galaxies and galaxy clusters, and with statistical measures of cosmic shear (weak lensing) that probes the gravitating (mostly dark) matter directly. Because the luminous (baryonic) matter is subject to non-gravitational forces, it dissipates energy and can change to different chemical and physical states. Most importantly, it is able to collapse to more compact structures such as galaxies and stars. The observed distribution of baryonic matter on scales from galaxies to galaxy clusters thus provides a wealth of information on the physics and evolution of the luminous as well as the dark Universe.

Galaxy clusters and groups are identified through optical observations as over-densities in space and / or colour, or through the X-ray emission of the hot intergalactic gas that accreted into potential wells of the dark matter. Dark matter wells are identified through the shear distortion of light from background galaxies. About ten thousand galaxy clusters (defined as structures with mass larger than 10^{14} Solar masses), and groups (with smaller masses) have been found. Several forthcoming experiments, in particular the Planck satellite, should expand this number significantly by discovering unbiased samples of clusters through a characteristic spectral distortion of the CMB as it scatters off the hot electrons in the intracluster medium, the so-called Sunyaev-Zeldovich effect. Sunyaev-Zeldovich surveys are more sensitive to distant clusters than X-ray surveys, and are expected to produce a sufficient number of the rare distant clusters beyond redshift one that are particularly interesting to constrain cosmological parameters and to study the main epoch of galaxy and supermassive black hole formation. By 2010 such surveys are expected to produce about 30 000 clusters, with hundreds of clusters at redshifts larger than one.

A major motivation of optical, Sunyaev-Zeldovich and X-ray cluster surveys is to establish the evolution of the large-scale structure mass spectrum in order to constrain the basic cosmological parameters, in particular the dark energy equation of state (§ 2.2). It is crucial to properly relate the observables to the baryonic and total mass, which requires a good understanding of the energetics of the intracluster medium as a function of time. There is evidence that at early epochs the energy density of the intracluster medium is strongly affected by merging, star formation and energetic outflows from active galactic nuclei. Conversely, the injection of energy and heavy elements into the cosmic gas provided a regulating feedback that significantly affected the formation of stars and galaxies, and possibly also of massive black holes. The tidal interaction or merging of galaxies and the ram pressure of the intergalactic medium on the galactic interstellar medium adds to the complexity of baryonic structure formation.

3.2.2 Key questions and experiments

The effects of the environment on the formation of galaxies are poorly observed and not well understood. How do star formation, supermassive black holes, galaxy encounters, magnetic field generation and heavy element enrichment affect the distribution and properties of galaxies in different environments, i.e., in clusters, groups, filaments, and voids? What are the physical properties of the diffuse gas and what is its relationship to the embedded galaxy populations?

Radio galaxies, quasars and starburst galaxies trace over-densities to higher redshifts than current galaxy cluster surveys. Studying the surroundings of such objects with Sunyaev-Zeldovich, sub-millimetre, or Balmer-line imaging provides information on the environmental dependence of galaxy formation at the earliest epochs of structure formation. Studies of starburst galaxies and dusty quasars require ALMA for the sub-millimetre regime, and Herschel for the far-infrared.

The growth of structures in low-density environments can be traced by absorption-line studies of low H I column density regions as demonstrated by numerical simulations of structure formation. The topology of the dark matter density field (sheets and filaments) and kinematics of the gaseous cosmic web can thus be derived from three-dimensional mapping of Ly-α absorbers with an ELT, using high-redshift quasars and compact, luminous galaxies as background targets. High-resolution ultraviolet (such as FUSE or HST) and X-ray spectroscopic capabilities (like the failed XIS instrument on SUZAKU) are needed to study the kinematics of the intracluster medium through spectroscopy of extended emission from ionic lines.

Large-area weak-lensing surveys with the VLT Survey Telescope (KIDS) will be complemented by near-infrared surveys with VISTA. Similar efforts in the USA include the Dark Energy Survey (DES), Pan-STARRS and LSST. Extension to all-sky surveys with depths of $R(AB) > 24$ and $K > 19$ is important to detect the galaxy concentrations of the rich clusters out to redshifts of 1.5. Deep, high-resolution multi-colour imaging with space telescopes will be required as well.

Sunyaev-Zeldovich survey instruments include APEX, the Atacama Cosmology Telescope (ACT, under construction), the South Pole Telescope (SPT, under construction), the Planck mission (launch 2008) and the interferometric projects SZ-Array and AMI. Several large-aperture single-dish (sub)-millimetre telescopes under construction or in the early planning stages will be able to provide the spatial resolution necessary for the joint modeling with X-ray and weak lensing studies, e.g., LMT and CCAT. The IRAM 30 m and the GBT could be equipped with high-resolution Sunyaev-Zeldovich imaging capabilities, and ALMA will provide key follow-up for the most interesting objects. The planned all-sky X-ray imaging survey with eROSITA will discover several hundred

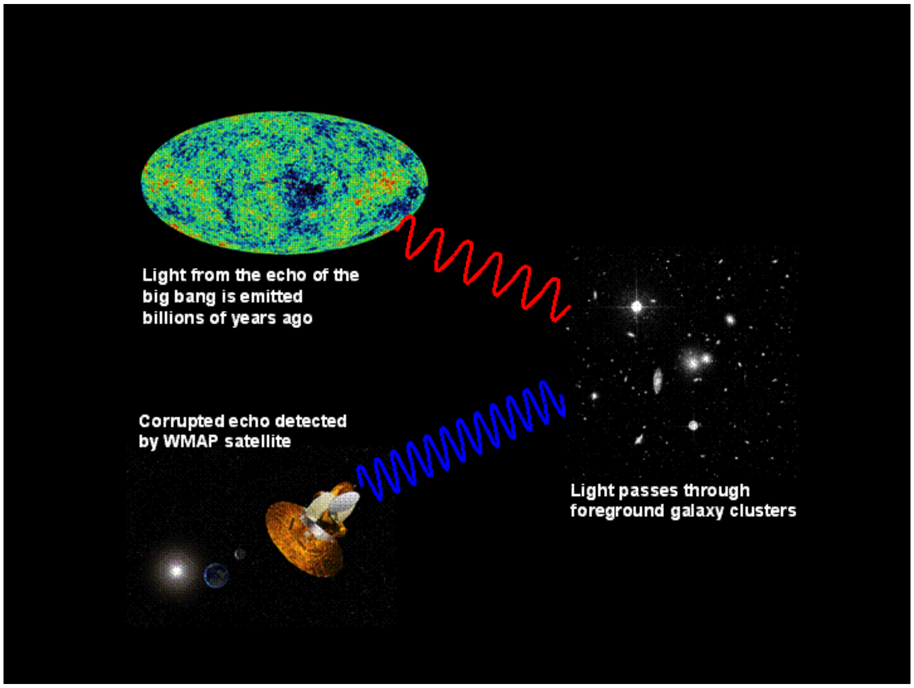

Figure 3.4: The Sunyaev-Zeldovich effect can be observed by WMAP or other instruments that look at the microwave background. The microwave background will be distorted by foreground hot electrons and this distortion can be measured. Interestingly enough, the Sunyaev-Zeldovich effect is distance independent and will therefore probe all intervening hot electron clouds.

thousand clusters and groups and thereby provide a database for statistical studies of large-scale structure. The continuous availability of the X-ray space telescopes XMM-Newton and Chandra is crucial to study the hot intracluster medium and active galactic nuclei in the large number of clusters that will be discovered.

These observational programs will need to be complemented by the next generation of numerical simulations of the structure and evolution of the cosmic web which include all the relevant physics. This will require substantial supercomputer resources.

3.3 Where are most of the metals throughout cosmic time?

3.3.1 Background

Heavy elements ('metals') are a prime tracer of star formation and feedback processes. They are produced in stars and dispersed by stellar winds and supernova explosions. To metal-enrich the intergalactic plasma on large scales requires powerful galactic outflows (superwinds) as shown by cosmological hydrodynamic simulations of galaxy formation. The pollution of the intergalactic medium (IGM) by heavy elements at early epochs is thought to be highly inhomogeneous. During the early phase of the reionization era of the Universe, i.e., before the formation of the first galaxies, the heavy elements should be distributed over small volumes in ionized bubbles around the first stars and stellar clusters. Theoretical models suggest that the first stars could have enriched the Universe to an average metallicity of about 0.01 per cent of the Solar value, comparable to that of the most metal-poor stars in the halo of our Galaxy (§ 3.5). During the peak of galaxy and black hole formation, at redshifts two to three, the mean metallicity of the IGM in photoionized regions had risen to about 0.1 per cent of the Solar value. Although there have been great advances in both numerical simulations and observations, several key questions are still open:

What is the inventory of heavy elements during the early reionization epoch? Are the voids and underdense regions pristine or already metal-enriched at redshift 2–3? Is the census of metals at the peak activity of galaxy and quasar formation, as well as in the local Universe, complete?

More than 90 per cent of the baryons are in the IGM. The present-day cosmic density of the baryons in various (diffuse and condensed) states is fairly well known, except for the warm-hot ($10^5 < T < 10^7$ K) phase of the intergalactic medium (WHIM), which could contain 40 per cent of the baryons, i.e. about half of the intergalactic gas. Heavy elements in the diffuse gas can be detected by their absorption signatures (mostly from resonance lines at ultra-violet wavelengths) in the spectra of luminous background sources. Spectroscopic inventories of baryons and metals in galaxies, the intracluster medium, and the general intergalactic medium as functions of redshift and environment provide important constraints on comological simulations and on the cycling of gas and metals between galaxies and the IGM in particular. Open issues are: Where are the metals at redshift 2–3? Where are the local baryons? What is the contribution of the WHIM to the metal budget in the local Universe?

3.3.2 Key questions

Reionization epoch. Cosmic reionization was most likely an extended process, starting with small isolated H`II` regions around the first sources of ultra-violet radiation, stars and / or intermediate mass black holes. During this early phase, luminous background sources should comprise gamma-ray bursts, supernovae and possibly extremely rare fairly massive black holes. With the formation of the first galaxies and black holes, pockets of ionized gas begin to percolate, a process completed by redshift six, and metallicity can then be traced over a large range of gas densities.

The onset of reionization occurs between redshifts 11 and 20, as deduced from the polarisation of the CMB detected by the WMAP satellite. Its end is revealed by the strong increase in ionization level of the IGM around redshift six. Some information on the metal enrichment of the IGM during this period is provided by observations of metal species at lower redshifts. Absorption-line studies suggest that the IGM was already metal-enriched at the end of the reionization era, and that its mean level of ionization increased strongly between redshifts six and 2.5. Before redshifts of seven, metals in the IGM should mostly be in the form of atoms and singly ionized elements.

The metal enrichment of the IGM can be measured by combining H`I` 21 cm absorptions in the spectra of luminous background radio sources with metal absorptions in the near-infrared spectra of these objects. This requires some knowledge of the spectral energy distribution of the ionizing radiation. A major challenge will be to find such very rare, bright sources at redshifts beyond seven. Large-area surveys with LOFAR may uncover a few such objects. High-resolution near-infrared spectroscopy alone of any bright background source will allow an estimate of the evolution of the O`I` and C`IV` mass densities, as well as the clustering of metal-rich sites and thus shed some light on the cosmic evolution of the luminosity and mass of the reionization sources.

Epoch of the peak activity of galaxy and quasar formation. At redshifts 2–3, less than 50 per cent of the metals produced by star formation activity in high-redshift galaxies has been detected in various sites: the galaxies themselves, the IGM and the damped Ly-α absorbers (proto-galaxies). Are the missing metals in galaxies or in gas expelled from star-forming galaxies? The answer is linked to the nature of the feedback processes and the inhomogeneous metal-enrichment of the IGM. In particular, the relative contribution to the cosmic metals of the general IGM versus metal-rich sites around star-forming galaxies and / or a WHIM (see above) is poorly known.

Galactic winds affect the temperature, density structure and metallicity of the IGM. Infall from the IGM can also trigger star formation. What should be done is to determine the metal content of the warm-hot gas and probe the galaxy-IGM interface. Current models that predict the amount of metals ejected by galaxies give conflicting results, and the extent of the regions affected by galactic (super-)winds is not yet well determined by numerical simulations.

Three-dimensional mapping of Ly-α and metal-rich absorbers probes the absorber-galaxy connection and constrains the sizes of the hot regions affected by galactic super-winds. Roughly one-third of the intervening CIV absorbers with substantial column densities are produced within $\sim$100 kpc of star-forming galaxies and may be up to $\sim$500 kpc for strong OVI absorbers. Probing the absorber-galaxy connection for small column densities of heavy ions will provide an estimate of the overall size of the regions affected by galactic winds, and will require ELT spectra of faint background targets at high resolution and signal-to-noise ratio. In underdense regions of the IGM, the expected column densities of heavy species are very small. Probing such low metallicity levels in regions with small HI column densities at redshifts 2–3 to constrain galactic super-wind models requires a gain of at least a factor 100 in the detection limit of individual metal absorption lines compared to the best results obtained with 8-10 m class optical telescopes.

The local Universe. The census of the baryons in the nearby Universe implies that about 40% of the cosmic baryons are still in the form of warm plasma in the IGM. The missing baryonic matter could reside in a WHIM ($10^5 < T < 10^7$ K) within diffuse large-scale structures in overdense regions, as suggested by cosmological simulations. Additional hot diffuse gas is in galaxy groups or galactic halos. When gas falls from the filamentary structures into the galaxy clusters at the intersections of the cosmic filaments, it can heat up to 10^8 K. The heavy element content of the WHIM is not well-determined observationally. The contribution of shock heated regions of the IGM to the cosmic metal density could be about 20 per cent. The expected metallicities of the photoionized IGM and the WHIM are similar.

Observations and simulations suggest that ultra-violet and X-ray absorbers trace multi-phase structures. If these absorbers are associated with large-scale shocks, there should be a spatial correlation with galaxy groups or overdensities as is the case for many OVI absorbers. Inferring metallicity limits for X-ray absorbers is even less straightforward than for ultra-violet absorbers, as mechanisms other than thermal broadening can be present, such as large-scale kinematic flows.

3.3.3 Future key facilities

The evolution of cosmic reionization can be probed by studying the HI 21 cm forest in the spectra of powerful background radio sources with a large-aperture radio telescope (SKA). Metallicity will be obtained by combining these observations with ELT near-infrared spectroscopy of quasars.

The very first, individual sources detectable at the onset of reionization should be transient explosive events (gamma-ray bursts and supernovae). Observing transient sources with ELTs, should be operationally feasible as the time-lag between discovery and ELT spectroscopy will be days (gamma-ray bursts) to weeks (supernovae). Unveiling this population will require all-sky monitoring near-infrared telescopes and next generation gamma-ray burst space missions. The first 'luminous' galaxies (and quasars) will be detected by JWST, but observing those of smaller mass will require an ELT. Very luminous radio sources will be detected by LOFAR (and later on SKA), and intermediate mass black holes will be found by large-aperture X-ray telescopes.

A next-generation X-ray observatory operating in the 0.15–1 keV range, with an effective area of several square meters at 1 keV, good angular resolution of 5 arcsec and equipped with a high spectral resolution instrument (1 eV or less), is essential for studying how the fraction of baryons and the metal content in the warm and hot phases evolve across cosmic time. High-resolution spectroscopy below 1 keV is needed to trace absorption lines of the most common ionic species expected in the warm / hot intergalactic medium, out to redshift one.

Solving the missing metals problem at redshifts 2–3 will require large samples of warm absorbers. This could be achieved with ELT spectroscopy of quasars at very high resolution and signal-to-noise ratio. Searches at lower redshift can be more easily done with a next generation ultra-violet satellites. Identification of the galaxies associated with the warm-hot absorbers is also necessary to constrain the strength of the galactic winds and is coupled to a general study of the IGM on large scales. The background targets could be Lyman break galaxies at redshift three together with rarer brighter quasars. Spectroscopy with ELTs should be made at intermediate resolution

Figure 3.5: The spider galaxy is a massive, young and growing galaxy at redshift z=2.2 (10.6 billion light-years away from Earth), where the attracted smaller galaxies are clearly visible as the web like structure. Image courtesy of G. Miley and HST.

with multi-integral field units. Detecting the hotter phase ($T > 10^6$ K) will remain a challenge for future large-aperture X-ray telescopes as it implies a high sensitivity in the 0.15–0.3 keV range.

Advances in the problem of missing baryons in the local Universe will continue after the next servicing mission of the HST (2008, installation of the Cosmic Origins Spectrograph) and with next-generation UV satellites and future large-aperture X-ray telescopes. The latter are mandatory for a full understanding of the local IGM, for probing the WHIM, the clustering and metallicity of the IGM as well as the Ly-α forest-galaxy correlation. Temperature estimates will require ultra-violet (and X-ray) spectra of high resolution, to minimize the effects of line blending.

3.4 How were galaxies assembled?

3.4.1 Introduction

One of the most spectacular successes of the past 15 years has been the opening of new observational windows to the study of the formation and evolution of galaxies. These range from deep observations of galaxies over most of the history of the Universe to unprecedented observations of the fossil records of stars in our Galaxy and its neighboring galaxies (§3.5). On the theoretical side the same cosmological picture that has revolutionized our understanding of the large scale structure and evolution of the Universe also provides a framework for understanding the physics of galaxy formation and evolution. In this picture small-scale density perturbations in the dark matter collapse and form the first generation of dark matter halos, which subsequently merge to form larger and larger structures such as galaxy groups, clusters and superclusters.

The component of the theory that traces the development of structure in the dark matter is well developed, and can now be modelled in exquisite detail using very large numerical simulations carried out on a new generation of parallel supercomputers. The largest such simulations track the assembly of dark matter halos with masses a few per cent that of our Galaxy in a volume comparable to that of the Sloan Digital Sky Survey, the largest galaxy redshift survey carried out to date. In spite of these advances, however, the theory is riddled with unsolved problems. The growth of the gaseous and stellar components of galaxies is not related in a simple way to the build up the of the dark matter, and the processes of gas accretion and cooling and star formation that drive these differences are not well understood, theoretically or observationally.

It is becoming increasingly clear that the growth of galaxies is regulated by a complex interplay of energy and matter exchange between different baryonic phases. When the dark matter collapses under gravity and virializes into gravitationally bound dark matter halos, gas is able to reach high enough overdensities to cool, lose pressure support and condense at the centre of the halo. The rate at which this occurs will depend critically on the temperature and density structure of the gas. Eventually the gas becomes cold and dense enough to form stars. The more massive stars end their short lives in supernova explosions. In strongly star-forming galaxies, the energy in these explosions is sufficient to heat the gas and drive many of the heavy elements synthesized by the supernovae out of the galaxy. These heavy elements will have a strong effect on the rate at which the gas can cool. It is also possible that some of the gas in the galaxy is also driven out in these explosions. Magnetic fields will play significant roles in all of these processes. Complicating these *in situ* processes within individual galaxies are the frequent interactions and mergers between protogalaxies and galaxies, which can result in large-scale angular momentum transfer of the interstellar gas and triggering of intense bursts of nuclear star formation or nuclear activity (or both). The role of the black hole accretion-triggered nuclear activity is especially important. The observed near-ubiquity of black holes in the centres of massive galaxies and the tight correlations between the mass of theses black holes and the masses of their parent galaxies offers compelling evidence for a strong connection between the formation of the black hole and its host galaxy. Additional support for this coupling comes from X-ray observations, which show that the evolution of active nuclei and star formation in massive galaxies appear to be very similar. Both developed during the same epochs in cosmic time, and both exhibit so-called 'cosmic downsizing', meaning that the strongest activity shifts from more massive to less massive galaxies as the Universe evolves.

The past decade has seen enormous progress in laying down the basic tenets of this picture and in providing the first reliable measurements of the star formation and nuclear accretion histories of the Universe. The more challenging task ahead is to build a physical picture that quantifies and incorporates a deeper understanding of these 'gastrophysical' processes of gas accretion on to galaxies and their central black holes, the subsequent star formation, and the feedback processes resulting from star formation and nuclear activities. The processes described above are highly complex and difficult to model from first principles. However they offer the path towards understanding the observed numbers, masses, structures, morphologies, and chemical compositions of galaxies, as well as the evolution of all of these properties with cosmic time.

Meeting this challenge will require a wide range of observational approaches, incorporating information from the local and the distant Universe. The former will include in-depth studies of the compositions, orbits, and ages of large samples of stars in our Galaxy (§3.5) and other types of galaxies in the local supercluster, along with direct observations of the physical processes that regulate the growth of galaxies and the exchange of matter and energy between the different baryonic phases of the Universe. Observations of distant galaxies will enter a new level of maturity as new instruments make it possible to trace the buildup of mass, cold atomic and molecular gas, dust, stars, metals, magnetic fields, and cosmic rays with time. High-quality observations of large samples need to be pressed to redshifts of two and beyond, where the cosmic star formation rate is at its peak. A multi-wavelength approach will be essential for both low-redshift and high-redshift studies, because the active phases in the growth of massive galaxies are enshrouded by interstellar dust, and observations over a wide wavelength span are essential for separating activities powered by black holes from those resulting from star formation alone. And we need to understand how this evolution changes as functions of a galaxy's type, mass, and environment.

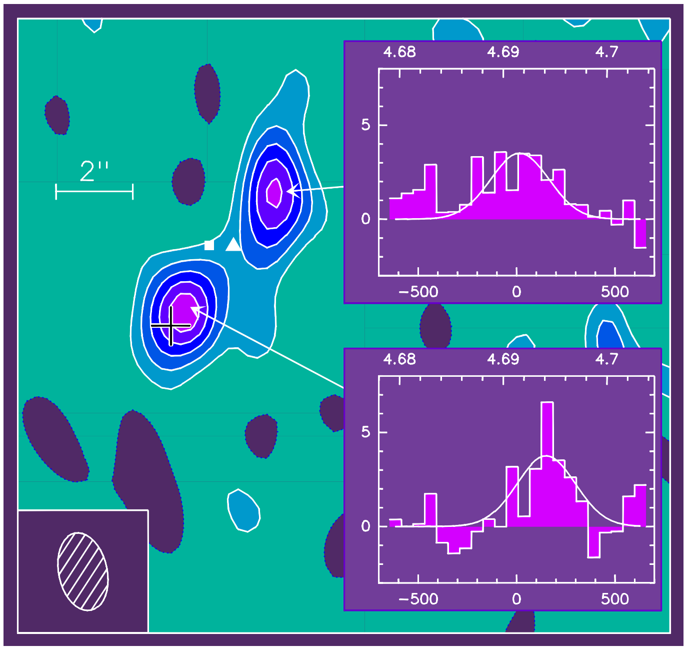

Figure 3.6: Dust and CO J=5-4 emission from the high redshift galaxy BR1202: a young merger at z = 4.7. Image taken with the IRAM interferometer.

3.4.2 Key questions and experiments

Gas cooling, accretion, and star formation. One of the chief handicaps in our current observational picture is the minimal amount of information available on the cold gas contents of galaxies beyond redshifts of order 0.1. Over the coming decade this subject will be revolutionized by ALMA, which will make it possible to produce spatially-resolved maps of the molecular gas in galaxies over a wide range of redshifts, and at the same time provide high-resolution measurements of the internal kinematics of these galaxies, critical for studying their mass distributions and the dynamical states of the gas discs. Continuum mapping at sub-millimetre wavelengths will provide detailed information on dust-obscured star formation and nuclear activity. ALMA offers the potential to probe these processes on scales down to tens of parsecs in nearby galaxies, over a much wider range of interstellar environments and host galaxy types than can be studied now. At the highest redshifts the low-level CO lines will shift out of the ALMA bands, but the lines would be accessible to a large centimeter-wavelength array such as the proposed SKA.

About half the cold gas in nearby galaxies is in atomic form. SKA will enable measurement of H I masses for Milky-Way mass galaxies out to redshifts of 2.5. At redshifts below one, SKA will establish the rate at which gas accretes onto galaxies, and clarify whether the accretion is in the form of lumps of gas that have already condensed in dark matter halos at some earlier epoch (so-called ’cold accretion’) or whether gas is able to cool from the hot phase in the form of a ’cooling flow’.

The combination of ALMA and SKA offers the promise of directly measuring the buildup of gas discs in galaxies. Coupling these observations to maps of the star formation rate will directly probe one of the main physical drivers of galaxy evolution. Observations of local galaxies reveal the existence of tight scaling laws relating the star formation rate to the surface density of cold gas (Kennicutt–Schmidt law). Numerical simulations and theoretical models routinely apply this local relation to young galaxies where the gas masses and star formation rates are often tens to hundreds of times higher than are typically found in the present-day Universe. These relations need to be tested and measured directly, both via direct observations of high-redshift objects and by higher-resolution observations of nearby galaxies. Observations of the kinematics of the neutral, molecular, and ionized gas will also probe the feedback processes in these environments.

Radio polarisation observations can image the components of the gas flow in the sky plane, complementing spectroscopic data, and can also trace galaxy interactions.

Some information on star formation distributions in distant galaxies are already available, either from observations of the redshifted ultraviolet continuum from HST or maps of Hα emission from deep imaging and integral-field spectroscopy on 8–10 m telescopes. An overarching limitation of these observations is the uncertain effects of dust extinction. Thanks to dedicated surveys with the ISO and Spitzer telescopes a comprehensive understanding of dust extinction in local galaxies is emerging, resulting in truly robust multi-wavelength diagnostics of star formation rates. The capabilities of this generation of infrared telescopes fades rapidly at higher redshift, as the peak dust emission redshifts out of the sensitive bandpasses of these instruments and sources are lost in confusion noise. The Herschel space observatory will make significant inroads in this area, and ALMA will be able to map the long-wavelength tail of the dust emission. But detecting dust-obscured star formation in the main population of distant galaxies will require a cold large-aperture far-infrared imaging observatory, as discussed in ESA's Cosmic Vision.

Radio continuum observations also probe star formation rates (and magnetic fields). The thermal radio continuum scales directly with the ionizing luminosity and the synchrotron radio luminosity exhibits a tight empirical scaling with far-infrared luminosity. Observations with LOFAR and SKA will provide a wealth of information on radio continuum emission of high-redshift galaxies.

The internal physics and dynamical evolution of high-redshift galaxies. We currently have no good idea about what is happening inside galaxies in the early Universe. This should be contrasted with our knowledge of galaxies in the nearby Universe, where we can observe the velocities of stars and gas in great detail. For the nearby galaxies we understand well how they maintain their equilibrium, and how they evolve with time. For galaxies in the distant Universe, we do not even know whether they are really in an equilibrium state: they could be undergoing repeated collisions with other galaxies or clumps, and hence be strongly evolving all the time.

Spatially resolved imaging and spectroscopy is needed to understand the internal physics of these distant galaxies. It will allow us to measure rotation and inflow in high redshift discs and obtain constraints on the fraction of the total stellar mass in a galaxy that is in a kinematically hot bulge component. It will also allow determination of the distribution of metals, dust and stars across the galaxies. In the optical and near-infrared, these goals can be achieved with integral-field spectroscopy. The galaxies are small, and very high spatial resolution is needed. This requires advanced adaptive-optics systems on a 30–50 m ELT. In order to study a sufficient number of galaxies, one needs to be able to observe tens of galaxies simultaneously.

The observed kinematics will also provide the masses of the galaxies. These are essential to 'calibrate' the techniques used for estimating stellar masses from models of the spectral energy distributions. The masses derived from kinematics will allow a determination of the applicability of these models. In the most actively star-forming galaxies we also expect to observe the signatures of galactic winds and star formation feedback in the kinematics of the emission lines. These measurements are critical to determine how the galaxies were assembled over time.

Linking star formation and metal enrichment with galaxy assembly. Low-mass stars have life times comparable to the age of the Universe, and retain in their atmospheres the elemental abundances from the gas at the time of their birth. By counting stars of different ages in a colour-magnitude diagram the rate at which stars are formed throughout time is obtained. Abundances can be measured from spectra of individual stars of known ages and thus the evolution of abundance of different chemical elements can be established. To date this type of analysis has been restricted to our Galaxy and its satellites and to some extent to the Andromeda galaxy M31. Enlarging this sample requires moving considerably further away in distance: the Sculptor Group and the M81 group (at about 2–3 Mpc) contain several more large spiral galaxies, while the nearest large elliptical galaxy is Centaurus A at 3.5 Mpc. The Virgo Cluster, located at a distance of 17 Mpc, is the real prize as it has over 2000 member galaxies of all morphological types. With an ELT of 40 m one will be able to image the brightest red giants at the distance of the Virgo cluster, a measurement that currently can only be carried out with long integrations from space with HST. Such a telescope should also be able to resolve the ancient main sequence turn-off stars in Centaurus A.

Spectroscopy provides the kinematic properties and chemical abundances of individual stars. Intermediate resolution ($R \sim 5000$) around the CaII triplet (at 860 nm) allows for a basic metallicity measurement, and a velocity accuracy of a few km / s, and is possible for red giant stars at a distance of the Sculptor group with a 40 m ELT. High resolution ($R \sim 40000$) will provide accurate abundances of numerous elements for red giant stars in M31, and perhaps as far as Centaurus A.

The motions of the stars are dictated by their host galaxy's gravitational potential, and the stellar orbits preserve an archaeological record of how the galaxy was assembled (§ 3.5). With an ELT multi-object spectrograph, the motions of millions of giant stars in galaxies sampling the full Hubble sequence lie within reach. These will shed light not only on the merger history of a galaxy, but also allow a measurement of its gravitational potential, the total mass, as well as the density profile and shape of its dark matter halo.

Mass, environment and redshift dependence of key physical processes. A major problem in trying to constrain models of galaxy formation is that in the present-day Universe many of the principal observable properties of galaxies, including their masses, ages, morphologies, and environments are all strongly correlated with each other. This makes it difficult to infer which of these parameters is fundamental in physically driving evolutionary trends. Observations of large and diverse samples are needed to isolate the effect of each variable individually.

The Sloan Digital Sky Survey made major inroads into this problem for the local Universe, but it lacks the depth and the infrared coverage needed to extend this work to redshifts 2–3 and beyond. Deep infrared surveys, e.g., with VISTA, will provide the identifications, magnitudes, and colours for the large samples that are needed. Infrared spectra of these faint objects with resolution and quality comparable to SDSS is required for accurate distance measurements together with estimates of stellar mass, star formation rate, mean stellar age, dust content, and metallicity. A large, multi-object infrared spectroscopic survey of 10 000 to 100 000 galaxies at redshifts 2–3 will provide sufficient statistics. Obtaining low-resolution spectroscopy for a sample of this size is feasible with JWST, but would require a few years on an 8 m class ground-based telescope. Of particular interest is a subpopulation of massive galaxies at redshifts greater than 2 with high stellar densities and no detectable emission lines. Such galaxies challenge current models of galaxy formation. High-resolution spectroscopy is required in order to place constraints on their star formation histories and kinematics.

For obscured starburst galaxies, which may trace the early stages of strong star-bursts, mid-infrared observations with the Herschel space observatory and radio continuum data will provide information on the total star formation rate and dust content, whereas sub-millimetre studies of molecules with ALMA will be necessary to measure their redshifts. Full characterization of this dusty population of young galaxies will require a large-aperture infrared space observatory.

3.5 How did our Galaxy form?

3.5.1 Background

Even though the hierarchical clustering cosmological model works well on large scales, it faces a large number of problems on small scales. Our Galaxy and the Local Group of galaxies are the ideal test-ground and a unique resource for constraining galaxy formation models, as many examples in the past century show. Evidence for extended dark halos surrounding galaxies was first made using Local Group observations. Surveys of stars in the halo have revealed evidence for hierarchical accretion, and have even uncovered new galaxies such as Sagittarius which is presently disrupting in the tidal field of the Galaxy's dark halo.

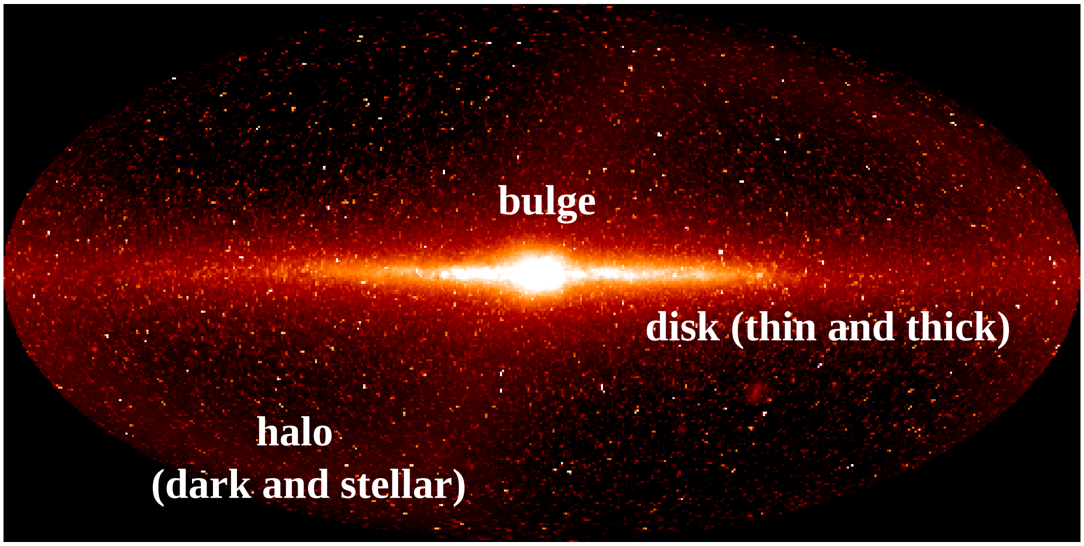

Figure 3.7: Our Galaxy and its different components.

Our Galaxy contains a complex mix of stars, interstellar gas and dust, as well as dark matter, distributed in a bulge, a halo (stellar and dark matter) and a thin and a thick disc. These constituents are widely distributed in age (reflecting their birth rate), in space (reflecting their birthplaces and subsequent motions), on orbits (determined by the gravitational forces due to their mass distribution), and also with different chemical element abundances (determined by the past history of star formation and gas accretion). Therefore the present-day structure and dynamics of our Galaxy are intimately linked to its assembly and evolution over the age of the Universe.

3.5.2 Key questions

Assembly and chemical history. The fossils of the formation process of the Galaxy can be retrieved in two complementary ways: through their dynamics and through their chemical element abundances. Stars with a common origin are expected to move together on similar orbits, and hence give rise to distinct substructures such as tidal streams. The discovery as well as the modelling of these substructures will provide answers to the key questions: How many mergers has our Galaxy experienced? When did these take place? What were the properties of the Galactic building blocks? How do they compare to the present-day nearby galaxies?

Low-mass stars live much longer than the present age of the Universe, and retain in their atmospheres a record of the chemical elements of the environment in which they were born. This means that the chemical abundance patterns present in their atmospheres provide strong constraints on the hisotry of star formation, of the initial mass function, and of the assembly of our Galaxy, in particular at very early times. Large samples of stars belonging to all Galactic components are necessary to find the rare pristine (very metal-poor) stars. These will enable us to answer questions such as: Are there any zero-metallicity stars? What was the typical mass of the first stars formed? When did the transition to the present-day initial mass function take place?

The Galactic bulge is the central component of our galaxy; it consists mostly of old, metal-rich stars, is spheroidal in shape, and is probably quite similar to an elliptical galaxy. Being only 8 kpc away, it can be studied in exquisite detail. The very strong obscuration of the bulge by dust has inhibited significant progress until recently, and most of what is currently known comes from studies of a few low-reddening windows with HST, with multi-object spectrographs such as FLAMES on the VLT, and from relatively shallow near-infrared photometric surveys such as 2MASS, all of which have begun to reveal the complexity and puzzling nature of this system. he

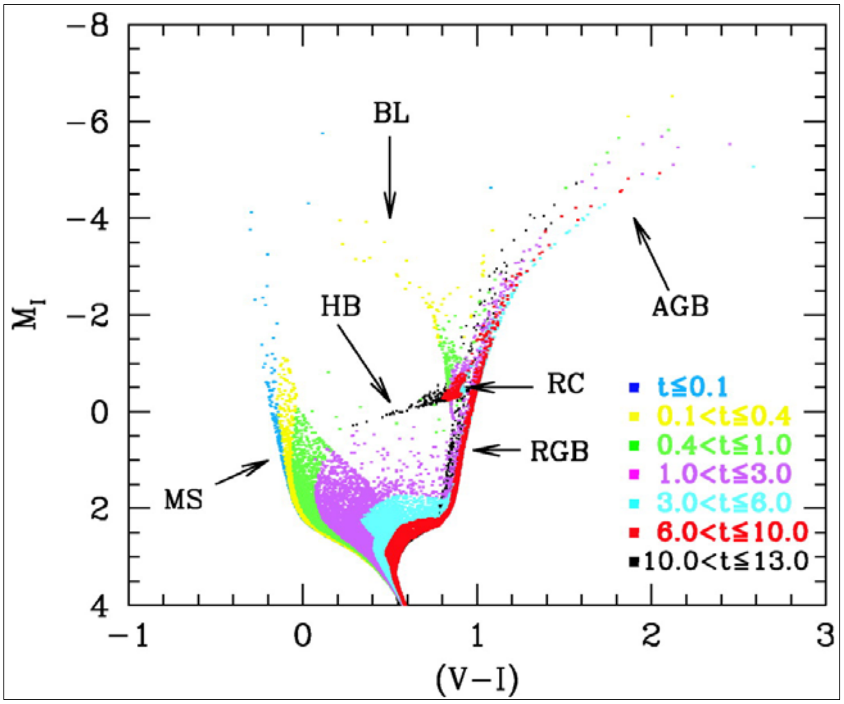

Figure 3.8: Synthetic Colour-Magnitude Diagram computed using constant star formation rate from 13 Gyr ago to the present and with metallicity linearly increasing from Z = 0.0001 to Z = 0.02. Stars in different age intervals are plotted in different colours, and the colour code is given in the figure, in Gyr. Labels indicate the different evolutionary phases (from Aparicio & Gallart, 2004 AJ 128 1465).

formation timescale of the bulge, its chemical enrichment history, and its relation to the galactic bar and disc are largely unknown.

The nature and distribution of dark matter. Dark matter is the dominant mass component of galaxies, yet its nature and distribution on galaxy scales are largely unknown. Studies of gas dynamics in external galaxies yields insights into how the dark matter is distributed, but they are generally confined to the inner regions of galaxies, where luminous matter makes its largest contribution. Distant (halo) tracers with good kinematics are strongly preferred as they allow a measurement of the total mass and density profile of our Galaxy. Tidal streams are very powerful probes both of the mass and shape of the dark matter halo, as well as of their time evolution, since they are formed by stars on parallel orbits. Different insights are provided by the satellites of our galaxy, which being the smallest yet the most dark-matter dominated systems in the Universe, can be used to put strong limits on the properties of dark matter particles (e.g., some candidates should annihilate and produce gamma rays; cores are expected for warm dark matter while cold dark matter gives rise to cuspy profiles).

3.5.3 Future experiments

To disentangle the surviving fossils of our Galaxy assembly process requires a vast catalogue of stellar properties, including ages and chemical abundances, spatial distribution (distance and location in the sky), and kinematics (proper motion and radial velocity, that is motion on the plane of the sky and along the line-of-sight, respectively). Large, multi-dimensional and unbiased samples are crucial to ensure significant progress. The stellar halo is a particular good example: it contains the most metal-poor stars known to date and it is believed to be the natural repository of debris from disrupted galaxies. However less than 1 per cent of the stars belong to this component. Any kind of criterion that would favour the selection of halo stars will introduce biases which are impossible to correct a posteriori. Furthermore, as shown in Figure XX, although positional information may be sufficient to discover tidal streams in the outer halo, it is certainly not enough to disentangle the fossils in the inner part of our Galaxy (where most of the stars are). Here the dynamical timescales are so short that only with full phase-space information (and possibly chemical abundances) will it be possible to disentangle the assembly history of the Galaxy.

Deep surveys (such as Pan-STARRS and LSST) will provide outstanding photometry for billions of stars, a unique resource for Galactic structure studies (as has been shown by e.g. SDSS despite its smaller sky coverage, poorer photometry and brighter magnitude limit).

A major milestone will be reached with ESA's Gaia mission, which will measure very accurately the positions, distances and motions of a billion stars all with spectrophotometry, and thereby revolutionize our knowledge of our Galaxy and its nearest neighbours. Its long list of science goals includes fundamental calibrations of the stellar luminosity functions and initial mass functions (§ 4.2), precise calibrations of distance indicators including Cepheid and RR Lyrae variable stars, accurate distances to stars of virtually every type and evolutionary phase, and the determinations of orbits and ages for hundreds of millions of stars, in effect a stellar census of our Galaxy. It is crucial to supplement the dataset with dedicated ground-based spectroscopic programs, in order to obtain radial velocity and detailed chemical abundances for fainter stars. Multi-object spectrographs on 8 m telescopes will play a critical role here. However the efficiency of this work will be greatly aided by the construction of wide-field multi-object spectrographs with both low-resolution and high-resolution capabilities, along the lines of the proposed WFMOS.

A near-infrared astrometric mission (such as the proposed Japanese mission JASMINE) to measure large samples of stars with space velocities and distances is crucial for studying the Galactic bulge. Deep near-infrared surveys and wide-field spectrographs with large multiplex capabilities are a necessary complement.

These surveys will ensure that a large number of tracers with precise kinematics become available to allow a measurement of the mass distribution of our Galaxy. In the case of the Galactic satellites, their mass distribution can be constrained with radial velocities of a thousand constituent stars, however, only with proper motion information can the modelling degeneracies be broken to allow a direct measurement of their mass and dark matter density profile.

The development of software for the analysis of large datasets, both for ground-based surveys as well as for the Gaia mission, and for the next generation of dynamical models is a prerequisite.

Chapter 4

What is the origin and evolution of stars and planets?

The life cycle of stars is a fundamental topic in astrophysics which is expected to be one of the most active in the coming decades. The cycle of matter from the interstellar medium into stars and then back is the basic engine that dictates the evolution of the baryons across the age of the Universe. Planetary systems like our own will form during the early phases of stellar evolution and the complex chemistry of the dense interstellar medium around newly formed stars may be a necessary process to produce the complex molecules that are the building blocks for life. Understanding the life cycle of stars is thus a fundamental step to find answers on the origins of our own Solar System and life on Earth as well as for other habitable exo-planets and life elsewhere in the Universe.

The discovery of exoplanets just over a decade ago has opened a new and fascinating front in astrophysics, also for the philosophical implications of the existence of planetary systems outside our Solar System. All extrasolar planetary systems discovered so far are very different from our own, due to the limitations of the techniques used so far for planet detection. Nevertheless, the discovery of such unexpected configurations has provoked a profound revision of our views of the formation and evolution of planetary systems. The search for extrasolar planetary systems similar to our own Solar System, of exo-Earths and of the signatures of extraterrestrial life is a long term goal for the astronomy of the 21st century. ESA has clearly acknowledged this fact by making it one of its recommended themes of research in Cosmic Vision 2015–2025. The detection of Earth-like planets and the first search for bio-markers as well as the direct study of planetary systems and analysis of the composition of planetary atmosphere also features prominently in the primary science case for Extremely Large Telescopes.

The key questions that will have to be addressed in the coming decades span a broad range of topics: the role of magnetic field, environment, multiple systems and clusters in star formation; the detailed shape and possible variations of the Initial Mass Function for stars; the mysteries of the internal structure of stars; the chemical processes in the interstellar medium; the final stages of stellar evolution and the feedback to the interstellar medium; the evolution of circumstellar discs leading to the formation of planetary systems; the diversity of planetary systems and the search for Solar System analogues.

Diverse techniques and facilities are required to attack such a broad range of topics. What we have learned in the recent past is the result of the tremendous effort invested by Europe in developing large astronomical facilities (including the VLT, IRAM, HST, XMM and ISO), new techniques (such as high precision spectroscopic radial velocity surveys), and a parallel development of an important theoretical understanding of the various processes. If we want to attack the key unknowns in our understanding of the origin of planetary systems and on the presence of Solar System analogues in the Galaxy, we now need to use a new generation of large facilities.

4.1 How do stars form?

4.1.1 Background

The sequence of events that leads to newly formed stars and stellar systems starting from the diffuse interstellar matter all go under the general line of 'Star Formation'. A detailed theory of star formation is the building block to understand both galaxy formation and planet formation throughout the Universe. We have only a rudimentary grasp of such a theory, so that the formation of galaxies and planetary systems is described with semi-phenomenological theories.

From clouds to stars. Despite being central to two fundamental issues in astrophysics – the evolution of the Universe and the origin of solar systems – star formation remains, literally, an obscure field: the main problems associated with a better understanding of the processes involved are linked to the deeply embedded nature of forming stars. Whereas the main sites of star formation (molecular clouds) and basic processes (gravitational instability, hierarchical fragmentation, angular momentum dissipation) have long been known, there are many difficulties in the details and we still lack a global theory with predictive power. One of the limitations in building such a theory is that solid observational constraints have been lacking for a long time.

The processes which trigger the initiation of star formation in molecular clouds are poorly known (e.g., when does collapse start, why are some clouds active and others not?). Once star formation has started, it appears to be fairly inefficient, in the sense that only about 30 per cent of the collapsing cloud core mass is eventually transformed into stars. This inefficiency is clearly linked to the fact that during its collapse over seven orders of magnitude in size a forming star has to shed considerable amounts of angular momentum. The process by which this occurs, cloud turbulence or magnetic braking, is still heavily debated, however. In the last decade, significant progress was made toward realistic (magneto-)hydrodynamic simulations of the dynamical evolution of the interstellar medium driven by turbulence created by large-scale processes such as supernovae, magnetorotational instabilities and gravitational instabilities.

The role of clusters. The current consensus is that most stars form in very dense environments ranging from about 40 stars / pc^3 to 10^7 stars / pc^3 in extreme cases, where the Solar neighbourhood has a density of 1 star / pc^3. Starbursts (10^7 stars / pc^3) are extremely bright and allow observational probing of star-formation over cosmological epochs. The physics of the formation and early evolution of such objects is largely unknown, but is crucially important for the stellar distribution in galaxies and for their morphological appearance, as well as for the formation of intermediate black holes and their transport to the centres of galaxies where they may merge (§ 3.1, 3.4).

A key issue is the physics of the formation of extremely dense stellar populations: how are the densities and pressures in the inter-stellar medium generated to induce the formation, and which gas- and radiation-transport mechanisms play a dominant role in forming the extremely dense star-gas mixture? Which physical processes are responsible for driving out the residual gas leaving a compact and bound stellar system? Which fraction of the formed stars follows the gas and leaves the dense system? Simplified calculations suggest that this fraction may be larger than 50 per cent, and that this fraction leaves with a velocity dispersion of tens of km / s. Thus, ultra-massive, ultra-dense star-bursts may throw out 10^7 stars with a few tens of km / s which could be one mechanism of forming thick discs of galaxies. This would be an independent mechanism to the usually postulated dwarf-galaxy mergers in a cosmological scenario, so the implications of clustered star formation become apparent – cosmological models of galaxy formation and evolution may be seriously flawed if these physical processes are not taken into account.

Observations show that giant molecular cloud complexes give rise to dense stellar clusters. Number counts of embedded clusters indicate a formation rate one order of magnitude larger than the birthrate for classical open clusters, suggesting a rapid cluster disruption timescale (few tens of Myr). It is clear, then, that many isolated stars were born in clusters. However, there is also evidence from isolated young stellar objects that cluster formation is not the only path to form stars.

The relative roles in star formation of large clouds versus small dense clouds is currently poorly determined. Dynamical considerations suggest that bound clusters may emerge only from molecular clouds with star formation efficiencies larger than 50 per cent. Such environments are rare in our Galaxy, but the possibility to study cluster formation in starburst galaxies may significantly improve our understanding of the formation of open and globular clusters.

Star clusters are also an important actor in a key unsolved issue of broad cosmological relevance: that of massive star formation. The high radiation pressure by hot stars strongly counteracts accretion during the short pre-main-sequence phase of such objects, hence preventing the growth of objects to more than some 20 Solar masses. The fact that all known young high mass stars are found within dense stellar clusters, suggests an intriguing alternative scenario, in which massive stars originate from coalescence of smaller objects. Some observations indicate, however, the possible presence of accretion discs also for very massive stars, suggesting that our current understanding of the radiation-hydrodynamic accretion process is flawed. The actual observation of massive-star formation in ultradense HII-regions has only become possible recently, thanks to the remarkable progress in infrared and millimetre astronomy at high angular resolution.

Formation of multiple stellar systems. Observations of individual stars within about 20 pc from the Sun have revealed that approximately 60 per cent of the stars are double. This result is valid for typical, fairly old stars of Solar mass or less. In dynamically-young nearby star-forming regions – within which the mutual gravitational forces are too small to significantly affect the stars – virtually all stars below a Solar mass and an age younger than about one million years have stellar companions. In star-forming regions in which the stellar density is so high that the stars have moved significantly, and encountered other stars within a few million years, the fraction of binary stars is again about 60 per cent. These observations suggest that virtually all stars, including perhaps even our Sun, are probably born in binary or triple stellar systems. Such a high initial binary proportion would also alleviate the angular momentum problem of star formation.

Because other observations have shown that most stars also form in star clusters, it is now thought that a large fraction of the initially multiple stellar systems break apart as a result of gravitational interactions between the systems in their birth cluster. This would imply that the varying binary fractions in different stellar populations are a result of somewhat different cluster properties, rather than differences in the initial binary properties.

The dynamical processes that reduce the initially high fraction of binaries to the levels seen in dense star clusters and in the Galaxy attract much interest. Observationally this is very challenging, because binary systems need to be found with orbital periods spanning from a few hours to millions of years. A wide variety of techniques is needed, such as searching for occultations of a star by its companion star, simple visual inspection to find companions, and detection of the slight wobbles of a star caused by the orbital motion of a planet by extremely accurate measurements of its radial velocity (spectroscopic observations) or of its proper motion (astrometric observations). Theoretically, encounters of binary systems in dense star clusters pose a significant computational burden, because binary systems with orbital periods of a few hours need to be followed over many millions, if not billions of years. Because tight binary stars can have a binding energy that is comparable to that of the whole star cluster, the evolution of entire star clusters may be significantly affected by highly energetic binary–star encounters that can lead to the ejection of individual stars away from the clusters with high velocities, up to a hundred km / s.

4.1.2 Key questions

The key open questions to be addressed in the formation and early evolution of stars and stellar systems concern the effects of diverse initial conditions and environment on star formation. The roles of turbulence and magnetic field in molecular clouds about to form stars are still to be clarified theoretically and constrained observationally. Similarly, the effect of metallicity on star formation has to be explored if we want to understand the star formation history across the age

Figure 4.1: Three star forming regions, from left to right: NGC3603, The eagle nebula and the extremely low metallicity Blue Compact Dwarf galaxy IZw18. Images courtesy of ESO and HST.

of the Universe. It is particularly important to understand if and how the formation of the most and least massive stellar objects depend on the environmental conditions (§ 4.2).

Conversely the feedback effect of newly formed stars on the parent environment and the forming stars is also an important issue to tackle. This is especially important in starbursts and dense clusters, where the extremely intense radiation field of many thousands of O stars will likely evaporate pre-stellar cores and remove material from forming stars – what effect does this have on the stellar initial mass function? Are intense star-bursts depleted in low-mass stars as a result of such processes? The subsequent dynamical evolution of the stellar clusters is also a critical process to understand as clusters are the likely birthsites of most stellar populations in galaxies.

A sizeable fraction of stars is expected to harbour planetary systems, and these systems will also form and evolve in dense stellar clusters. Indeed, the formation of a planetary system may either be hindered or helped by a passing sibling star (§ 4.5). In dense star clusters, where the encounters between stars are relatively frequent, close passages may disrupt the circum-stellar discs around the stars so that further planet formation may not be possible. Or, already formed planets may be scattered into interstellar space. The weak, distant and more frequent encounters, on the other hand, may actually help the formation of planets in circum-stellar discs. The weak perturbations of the outer regions of such discs may lead to compression and rarefaction of the gas and dust which can locally lead to faster coagulation of the dust into larger bodies. It is essential to address theoretically, computationally and observationally how the collisional environment affects the formation of planetary systems. Is planet formation induced, or inhibited, or are already formed planetary systems brought into disorder through multiple encounters with sibling stars?

4.1.3 Strategy for the future

In the next few years, a suite of new tools will become available to probe the obscured cores of molecular clouds at the required angular resolution and sensitivity (e.g., ALMA, Herschel, JWST), while complementary and more advanced ones are being planned (e.g., SKA, and a far-infared interferometer in space). This observational progress has to be matched with improved theoretical models and numerical simulations of the star formation process. These models will be tested by high spatial and spectral resolution data of the clump-mass spectrum and velocity structure of clouds using ALMA and single-dish far-infrared and (sub)millimetre telescopes.

Magnetic fields in clouds will be imaged with SKA through radio emission of background galaxies polarized by the intervening interstellar matter. Combined with images of polarized synchrotron emission, three-dimensional maps of the magnetic field in the Galaxy can be made on the scales of molecular clouds. On smaller scales of star-forming cores and discs, the magnetic field structures can be constrained using ALMA through polarized dust emission and the Zeeman splitting of molecules. This knowledge of the magnetic field will be fed back into theoretical simulations to clarify the role of magnetic field in controlling the assembly of stars.

Star formation is a collective process, so the evolution of dense stellar systems plays a central role in shaping the output of star formation and merging the newly formed stellar populations within the Galaxy. Most of the theoretical issues cannot currently be dealt with adequately. The basic problem is that the computational time scales as the square of the number of stars. While special-purpose machines have been developed to treat up to 105 stars with high precision, these are currently inadequate to handle a large binary-star population. Calculation of the evolution of rich stellar clusters over billions of years requires sophisticated computer codes and the fastest special-purpose machines. Similarly, it is crucial to address how the formation of planetary systems, and their survival, depends on where the parent star is formed. These problems require technological advances in supercomputing machines, as well as mathematical breakthroughs in the methods used in performing the space-time transformations of highly energetic multiple stellar systems.

On the observational side, the distribution of stellar masses, binary fractions and other key properties of dense young stellar clusters will require a significant advance in optical and infrared telescopes. To appropriately investigate the cores of the most dense and massive young stellar clusters of our own Galaxy will require an ELT with a good adaptive optics system. To explore the properties 'at birth', before dynamical evolution, even younger clusters will need to be explored at high angular resolution in the infrared and (sub-)millimetre with the JWST and ALMA.

A major breakthrough in our observational characterization of binary stars in a variety of environments is expected from the Gaia space mission, which will measure tens of millions of binaries over the entire sky. Gaia's extreme sensitivity to non-linear proper motions will reveal large numbers of astrometric binaries with periods in the range 0.03–30 year and will provide photocentre orbits when the period is less than about eigth years. Radial-velocity observations of stars brighter than 15th magnitude will define large numbers of shorter-period binaries. Gaia will also observe millions of eclipsing binaries, mostly too faint for radial-velocity observations, and will resolve individual components of all binaries with separations above some 20 milli-arcsecond which have moderate magnitude differences between the components.

4.2 Is the initial mass function of stars universal?

4.2.1 Background

Our Sun is an average, inconspicuous star, one among billions in our Galaxy. But when looking at all the stars in a volume with radius of, say 10 pc, surrounding the Sun, we discover that our Sun is very much the exception: among the few hundreds of stars only a handful are as massive and bright as the Sun. The vast majority are faint red-dwarf stars with masses much smaller than that of the Sun. Roughly the same number of brown-dwarfs probably reside in this small volume, which is therefore almost exclusively populated either by red dwarfs or brown dwarfs that are of too small mass to be supported by thermonuclear fusion reactions in their interiors and have therefore faded beyond our current observable reach.

Is this mixture of stars and brown dwarfs representative of all stellar populations? A related but more general question is how many stars of a given mass exist: what is the mass-distribution of stars? Answering this question is of paramount importance to much of astrophysics and cosmology, because it is in stars that normal matter, which originally only came in the form of hydrogen and some helium, is transformed into heavier elements from which, ultimately, also life develops. The stars are the shining arrangements of matter with the help of which we can map most of the structure of the Universe, and through their radiation heat the interstellar medium of galaxies and drive galaxy-wide weather-patterns which locally can lead to further star formation. These processes are strongly dependent on the mass of a star. For example, a star which is one-hundred times heavier than the Sun is more than a million times brighter but only lives a few million years

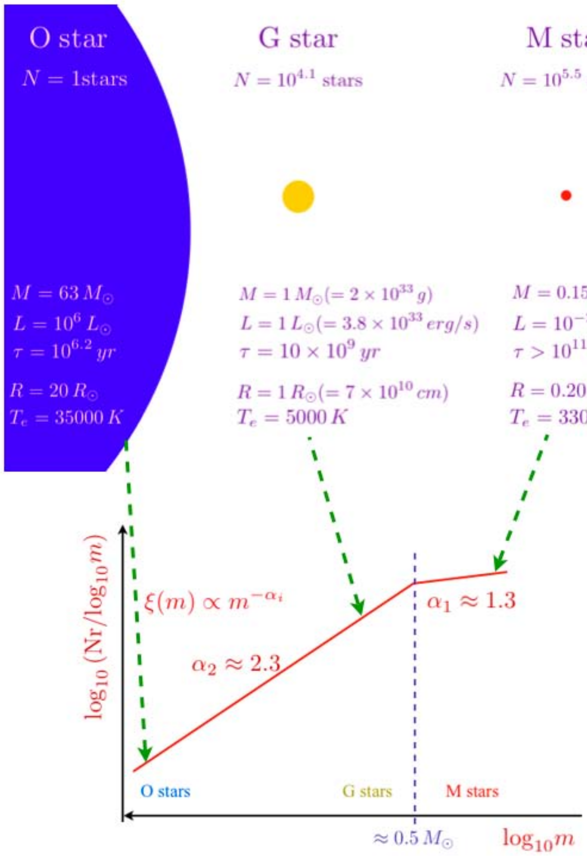

Figure 4.2: Three typical stars are shown in the top of the figure, a 63 $M_\odot$ O star, a 1 $M_\odot$ G star and a 0.15 $M_\odot$ M star, with some physical parameters (luminosity, life-time, radius and effective temperature). The relative number of these stars is also given, resulting from the currently known shape of the IMF, $\xi(m)$ = number of stars per unit mass interval, is indicated schematically in the lower part of the figure. The exact shape of the IMF, $\xi(m)$, and the possible existence of discontinuities as a result of different physical regimes of star formation, remains disputed.

which is a thousand times shorter than the Sun's life. Such stars explode as supernovae throwing most of their mass enriched with higher elements back into the galactic interstellar medium with tremendous energy. The low-mass red dwarfs, on the other hand, have effectively no influence on galactic weather and live to be as old as the age of the Universe. Their importance therefore comes in their large numbers. Cumulatively they account for most of the stellar mass in a galaxy and so they are important in defining the gravitational potential on galactic scales.

Star counts show that massive stars are very rare in our own Galaxy, while low-mass red-dwarf stars are abundant. The mass distribution is approximately an inverse square-law, with a turnover to a shallower slope at a few-tenths of a Solar mass. Such surveys rapidly become extremely challenging, because the largest telescopes are needed to see the faintest stars, while all-sky surveys are needed to map the structure of our Galaxy in order to assess the spatial distribution of the stars. The very massive stars are extremely rare, but, due to their very short life-times are concentrated in certain regions of the Galaxy where star-formation occurred recently. The long-lived faint stars, on the other hand, have enough time to diffuse away from their birth sites and are spread smoothly throughout the Galaxy. In order to connect the smooth counts of faint stars with those of massive stars, complicated corrections for their different spatial distribution are required.

These surveys show that the form of the mass distribution, the stellar initial mass function (IMF), appears to be very similar anywhere astronomers have looked, with, thusfar, only a moderate dependence on the environment. If confirmed, this would simplify the understanding of how galaxies change with time, but it challenges seemingly well-understood physical principles according to which the IMF should depend on the conditions of star formation. For example, when the galaxies were still very young their gas content had few elements heavier than helium and the physical theory of star formation would predict there to be many massive stars and fewer low-mass red dwarfs (§ 3.1). Instead, the uniformity of the IMF, as observed even in the relatively low-metallicity environments in the local Universe, would suggest there to have been few massive stars. This is surprising. Furthermore, observations and star counts indicate that no stars seem to form with masses above about 150 Solar masses. Despite a century of work, this physical mass limit is not understood. Perhaps more massive stars do form but quickly implode to black holes? At the other extreme of the mass function, brown dwarfs seem to form together with stars, although other formation mechanisms for these objects are possible and need study.

4.2.2 Key questions and strategy for the future

One of the key questions is to define the shape of the local IMF, including its limit at the low mass end. The other important issue that needs to be settled is to what extent the specific shape of the IMF depends on environment, in particular on the metal content in the gas out of which stars form. These questions are connected to the more fundamental one of the origin of the stellar IMF. A recent development that still needs to be explored is the possibility that the stellar IMF is determined by the cloud fragmentation process in the early stages of star formation.

The observational challenges are to test the uniformity of the IMF by observing many more young stellar populations and also entire populations of stars of whole galaxies. Especially challenging are observations of the lowest-metallicity stellar populations in the dwarf galaxies and (extragalactic) globular clusters in the Local Universe. These will require a substantial jump both in terms of sensitivity and angular resolution as compared to the current generation of adaptive optics assisted large optical and infrared telescopes, that only a fully adaptive ELT may be able to offer. On the theoretical side the physical description of the star-formation process needs to be improved, which will require powerful computers and codes in order to allow complete simulation of the physical processes occurring during the formation of stellar systems.

It is equally important to improve our knowledge of the stellar IMF in the Galaxy. Gaia will provide a complete and homogeneous census of all stars to magnitude 20, and an accurate determination of the mass-luminosity relation (enabling a major improvement in the transformation from luminosity functions to mass functions). For the Solar neighbourhood, this will enable a statistically significant study of the luminosity and mass distributions in a distance-limited volume. The similar census of the stellar content of a large number of clusters, associations and moving groups, will provide the IMF in star forming regions.

Studies of clouds which are about to form stars are needed for understanding the origin of the stellar IMF and its relation to the cloud fragmentation process. Major advances in this area require large multi-frequency and high spatial resolution surveys with Herschel, JWST and ALMA.

4.3 What do we learn by probing stellar interiors?

4.3.1 Background

The basic understanding of stellar structure and evolution is well-established through decades of extensive model computations and comparisons with observed properties of stars. At a superficial level the models and the observations appear to agree in, e.g., the Hertzsprung–Russell diagrams of stellar clusters. Yet this apparent success hides a great deal of uncertainty in the

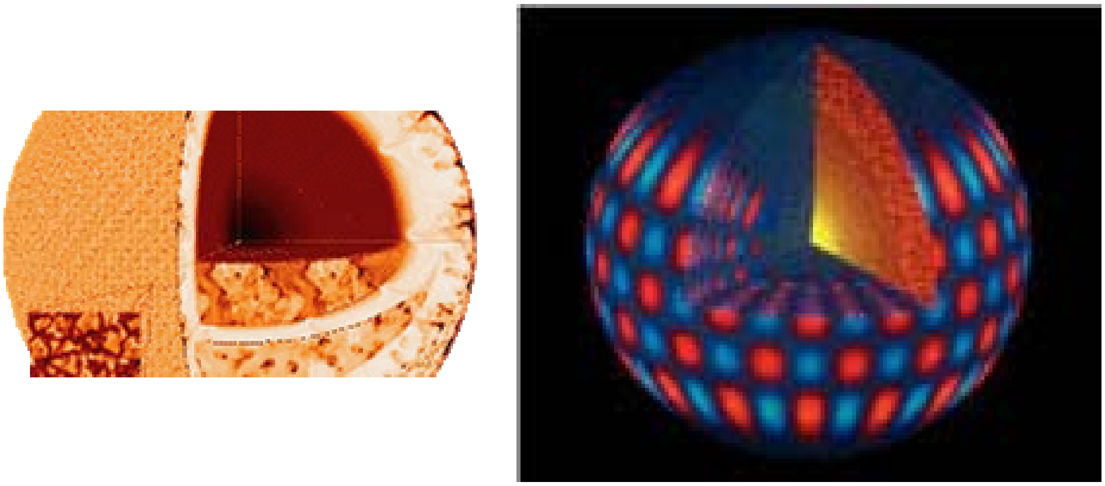

Figure 4.3: Oscillations generated at the inside a star will travel outwards and will become visible as parts of the surface moving up and down. The mode, frequency and intensity of these oscillations give valuable information about the inside of the star. It can be compared to the information that we get about the inside of the earth from the propagation of the seismic waves below the crust.

modelling, which in a fundamental way affects our understanding of stellar evolution and its use in other areas of astrophysics. The basic problem is that ‘classical’ photometric and spectroscopic observations have limited sensitivity to the internal structure of the stars.

Asteroseismology, based on observations of stellar pulsations, provides information about the internal properties of stars. The oscillation frequencies can be determined with exquisite accuracy, and provide measures of the structure and dynamics of the stellar interior. With a sufficient number of observed modes, as available in many types of stars, specific information can be extracted which provides a much more direct test of stellar models than do the ‘classical’ observations. Pulsating stars are found in essentially all parts of the Hertzsprung–Russell diagram, and hence asteroseismology can test and extend our understanding of all phases of stellar evolution.

Helioseismology has provided detailed information about the Solar interior (§ 5.1). Such results are crucial for understanding the origin of Solar activity, and its impact on the Earth (§ 5.3).

4.3.2 Key questions

The main uncertainties in stellar modelling concern hydrodynamic phenomena in the interiors. Convection, and penetration of motion beyond convectively unstable regions, remains a serious problem, affecting the internal composition structure and the later evolutionary stages, as well as the observed surface abundances. Even more serious are the generally neglected effects of stellar rotation. Little is known about the evolution of rotation resulting from the internal transport of angular momentum or the loss of angular momentum to stellar winds, and circulation and instabilities associated with rotation are likely to play an important role in mixing of chemical elements, again affecting the evolution and the observed surface abundances. Modelling of rapidly rotating stars will have to move beyond the current treatment of rotation as a perturbation around a spherically symmetric state, to a full two-dimensional calculation of stellar structure and evolution. Asteroseismology could provide a critical measure of the mixed (convective) core of main sequence stars more massive than 1.2 Solar mass. The size of the core directly determines the lifetime of these stars, thus affecting the age dating in particular of high redshift galaxies (hence their formation epoch) through stellar population synthesis methods based on stellar models.

Extrasolar planets appear to be preferentially associated with stars with relatively high surface abundances of heavy elements. It is not known whether this is a result of the stars being born from matter especially enriched in heavy elements or whether it is the result of the pollution of the outer layers by infalling material, e.g., from planets being swallowed by the star. Asteroseismology has the potential to distinguish between these two possibilities.

A better understanding of stellar activity would certainly improve our understanding, and perhaps our prospects for predicting, Solar activity (§ 5.3).

4.3.3 Strategy for the future

The observational prospects for asteroseismology are promising. The development of extremely stable spectrographs for the search for exo-planets has led to a revolution of the study of Solar-like pulsations in main sequence stars. The WIRE and MOST satellites provide excellent data on stars with relatively large amplitudes. The CoRoT and Kepler missions will allow detailed asteroseismic investigations of a large number of stars through very precise space-based photometry. Experience from helioseismology shows that full utilization of the potential of asteroseismology requires high signal-to-noise Doppler-velocity observations of carefully selected stars, extending over several months or even years. This will require networks of dedicated telescopes. In addition to asteroseismology such networks would also provide very valuable data on extrasolar planets.

Asteroseismic analyses have already provided significant constraints on stellar modelling for a broad range of stars, including Solar-like stars, massive main-sequence stars and white dwarfs. The coming years will show increasingly reliable asteroseismic determination of stellar ages, of great value to investigations of the evolution of the Galaxy. Information will be obtained on the extent of convective envelopes and cores, and on the importance of mixing processes which affect the composition and hence the oscillation frequencies. In addition, constraints will be placed on the internal rotation of stars from observations of rotational splitting of oscillation frequencies, and hence on the modelling of the evolution of stellar rotation. Finally, as in the case of helioseismology, information will be obtained on the physics of stellar matter, including equation of state and opacity, under conditions that are substantially different from those in the Sun, extending the concept of using stars as laboratories for physics under extreme conditions (see § 5.1).

A crucial complement to the information provided by asteroseismology is an accurate knowledge of stellar global parameters. Many results were obtained from the astrometric data provided by the Hipparcos satellite (fine structure of the Hertzsprung–Russell diagram, age of local halo stars – solving for the first time the discrepancy between the age of the oldest objects in the Galaxy and the expansion age of the Universe). Its successor, Gaia, will measure accurate global parameters for a billion stars covering all types and populations, reaching the rarest and the most rapidly evolving objects. It will provide membership, extremely accurate distances and photometry of nearly many galactic open and globular clusters, providing a dramatic improvement in quantifying the effects of age, metallicity and He content. The simultaneous availability of seismological measurements and Gaia absolute magnitudes will also, for example, allow separation of the effects of rotation and overshooting on the position of a star in the Hertzsprung–Russell diagram.

A long-term goal is to carry out asteroseismology with sufficiently high spatial resolution to resolve stellar surfaces and study modes of relatively high spherical-harmonic degree. This will allow analysis of the internal dynamics, including the generation of stellar magnetic fields through a rotation-convection dynamo. Such observations will likely require space interferometry.

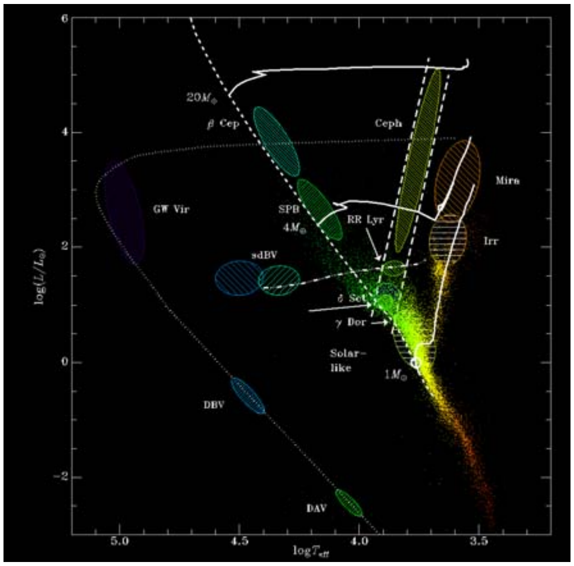

Figure 4.4: The observational Hertzsprung-Russell diagram, absolute luminosity versus temperture, for the 16 631 Hipparcos single stars with an error in the parallax smaller than 10%, and with the additional constraint that the error in (B-V) smaller than 0.025 mag. The evolutionary tracks have been indicated for a star of 1, 4 and 20 M_{odot}. The position of several types of stars have been indicated by the filled areas. The colors indicate the color of the stars.

4.4 What is the life-cycle of the Interstellar Medium and Stars?

4.4.1 Background

The nuclear reactions inside stars are the main mechanism that drive the chemical evolution of baryons in the Universe and, starting from the primordial gas of light elements, allow for the creation of heavier atoms that are at the base of our biology. However, stellar nucleosynthesis is only one of the many steps in the life cycle of matter. The nucleosynthesis of the heaviest elements is only possible during supernova explosions and the combination of elements in large molecules is only possible in dark clouds. Chemistry in cores and circumstellar discs during the formation of stellar and planetary systems may also be responsible for the formation of some of the complex molecules that are building blocks of life.

From simple to prebiotic molecules. In the densest and coldest cloud cores, the bulk of the heavy species freeze out onto the grains and form ices, resulting in an enormous reservoir of water ice. Some of these ices will be transported directly into icy planetesimals like comets and Kuiper Belt Objects (§ 5.5), as revealed by their isotopic enrichment inherited from the cold phase (particularly in deuterium). Once the protostar is formed and heats up its surroundings, the ices evaporate back into the gas where they form the basis for a high temperature chemistry leading to large organic molecules. Complex species such as ethers, sugars, acids and alcohols have been found with greatly enhanced abundances in the so-called ’hot cores’ associated with low- and high-mass star formation. Even larger polycyclic aromatic hydrocarbons (PAHs) are seen as well.

Some of this material will be incorporated into protoplanetary discs where ultraviolet, X-ray and thermal processing can further modify the composition. Chemical studies of discs are currently limited to the outer disc (beyond 50 AU). Little is known of the chemistry in the planet-forming zones. Determining the chemical evolution of protostellar and protoplanetary material inside 10 AU for a wide range of stellar masses, luminosities and environments is a major scientific goal.

Dust in the interstellar medium consists of silicates and carbonaceous material, mostly in amor-

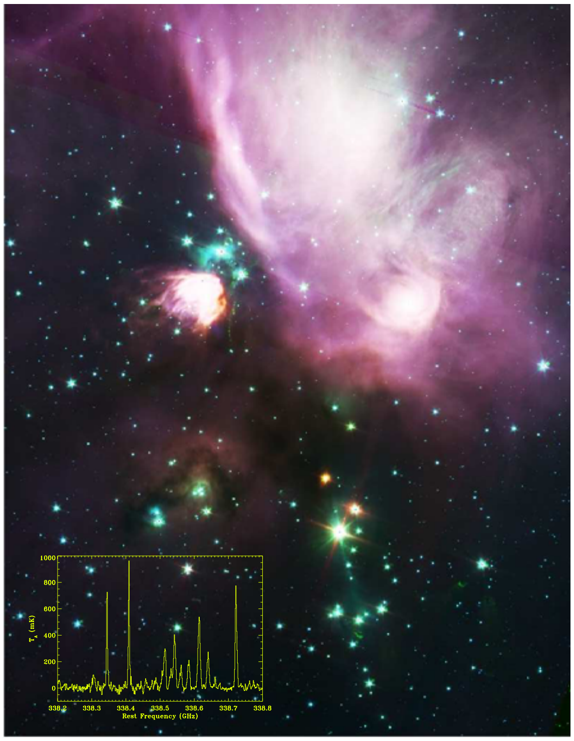

Figure 4.5: Detection of hot methanol in the young protostar IRAS 16293-2422 in Ophiuchus.

phous form. A key surprise was the discovery that a significant fraction of the dust in discs around both young and old stars can be in a more ordered, crystalline form. This change in dust structure must be related to the physical processes in discs, i.e., grain coagulation, settling, transport and annealing. Thus, mapping the dust features in discs provides a direct probe of disc evolution.

Late stages of stellar evolution and the feedback to the interstellar medium. Stars spend most of their life in relatively stable configurations during which nuclear reactions take place in their innermost regions. Before terminating the production of nuclear energy, low mass stars (between 0.8 and 8 Solar masses) pass through a crucial and complex stage called the Asymptotic Giant Branch phase. This is characterized by a poorly-understood thermally pulsing phase during which the star ejects matter at an increasingly rate (up to 10^{-4} Solar mass per year) which ultimately causes the end of nuclear burning and the transition to the spectacular Planetary Nebula phase, after which the star becomes a white dwarf. When the white dwarf accretes material either from a companion star or merges with another white dwarf, it can reach the Chandrasekhar limit and explode via a thermonuclear runaway producing a Type Ia supernova (§ 2.4). As a result, these low-mass stars feed back the light elements (He, CNO) into the interstellar medium where they are then available for future generations of stars. Understanding the complex structure of Planetary Nebulae is necessary to address many of the main open questions regarding the mass loss history, the wind interaction, the ionization processes, the effect of magnetic fields, of the binarity of the central star and microstructures, the formation and evolution of dust, and the synthesis of complex molecules.

The evolution of stars more massive than 8 Solar masses is very different, and results in the collapse of the stellar core. After helium burning, neutrino losses dominate over radiative and convective transport and the star evolves so quickly that the outer layers are unable to keep up. Reliable empirical mass loss rates and quantitive theories able to provide a self-consistent picture are missing. Computations show that rotation during hydrogen and helium burning increases the mass of the CO cores with respect to non-rotating models and causes additional mixing. During the late stages the interplay of mass-loss and rotation dominates the extraction of angular momentum from the core. Models predict large angular momenta which are inconsistent with the slow rotation rates of pulsars but, on the other hand, are required by gamma-ray burst models (§ 2.4).

The elements returned to the interstellar medium through stellar winds and supernovae, together with the energy released, are important ingredients in galactic evolution. They modify the heavy element content, the mass inflow and outflow from the galactic central regions and lead to the formation of large-scale structures (§ 3.2). Under the action of such agents the structures of the interstellar medium may become unstable and form new molecular clouds which eventually lead to new star formation episodes.

4.4.2 Key questions

From simple to prebiotic molecules. One of the main challenges for the next decade is to chart the composition of gaseous and solid materials at the different stages of star- and planet formation, some of which may eventually form the basis for life in other planetary environments. An equally important goal is to exploit molecules and dust features as diagnostics of the physical processes associated with forming stars and planets, since their infrared and submillimetre bands are the only observables in highly extincted regions.

Late stages of stellar evolution and the feedback to the interstellar medium. For low-mass stars the available stellar evolution models are not able to reproduce the composition of the surface layers of stars on the asymptotic giant branch in the Hertzsprung–Russell Diagram, expelled during mass loss episodes, and the dependence of mass loss on time, stellar mass and metallicity. The final stages of the evolution of massive stars remain surrounded by many uncertainties (§ 2.4). The rate at which mass is lost, and how it depends on metallicity, is uncertain, in particular for Wolf–Rayet stars. The role of rotation must be understood in all phases.

4.4.3 Strategy for the future

From simple to prebiotic molecules. Enormous advances can be expected in this area. ALMA will have the sensitivity and spatial resolution to image both simple and complex gaseous molecules in the planet-forming zones of discs for the first time, on scales down to 0.1 arcsecond or better inside 10 AU in the nearest star-forming regions). It can also resolve the chemical structure of shocks due to outflows, infall or winds, and pinpoint the sputtering of grain cores and liberation of ices. ALMA will allow deeper searches for the most complex prebiotic organic molecules in hot cores by spatially resolving the emission and filtering out confusing lines from more extended species. The largest molecules have their strongest transitions at cm wavelengths where SKA will excel. A key molecule, water, will be surveyed by Herschel, but higher spatial resolution far-infrared data will be needed in the future. Large ground-based optical telescopes and JWST equipped with mid-infrared instruments, together with Herschel and future far-infrared missions, will be essential to probe the solid-state component (silicates, ices, PAHs) as well as hot organic gases not observable by ALMA on similar scales. They can also map ices in cold cores on less than 1000 AU scales (comparable to gas-phase maps) through absorption studies against background stars. X-ray missions, optical and infrared data probe the elemental composition of material returned by (super)novae and winds of dying stars back into the interstellar medium.

This observational progress should be accompanied by similar progress in our understanding of the basic physical and chemical properties of molecules and solids, by means of theoretical calculations and laboratory experiments. For example, frequencies of many complex molecules in the ALMA wavelength range are still lacking and the Herschel THz range is still largely unexplored spectroscopically. Also, basic collisional rate coefficients to analyze ALMA and Herschel data are highly incomplete, and laboratory studies of PAHs and gas-solid processes, important to interpret JWST data, are still in their infancy. The scientific return of these instruments will be greatly enhanced by relevant studies of basic molecular processes.

Late stages of stellar evolution and the feedback to the interstellar medium. Fundamental tests of stellar models are provided by the study of non-radial oscillations of stars. If important information can be obtained for the brightest stars through high-resolution spectroscopy from ground, a dramatic impulse to asteroseismology is expected in the near future form the space massions like Corot, expected to be launched at the end of 2006, and, subsequently, Kepler.

Space missions are also the only way to access the ultraviolet and X-ray window, on one side, and mid- and far-infrared, on the other. Both domains are fundamental for the study of both the fast and slow winds from evolved stars. Spitzer and Herschel will play a major role at long wavelengths along with millimetric observations from Alma. At shorter wavelengths after the current workhorses HST, XMM and Chandra the situations is still undefined.

4.5 How do planetary systems form and evolve?

4.5.1 Background

Circumstellar discs are ubiquitous around low-mass pre-main-sequence stars. Discs appear to exist around more massive stars, but are more difficult to separate from the surrounding molecular cocoon, because of the short evolution time scale of these objects. These discs are believed to play a major role in the assembly of the central star and in the formation of a planetary system.

Disc evolution is controlled on one side by the decline of the accretion rate, and on the other by changes in dust properties. High spatial resolution studies of discs at millimetre frequencies have shown that grains in the outer disc (at radii larger than 30–50 AU from the star) have grown in many cases to very large sizes (up to few cm); mid-infrared spectroscopy, on the other hand, indicates that there is a residual population of micron-size silicates on the disc surface, at least within a few tens of AU from the star. Much smaller grains (less than 5–10 nm) and macromolecules (PAHs) are also present on the disc surfaces, even in rather evolved objects, and have an important role in determining the gas physical and chemical conditions.

The properties of the dust population, including its size distribution, are very likely the result of sedimentation and coagulation, which, in turn, depend on the gas motions. Grains in the disc midplane may form larger aggregates that eventually lead to planetesimal formation, which, in turn, are the building blocks of the rocky core of planets.

When the planetary system is assembled, the disc is almost completely devoid of gas and the dust component is reduced to a few lunar masses. In these discs, the dust is not pristine, but represent the debris of the planetary formation process, for this reason these are called ‘debris discs’. The study of debris discs is a powerful indirect proble of the planetary system they host.

4.5.2 Key questions

The evolution of the dust-to-gas ratio in discs is a major unsolved issue. The main constituent of the gas, molecular hydrogen, is essentially unobservable, except in a small and warm region of the disc where the mid-infrared lines of H_2 can be produced. The study of the gas properties will mostly remain a very indirect process which must rely on observations of much less abundant species, such as CO or molecular ions. We have only a rudimentary grasp of very basic properties such as the initial gas and dust disc mass, disc sizes, and the distribution of material in the disc.

The nature of the interaction of the disc with the environment is an open issue. There is evidence that discs, accretion onto the central star, and the ejection of material in powerful jets are connected, but there is no self-consistent model of a disc-star-jet system. This is partly due to the lack

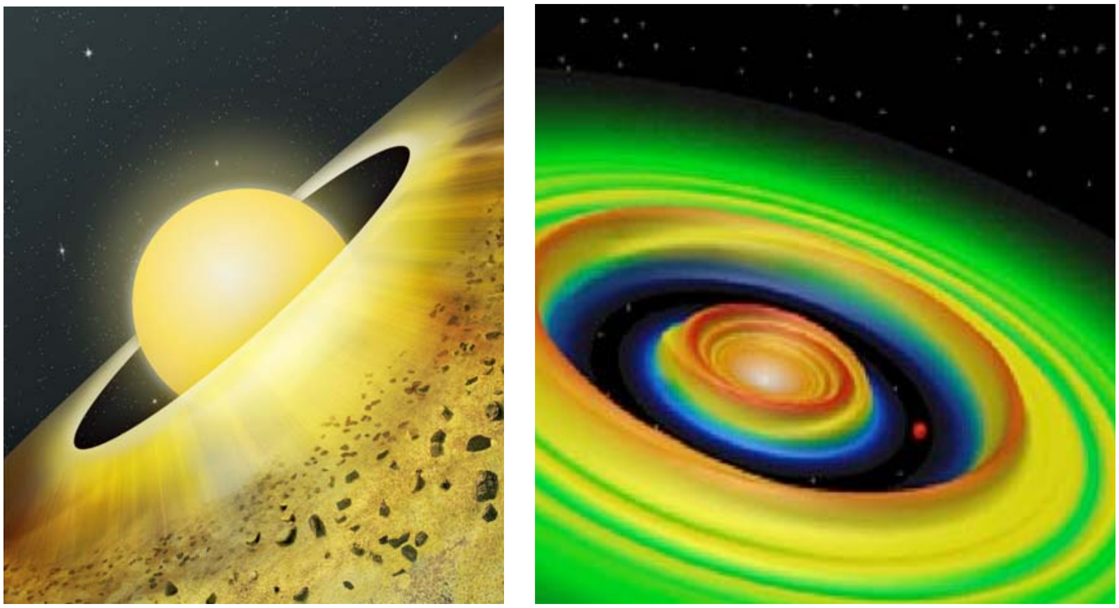

Figure 4.6: Dust evolution and planets formation in circumstellar discs. Left: artistic view of the formation of pebbles in circumstellar discs as suggested from millimeter wave observations of the TW Hydrae system. Right: simulation of the formation of a gap in a disc around a young star due to the gravitational effect of a newly formed giant planet.

of detailed observations of the inner regions of these systems where most of the interaction occurs and partly due to our limited understanding of the physical processes in discs (e.g., viscosity).

Understanding the physics of discs would progress significantly if an evolutionary sequence could be drawn. Given the difficulties in constraining the ages of pre-main-sequence stars, this can only be addressed through a statistical approach, and will be a major undertaking: so far only a handful of objects have been studied in terms of gas content, indirectly through CO. This evolutionary sequence may well be different in star forming regions of various environments. Tidal processes in dense stellar environments are expected to affect the disc properties. For example, current observations suggest that circumstellar discs disappear within about 10 million years, while Uranus, Neptune and Kuiper-Belt objects require about 100 million years to form (§ 5.4).

While the idea that planets form in protoplanetary discs in the early phases of the stellar life is generally accepted, many key questions are still unanswered: is the formation of planetary systems a robust and common process (how frequent is the formation of a planetary system)? What are the demographics of planetary systems (is our own Solar System a common product of the planetary formation process)? How strongly is the formation of planetary systems influenced by the properties and evolution of the central star (are the conditions that led to the formation of our own Solar System ‘special’)?

The actual mechanism of planetesimal formation, the first step in planet formation, is also still debated. A tail of smaller grains is predicted, either as a leftover of the initial solid population, or as the result of collisional fragmentation of larger bodies. The sequence of events depends on a number of assumptions, both on the disc properties (e.g., the relative velocity of colliding particles) and grain properties (e.g., the sticking probability). Calculations have been performed only for the conditions of the primordial Solar nebula, and need to be expanded and improved. The observational results on the properties of the grain population, which have revealed grain sizes up to few centimeters can provide important constraints to the models. Another potentially strong constraint could come from the spatial distribution of the large grains within the disc.

4.5.3 Strategy for the future

High angular resolution observations of dust and gas are required to constrain disc properties. Both dust and gas can be studied through the infrared to the millimetre and even longer wavelengths. A very large wavelength coverage is critical. For the youngest, most massive discs, the dust opacity is low enough only at the longest wavelengths to allow penetration into the inner regions (10 to 20 AU). When the disc starts to dissipate, the opacity decreases and the infrared regime becomes progressively more appropriate and more sensitive.

The evolution of large grains (cm-sized) into planetesimals (km-sized) and rocky cores can only be explored numerically as these phases of the evolution are not accessible to direct observations. A related question concerns the formation of the gaseous planets. Both observations and models are required to address whether gas accretion on rocky cores or gravitational instabilities within gaseous protoplanetary discs is the formation mode. The latter scenario seems to predict cores for the giant planets in our Solar System that are inconsistent with the most recent estimates of the inner structure of these planets (§5.4). The mechanism of gas accretion onto the core is favoured by two main observations: the fact that the present composition of giant planets gas is depleted in volatiles (H, He) compared to the Solar composition, and the existence of a large rocky core. These facts seem to indicate that giant planet formation took place in an evolved disc. On the other hand the effect of competitive accretion between forming planets and the detailed interaction of the accreting planet with the protoplanetary disc still needs to be fully addressed.

Constraints on the different planetary formation theories can be provided by observing planet-forming discs at the required resolution (of order a few AU). Oservations of structures in discs (spirals, gaps, density enhancements) and the newly formed planets within discs are all powerful tools that require the next generation facilities that will offer high angular resolution and sensitivity (ALMA, JWST, SKA, ELT and a far-infrared interferometer in space).

4.6 What are the demographics of planets in the Galaxy?

4.6.1 Background

Since 1995, the year of the discovery of the first Jupiter-sized object orbiting a star other than the Sun, the number of detected extrasolar planets have steadily increased each year. The vast majority of these discoveries have been accomplished by various high precision Doppler surveys on samples of more than a thousand nearby stars. This is illustrated in Figure 4.7 where current and planned planet detection techniques are summarized, together with their successes.

The unexpected properties of the extrasolar planets found so far have sparked much new theoretical work, with the aim to move from a set of models describing separate aspects of the physics of the formation and evolution of planetary systems to a plausible, unified theory, capable of making robust and testable predictions. After a decade of extrasolar giant planet discoveries, the only idea that has not yet undergone significant revision is the paradigm that planets form within gaseous discs around young T Tauri stars. Many old ideas were revisited or revived, and a number of new ones were proposed in an attempt to explain the observational data on extrasolar planets.

The extrasolar planet sample exhibits many interesting and surprising orbital characteristics compared to the giant planets of our Solar system. The most striking feature is the presence of giant planets on very short orbits. This has profound implications for our understanding of planet formation. Interactions between the planet and the disc at the early stage of formation can affect the orbit of the planet. Resonant interactions of a planet with a disc of planetesimals inside its orbit, dynamical friction between a planet and a planetesimal disc as well as tidal interaction between a gaseous disc and an embedded planet can lead to the migration of the planet up to a very short orbital distances of its star. However, the scenarios for giant protoplanet migration in gaseous discs are not without problems. Timescales for migration are very short, much shorter than typi-

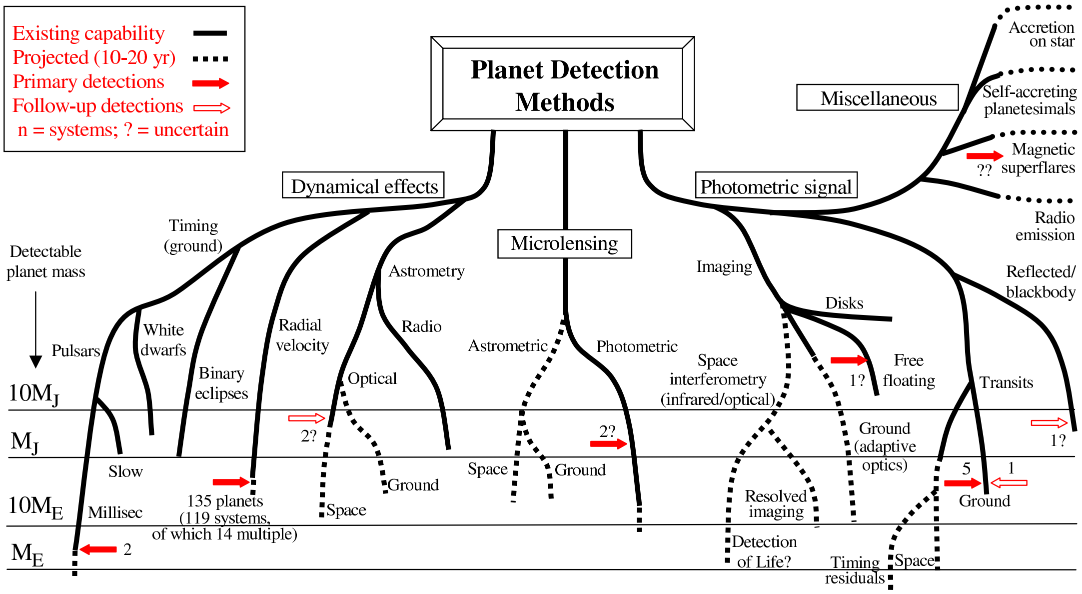

Figure 4.7: Detection methods for extra-solar planets. The lower extent of the lines indicates, roughly, the detectable masses that are in principle within reach of present measurements (solid lines), and those that might be expected within the next 10-20 years (dashed). The (logarithmic) mass scale is shown at left. The miscellaneous signatures to the upper right are less well quantified in mass terms. Solid arrows indicate (original) detections according to approximate mass, while open arrows indicate further measurements of previously-detected systems. ? indicates uncertain or unconfirmed detections. The figure takes no account of the numbers of planets that may be detectable by each method. Picture taken from ESA-ESO report nr 1: Extra-Solar Planets.

cal disc and planet formation lifetimes. Furthermore, a stopping mechanism must be devised in order to prevent the migrating protoplanets from plunging into the central star, and to reproduce the observed pile-up of planets on few-day orbits.

The typical planet frequency in the giant planet mass range and with semi major axis shorter than 3 AU, is estimated to be about 5–7 per cent. The planet detection rate increases strongly for stars with metallicity higher than the Sun. This is likely due to the initial enrichment of the nebulae but its precise link to planet formation mechanisms is a matter of debate.

With the large number of planets detected, many statistical studies of the orbital parameters and masses have been carried out. The mass distribution of extrasolar planets extends to above ten Jupiter masses and the numbers rise steeply towards the low mass planets. Smaller host stars may exhibit a lack of massive planets on short orbits and a decrease in typical planetary mass.

The distribution of eccentricities of the orbits of extrasolar planets is a major puzzle. Dynamical interaction of a migrating planet in a disc of planetesimals and resonant interactions between migrating planets can excite eccentricities. Close encounters between planets may lead to highly eccentric orbits as well. Finally, secular interactions with a distant companion star out of the planet's orbital plane could excite a planet's eccentricity. Although all contribute, none of these mechanisms seems to dominate the planet formation process.

The number of known multiple-planet systems is increasing both with the extension of the duration of surveys and the improvement of the measurement precision. Current findings suggest that the fraction of such systems containing giant planets may be high in comparison with single-planet systems. The possibility of resonances and gravitational coupling in these systems has revived interest in orbital dynamics. Due to lack of information on planet masses and relative inclination angles in multiple systems, general conclusions on the architecture, orbital evolution and long-term stability of the newly discovered planetary systems are difficult to derive.

Transiting planets have been detected by photometric surveys. From these transit measurements in combination with radial velocity measurements, planetary densities have been obtained which provides information on their internal structure. For hot Jupiter planets the measured densities confirm their gaseous nature but deviate from predictions of internal structure models. For smaller planets, transit information provides a way to test whether the planet possesses a core made of rocky material, constraining scenarios of planet formation and evolution. The special geometry of a transiting planet also permits interesting follow-up studies, such as searches for planetary satellites and studies of features in the planetary atmosphere by transmission spectroscopy during the transit. Very recently, the detection of secondary eclipses on HD 209458 and TrES-1 with Spitzer have revealed the flux emerging from these planets themselves.

4.6.2 Key questions

Searches for extrasolar planets have revealed a large diversity amongst planetary systems. Now that surveys are becoming sensitive to systems with giant planets in the same distance range as in our own Solar system, an important question to be addressed is the frequency of Solar System analogues. The relation between planetary systems and the properties of the host stars needs further work, while the study of exoplanet atmospheres exoplanets is an almost unexplored topic.

The ultimate goal of planet searches is the direct detection and characterisation of Earth-like planets in habitable zones of Solar-type stars. This is an unknown territory up to now, as current instrumentation does not have the capabilities of detecting such planets. The goals for the coming decades will be to detect Earth-like planets around other stars, to estimate the frequency of their occurrence and possibly to obtain direct images of some of these. This may then allow a future spectroscopic characterization for the search for extrasolar life (§ 4.7).

4.6.3 Strategy for the future

We are in an exciting era for exoplanet searches, with the first spectroscopic surveys being extended to detect smaller giant planets in larger orbits, approaching conditions more similar to our own Solar System, and new search methods and missions are being planned. The Corot, Gaia and Kepler satellites will provide an enormous enhancement in our capabilities of detecting planets via transit, microlensing and astrometry. Thanks to its extreme sensitivity to non-linear proper motions, Gaia will monitor hundreds of thousands of stars within 200 pc from the Sun for the presence of planets. This is expected to reveal of the order of 10 000 exoplanets, with periods smaller than ten years, mostly Jupiter mass, going down to a few tens of Earth masses for the nearest stars. Complete characterisation of the host stars (luminosity, gravity, effective temperature, abundances, rotational velocity), will give basic information on the conditions necessary for planetary system formation. The Gaia photometry will also uncover about 5000 planetary transits. The requirements for exoplanets detections based on orbital motion (radial velocity, astrometry, transits) are shown in Figure 4.8.

Detection and characterisation of exo-Earths, even around nearby stars, is extremely challenging. The required technological effort will lead to facilities able to provide a much better characterisation of more massive exoplanets. The frequency of occurrence of Earth-mass planets around a variety of stars may be derived with appropriate radial velocity and transit search programmes.

High-resolution spectrographs with the required long-term stability to produce measurements of accuracy better than 0.1 m/s (the amplitude induced by an Earth-mass planet at 1 AU around a Solar-type star) on 4–10 m telescopes should be developed for systematic searches of nearby stars. Observations should first prove the stability of the stellar photospheres to the required level and then last as long as required to explore the habitable zones of the target stars. Ssimilar instruments attached to much larger telescopes should extend the survey to more distant Solar-type stars and also to the nearby coolest stars (exploiting the near-infrared). The outcome of this effort is likely

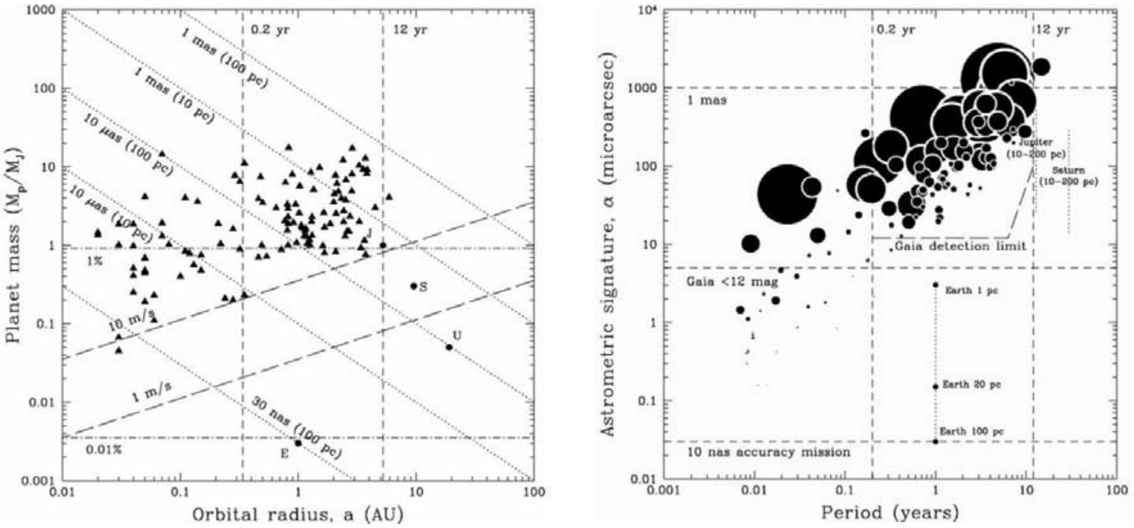

Figure 4.8: Left: Detection domains for methods exploiting planet orbital motion, as a function of planet mass and orbital radius, assuming $M_* = M_\odot$ Lines from top left to bottom right show the locus of astrometric signatures of 1 milli-arcsec and 10 micro-arcsec at distances of 10 and 100 pc; Vertical lines show limits corresponding to orbital periods of 0.2 and 12 years, relevant for Gaia (where very short and very long periods cannot be detected). Lines from top right to bottom left show radial velocities corresponding to K = 10 and K = 1 m s^{-1}. Horizontal lines indicate photometric detection thresholds for planetary transits, of 1% and 0.01%, corresponding roughly to Jupiter and Earth radius planets respectively. The positions of Earth (E), Jupiter (J), Saturn (S) and Uranus (U) are shown, as are the lower limits on the masses of known planetary systems as of December 2004 (triangles). Right: Astrometric signature, α, induced on the parent star for the known planetary systems, as a function of orbital period. Circles are shown with a radius proportional to M_p sin i. Astrometry at the milli-arcsec level has negligible power in detecting these systems, while the situation changes dramatically for micro-arcsec measurements. Short-period systems, to which radial velocity measurements are sensitive, are difficult to detect astrometrically, while the longest period systems will be straightforward for micro-arcsec positional measurements. Effects of Earth, Jupiter, and Saturn are shown at the distances indicated. Both graphs are taken from ESA-ESO report nr 1: Extra-Solar Planets.

to provide detection of the nearest Earth-mass planets and crucial input for subsequent direct imaging and spectroscopic characterization programmes.

Monitoring of several million stars with high precision photometry (0.01 millimag) over a period of five to six years from space will be key to establish the frequency of Earth-like planets via transits. The same approach could also perform a microlensing search (provided the number of targets in the field is sufficiently large). Both techniques will provide a more solid statistical basis and will allow a comparison of planet statistics in a variety of environments of the Galaxy.

Facilities with extreme high contrast and large spatial resolution are required for direct detection (imaging) of exo-Earth candidates. Increased spatial resolution can be provided by two complementary facilities: an extremely large optical / near infrared telescope from ground with sufficient collecting area to capture the reflected light of planets at 1–2 AU from the parent star even at distances of 50–100 pc and by a sufficiently sensitive mid-infrared interferometer in space.

In both cases it is essential to achieve very high contrast imaging at the level of 109. For ground based telescopes this requires major developments in the field of extreme adaptive optics, coronagraphy and differential imaging, as this level is three orders of magnitude beyond the capability of current imaging systems. Space-borne nulling interferometry in the mid-infrared is an exciting alternative crucial for characterisation of Earth-like planets in a spectral domain very difficult to observe from ground. Any effort undertaken to prove new concepts and techniques for high contrast imaging both in ground and space-based telescopes will be extremely valuable.

4.7 How do we tell which planets harbour life?

4.7.1 Background

The first direct detection of the atmosphere of a giant hot planet orbiting a star outside our Solar system was performed by HST when the planet passed in front of its parent star, allowing to see light from the star filtered through the planet's atmosphere. This unique observation demonstrated that, under fortunate conditions, it is possible to measure the chemical composition of exoplanet atmospheres even with current instruments.

In general, the characterization of the atmospheres requires challenging, direct observations of very faint objects in the glow of bright stars (contrasts of order 10^9). Direct imaging allows, in addition to detection, determination of the orbital parameters, as it is believed that a low eccentricity stable orbit is a pre-requisite for the development of life. Spectroscopy provides information about the presence and composition of the atmosphere. Polarization gives hints on the structure of the planet atmosphere and the presence of dust. Light curves might provide the planet period of rotation around its own axis, information about the presence of satellites and rings, or even about the presence of clouds and possible structures on the surface. If we could reasonably expect to characterize gaseous giants with the instrumentation already envisaged for the coming decade, the direct observation of rocky planets, according to our current understanding the only ones able of host life, requires a new generation of ground- and space-based instrumentation.

Based on the knowledge of carbon-based life on Earth, which requires water for its chemical reactions, the habitable zone has been defined as the distance range over which liquid water is likely present on a planet surface, and the continuously habitable zones are those regions in which liquid water is expected to be present over a significant fraction of the star main-sequence lifetime. For this reason the search for habitable planets will be concentrated on rocky planets in low-eccentricity orbits around Sun-like stars at about 1 AU distance. The position and extent of the habitable zone depend on the stellar luminosity and age, on the planetary atmDetection methods for extra-solar planets. The lower extent of the lines indicates, roughly, the detectable masses that are in principle within reach of present measurements (solid lines), and those that might be expected within the next 10-20 years (dashed). The (logarithmic) mass scale is shown at left. The miscellaneous signatures to the upper right are less well quantified in mass terms. Solid arrows indicate (original) detections according to approximate mass, while open arrows indicate further measurements of previously-detected systems. ? indicates uncertain or unconfirmed detections. The figure takes no account of the numbers of planets that may be detectable by each method. Picture taken from ESA-ESO report nr 1: Extra-Solar Planetsosphere and on possible internal heat sources. The presence of life can be inferred by the detection of life-related compounds like O_2 and O_3, and more robustly by the simultaneous presence of O_3 and N_2O; all these molecules are difficult to detect.

4.7.2 Key questions and strategy for the future

The long-term goal is to spectroscopically study extrasolar planetary systems which contain at least one rocky planet orbiting in the habitable zone in order to detect signatures of the development of life on the planetary surface. While this is a highly ambitious goal and possibly not within the foreseen capabilities of the facilities that will become available in the coming decades, extrasolar planetary systems search programs and technological developments for innovative instrumentation concepts should be carried out in the future with this as ultimate goal.

Life may develop in conditions which differ from those on the Earth. The detection techniques of biomarkers must therefore foresee the search for by-products of metabolism that may be unfamiliar to us. It is important to explore in great detail, both theoretically and in the laboratory, those signs of life that might be relevant in planets of different ages and compositions (§ 5.6).

The search for life will be based on improved radial velocity and astrometric surveys which will characterise the exo-planetary systems most likely to provide the conditions under which life might have developed; on space-based transit surveys which will identify the frequency of occurrence of Earth-like planets in Earth-like orbits; and, further into the future, space-based infrared interferometers capable of measuring the specific biosignatures indicated by these studies.

Chapter 5

How do we fit in?

The Solar System is unique in its role as the vantage point for humankind to explore the Universe. The Sun and the Solar System bodies – planets, satellites, asteroids and comets – play a critical role for unraveling the secrets of stellar physics and of the formation of planetary systems orbiting our own star as well as others. Here, we can study a star and its planetary system in exquisite detail. However, these studies are only a snapshot in its life that stretches billions of years. To understand the physics of the Sun, and the past and the future of our Solar System, we need to compare it with other stars and their planetary systems. The key questions considered by Panel D are:

- *What can the Sun teach us about fundamental astrophysical processes?* Observations of the Sun reveal intricate patterns of magnetic fields and the complex dynamics of a stellar atmosphere at the physically relevant spatial scales. Probing the interior of the Sun with helioseismology is the litmus test for all theoretical models of stellar structure and evolution.
- *What drives Solar variability on all scales?* The Sun varies on a wide range of spatial and temporal scales, displaying important energetic phenomena over the whole range. We do not fully understand and cannot accurately predict basic aspects of Solar variability.
- *What is the impact of Solar activity on life on Earth?* Solar magnetic activity variations induce terrestrial changes which can affect millions of humans on short and long time scales. We need to predict disturbances of the space environment which are induced by the Sun and to understand the links between the Solar output and the Earth's climate.
- *What is the dynamical history of the Solar System?* The formation of planets by accretion within a rotating disc is a common phenomenon. The underlying physical processes, including the source of disc viscosity, the protoplanet-disc interaction, and the time scales on which our Solar System as well as exoplanet systems formed, remain largely a mystery.
- *What can we learn from Solar System exploration about its formation and evolution?* Space exploration has unveiled an amazing diversity of objects in the Solar System, and we expect a similar variety in the nature of exoplanets. This richness needs to be understood.
- *Where should we look for life in the Solar System?* Although we still do not know how life appeared on Earth, it must have benefitted from the presence of liquid water. Searching for liquid water in the atmospheres of the terrestrial planets and the interiors of outer satellites is a major objective for future planetary exploration.

Progress will require a coordinated effort to develop new capabilities in a broad range of disciplines that take a systems-level look at the Sun and its planetary system. This effort will need to combine theory, simulations, observations, laboratory experiments, and *in situ* exploration.

5.1 What can the Sun teach us about astrophysical processes?

5.1.1 Background

The Sun is the closest and therefore by far the best star for testing many current astrophysical paradigms and laboratory physics experiments. Solar research has contributed significantly to a diversity of fundamental physical problems, including atomic physics, nuclear fusion, particle properties and the Solar neutrino flux problem. Current fundamental questions range from the recently discovered atomic polarization through optical pumping to the formation of stellar magnetic fields by dynamo phenomena, and from the finding of a realistic equation of state for the gas in stellar interiors to the prominent role of magnetic fields for the energy balance of stellar coronae. The Sun is of paramount importance for astronomy and for physics and may help solving even the dark matter puzzle: by directly measuring Solar axions or by measuring neutrinos which have been produced through interaction of neutralinos with nuclei in the Solar interior.

Magnetohydrodynamics (MHD) – the science which describes the hydrodynamics of a plasma – is the physical fundament of much of contemporary Solar physics. The equations which govern MHD are difficult to solve, but a comprehensive model of a star requires their consideration everywhere. Relevant scales of MHD processes range from the Solar or stellar radius and the thickness of a convection zone to the pressure scale-height in the atmosphere and the dimensions of boundary layers between domains of differently magnetized plasma. This comprises a spatial range from 700 000 km to just a few km and a temporal range from decades to seconds for the case of the Sun. Turning to other stars, the discs of only a handful of giant stars can be resolved even with the largest telescopes, and imaging complex surface structures with optical interferometry may be forbiddingly difficult for some time to come. Many stellar phenomena like starspots, activity cycles and stellar coronae use the Solar paradigm to describe their likely nature, although they manifest themselves only by a variation of total flux in various spectral regions and time domains. Observing these structures on the Solar surface at high resolution has become possible in the last decade mostly with the commissioning of meter-class ground-based telescopes and the successful development of Solar adaptive optics.

Magnetic fields play a role everywhere there is a plasma, which means practically everywhere in the Universe. These fields are exceedingly difficult to observe from a distance due to their complex configuration, which is probably why they are rarely considered. Understanding Solar and stellar magnetism may help understanding MHD processes elsewhere in the Universe.

5.1.2 Experiments

Magnetic field emerges on the surface of the Sun at a variety of scales: from tiny flux tubes of a size smaller than the spatial resolution of current Solar telescopes to large, complex active regions. Most of the flux is attributed to the production by a global dynamo mechanism, which derives its energy from the Solar rotation and the large-scale velocity field in the Solar convection zone, and which is responsible for the overall Solar magnetic field and its variation through the 22-year Solar activity cycle. Understanding global dynamos is essential for understanding stellar magnetic cycles and the origin of Solar variability. Theories exist which explain a global dynamo based on the differential rotation within the Solar convection zone, but the current knowledge about the internal state of the Sun is too limited to be able to discriminate between competing models and to constrain model parameters in a meaningful way. In particular, it is not clear which dynamo processes are at work and to which extent several processes occur concurrently. Progress in this area will depend on the capability to develop very detailed MHD models of entire stars and to measure global velocity fields with a precision of considerably less than a meter per second throughout the Solar convection zone through local helioseismology.

The Solar surface rotates unevenly, with the rotation rate being much larger at the equator than at the poles. Global helioseismology has established that differential rotation occurs throughout the

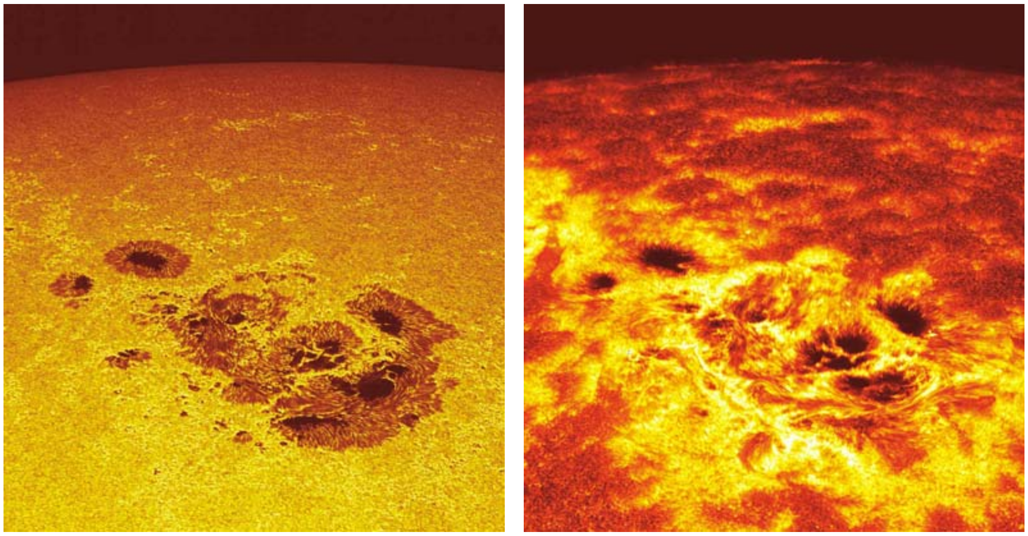

Figure 5.1: A sunspot region observed in November 2003. The left picture shows the lowest visible layer of the solar atmosphere, the photosphere, in the light of CH molecules, and the right picture shows the 1000 km higher layer, the chromosphere, in the Ca II H line of once-ionized calcium. In the photosphere, magnetic fields suppress the convective energy transport from the solar interior, which makes sunspots dark. Solar magnetic fields become the dominant force in the chromosphere, and they become brighter due to magnetic heating processes. Picture courtesy Dutch Open Telescope.

convection zone which comprises the outer 23 per cent of the Solar radius while the Solar interior rotates like a solid body. Differential rotation implies a slow meridional flow field which transports angular momentum from the poles to the equator; such a flow field is just about detectable at the Solar surface. The origin of differential rotation must be related to an interaction between rotation and convection within the Solar convection zone, but is poorly understood. There is evidence that other stars also show differential rotation. How does differential rotation depend on the stellar rotation rate? Which are the rotation laws of rapidly rotating stars? How are stellar rotation and stellar magnetic activity interrelated? Resolving these questions requires, besides detailed hydrodynamical numerical models of entire stars and local helioseismology as above, high spectral resolution surveys of large samples of nearby Solar-type stars.

Some of the Solar magnetic flux may be caused by induction effects which are driven by the small-scale velocity field, predominantly near the Solar surface (‘local dynamo’). Although small-scale magnetic fields have been discovered in the Solar photosphere, a detailed study of their origin and evolution is very difficult and requires much more spatial resolution and light collecting power than presently available. Progress will require ground-based telescopes with apertures of several meters to resolve scales of the order of 10 km in the Solar photosphere at visible wavelengths with sufficiently high sensitivity to do spectropolarimetry with a precision of one part in ten thousand, and assisted by high-order multi-conjugate Solar adaptive optics.

Observing the Sun from the Earth or near-Earth space (including SOHO at the Sun–Earth Lagrange point L1) limits detailed studies to the Solar equator region. The Solar poles are *terra incognita* to a considerable degree. In particular, the magnetic, thermal and velocity structure at the Solar poles is unknown. Little is known about the configuration and evolution of the Solar polar magnetic field and its fine structure, leaving a considerable fraction of the Solar surface and its role in the generation and processing of magnetic field unexplored. The Ulysses mission observed the Sun from a polar orbit and made important discoveries about the Solar wind at a distance from the Sun which corresponds to Jupiter’s orbit. The next opportunity to study the poles in detail will be the Solar Orbiter mission which should include remote sensing instruments and observe the Sun and its surface from an inclined orbit, much closer-in than Ulysses.

5.2 What drives Solar variability on all scales?

5.2.1 Background

When looked at from a distance, the Sun is a remarkably stable star. Its luminosity varies over time scales of weeks by less than one part in a thousand, and over decades by less than one part in ten thousand, presumably one of the conditions for a stable climate on Earth. The picture changes when one takes a close-up view. The Solar output is modulated by the magnetic field at the surface, giving rise to variations of the Solar radiation, most prominently in the UV and X-ray spectral regime, over time scales ranging from seconds to decades, and even centuries.

The overarching process causing Solar variability is the interaction of the Solar magnetic field with the Solar atmosphere. Magnetic fields manifest themselves in the lowest layer most prominently in the form of active regions, concentrated in dark sunspot groups at latitudes which vary with the activity cycle. Concentrated magnetic fields have profound effects on convection, causing a variety of small-scale structures in sunspots and their vicinity, some of which challenge the spatial resolution of today's largest Solar telescopes. Active regions are responsible for the variability of the Sun at visible wavelengths for time scales of weeks, manifested by a decrease of the Solar luminosity when a sunspot group transverses the visible disc. However, there is a positive correlation of Solar luminosity and activity, implying that the Sun is on the average brighter during Solar activity maximum, which is counterintuitive. The source for excess brightness is not fully established, but appears to be connected with an increase of small-scale magnetic structure ('plages'), also during activity maximum. Given sufficient polarimetric sensitivity, small-scale magnetic field can be detected everywhere in the photosphere, independent of the activity cycle. The nature of this distributed field and its role for the Solar dynamo are not clear.

The extended Solar atmosphere is a physical system of extremes, covering a range of ten orders of magnitude in particle density from the photosphere to the outskirts of the heliosphere, and including the coolest parts of the Sun – the temperature minimum between photosphere and chromosphere at a few thousand Kelvin – and the hottest – the corona which can reach a hundred million Kelvin. The magnetic field connects the entire atmosphere and drives many of its physical phenomena, causing a highly structured and dynamic system. Because of the high temperatures and the very low density of the plasma, the Solar atmosphere changes constantly and rapidly when observed in the extreme ultra-violet, showing a rich spectrum of variability phenomena. The variability of the Solar output is most extreme at short wavelengths. Energetic events such as flares increase the Solar X-ray flux by orders of magnitude over time scales of minutes.

While the lower layers, photosphere and chromosphere, can be observed with optical telescopes at visible and infrared wavelengths from the ground, a full characterisation of the chromospheric and coronal plasma is only possible with ultra-violet and X-ray telescopes from space. Transient events in the upper atmosphere are observed from space and with radio telescopes. *In situ* measurements of the composition and magnetic field of the Solar wind and its relation to transient events near the Solar surface require instruments on board of spacecraft.

5.2.2 Experiments

Magnetoconvection is the interaction between the convective velocity field and the magnetic field in a plasma. It is strongest in the Solar photosphere because magnetic, thermal and kinetic energy density are similar, and determines in these layers the appearance of sunspots and pores. The three-dimensional structure of sunspots is to a large degree unknown, their detailed thermal structure remains to be explained. What is the nature of magnetoconvection and how does it structure the Solar photosphere? Why are umbrae so dark and penumbrae so bright? What determines the energy transport and the evolution time scales in sunspots?

Together with the high conductivity of the plasma, magnetoconvection results in a highly inter-

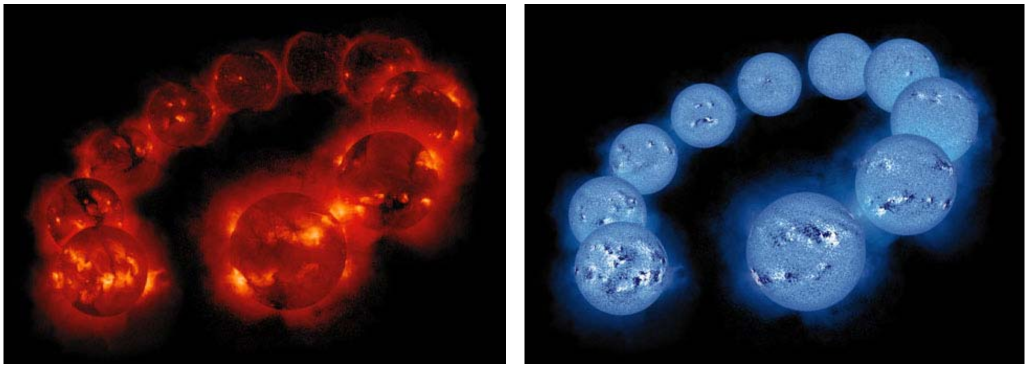

Figure 5.2: Sunspots and the corresponding magnetic fields follow an 11-year cycle, the sunspot cycle. The composite image on the left shows ten magnetic maps of the Sun approximately one year apart, from one maximum to the next. As the cycle fades, the large regions disappear. As the next cycle of activity picks up, the magnetic polarities of the sunspot regions are reversed. Wherever strong magnetic fields penetrate the solar surface, the outer atmosphere is heated to several million degrees. The composite image on the right shows the X-ray emission from this hot corona. As the cycle fades out, the X-ray emission becomes weaker and more diffuse, to brighten again as the next cycle starts. Picture courtesy Lockheed Martin Solar and Astrophysics Laboratory, original images from Kitt Peak Vacuum Telescope of the National Solar Observatory (magnetic maps) and the Japanese-American YOHKOH satellite (X-ray emission).

mittent distribution of magnetic field, which has the tendecy to concentrate at the loci of convective downflows towards the Solar interior. A continuous change of magnetic topology, driven by the convective velocity field, is transferred to higher atmospheric layers where magnetic forces dominate. A transfer of mechanical energy from deeper layers into the upper atmosphere is connected with the topology change, which is the likely source of energy for coronal heating and energetic events like flares. How is energy transferred from the photosphere to the upper layers? Which is the main non-radiative mechanism for such a transport, which should be responsible for the energy transport from the photosphere to the chromosphere and corona? Several types of waves have been proposed that are thought to travel through magnetic flux tubes. However, the exact characteristics of such waves, the mechanisms by which their energy is transferred to the magnetic field, and subsequently released in the corona, are largely unknown.

Understanding magnetoconvection at the photospheric level is a prerequisite for understanding Solar variability of the upper atmosphere. Such an understanding requires large-scale numerical MHD models of stellar atmospheres which are capable of resolving the spatial scales of magnetic concentrations within an adequate volume, in conjunction with spectro-polarimetric observing capability at the highest spatial and temporal resolution and sensitivity. Such observing capability implies ground-based telescopes with apertures of several meters to achieve the required spatial resolution and light collecting power as well as adaptive optics. In order to investigate the three-dimensional structure of active regions, the tools of local helioseismology need to be improved and access to helioseismology observatories in space and on the ground must be ensured.

How much magnetic flux is present on the Sun and how is it distributed? How does the distribution change on short and long time scales? How does magnetic flux emerge at the Solar surface and how is it removed? Although synoptic observations of the Solar magnetic field have been carried out for decades, only the newest generation of detectors and computer facilities make it possible to measure weak, distributed flux with sufficient sensitivity. This is an important task for a network of medium-sized synoptic telescopes which continuously observe the Sun and provide vector magnetograms with a resolution of better than an arcsecond and with high cadence.

What is the structure of magnetic fields in the chromosphere and the corona? How does it evolve with time? So far, only photospheric magnetic fields are accessible to accurate measurements. Spectropolarimetric diagnostics are mostly based on visible and infrared lines which to a large part originate in the photosphere. Very recently, Hanle-effect based diagnostics (which unveil a genuine ‘second Solar spectrum’ in linear polarization) as well as several infrared spectral lines formed above the photosphere start providing quantitative information on chromospheric magnetic fields. By far, the most promising magnetic diagnostic in the transition region and the corona will result from the Hanle effect measurements in permitted ultra-violet lines like the Lyman series of hydrogen, or ionized OVI, NV, and CIV lines that are known to be excited by radiation from the Solar disc and, therefore, should be polarized by scattering.

Coronal magnetic fields dominate the dynamics, determine the structure and deliver the energy for flares and coronal mass ejections. The processes which convert the energy stored in magnetic fields nearly instantly into heat, kinetic energy of particles, and a flash of radiation covering almost the entire electromagnetic spectrum, are poorly understood. The driving mechanisms of coronal mass ejections are similarly unclear. The main obstacle is the poorly-known topology of coronal magnetic fields. New approaches for its quantitative measurement require ultra-violet polarimetry from space, as well as ground-based infrared observations and radio interferometry.

What are the Solar oxygen and other metal abundances? Chemical abundances measured for the Sun are often used as a reference for other astronomical objects. However, the accuracy of Solar determinations is not put into debate whilst many other unknowns might be jeopardizing the measurements. Sophisticated inversion codes of the radiative transfer equation, taking into account the full physics of line formation, are likely the means to reach the goal.

Recent observations of the corona have revealed an elemental abundance enhancement of species with a low first ionization potential. This effect has been detected on some stars and not on others but its origin is unknown. It may be providing important clues for the magnetic coupling of the different layers of the Solar atmosphere.

5.3 What is the impact of Solar activity on life on Earth?

5.3.1 Background

The term ‘space weather’ encompasses all conditions on the Sun, in the Solar wind and the Earth’s magnetosphere which can influence infrastructure in space and on the ground, as well as human life, and has given rise to an emerging field of space research. The goal is to understand physical mechanisms which originate on the Sun and which infuence the near-Earth space environment, to predict – or at least to recognize – potentially hazardous situations timely enough to enable protective measures. SOHO enabled much progress in this direction because of its continuous surveillance of the Sun, and much has been learned about recognizing Earth-directed coronal mass ejections and determining their travel time. The main goal of research in this area is the ability to predict energetic events on the Sun earlier to allow for more time to react. In particular, satellites and astronauts are put at risk by those events.

Solar variability has had a noticeable influence on the Earth’s climate over time scales of decades. Potential mechanisms involve the variation of Solar luminosity, the strong variability of the Solar ultra-violet radiation which affects the upper layers of the atmosphere, and a cycle-dependent variation of the Solar wind which modulates the galactic cosmic radiation reaching the Earth. None of these processes is well understood, mostly because direct measurement of relevant quantities has been possible only for a short time. A substantial effort is therefore invested in deriving proxies for the climate and the Solar output from historic records and geological data.

Solar variability is caused by its magnetic activity, from short term variations over the 11 year sunspot cycle to long-period variations which modulate the activity cycle maxima, and may result in extended minima of activity like the Maunder minimum in the 17th century when sunspots were absent for some 60 years. The Maunder minimum coincides with the ‘small ice-age’ when extremely cold winters occurred in Europe. A detailed understanding of the Solar dynamo is clearly needed in order to predict Solar activity for longer periods.

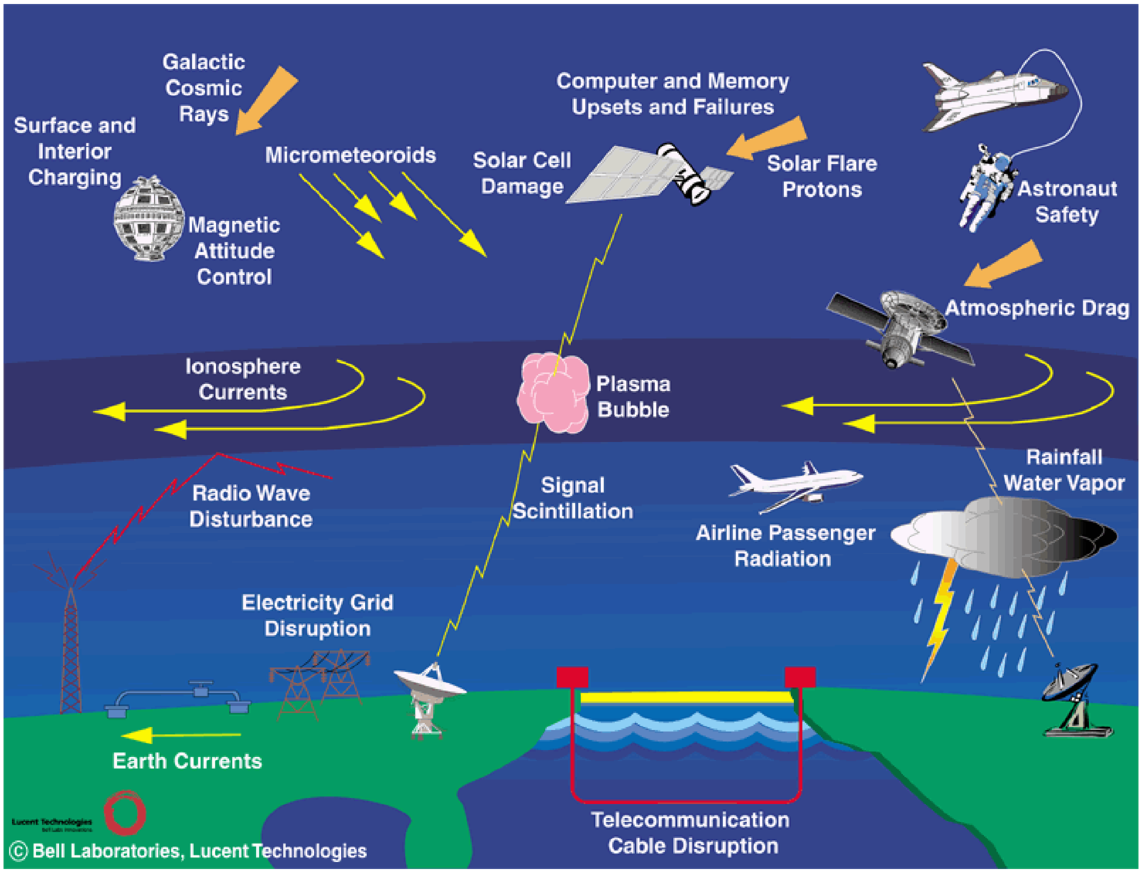

Figure 5.3: Solar eruptions such as flares and coronal mass ejections, triggered by changes in the magnetic field in the solar atmosphere, accelerate charged particles and magnetic fields towards Earth, creating space weather. This picture summarizes the various hazards created by space weather events. Picture courtesy Lou Lanzerotti, Bell Laboratories.

5.3.2 Experiments

In order to protect mankind's investments and activities in space, space weather prediction needs to be developed and improved. An essential ingredient will be better prediction of Solar activity at all relevant temporal scales, which requires improved understanding of the underlying physical processes which are the subject of the previous sections. Specific to space weather forecast is an improved prediction of the configuration of the Solar magnetic field and its changes, the early recognition of configurations which could lead to the emergence of an energetic event or an Earth-directed coronal mass ejection, and a reliable prediction of their magnitudes.

Forecasts need a dense network of observing stations to be effective. Continuous monitoring of full disc photospheric velocity and vector magnetic fields will provide important boundary conditions, as will proxies for the chromospheric and coronal magnetic field. Suitably equipped space missions like the Solar Dynamics Observatory as well as ground-based synoptic facilities like SOLIS are likely to provide most of the synoptic data. Detailed high-resolution investigations with the recently launched Japanese Hinode (Solar B) mission and large aperture ground-based facilities will lay the groundwork for a better understanding of the emergence of energetic events.

A serious limitation is the fact that only one Solar hemisphere can be observed directly from near-Earth. The magnetic field on the far side is currently not accessible at all. Furthermore, any coronal activity directed towards the Earth appears in projection against the Solar disc and is difficult to detect and to characterise. This situation will improve once the two STEREO spacecraft, which will observe the Sun from positions on the Earth's orbit leading and trailing the Earth, are launched. They provide a stereographic view of the Solar environment pointing at the Earth, permitting a much better characterisation of earthward moving ejecta. Another emerging approach is farside imaging of the Solar surface using helioseismology. This method has passed proof-of-principle tests and has great potential for space weather forecasting.

5.4 What is the dynamical history of the Solar system?

5.4.1 Background

Planets form inside discs of dust and gas orbiting newly-born stars. This paradigm was developed from centuries of studies of the Solar System. With the discovery of the first extra-solar planet (51 Peg *b*) in 1995, and the additional 200 extra-solar giant planets discovered since, our ideas about planetary systems have changed drastically. Even if our current instruments still do not allow us to detect planets as small as the Earth, we know now that planetary systems are much more diverse than originally predicted by theory (§ 4.6). The lack of a good theoretical understanding of this diversity creates two major problems. First, we cannot explain what is observed and second, we cannot predict what should be observed. The formation and evolution of planets is a difficult subject involving large changes of scales, long time-scales and many non-linearities and feed-back mechanisms. This multi-disciplinary topic involves astronomy, cosmo-chemistry, material sciences, planetary sciences, climate physics, and exo-biology.

While it is generally accepted that the terrestrial planets (and possibly the cores of giant planets) formed from the collisional accumulation of planetesimals, the formation of these planetesimals is being debated extensively. The dynamics and early growth of bodies too small for self-gravity to play a significant role is determined by gas drag. Sticking between small ($\approx$ 1- 100 micrometer) dust grains over a relatively narrow range of temperatures and impact velocities has been demonstrated in the laboratory. Beyond this size, and until gravity dominates, the sticking mechanisms are much less well determined. Suggestions for growth involve mechanical properties (porosity), trapping in vortices or gravitational instability in spiral arms among others.

Studying the collisional growth of a swarm of billions of solid km-size bodies embedded in a differentially rotating gaseous disc for millions of years has turned out to be a challenging problem. Because of the large changes of scales, the non-linearities, and the long time scales involved, it has so far remained beyond the capabilities of even the fastest direct integration methods. Thus, statistical methods have been used in which the time evolution of a binned mass distribution and mean orbital parameters are calculated.

Two scenarios for giant planet formation are currently available: the direct-collapse model and the core-accretion model. Unfortunately, the challenging multi-dimensional simulations involving complex physics required to model the disc instability leading to collapse still prevent quantitative comparisons to observations. This is not the case for the core-accretion scenario. In this model, a solid core is formed first by the accretion of planetesimals in a similar way as described above. As the core grows, a gaseous envelope is accreted slowly at first but in a runaway fashion as the core reaches a critical mass leading to a rapid build up of a massive gaseous planet (§ 4.5).

The core-accretion scenario seems to be favoured on the basis of the enrichment in heavy elements observed in the giant planets (§ 5.5), but it is not without problems. A major difficulty is related to the time scale required to form a giant planet. Based on astronomical observations, protoplanetary discs are believed to have a lifetime of up to 10 million years. This lifetime is of the same order, if not smaller, than the giant planet formation time scale in the standard core accretion scenario. More recent models based on an extended version of the core-accretion scenario including disc evolution and planetary migration have shown once these effects are taken into account, planets can form through core accretion well within disc lifetimes. It remains to be seen if these models can explain in a statistical sense all the properties of the currently known extra-solar planets.

5.4.2 Key questions and opportunities

Major unsolved questions in the formation of the Solar System include: What is the formation mechanism of the first km-sized planetesimals? Do they form through collisional processes or through gravitational instabilities? How does the Solar nebula evolve? Are there vortices and spiral waves? What are the interactions of planets and discs? In which case can they lead to migration of planets? Can they account for large orbital eccentricities? What are the mechanisms of core formation? What is the possible role of magnetic fields? What are the effects of giant impacts? In particular, can they account for the origin of the Moon and Mercury?

It is important to emphasise the role of theory. For example, exploring the effect of a giant impact on the origin and evolution of the Moon and Mercury will require further theoretical work. In the case of the Moon, it is not clear whether the giant impact can explain its detailed elemental and isotopic chemistry (for example the oxygen isotopes). In the case of Mercury, one needs to estimate the chemical and isotopic fractionation originating from a giant impact blowing away most of the mantle. This requires the coupling of a hydrodynamics code with a chemical network in order to compute the expected signature, with extremely high numerical resolution and a suitable thermodynamical description including non-equilibrium chemistry. The theoretical predictions will be compared to data obtained by the future missions Messenger and BepiColombo.

5.5 What can we learn from Solar System exploration?

5.5.1 Background

Terrestrial planets. Mercury, Venus, Earth, and Mars consist mostly of iron-rich cores and silicate mantles. The outermost layer of these rocky planets, the crust, consists of the crystallized low-melting component of the mantle silicates and is a product of mantle melting. Although largely similar in composition and size, these bodies form vastly differing worlds: from the sunbathed stony Mercury and the hot and acid atmosphere of Venus, to Earth with its life and plate tectonics and the cold desert of Mars that may have had a more habitable past. The Moon, although not a planet in its own right, is best considered together with these bodies and offers another extreme: it is most likely the product of a giant impact on the Earth and is thought to stabilize habitable conditions on Earth by stabilizing the rotation axis through tidal interaction.

Earth has a self-sustained magnetic field generated by a geodynamo in the iron-rich core. The Mariner 10 fly-by data indicate that Mercury also has a self-sustained magnetic field but the data are incomplete. A magnetic field has not been detected at Venus while Mars and the Moon have a fossil magnetic field in the crust. For Mars this is taken as evidence for an early magnetic field existing for about 500 million years up to 4 billion years before the present, while for the Moon other explanations are possible, including magnetization by plasma clouds generated by major impacts.

The tectonic styles of the Earth and the other terrestrial planets differ substantially. Earth has plate-tectonics; its oceanic crust is only a few million years old, while its continental crust is two billion years old, on average. The crusts of Mars, Mercury and the Moon are older, 3.5 billion years or more. From the relative scarcity of craters on Venus' surface, it is speculated that the Venusian crust may be much younger.

The terrestial planets appear to be differentiated, with the heaviest components in the center and a gaseous shroud, the atmosphere, forming the outermost layer. The study of short-lived isotopes suggests that this differentiation occurred during or shortly after the formation of the planet. The Earth and Mars show clear evidence for a central core and the high density of Mercury is difficult to explain without a large iron core. The similarity in size between Venus and the Earth and Venus' apparently young surface also suggest that both may be differentiated in the same way, although the dichotomy between Ganymede and Callisto may provide a caveat.

Figure 5.4: The caldera of Olympus Mons, the highest volcano on Mars, observed from a distance of a few hundred km with the high-resolution stereo camera (HRSC) of Mars Express. The HRSC images are being used to obtain a precise age of the different types of surfaces, by measuring the density of their impact craters. This analysis has revealed that many terrains, including the summit of Olympus Mons, are significantly younger than previously thought, with ages of only a few million years.

The Earth's core is liquid with a solid inner region. It is not known whether the cores of the other planets are liquid, have solid inner cores or are completely frozen, but it is widely held that the cores of Venus and Mars are liquid, while Mercury has a solid inner core. This conjecture results from theoretical considerations and provides a means of understanding why Mercury and the Earth have magnetic fields while Mars and Venus have not. Fractional crystallization of the core releases buoyancy that may efficiently drive a present-day dynamo. Calculations suggest that thermal buoyancy in the core decreases rapidly during the first few hundred million years. Thermal buoyancy may therefore explain early magnetic fields and fossil magnetic fields in the crust.

Mercury is too small and too close to the Sun to retain even the heaviest gases. The permanent atmospheres of Venus, the Earth and Mars are characterized by very different physical conditions, with surface pressures ranging from almost 100 bars (Venus) to less than 0.01 bar (Mars), and surface temperatures ranging from 730 K on Venus down to 150 K in some places on Mars. In contrast, their primitive atmospheric composition, dominated by CO_2, H_2O and N_2, shows more similarity, if one considers that the primitive atmosphere of the Earth was probably like those of Venus and Mars; oxygen appeared on Earth as a consequence of the development of life. Apart from oxygen, the main difference between the three planets is the evolution of water. On Venus, closer to the Sun, water was gaseous and probably disappeared through photodissociation and escape; the large amounts of gaseous CO_2 led to a strong greenhouse effect, responsible for the high surface temperature observed today. On Earth, the temperature was such that water could stay liquid, allowing gaseous CO_2 to be trapped in oceans, so that the greenhouse effect remained moderate. On Mars, water is now trapped under the surface as ice and / or permafrost.

Venus and Mars have been studied extensively by *in situ* and orbiting space missions, as well as by ground-based observations from Earth. Venus is covered with a very thick cloud deck, mostly composed of H_2SO_4 particles. Mars, in contrast, has very few H_2O and CO_2 clouds, but exhibits two polar caps of CO_2 and H_2O, with strong seasonal exchanges of the CO_2, H_2O and dust. Both planets, as well as the Earth, exhibit intense dynamical activity, including a Hadley circulation.

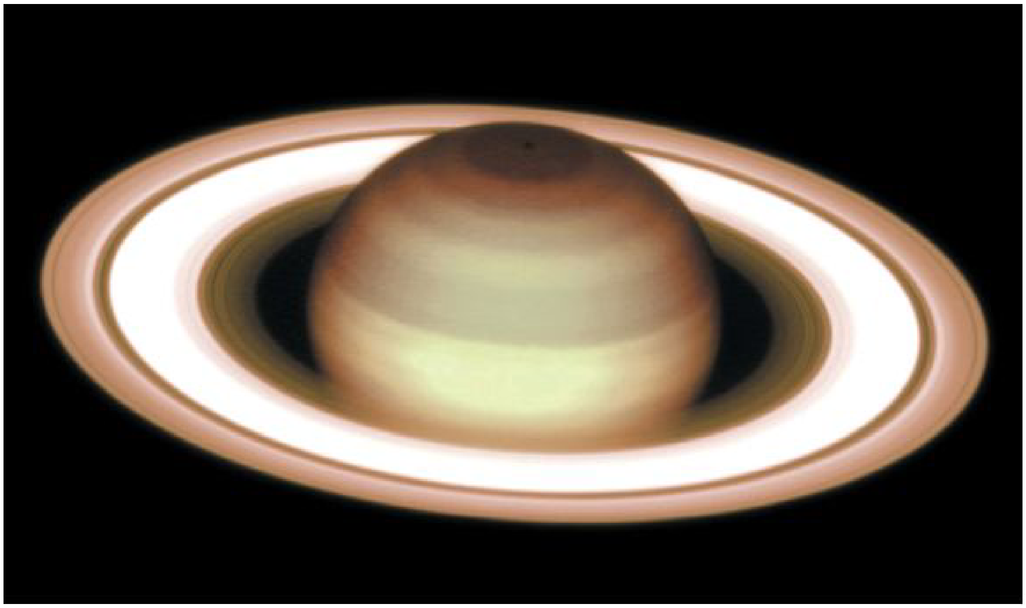

Figure 5.5: With the use of adaptive optics, ground-based diffraction-limited images can now be obtained in the near-infrared range. This image of Saturn was obtained at the VLT using the NAOS-CONICA system. The angular size of the Saturn disc is about 15 arcsec. The image quality is sufficient for the ring systems to be separated. Bands and zones are also clearly resolved on Saturn's disc.

Giant planets and their satellites. The chemical composition of giant planets, dominated by hydrogen, helium and minor species in reduced form, has been determined from space and ground-based spectroscopy. In addition, in-situ measurements of the atmospheric composition of Jupiter were obtained by the Galileo probe. All data show an enrichment in heavy elements relative to hydrogen as compared to the Solar values. This enrichment increases from Jupiter to Uranus and Neptune and provides strong support for the nucleation model according to which giant planets formed from an icy core and the subsequent collapse of the surrounding protosolar nebula (§ 5.4.1). The internal planetary structure is much less understood. The gravitational moments have been determined from spacecraft flybys and orbiters, but the equations of state are poorly known. As a result of their fast rotation, giant planets exhibit strong dynamical systems, including convective Hadley-type belt-zone systems and more localized phenomena, such as the Great Red Spot on Jupiter and more variable spots on all planets. In contrast with the gaseous phase, the nature of aerosols is still poorly known. The main cloud structure is derived from thermochemical modelling (with, in particular, NH_3, NH_4SH and H_2O for Jupiter and Saturn), however the exact nature of the condensibles in Jupiter's Great Red Spot, for instance, is still unknown.

The giant planets have numerous 'regular' satellites, mostly on circular and concentric orbits close to the planet's equatorial plane. Other satellites, with high obliquities and inclinations, have been captured by the giant planets' gravity field. This general property is a natural consequence of the formation scenario (§ 5.4.1): regular satellites formed within the equatorial disc resulting from the collapse of the surrounding subnebula around the initial planetary icy core. There is a spectacular diversity among the outer satellites, from Jupiter to Neptune, but also, within each system, as a function of the planet's distance. The Galilean satellites display a clear density gradient, with Io's surface exhibiting active volcanism and the three others being mostly covered with water ice. Due to tidal forces associated to resonances with Io and Ganymede, Europe is believed to host an internal water ocean, which has important potential implications for astrobiology (§ 5.6). Titan, Saturn's largest satellite, is unique with its thick nitrogen-dominated atmosphere where prebiotic chemistry might be at work. Enceladus, a smaller icy satellite of Saturn, has recently shown evidence for active cryovolcanism, as detected by the Cassini spacecraft. Cryovolcanism was also revealed earlier on Triton at the time of the Voyager 2 flyby.

Figure 5.6: The nucleus of comet Halley from a distance of about 500 km. The picture was taken on March 13, 1986, with the Halley Multi-Colour Camera on board the ESA satellite Giotto, and it was the first detailed image of a cometary nucleus. The image distinctly shows an irregular and inhomogeneous shape, very different from previous expectations, with bright jets (associated with water outgassing) and black regions of very low albedo. In addition, the other instruments aboard Giotto made many unexpected discoveries, including evidence for carbonaceous compounds, and the detection of several minor parent molecules.

Small bodies and extraterrestrial matter. Our understanding of cometary nuclei has advanced enormously during the 20 years since the exploration of Halley's comet. At that time even the verification that the solid nucleus did exist was considered a major advance. Nowadays we have reached a fairly good picture of the relative abundances of the molecules that are outgassed from the nuclei of active comets near the Sun, including the typical ranges of variation between different comets. Comparison with the abundances observed in molecular clouds and star-forming regions are being pursued. This also holds for cometary dust, where the occurrence of large quantities of crystalline silicates has been established via infrared spectrometry and by the Stardust sample return. The compositional dichotomy between the retention of extremely volatile species in the ice phase and very high temperature condensates in the dust awaits a proper explanation.

Advances in the modelling of outgassing and interpretation of nongravitational orbital effects have established the highly porous nature of cometary nuclei with mean densities ~ 0.5 g/cm^3. An extremely low material strength became apparent with the studies of the splitting of comet Shoemaker–Levy 9 and was confirmed by the Deep Impact experiment. Microporosity appears to be more important for cometary nuclei than macroscopic rubble-pile structure. Porosity and gas diffusion are hence important concepts of recent, global thermal models of comets, and the structure of the ice is often assumed to be amorphous. This assumption is supported by both laboratory work and the observed outgassing of comets, but remains to be proven.

The exploration of the trans-Neptunian population of objects is in a fascinating stage, with surprising discoveries following each other, and theoretical studies making rapid progress. This population – unknown with the sole exception of Pluto as recently as 15 years ago – is known to be made up of a 'classical' Kuiper Belt with dynamically hot and cold components, resonant groups, a scattered disc that extends to large distances and forms a bridge to the Oort Cloud, and a component with perihelia detached from the planetary system. The latter may provide evidence for the formation of the Solar System within a tight stellar group that was soon dissolved. The scattered disc is partly fed by resonant transfer from the Kuiper Belt and partly a fossil remnant of a planetesimal disc associated with the formation of the giant planets and the migration of Neptune. It appears to be the main source for captures into the Centaur and short-period comet populations, although some contribution by captures from the Oort Cloud is not ruled out. The size distribution, the collisional evolution, compositional trends and spectral diversity all have a bearing on the formation of this major structure and its relation to the extended dust discs around other stars.

The structure of the asteroid main belt including gaps at mean motion and secular resonances is well understood in terms of orbital dynamics. The statistics of spin rates verified the expectation that all but the smallest objects should be gravitational rubble-piles resulting from major collisions that led to fragmentation without complete dispersal. The recent Hayabusa mission to the 0.5-km asteroid Itokawa indeed indicated such a structure even for this small body. The realization of the importance of the Yarkovsky effect in main-belt asteroid dynamics has been a major step forward. It explains the delayed delivery of objects into the resonances and the existence of large and old near-Earth asteroids such as Eros.

Although Europe has no large-scale programme for the continued detection of asteroids in general and near-Earth objects in particular, Gaia will provide taxonomic classification and highly accurate orbits for many tens of thousands of the larger objects, in principle permitting potential impact orbits to be calculated even decades in advance. Through accurate proper motion measurements, Gaia will also track nearby stars which have passed relatively close to the Sun over the past millions of years, events suspected in part to be responsible for enhanced cometary activity due to perturbations of the Oort Cloud. These events may have been correlated with past mass extinctions, and similar events in the future could in principle now be predicted.

Finally, laboratory analysis of extraterrestrial matter provides a unique way to address fundamental questions about the origin of Solar System matter, of planets, and of life in the Solar System. Dating of lunar rocks could only be achieved by laboratory measurements, and allowed us to understand the timing and processing of planetary building by accretion. Analysis of the trace element composition of the lunar crust led to the identification of major differentiation processes in terrestrial planets, such as magma ocean stages and large-scale volcanism. The extremely precise analysis of the oxygen isotope composition of lunar rocks established the genetic relationship between Earth and Moon due to a giant impact. The isotopic composition of light elements in the Sun could be finally measured with precision in lunar soils, leading to tremendous advances in our understanding of how the Sun changed through time. So far, 380 kg of lunar soil and rocks, the Solar wind sampled during 27 months by the Genesis spacecraft, and tiny grains from comet P / Wild 2 have been returned to Earth. Since the Apollo missions, scientists have increased their ability to analyse tiny samples. These advances have revealed the nature and composition of cometary grains, and the isotopic composition of some key elements of the protosolar nebula.

5.5.2 Key questions and opportunities

Terrestrial planets. There are many unsolved questions related to the structure, composition, and atmospheres of the terrestrial planets: What is their interior structure, and how big are their cores? How do the geodynamo and the planetary heat engine work? What is the chemical composition and how does it relate to the Solar composition? Are there recycling mechanisms between the atmosphere, the surface and the interior? Is there active volcanism on Venus? What is the history of the atmospheres and of the water inventory? Did liquid water stay on the surface of Mars in its past history, and for how long? Could life have appeared and developed on Mars at that time? If so, could we hope to find fossil traces of it, and where should we look for it (§ 5.6).

These questions can be answered through combinations of theoretical work, laboratory experiments and in-situ exploration. Until recently, planetary space exploration was mostly carried out by NASA with an important contribution from the (former) Soviet Union. Over the past few years, Europe contributed significantly to the exploration of the Solar System with the Mars Express, Venus Express and Smart 1 missions. Future missions include ExoMars and Bepi Colombo.

The next steps of Mars exploration should include networks of small stations with geophysical

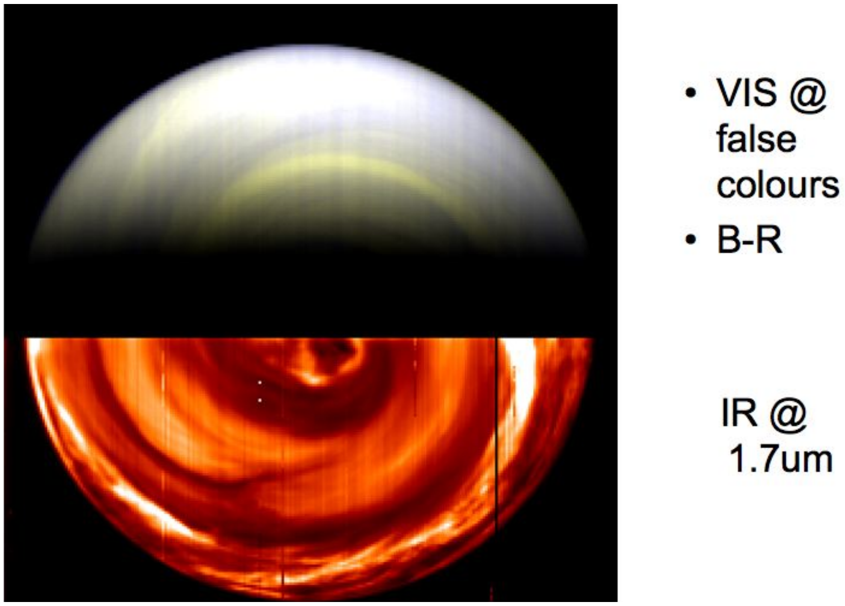

Figure 5.7: The south pole of Venus taken by the VIRTIS instrument on board Venus Express in June 2006 one day after the spacecraft entered the Venus Insertion Orbit. The upper image (in false colours) is taken in visible light, while the lower image is taken in the near-infrared range, at 1.7 micron. At this wavelength, the cloud deck of Venus is relatively transparent, which allows the thermal flux from the lower atmosphere to be detectable. The nature of the dynamical motions and the origin of the polar vortex are still a puzzle.

and meteorology packages. The networks should be followed by sample return missions. These missions will allow calibration of crater chronologies and provide geochemical and isotope data, as well as atmospheric compositional data. The isotope data will allow the dating of major events in the history of the planet. It is also important to study the evolution of the atmosphere and its relation to the magnetic field history. The search for life on Mars is the ultimate driver.

Giant planets and their satellites. A major question regarding the giant planets is: when, where and how did they form, and what was the nature of their planetesimals? Indeed, Galileo measurements of elemental abundances in Jupiter (in particular nitrogen and argon) suggest that Jupiter's planetesimals must have been formed at very low temperature (i.e., below 40 K). This is a very stringent constraint for formation models (§ 5.4). Another important question concerns the differences among giant planets, and especially between the two icy giants: Neptune has a strong internal energy source and is subject to strong dynamical motions, while Uranus has no internal energy source and is much more sluggish. Are these differences due to different formation conditions, or to different evolution processes? Other questions deal with the nature of the giant planets' complex dynamical systems, and with the chemical nature of their condensates.

Regarding the outer satellites, there is the long-standing question about the origin of Uranus' high obliquity. How fast was this process? If the obliquity acquisition was slow, the satellite system may have followed the equatorial plane in its evolution. If, in contrast, the process was fast (as a result of a collision for instance), then the event must have occurred very early, before the formation of the satellite system. There are also many open questions concerning the physical properties of the individual outer satellites. How deep is the ocean on Europa? How is methane replenished on Titan? Are lakes present on the surface, and what is the nature of cryovolcanism? Is there liquid water below the surface of Enceladus, and, if so, at which depth? More generally, which outer satellites could host liquid water below their surface, and exhibit cryovolcanism?

The *in situ* exploration initiated by NASA with Pioneer 10, Voyager 1 and 2 and Galileo, and later by NASA and ESA with Cassini-Huygens, should be followed by space missions towards Jupiter and its system (in particular Europa), Saturn (including a probe) and the Saturn system (including

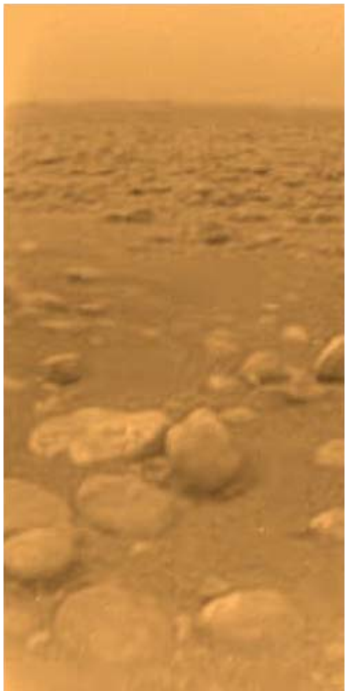

Figure 5.8: This image became famous all over the world when the Huygens probe landed on Titan's surface on January 14, 2005. Conceived and developed by the European Space Agency as part of the Cassini mission, operated by NASA and ESA, the Huygens probe achieved the first in-situ analysis of Titan's atmosphere and surface. The Huygens imaging spectrometer revealed a flat surface, probably made of hydrocarbon solid material, partly covered with boulders which are likely to be water ice. Liquid hydrocarbons could be present in the form of lakes, but this remains to be confirmed.

in particular Titan and Enceladus). Further in the future, the exploration of Uranus and Neptune by atmospheric probes will provide key diagnostics on their formation processes. Astronomers will also take advantage of large ground-based facilities (VLT, ALMA, and in the future, an ELT).

Small bodies and extraterrestrial matter. The number of km-size near-Earth asteroids is believed to be known to about 10 per cent, based on knowledge of their dynamical origin and the biases involved in the search programs. However, the occurrence of binaries poses some problems due to the expected break-up of such binaries at close encounters with terrestrial planets, and the estimated infeed rate from the source regions may be affected, and needs to be established.

The estimated rate of impacts with the Earth by $\sim$100 m objects is thought to be significantly smaller than one would guess based on the occurrence of the 1908 Tunguska event, or the upcoming very close encounter with asteroid Apophis, but it is not known whether the population size has been underestimated. Another uncertainty concerns the cometary fraction of impacts on the Earth, although it is commonly agreed that this is minor compared to the asteroids, except for the largest and rarest impacts. Major advances have been made in the ability to predict the circumstances of future returns of observed asteroids and to estimate the risk of an impact in cases where this is non-zero. A lack of proper procedures for communicating such predictions to the public and media, as well as providing advice to the responsible governmental or international agencies has been identified as a problem, which is likely to grow as future, deep searches will reveal an increasing number of small objects on a possible (though unlikely) collision course.

Analysis of extraterrestrial samples from planets and small bodies would allow us to address the following fundamental questions: What is the origin of organics? What is the origin and processing of cometary constituents? What is the relationship between asteroids and meteorites? What is the origin of giant planets and their satellites? What is the early and long-term evolution of terrestrial planets? As a first step, the analysis of samples from Mars would provide key information about the planet's formation and evolution processes.

Many of the rapid advances in our understanding of the small bodies of the Solar System came from a truly international network of collaborations. This holds both for work on the theoretical side and for the very successful space missions, where for instance the ESA Rosetta / OSIRIS observations of the NASA Deep Impact event on comet Tempel 1 in 2005 is a beautiful example of the synergies that come from such collaborations. The continued exploration that is foreseen for the coming decade should not be much different, since, e.g., the Rosetta project will both serve and involve the whole international community of researchers. The use of ALMA to study cometary outgassing in more detail than ever before will likely involve international teams as well. It would be of great importance for space exploration, if the ESA science programme could indeed move forward with smaller and more precisely defined missions as necessary supplement to the large and comprehensive cornerstone missions like Rosetta. Finally, regarding in particular the near-Earth objects, even though most new discoveries will undoubtedly be made by the large US search programs, an international program will be required to secure follow-up observations and analyze the orbits for a proper reaction to possible impact threats. Europe has a key role within an international framework and must maintain and strengthen the expertise already developed.

5.6 Where should we look for life in the Solar System?

5.6.1 Background

Ground-based (sub)millimetre observations have revealed that the interstellar medium is rich in complex prebiotic molecules, with as many as 13 atoms, most of them being hydrocarbons and nitriles. About twenty of these molecules have been detected in comets as well. The atmosphere of Titan, dominated by N_2 and CH_4, is another medium where prebiotic chemistry is at work. Some of the nitriles found in Titan are also found in laboratory experiments of prebiotic synthesis. Even more complex prebiotic molecules, including about 20 amino-acids, have been found in a primitive meteorite, Murchinson, with relative abundances comparable to those found in laboratory simulation experiments in which, following the pioneering work of Miller and Urey in 1953, amino-acids are synthesized from a reducing mixture of organic ices, irradiated by a high-energy source. The presence of amino acids in some meteorites demonstrates that a complex prebiotic chemistry was taking place already in the early stages of Solar System history.

Although we do not know yet how life appeared on Earth, liquid water was most likely essential for its appearance and development because it provides a medium in which molecules can dissolve and react together. Water is relatively abundant in the Universe, as it is formed from two atoms with large cosmic abundances. It is found in liquid form over a large range of temperature. Water is also the only molecule for which the solid state is lighter than the liquid phase, thus allowing conservation of life in the deep oceans if a cooling episode occurs in a planet's history.

The phase diagram of water indicates that there are two main areas where we can look for water: in the atmospheres of Earth-like planets, at temperatures ranging between 270 and 370 K, and in the interiors of outer satellites, at high pressures and temperatures. Venus is too hot to retain liquid water as a result of its huge greenhouse effect, leaving Mars as the only possible candidate (besides Earth). The Viking mission uncovered several indications in favour of the presence of liquid water in the past history of Mars: ramified valley networks, outflow channels, and lobate ejection craters on the surface. The Mars Global Surveyor mission demonstrated that Mars had an intrinsic magnetic field during the first million years; most likely, the Martian atmosphere was denser and warmer than today, probably warm enough to allow the presence of liquid water during some periods. Results from Mars Express indicate that hydrated minerals are localized in the most ancient terrains, which suggest the presence of liquid water in the very early stages. Could life have appeared and developed at that time? This is still an open question, and the major objective for the future space exploration of Mars.

Since the two Voyager flybys in 1979 and 1980, Europa is known to have a complex network of structures on its water-ice surface, which suggests the possible presence of a viscous or possible liquid layer, most likely made of salty liquid water. This hypothesis is supported by the detection of a magnetic field by the Galileo magnetometer. This field could be induced by a dynamo effect within a salty liquid ocean. The energy needed to heat the interior would come from the tidal forces induced by the geometry of Europa, in resonance with Io and Ganymede. Finally, another outer satellite has raised interest since the beginning of the Cassini exploration. Cassini's instruments have revealed that Enceladus' south pole is significantly warmer than the rest of its surface, with evidence for cryovolcanism and ejection of plumes. Other satellites with similar properties remain to be discovered.

5.6.2 Key questions and opportunities

There are several key questions to be addressed in the frame of future Solar System exploration. Are amino-acids present in comets? Did the Martian surface and / or sub-surface host liquid water in the planet's past history? If so, did life appear and develop, and could we find traces of it? Are there life products on the surface of Europa that could have been brought to the surface either by tectonic activity or meteoritic impacts? How deep is the icy crust of Europa? Are there other outer satellites which host a water ocean under their surface, and, if so, at what depth?

The question of cometary composition and the possible presence of amino-acids will be addressed very soon with the analysis of the Stardust samples, and later in the future by ESA's Rosetta mission to comet Churiumov–Gerasimenko. Additional information will come from millimetre and submillimetre spectroscopy using Herschel and ALMA. The space exploration of Mars will remain a priority with, in particular, the search for places where liquid water might have been present. Regarding the exploration of the outer satellites, astronomers from Europe, US and other countries are actively working on the definition of future space missions toward the outer solar system. Present concepts include the exploration of Europa and the Jovian system, a return mission to Titan, and possibly an exploration mission toward Enceladus.

Chapter 6

Recommendations

The previous chapters of this report describe the main science goals in the four thematic areas, and identify certain key technologies and facilities that are needed to achieve these in the next two decades. The main scientific goals and required facilities are summarized here, with the list of facilities divided into two classes: *principal facilities*, where the strongest single motivation for the facility lies within the domain of the corresponding panel, and *supporting facilities*, which have their prime motivation elsewhere, but which also have a critical role to play in this area of interest.

6.1 Do we understand the extremes of the Universe?

The questions posed in this section are a mixture of astrophysics and fundamental physics, and lie at the very foundation of our understanding of the Universe, its formation, and composition. The specific science goals reflect this mixture:

- Measure the evolution of the dark-energy density with cosmological epoch, to search for deviations from a cosmological constant;
- Test for a consistent picture of dark energy using independent probes that are sensitive to its effects in complementary ways, thus either verifying General Relativity or establishing the need for a replacement theory;
- Measure the polarization of the microwave background at ten-degree scales, to search for the signature of relic gravitational waves;
- Establish the direct detection of astrophysically-generated gravitational waves to measure strong-gravity effects, in particular arising from black-hole coalescence;
- Make direct images of the regions near the event horizon of supermassive black holes in galactic nuclei, to understand how large-scale relativistic jets are launched;
- Understand the astrophysics of compact objects and their progenitors, particularly the functioning of the supernova explosion mechanisms;
- Understand the origin and acceleration mechanism of cosmic rays and neutrinos, especially at the highest energies.

These goals map into a clear set of facilities, although not always in a one-to-one manner. A number of attractive potential future facilities exist that are capable of addressing several of the questions of prime interest, in whole or in part. This multiplicity is highly beneficial in independently cross-checking the reality of some of the more subtle phenomena at issue.

Principal facilities. This list is headed by a variety of survey experiments, since future progress in cosmology rests in large part on major statistical studies that require mapping of a large part of the celestial sphere:

- An optical/infrared multi-colour imaging survey across a major fraction of the sky would provide data for some of the principal tests of dark energy: gravitational lensing, baryon oscillations in redshift shells, and new supernova samples. Near-infrared photometry is essential, and imaging in space offers huge advantages due to the low noise background. Space-based optical imaging will also offer high image fidelity over an extremely wide field that cannot be matched from the ground;
- Beyond ESA's Planck mission, the next major step is to measure the polarization signal, which is sensitive to primordial gravitational waves from inflation. This effect is potentially detectable by Planck, since simple inflation models combined with data from the WMAP satellite predict a detectable level of gravitational waves. A next-generation polarization satellite would probe this signature in detail, providing a direct test of the physics of inflation and thus of the fundamental physical laws at energies $\sim 10^{12}$ times higher than achievable in Earth-bound accelerators;
- An additional method of probing the earliest phases of cosmology is to look for primordial gravitational waves at much shorter wavelengths. The space interferometer LISA has the potential to detect this signature by direct observation of the gravitational-wave background according to some models;
- Studies of astrophysical gravitational waves are also of great importance, since they are expected to arise from some of the most extreme environments imaginable, such as the merging and coalescence of two black holes;
- To determine the origin of high-energy cosmic rays will require huge air shower arrays and air fluorescence detectors (possibly satellite-borne), as well as Cerenkov and radio detectors for high-energy neutrinos employing very large volumes of water and ice. The next generation gamma-ray air Cerenkov telescopes such as the CTA will provide a much improved understanding of the acceleration of relativistic particles in supernova remnants and active galactic nuclei.

Supporting facilities. These include:

- Emission processes close to the inner edge of black hole accretion flows can be probed by future X-ray and gamma-ray satellites, and also by large submillimetre interferometers. The signals from high-redshift material will be weak and must be studied in high detail. This will require missions with large collecting area, such as the gamma-ray experiment GLAST and the proposed X-ray missions XEUS and Constellation-X. Such experiments can also probe the nature of dark matter, by searching for the annihilation between dark particle-antiparticle pairs;
- Some probes of cosmological geometry are also possible using an extremely large radio telescope at centimetre wavelengths. Improved tests of General Relativity, for example, can be expected from the large sample of new pulsars expected from the proposed SKA, which will also address various astrophysical problems ranging from cosmic reionization to the formation of galaxies, stars, black holes, and magnetic fields;
- Fundamental cosmology will profit from spectroscopic classification of supernovae at extreme redshifts, and also through hyper-accurate quasar spectroscopy, leading to limits of fundamental constants. Detailed quasar spectroscopy can also limit the nature of dark matter by searching for a small-scale coherence length in the mass distribution, and measure directly the acceleration of the Universe. The understanding of thermonculear and core-collapse supernovae as well as their connection to GRBs, calls for detailed spectroscopic observations in the optical and near infrared of the debris. All the above require an optical-infrared telescope of significantly larger collecting area than is currently available.
- Substantial high-performance computing facilities will be mandatory, both for the analysis of the extensive observational data, and for the careful comparison between these datasets and detailed predictions from theory and simulations.

6.2 How do galaxies form and evolve?

The challenge of understanding galaxy formation and evolution is fundamentally intertwined with the challenge of understanding the structure and evolution of the Universe itself. To address these challenges several important goals can be identified:

- Map the growth of matter density fluctuations in the early Universe, both during and after the dark ages;
- Detect the first stars, black holes, and galaxies, and thus establish the nature of the objects that reionized the Universe and discern the first seeds of galaxies;
- Determine the evolution of the galaxy cluster mass function and constrain the equation of state of the dark energy;
- Make an inventory of the metal content of the Universe over cosmic time, and connect its evolution to detailed models of star formation, and the subsequent metal production and ejection from galaxies by superwinds;
- Detect and measure the metallicity of the warm-hot phase of the intergalactic medium in the local Universe and solve the missing baryons problem;
- Measure the build up of gas, stars, metals, magnetic fields, and masses of galaxies and the evolution of the Hubble sequence with cosmic time;
- Obtain a comprehensive census of the orbits, ages, and compositions of stars in our own Galaxy, aiming to produce a complete history of its original collapse, early formation, and subsequent evolution.

Achieving these goals requires substantial improvements in both observation and theory, including enhanced numerical simulations covering large cosmic volumes, taking into account detailed physical processes and feedback mechanisms. This work will also draw on new facilities covering wavelengths ranging from the low-frequency radio to the gamma-ray domains.

Principal facilities. Some of the most important breakthroughs will be made with telescopes with large collecting area in the radio, visible and near-infrared, and X-rays. Each of these will make major contributions to several of the goals highlighted above:

- An ELT with adaptive optics and high-resolution imaging cameras and spectrometers ($R \sim 5 \times 10^4$), as well as highly multiplexed near-infrared spectrographs (including multiple integral-field instrumentation) will characterize the evolution of large-scale structures over cosmic time, decipher the internal physics of high-redshift galaxies, directly resolve stellar populations in the local supercluster, and provide a powerful complement to the revolutionary information that will come from JWST and Gaia;
- A large-aperture X-ray space mission with moderate-resolution spectroscopic capability ($R \sim 1000$) will address key problems relating to the intergalactic medium, missing baryons, black hole evolution, and galaxy assembly;
- Important breakthroughs will be enabled by an extremely large collecting area telescope in the centimetre radio domain, covering wide dynamical and spectral ranges, such as the proposed SKA. This instrument will address problems ranging from cosmic reionization to the formation of galaxies, stars, black holes, and magnetic fields, and build on the foundation of submillimetre to centimetre work that will be opened up by ALMA and LOFAR;
- A cooled 4–8 m class infrared space telescope will trace dust-obscured galaxy formation, star formation, and black hole formation and growth back to the reionization epoch;
- A 4–8 m class ultraviolet-optimized space telescope could obtain high-resolution imaging and spectroscopy of galaxies and background quasars over thousands of sightlines in the Universe, and trace the evolution of intergalactic baryons and the exchange of matter and metals between galaxies and the intergalactic medium over cosmic time.

Supporting facilities. These include:

- A wide-field optical/infrared telescope designed to constrain dark matter and energy;
- A next-generation gamma-ray mission such as the ACT space mission concept;
- A space-based gravitational wave mission such as LISA;
- Over the longer term, a far-infrared space interferometer could carry out groundbreaking observations of H_2 molecules at high redshift (dust-obscured and shock-heated regions), an ideal complement to ALMA.

6.3 What is the origin and evolution of stars and planets?

Understanding the formation and evolution of stars is at the very foundation of explaining the past evolution and present structure of our Galaxy and the Universe as a whole. Equally relevant is understanding the evolution of circumstellar discs leading to the formation of planetary systems, the search and study of exoplanetary systems, including possible life-hosting systems, as they are all milestones towards putting our Solar System in context and determine our place in the Universe. To address these challenges several important goals can be identified:

- Determine the initial physical conditions of star formation, including the evolution of molecular clouds, and the subsequent development of structure in general, like the formation of single, binary or multiple stellar systems and stellar clusters;
- Understand the nature, dependence on environment and evolution through cosmic time of the Initial Mass Function of stars;
- Probe a wide range of stellar interiors via asteroseismology, including Solar-type stars and white dwarfs;
- Determine the process of planet formation, aiming for a full understanding of the timeline for the formation of planets and the chemical evolution of the material that will eventually end up in exo-planets.
- Establish the demography of exo-planets in a wide mass range from giants to Earth-like, to characterise the diversity of planetary systems;
- Determine the frequency of Earth-like planets in habitable zones and push towards their direct imaging with the long-term goal of spectroscopic characterization including the detection of biomarkers in their atmospheres;

This ambitious programme requires substantial technological development in the coming decades. The techniques to be developed include coronography, extreme adaptive optics, high-contrast imaging systems, extremely large telescopes, infrared interferometry in space, very large effective area millimeter and radio interferometers, transits, microlensing, radial velocity and astrometric searches, laboratory experimental facilities and supercomputing machines.

Principal facilities. A combination of techniques, supplemented by detailed theoretical modeling, will be needed to achieve these goals, and a number of facilities will be required. Very important breakthroughs would come from:

- Near- and mid-infrared imaging and spectroscopy at high spatial resolution and sensitivity provided by an ELT with high-performance adaptive optics, as a perfect complement (or follow-up) to the capabilities of JWST and Herschel;
- Near- to far-infrared observations at high angular resolution of a few milli-arcseconds, with very high contrast. A combination of substantial collecting area and baseline will be required. Dedicated space missions have been widely considered as the way forward;

- High angular resolution and high sensitivity millimeter continuum and line spectroscopy with ALMA and its future upgrades will be key to understand the physical and chemical evolution of dust and gas in the early phases of planet formation; complementary data addressing a subset of these goals can be obtained at centimeter wavelengths with SKA.
- The provision of high spectral resolution in the near-to-far infrared, as the next step in the study of gas dynamics and organic content in the planet-forming zones of discs and in the characterization of exo-planetary atmospheres over a range of important spectral lines, seems only possible using a fully adaptive ELT or interferometry from space;
- At the levels of accuracy and sensitivity demanded, high-precision photometry, with long-term continuum monitoring, can only be expected from a dedicated space platform. This would allow much needed advances in asteroseismology, and in the characterization of low-mass exo-planetary systems through transit surveys;
- Higher accuracy radial velocity experiments from ground will quantify the existence and characteristics of exo-planets over a much wider range of mass and orbital periods;
- Higher accuracy astrometric measurements from ground and space will resolve the degeneracy in mass and orbital inclination, quantify the relative orbital inclination of multiple exo-planetary systems, and extending the measurements to Earth-mass systems.

Supporting facilities. These include:

- Measurement of stellar rotation and velocity fields from small to large scales, and imaging polarimetry at millimetre and centimetre wavelengths to derive magnetic field strength and geometry, will be needed to advance the field of star formation;
- Single-dish millimeter and centimeter facilities to rapidly survey large areas of the sky and find sources for detailed studies of star- and disc formation and evolution with ALMA, SKA and VLBI.
- Asteroseismology and exo-planet transits will benefit from coordinated programmes of ground-based ‘whole-earth’ telescopes.
- The continuing availability of X-ray observatories for deep surveys and detailed spectroscopy of individual objects will be an essential asset to study the origin of stellar activity through all phases of stellar evolution.

6.4 How do we fit in?

The Sun and its Solar System bodies constitute our close neighbourhood in a vast Universe that is directly accessible to *in situ* measurements and detailed remote observations, and they directly influence life on Earth at all times. To better understand the formation, evolution, and detailed properties of the Solar System, the following goals have been identified:

- Measure the magnetic and velocity fields in the Sun’s convective interior and atmosphere over extended time periods, to establish the basic mechanisms of magnetic field generation, and to forecast transient events that directly or indirectly affect life on Earth;
- Detect and identify the mechanism by which energy is transferred from the Solar surface that heats the upper Solar atmosphere and eventually accelerates the Solar wind;
- Understand the role of turbulence and magnetic field in the evolution of the primordial nebula, the mechanism of particle growth, and the observed elemental and isotopic ratios in these nebula, as well as in Solar System bodies;
- Determine the dynamical history and the composition of trans-Neptunian objects and asteroids, and the rate of large potential impactors in the near-Earth asteroid population; search for complex molecules in comets and study the link between comets and interstellar matter;
- Constrain the models of internal structure of planets and satellites and the origin of their internal heat, the surface-atmosphere interactions and the recycling mechanisms in the terrestrial planets and outer satellites;

- Understand the origin and evolution of Titan's atmosphere, searches for liquid water at the surface and subsurface of Mars, and for liquid water oceans below the surface of Europa and other outer satellites.

Advances in studies of the Sun will require new observational facilities covering a substantial range in wavelengths, numerical simulations of the time-dependent three-dimensional magnetic fields, and advances in the theory of magneto-turbulence. The planetary science objectives will require a combination of theoretical work, *in situ* exploration, ground-based observations, and laboratory experiments.

Principal facilities. Important breakthroughs will be made with Solar telescopes on ground and in space that achieve higher spatial or temporal resolution, and cover much larger fields over longer times. Various space missions will be required to address the planetary science objectives:

- A large-aperture (3–5 m) Solar telescope with adaptive optics and integral-field spectro-polarimeters to observe astrophysical processes at their intrinsic scales, and thereby observe the interaction of magnetic fields and plasma motions in the Solar atmosphere;
- A space mission such as Solar Orbiter that can make *in situ* measurements close to the Sun and provides remote observing capabilities of the polar regions to complement the earlier Ulysses mission and improve knowledge of the inner Solar wind regions, and the magnetic and velocity field around the poles;
- A medium-aperture (1–2 m) (extreme-)ultraviolet satellite with X-ray capabilities to study the processes occurring in the Solar corona that cannot be studied from the ground;
- The next generation of space missions to the outer Solar System should include exploration of the Jovian system with special emphasis on Europa, and of the Saturn system with special emphasis on Titan and Enceladus;
- A Mars sample return mission should be considered as a follow-up to the recent orbiter, lander and rover missions;
- Further in the future, other dedicated missions would include the exploration of a near-Earth asteroid, including a lander and preferably a sample return; a Saturn probe, for a direct measurement of its elemental and isotopic ratios; a mission to Venus, including in particular a study of its atmospheric escape and the surface-atmosphere interaction.

Supporting facilities. These include:

- A network of ground-based, synoptic instruments that monitor the full-disc Solar magnetic and velocity fields on time-scales from minutes to decades will provide observations for space-weather forecasts as well as understanding magnetic field generation and destruction;
- Several space missions are presently in operation (Mars Express, Mars Odyssey, Mars Reconnaissance Orbiter, Mars rovers, Venus Express) or having their samples analysed (Stardust and Genesis). NASA's Mars Science Laboratory and the European ExoMars will be launched soon;
- In the near future, significant achievements are expected from the scientific exploitation of Cassini, Rosetta, and Bepi-Colombo;
- Planetary science will benefit from large space and ground-based facilities, covering wavelengths from ultraviolet to the radio. Herschel and ALMA will be crucial for studying the composition of comets and planetary atmospheres, and for exploring the Kuiper Belt;
- Radio telescopes spanning sub-millimetre to metre wavelengths (e.g., ALMA and LOFAR) will be able to study the Sun and the heliosphere with spatial resolution inaccessible at optical wavelengths;
- High-resolution multi-object spectrographs at 4–8 m class telescopes to routinely monitor a large number of Solar-like stars to test theories of magnetic and velocity fields on stars that rotate faster or slower than the Sun and / or show different amount of magnetic flux;

- JWST as well as an ELT will allow, in particular, imaging spectroscopy of planetary atmospheres and surfaces, comets, and trans-Neptunian objects;
- Laboratory studies will be required for determination of the equations of state at high temperature and pressure; study of phase diagrams at low temperature; chemical reactions at low temperature under various environmental conditions; and analyses of meteorites and extra-terrestrial samples from the Moon, comets, and Mars;
- Specific theoretical work is required in areas including evolution scenarios of protoplanetary discs with turbulence; dynamical evolution of small bodies in the Solar System; modeling of planetary interiors; and modeling of climate and climate evolution on long time-scales.

Contributors

The tables below list the membership of the Science Vision Working Group and its four supporting panels.

Science Vision Working Group

Tim de Zeeuw	chair	Sterrewacht Leiden	Netherlands
John Peacock	chair panel A	Royal Observatory Edinburgh	UK
Claes Fransson	co-chair panel A	Stockholm Observatory	Sweden
Jacqueline Bergeron	chair panel B	Institut d'Astrophysique	France
Rob Kennicutt	co-chair panel B	Institute of Astronomy Cambridge	UK
Leonardo Testi	chair panel C	Osservatorio Astrofisico di Arcetri	Italy
Rafael Rebolo	co-chair panel C	Instituto de Astrofísica de Canarias	Spain
Oskar von der Luhe	chair panel D	Kiepenheuer-Institut für Sonnenphysik	Germany
Therese Encrenaz	co-chair panel D	Observatoire de Paris	France
Reinhard Genzel		MPI-E Garching	Germany
Michael Perryman		European Space Agency	Netherlands
Alvio Renzini		INAF-OAPD Padova	Italy
Rashid Sunyaev		MPI-A Garching	Germany
Catherine Turon		Observatoire de Paris	France
Michael Bode	WP-3 leader	Liverpool John Moores University	UK

Panel A: Do we understand the extremes of the Universe?

John Peacock	chair	Royal Observatory Edinburgh	UK
Claes Fransson	co-chair	Stockholm Observatory	Sweden
Juan Garcia-Bellido		Universidad Autónoma de Madrid	Spain
Francois Bouchet		Institute d'Astrophysique de Paris	France
Andrew Fabian		Cambridge	UK
Bruno Leibundgut		European Southern Observatory	Germany
Subir Sarkar		Oxford University	UK
Peter Schneider		Universität Bonn	Germany
Ralf Wijers		Sterrenkundig Instituut Anton Pannekoek	Netherlands
Bernard Schutz		MPI-G Potsdam	Germany

Panel B: How do galaxies form and evolve?

Jacqueline Bergeron	chair	Institute d'Astrophysique de Paris	France
Rob Kennicutt	co-chair	Cambridge University	UK
Xavier Barcons		CSIC-UC	Spain
Frank Bertoldi		Universität Bonn	Germany
Andrea Ferrara		SISSA / ISAS Trieste	Italy
Marijn Franx		Sterrewacht Leiden	Netherlands
Amina Helmi		Kapteyn Instituut	Netherlands
Guinevere Kauffmann		MPI-A Garching	Germany
Ian Smail		Durham University	UK
Matthias Steinmetz		Astrophysikalisches Institut Potsdam	Germany

Panel C: What is the origin and evolution of stars and planets?

Leonardo Testi	chair	Osservatorio Astrofisico di Arcetri	Italy
Rafael Rebolo	co-chair	Instituto de Astrofísica de Canarias	Spain
Jørgen Christensen-Dalsgaard		Århus University	Denmark
Ewine van Dishoeck		Sterrewacht Leiden	Netherlands
Stephane Guilloteau		Observatoire de Bordeaux	France
Pavel Kroupa		Universität Bonn	Germany
Didier Queloz		Observatoire de Genève	Switzerland
Massimo Turatto		Osservatorio Astronomico di Padova	Italy
Christoffel Waelkens		Universiteit Leuven	Belgium

Panel D: How do we fit in?

Oskar von der Luhe	chair	Kiepenheuer-Institut für Sonnenphysik	Germany
Therese Encrenaz	co-chair	Observatoire de Paris	France
Willy Benz		Universität Bern	Switzerland
Michele Dougherty		Imperial College London	UK
Richard Harrison		Rutherford Appleton Laboratory	UK
Artie Hatzes		Thringer Landessternwarte Tautenburg	Germany
Christoph Keller		Sterrekundig Instituut Utrecht	Netherlands
Hans Rickman		Uppsala Astronomical Observatory	Sweden
Tilman Spohn		Deutches Zentrum für Luft- und Raumfahrt	Germany
Jose Carlos del Toro Iniesta		Instituto de Astrofísica de Andalucía	Spain

Additional contributions and comments were provided by Rainer Beck, Heino Falcke, Michael Kramer and Frank Molster.

List of abbreviations

Below a listing of all the facilities and acronyms which appear in the main text.

2MASS: the 2 Micron All Sky Survey used two highly-automated 1.3-m telescopes, one at Mt. Hopkins, Arizona, and one at CTIO, Chile. Each telescope was equipped with a three-channel camera, capable of observing the sky simultaneously at J (1.25 microns), H (1.65 microns), and Ks (2.17 microns). The survey was started in 1997 and finished in 2001.
`http://www.ipac.caltech.edu/2mass/`

ACES: Atomic Clock Ensemble in Space is a programme to test the performance of a new type of atomic clock that exploits and depends upon microgravity conditions.
`http://www.spaceflight.esa.int/users/downloads/factsheets/fs031_10_aces.pdf`

ACT: The Atacama Cosmology Telescope is designed specifically for high-sensitivity large-area surveys of the sky requiring dedicated observations for months at a time. It will map the CMB temperature anisotropy over 100 square degrees and expects to see first light in late 2006.
`http://www.hep.upenn.edu/act/`

Adaptive Optics: A technique that uses deformable mirrors to correct for the atmospheric blurring of astronomical images.

ALMA: The Atacama Large Millimeter Array is a Joint Europe–North America–Japan project to build a synthesis telescope of fifty 12 m antennae that will operate at millimeter and submillimeter wavelengths at the 5000 m Chajnantor site in Northern Chile.
`http://www.eso.org/projects/alma/`

AMI: the Arcminute Microkelvin Imager will make images of the Sunyaev-Zel'dovich effect in the cosmic microwave background radiation with an angular sizes of arcminutes, rather than the degree scales of most CMB instruments.
`http://www.mrao.cam.ac.uk/telescopes/ami/index.html`

AMS: The Alpha Magnetic Spectrometer, is a space borne experiment scheduled for a three years mission on the International Space Station. It is designed to study: the presence or absence of antimatter in distant galaxies through the detection of cosmic anti-nuclei as anti-Helium and anti-Carbon, the origin and structure of dark matter which is believed to make up 90% of the known Universe and the origin and the composition of charged cosmic rays.
`http://ams.pg.infn.it/`

ANAIS: A large-mass dark matter search experiment with NaI scintillators at the Canfranc Underground Laboratory.
`http://www.unizar.es/lfnae/ipaginas/ip0400.html#manais`

ANTARES: Astronomy with a Neutrino Telescope and Abyss environmental RESearch; a 30 million cubic meter seawater, 1000 photomultiplier neutrino telescope for particle physics and astrophysics applications under construction at a depth of 2700 m in the Mediteranean off the coast of Toulon. `http://antares.in2p3.fr/`

APEX: Atacama Pathfinder EXperiment. 12 m submillimeter telescope on the ALMA site at Chajnantor in Chile, operated by MPIfR Bonn, Sweden and ESO Observations started in 2005.
`http://www.mpifr-bonn.mpg.de/div/mm/apex.html`

ArDM: Argon Dark Matter experiment.
`http://neutrino.ethz.ch/ArDM/ardm.html`

ASPERA: is an FP6 ERA-net program which started in July 2006, and comprises 16 national funding agencies in Europe together responsible for funding astroparticle physics research.

ASTRONET: ERA-NET project which was created by funding agencies and ministries from France, Germany, Italy, Netherlands, Spain, UK plus ESA, ESO and NOTSA in order to establish a comprehensive long-term planning process for the development of European astronomy.
`http://www-astronet-eu.org`

ATNF: Australia National Telescope Facility.
`http://www.atnf.csiro.au/`

AU: Astronomical Unit, the mean distance from the Earth to the Sun.

CCAT: Cornell Caltech Atacama Telescope is a large sub-millimeter telescope in the high Andes of northern Chile. This instrument will be used to address fundamental questions regarding cosmic origins.
`http://www.submm.caltech.edu/∼sradford/ccat/`

CCD: Charge-Coupled Device, highly efficient solid-state device used as detector for astronomical observations in the optical wavelength regime.

Chandra: An imaging X-ray space observatory launched in 1998 (formerly AXAF). This NASA Great Observatory contains an X-ray mirror with a diameter of 1.2 m and a focal length of 10 m, and provides unprecedented 0.5 arcsecond resolution in X-rays up to 10 keV. This mission is complementary to ESA's XMM–Newton.
`http://asc.harvard.edu/`

CMB: Cosmic Microwave Background

CNO: Carbon-Nitrogen-Oxygen (cycle). Nuclear fusion cycle in stars.

CODALEMA: Cosmic ray Detection Array with Logarithmic ElectroMagnetic Antennas is meant to detect cosmic showers by radio waves. A demonstrator of CODALEMA is located at the radio observatory Nancy in France.
`http://www-subatech.in2p3.fr/%7Eastro/codalema.html`

CODEX: Cosmic Dynamics Experiment is an instrument concept for high resolution spectroscopy with an ELT.

Constellation-X: The Constellation-X Observatory is a proposed combination of several X-ray telescopes in space working in unison to generate the observing power of one giant telescope. With this space observatory, scientists will investigate black holes, Einstein's Theory of General Relativity, galaxy formation, the evolution of the Universe on the largest scales, the recycling of matter and energy, and the nature of dark matter and dark energy.
`http://constellation.gsfc.nasa.gov/`

CREAM: The Cosmic Ray Energetics And Mass project will investigate ultra high energy (10^{12} to $> 5 \times 10^{14}$ eV) cosmic rays over the elemental range from protons to iron utilizing a series of Ultra Long Duration Balloon flights. The goal is to observe spectral features and/or abundance changes that might be related to a supernova acceleration limit.
`http://cosmicray.umd.edu/cream/cream.html`

CTA: the Cherenkov Telescope Array, is a future project where several cherenkov Telescopes will be coupled together to look at air showers due to high energy cosmic particles ($GeV to several TeV$).

DAMA: particle DArk MAtter searches with highly radiopure scintillators at Gran Sasso.
`http://people.roma2.infn.it/~dama/web/`

DAMA/LIBRA: A new 250 kg second generation highly radiopure NaI(Tl) set-up for DAMA.
`http://www.edpsciences.org/articles/epja/abs/2006/03/10050_2006_Article_608007/10050_2006_Article_608007.html`

DES: the Dark Energy Survey is a high precision multi-bandpass wide area survey, designed to produce photometric redshifts from $0.2 < z < 1.3$. The instrument for this survey will be mounted at the prime focus of the Blanco 4-meter telescope at CTIO. The survey data will cover 5000 sq-degrees, with 4000 sq-degrees overlapping the Sunyaev-Zeldovich CMB survey being conducted by the South Pole Telescope.
`http://www.darkenergysurvey.org/`

(E-)ELT: (European) Extremely Large Telescope.
`http://www.eso.org/projects/e-elt/`

ERA-nets: the European Research Area Networks, actively contribute to the process of unifying research across Europe. The long-term goal is to keep European science competitive. See also ASTRONET.

eROSITA: the extended ROentgen Survey with an Imaging Telescope Array will be an instrument on the Russian satellite Spectrum-X-Gamma. eROSITA should 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. Launch is expected in 2009.
`http://www.mpe.mpg.de/projects.html#erosita`

ESA: European Space Agency.
`http://www.esa.int/`

ESO: European Southern Observatory, with headquarters in Garching bei München. ESO operates observatories on Cerro La Silla and on Cerro Paranal and, in the future, on Chajnantor in Chile.
`http://www.eso.org/`

EURECA: European Underground Rare Event Calorimeter Array is a proposal for a tonne scale dark matter detector.
`http://www.pparc.ac.uk/roadmap/rmProject.aspx?q=45`

EUSO: the Extreme Universe Space Observatory is a space mission devoted to the investigation of cosmic rays and neutrinos of extreme energy ($E > 5 \times 10^{19}$ eV), the detection being performed by looking at the streak of fluorescence light produced when such a particle interacts with the Earth's atmosphere. EUSO will be accommodated, as an external payload of the Columbus module, on the International Space Station.
`http://www.euso-mission.org/`

FLAMES: The Fiber Large Array Multi-Element Spectrograph is an Australian-built multi-object fiber feed on the VLT, enabling simultaneous spectroscopy of 130 sources, or fifteen $2'' \times 3''$ integral fields, or one larger $12'' \times 7''$ integral field. Fibers can be positioned robotically over a $25'$ diameter field.
`http://www.hq.eso.org/instruments/flames/`

FUSE: the Far Ultraviolet Spectroscopic Explorer is a NASA mission launched in June 1999 to take high-resolution ($\lambda/\Delta\lambda = 20000$) spectra of objects from 905 to 1187 Å.
`http://fuse.pha.jhu.edu/`

Gaia: An ESA space observatory which will measure the distances and motions of a billion stars in the Galaxy with extraordinary precision. It will allow astronomers to determine the Galaxy's three-dimensional structure, the space velocities of its constituent stars and, from these data, to understand the Galaxy's origin and evolution. Gaia will also obtain multi-color photometry as crucial diagnostic data for all stars observed, along with radial

velocities for the brighter objects to complete the kinematical data. Launch is planned in 2012.
`http://astro.estec.esa.nl/GAIA/`

GBT: The Green Bank Telescope is a movable 100 meter single dish radio telescope located at the National Radio Astronomy Observatory's site in Green Bank, West Virginia.
`http://www.gb.nrao.edu/GBT/GBT.shtml`

GLAST: Gamma-ray Large Area Space Telescope will study the cosmos in the energy range 10 keV – 300 GeV. GLAST is expected to be launched in the fall of 2007.
`http://glast.gsfc.nasa.gov/`

Herschel: ESA's Fourth Cornerstone Mission is a Far-Infrared and Submillimeter Telescope (formerly called FIRST). The satellite, to be launched in 2008, will provide a natural extension of the research done with the ISO satellite, and will form an excellent complement to ALMA.
`http://sci.esa.int/home/herschel/index.cfm`

HESS: The High Energy Stereoscopic System is a system of Imaging Atmospheric Cherenkov Telescopes for the investigation of cosmic gamma rays in the 100 GeV energy range. HESS is located near the Gamsberg in Namibia.
`http://www.mpi-hd.mpg.de/hfm/HESS/HESS.html`

HiRes: the High Resolution Fly's Eye is an experiment to study the highest energy cosmic rays to determine the energy, direction, and chemical composition of the incident particle. HiRes is located at the University of Utah.
`http://hires.physics.utah.edu/`

HR diagram: the HertzsprungRussell diagram shows the luminosity of a star plotted against its color. Most stars evolve along specific tracks in this diagram. The positions on these tracks are specific for the age and mass of a star.

HST: The Hubble Space Telescope of NASA (with 15% ESA participation) is a 2.4 m optical telescope, launched in 1990. The HST is unique because of its high spatial resolution over a wide field, its performance in the ultraviolet and near-infrared, and the regular upgrades with new instrumentation during servicing missions with the space shuttle.
`http://www.stsci.edu`

IceCube: The IceCube Neutrino Detector is a neutrino telescope currently under construction in deep Antarctic ice by deploying thousands of spherical optical sensors at depths between 1450 and 2450 meters. The main goal of the experiment is to detect neutrinos in the high energy range, spanning from 10^{11} eV to about 10^{21} eV.
`http://icecube.wisc.edu/`

IceTop: A km^2 array of particle detectors that is currently being installed at the South Pole, right above the IceCube neutrino telescope. It is used to detect extended particle showers induced in our atmosphere by high energy cosmic rays. Its spacing (80 stations = 160 Cherenkov ice tanks, 125 m mean distance) allows the observation of cosmic rays with energies between 10^{14} eV and 10^{17} eV.
`http://www-zeuthen.desy.de/nuastro/exp/icetop_e.html`

ICM: Inter Cluster Medium

IGM: Inter Galactic Medium

IMBH: Intermediate Mass Black Holes are black holes with a mass between several tens to a few thousand solar masses.

IMF: the Initial Mass Function describes the relation between the number of stars born and their mass.

IRAM 30 m: The 30 m single dish millimeter telescope from the Institut de Radioastronomie Millimétrique, located on Pico Veleta in Spain.

`http://iram.fr/IRAMES/index.htm`

ISO: ESA's Infrared Space Observatory operative from 1995-1998, provided high sensitivity imaging and spectroscopic observations in the mid- to far-infrared.
`http://www.iso.vilspa.esa.es/`

JASMINE: Japan Astrometry Satellite Mission for INfrared Exploration is a proposed scanning astrometric satellite to measure parallaxes, positions, and proper motions with a precision of 10 μarcsec at 15.5 mag at 9000 nm.
`http://www.jasmine-galaxy.org/index.html`

JDEM: the Joint Dark Energy Mission will be a NASA / DOE mission to investigate the properties of dark energy. At the moment 3 proposals are competing ADAPT, Destiny and SNAP.
`http://universe.nasa.gov/program/probes/jdem.html`

JWST: James Webb Space Telescope: the planned successor to HST, optimized for observations in the near infrared out to 28 μm, with a 6.5 m mirror.
`http://www.jwst.nasa.gov/`

KIDS: the Kilo-Degree Survey is a large (1500 square degrees in four bands) survey to be performed with the VlT Survey Telescope which targets two areas of the sky where large redshift surveys have taken place, and where near-infrared surveys are soon to begin: an equatorial strip on the North Galactic Cap, and a patch near the South Galactic Pole.
`http://www.eso.org/observing/webone.html`

LHC: Large Hydron Collider at CERN is designed to collide two counter rotating beams of protons or heavy ions. Proton-proton collisions are foreseen at an energy of 7 TeV per beam with a planned start-up in 2007.
`http://lhc.web.cern.ch/lhc/`

LIGO: Laser Interferometer Gravitational Wave Observatory, ground-based interferometric gravitational wave detectors located at two separate sites in the USA. LIGO may detect merging stellar-mass black holes in other galaxies. An advanced version, LIGO-2, is currently under study.
`http://www.ligo.caltech.edu/`

LISA: Laser Interferometer Space Antenna. Planned joint ESA / NASA interferometric gravitational wave detector, to consist of three free-flying spacecraft that form an equilateral triangle 5 million km across, orbiting the Sun in the Earth's orbit, trailing the Earth by 20 degrees. Will detect gravitational waves from thousands of galactic binaries, and likely from merging supermassive black holes.
`http://lisa.jpl.nasa.gov/`

LMT: The Large Millimeter Telescope is a 50 m diameter single-dish telescope optimized for astronomical observations at millimeter wavelengths (0.85 mm $< \lambda <$ 4 mm) under construction on Volcán Sierra Negra in Mexico. The LMT Project is a bi-national collaboration between Mexico and the US.
`http://www.lmtgtm.org/`

LNGS: The Laboratori Nazionali del Gran Sasso is an underground laboratory for (astro-)particle physics and houses experiments like DAMA and XENON.
`http://www.lngs.infn.it/`

LOFAR: The Low Frequency ARray is high resolution and high sensitivity radio interferometer working at frequencies between 30 MHz and several hundred MHz; a collaborative initiative of the Netherlands, Germany, Sweden, France, the UK and the USA. Stations will be located in the Netherlands, Germany, France, UK Denmark and Sweden.
`http://www.lofar.org/`

LOPES: The LOFAR Prototype Station, built to test and clarify the nature and properties of the radio emission from air showers. It is located at the same site as the KASKADE-Grande

observatory.
`http://www.astro.ru.nl/lopes/`

LSST: the Large Synoptic Survey Telescope is a proposed ground-based 8.4-meter, 10 square-degree-field telescope that will provide digital imaging of faint astronomical objects across the entire sky, night after night. In a relentless campaign of 15 second exposures, LSST will cover the available sky every three nights, opening a movie-like window on objects that change or move on rapid timescales: exploding supernovae, potentially hazardous near-Earth asteroids, and distant Kuiper Belt Objects. The superb images from the LSST will also be used to trace billions of remote galaxies and meaasure the distortions in their shapes produced by lumps of Dark Matter, providing multiple tests of the mysterious Dark Energy.
`http://www.lsst.org/lsst_home.shtml`

MAGIC: MAGIC is an IACT that is located at La Palma and has started measuring since late 2004.
`http://magic.mppmu.mpg.de/`

Mars Express: An ESA built orbiter for Mars. It has 7 instruments on board to study the surface and atmosphere of Mars. It was launched in July 2003, arrived five months later at Mars, and will operate until at least 2007.
`http://www.esa.int/SPECIALS/Mars_Express/index.html`

MHD: Magneto-Hydro-Dynamics

Millennium run: The Millennium Simulation run is the largest cosmological N-body simulation ever carried out and has been used to construct sophisticated semi-analytic models of galaxy information that cover the complete galaxy population in a representative piece of the Universe down to luminosities substantially below L*.
`http://www.mpa-garching.mpg.de/galform/virgo/millennium/index.shtml`

MOND: Modified Newtonian Dynamics, is a modification of the usual Newtonian force law hypothesized in 1983 by Milgrom as an alternative to Dark Matter.
`http://www.astro.umd.edu/ ssm/mond/`

MOST: the Mircovariability & Oscillations of Stars is a suitcase-sized microsatellite designed to probe stars and extrasolar planets by measuring tiny light variations (up to 1ppm).
`http://www.astro.ubc.ca/MOST/`

MWA: The Mileura Widefield Array is a project to create two complementary and co-located but substantially independent scientifically-capable demonstration instruments, one led by MIT in the 80-300 MHz frequency range, and the other led by ATNF in the 800-1600 MHz range.
`http://www.haystack.mit.edu/ast/arrays/mwa/index.html`

NASA: National Aeronautics and Space Administration.
`http://www.nasa.gov/home/index.html`

OSIRIS: the Optical, Spectroscopic, and Infrared Remote Imaging System is a scientific imaging system on the orbiter of ESA's Rosetta mission to comet P/Wirtanen.
`http://www.mps.mpg.de/projects/rosetta/osiris/`

OWL: Orbiting Wide-angle Light-collectors is an Earth-orbiting system to study air showers initiated by $> 10^{19}$ eV particles.
`http://owl.gsfc.nasa.gov/`

PAH: Poly-Aromatic-Hydrocarbon

PAMELA: The Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics is a powerful particle identifier using a permanent magnet spectrometer with a variety of specialized detectors that will measure spectra of cosmic electrons, positrons, antiprotons and light nuclei over a very large range of energy from 50 MeV to hundreds GeV, depending on species. It will complement also the measurements of electromagnetic radiations that will be carried out by AGILE and GLAST space missions. It was launched in June 2006.

`http://wizard.roma2.infn.it/pamela/`

PAN-STARRS: the Panoramic Survey Telescope & Rapid Response System – is a wide-field imaging facility being developed at the University of Hawaii's Institute for Astronomy. It will be able to observe the entire available sky several times each month. The immediate goal of Pan-STARRS is to discover and characterize Earth-approaching objects, both asteroids and comets, that might pose a danger to our planet.
`http://pan-starrs.ifa.hawaii.edu/public/`

PaST: The Primeval Structure Telescope Consisting will consist of an array of some ten-thousand log-periodic antennae spread over several square kilometers, PaST will capture a detailed radio image of the sky in the range of 50–200 MHz. Current sites under consideration include Ulastai, China and Amundson-Scott South Pole Station, Antarctica.
`http://web.phys.cmu.edu/~past/`

Piere Auger Observatory: The Piere Auger Observatory is a hybrid detector, employing two independent methods to detect and study high-energy cosmic rays. One technique is ground-based and detects high energy particles through their interaction with water. The other technique tracks the development of air showers by observing ultraviolet light emitted high in the Earth's atmosphere. It is located in the Pampa Amarilla in western Argentina. The Auger Observatory is in the final stages of construction and has begun to collect data.
`http://www.auger.org/`

Planck: Space mission designed to image the anisotropies of the Cosmic Background Radiation Field over the whole sky, with unprecedented sensitivity and angular resolution. Planck is planned to be launched in 2008, together with the Herschel satellite.
`http://www.rssd.esa.int/index.php?project=Planck`

Rosetta: ESA's rendez-vous mission with a comet, the third Cornerstone Mission of the Horizon 2000 program. The satellite was launched in 2004 and will meet with comet Churiumov-Gerasimenko in 2014. Scientific measurements will be performed from orbit and with a surface science package which will be landed on the surface to take insitu data.
`http://www.estec.esa.nl/spdwww/rosetta/html/index.html`

SDSS: Sloan Digital Sky Survey
`http://www.sdss.org/`

SKA: Square-Kilometre Array. A proposed giant radio telescope with an effective collecting area of 1 km^2. Construction could start after 2010.
`http://www.skatelescope.org/`

SOHO: the Solar and Heliospheric Observatory has been launched in December 1995. It studies the Sun, from its deep core to the outer corona, and the solar wind. It has 12 instruments on board.
`http://soho.esac.esa.int/`

SOLIS: Synoptic Optical Long-term Investigations of the Sun is a synoptic facility for solar observations over a long time frame.
`http://solis.nso.edu/`

SPASE: The South Pole Air Shower Experiment is a large-area air shower array established at the geographic South Pole for the detection of cosmic rays with primary energies above 50 TeV.
`http://www.bartol.udel.edu/spase/`

SPITZER: Spitzer (formerly SIRTF, the Space Infrared Telescope Facility) is an infrared space telescope. It was launched into space on 25 August 2003 and is still operating today. Spitzer has 3 instruments which enables it to take images and spectra between 3 and 180 microns.
`http://www.spitzer.caltech.edu/`

SPT: The South Pole Telescope is designed for conducting large-area millimeter and sub-millimeter wave surveys of faint, low contrast emission, as required to map primary and secondary

anisotropies in the cosmic microwave background.
`http://spt.uchicago.edu/`

STEREO: Solar TErrestrial RElations Observatory is a two-year mission using two nearly identical observatories, one ahead of Earth in its orbit and the other trailing behind. The duo will provide three-dimensional measurements of the Sun and its flow of energy, enabling scientists to study the nature of coronal mass ejections and why they happen.
`http://stereo.jhuapl.edu/`

Sunyaev-Zeldovich (effect): The distortion of the spectrum of the cosmic microwave background radiation caused by an intervening hot electron cloud.

SUZAKU: is a Japanese X-ray satelite that was launched in July 2005. It performs observations from $0.4 - 600$keV.
`http://www.astro.isas.ac.jp/suzaku/index.html.en`

SVWG: Science Vision Working Group

Telescope Array: the Telescope Array project, a collaboration between universities and institutes in Japan, Taiwan, China and the United States, is designed to observe cosmic-ray-induced air showers at extremely high energies (10^{19} eV and up) using a combination of ground array and air-flurescence techniques. The cosmic rays are observed at three fluorescence sites and a separate ground array consisting of 576 detectors. It is being deployed in the high desert in Millard County, Utah, USA. First data from the Telescope Array is expected in the spring of 2007.
`http://www.telescopearray.org/`

TUNKA: A cosmic air shower detector that makes use of Cerenkov light emission. It consists now of 25 wide angle integral detectors. The detectors are deployed in a square of 340x340 m^2 in the Tunka Valley, Siberia, Russia.

Venus-Express: Venus Express is the first ESA mission to VENUS. It studies the atmosphere, the plasma environment, and the surface of Venus.
`http://www.sci.esa.int/venusexpress/`

VIRGO: The Virgo detector for gravitational waves consists of a Michelson laser interferometer made of two orthogonal arms being each 3 km long. Multiple reflections between mirrors located at the extremities of each arm extend the effective optical length of each arm up to 120 km.
`http://wwwcascina.virgo.infn.it/`

VISTA: Visible and Infrared Survey Telescope for Astronomy. 4 m telescope to be completed at ESO's Paranal Observatory in 2006. Initially this telescope will have a square-degree wide-field camera for the near-infrared JHK bands. `http://www.vista.ac.uk/`

VLA: The Very Large Array, consists of 27 radio antennas in a Y-shaped configuration on the Plains of San Agustin, New Mexico. The largest possible baseline is 36 km.
`http://www.vla.nrao.edu/`

VLBI: Very Long Baseline Interferometry: a technique used in radio astronomy, in which signals collected simultaneously by telescopes at different locations in the world are correlated afterwards. This effectively allows them to operate as one giant telescope with extremely high spatial resolution.

VLT: ESO's Very Large Telescope, comprising four 8.2 m diameter optical/infrared telescopes on Cerro Paranal in Northern Chile. First light in mid-1998, and all four telescopes completed by 2000.
`http://www.eso.org/observing/vlt/`

VLTI: The Very Large Telescope Interferometer, which combines the light of the four unit 8.2 m telescopes of the VLT and the auxilary telescopes at a special VLTI focus, to obtain diffraction-limited imaging at milli-arcsecond resolution in the near- and mid-infrared. This is a unique

feature of the VLT.
`http://www.hq.eso.org/projects/vlti/`

WARP: WIMP Argon Programme intends to search for Cold Dark Matter in the form of weakly interacting, massive sub-atomic sized particles known as WIMP's.
`http://warp.pv.infn.it/`

WHIM: Warm-Hot Intergalactic Medium

WIMP: Weakly Interacting Massive Particles

WIRE: Wide Field Infrared Explorer. It was meant to study galaxy evolution at high redshift. Soon after the launch in 1999, control was lost and by the time it was regained the coolant was evaporated. Now WIRE is used to conduct astroseismology investigations and as a test-bed for other science opportunities, technology infusion, risk management, educational outreach, and training.
`http://sunland.gsfc.nasa.gov/smex/wire/`

WMAP: Wilkinson Microwave Anisotropy Probe, is a NASA Explorer mission measuring the temperature of the cosmic background radiation over the full sky with unprecedented accuracy. It was launched in 2001 and will operate until 2009.
`http://map.gsfc.nasa.gov/`

XENON: A direct dark matter detection experiment using liquid xenon as the detector medium. The goal is to detect the small charge and light signal after a dark matter particle interacts with a xenon nucleus. The first module is to be operated at the Gran Sasso Undeground Laboratory (LNGS).
`http://www.phys.ufl.edu/xenon/`

XEUS: Proposed follow-on to ESA's Cornerstone X-Ray Spectroscopy Mission (XMM-NEWTON). The mission is under study as envisaged by the Horizons 2000 Survey Committee. XEUS will be around 200 times more sensitive than XMM-Newton.
`http://sci.esa.int/science-e/www/area/index.cfm?fareaid=25`

XIS: XIS is an instrument on Japanese 5-th X-ray satellite Suzaku. It consists of four CCD X-ray imaging spectrometers, which provide images of X-ray sources.
`http://space.mit.edu/XIS/`

XMM–Newton: X-ray Multi-mirror spectroscopy Mission (officially known as XMM-Newton). The second Cornerstone Mission of ESA's Horizon 2000 plan, launched in 1999. The satellite observatory has a projected life time of more than a decade. With three large X-ray mirrors with a total effective area of 6000 cm^2 XMM is the high-throughput spectroscopic complement of NASA's Chandra.
`http://sci.esa.int/home/xmm-newton/`

ZEPLIN: ZonEd Proportional scintillation in LIquid Noble gases for direct dark matter searches.
`http://www.shef.ac.uk/physics/research/pppa/research/dm/zeplin.php`