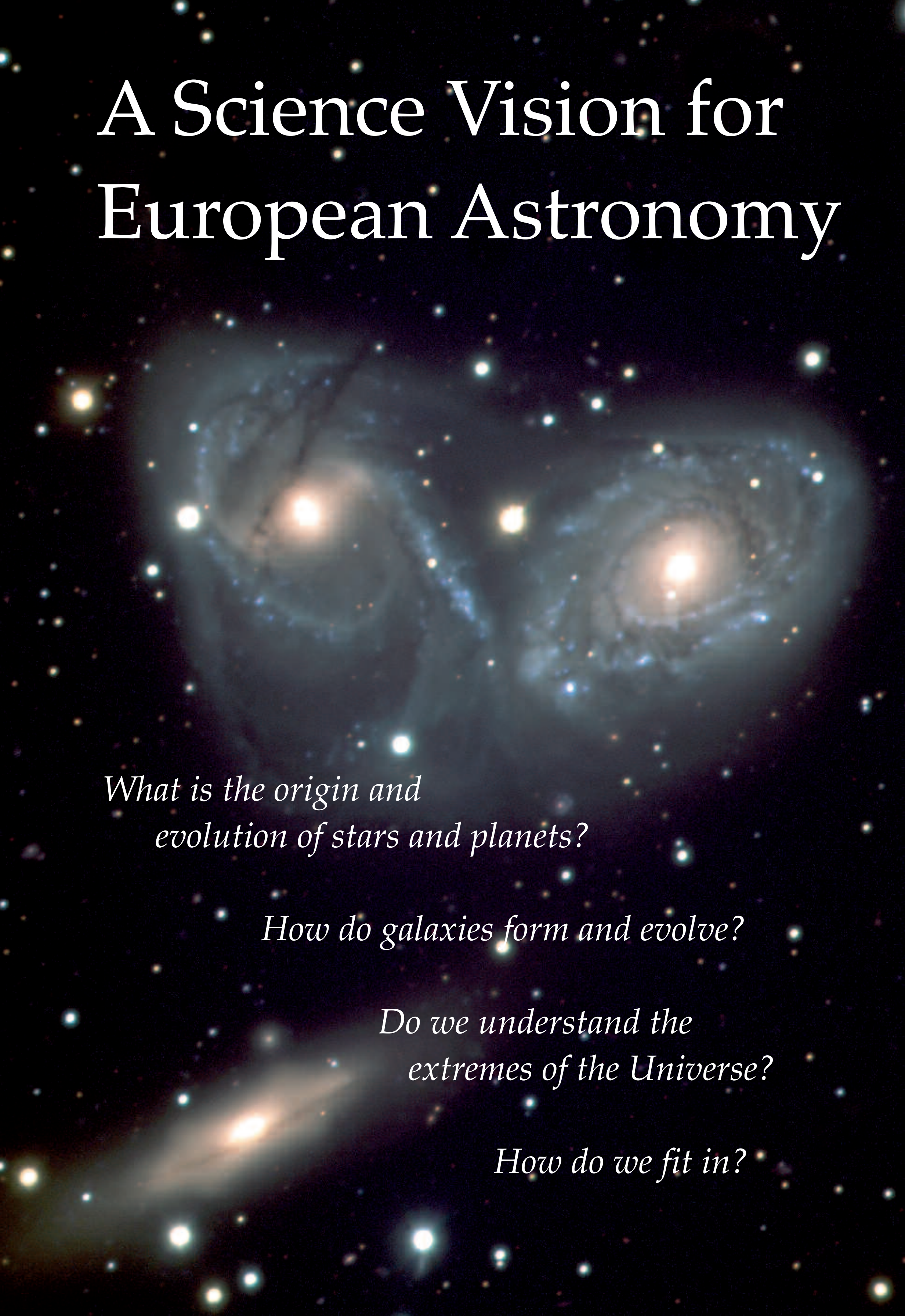
# A Science Vision for European Astronomy



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

*How do galaxies form and evolve?*

*Do we understand the extremes of the Universe?*

*How do we fit in?*

# A Science Vision for European Astronomy

This document has been created by the Science Vision Working Group under the auspices of ASTRONET, acting on behalf of the following members:
BMBF (DE), CNRS/INSU (FR), DFG (DE), ESA (INT), ESO (INT), ETF (EE), FWF (AT), GNCA (GR), HAS (HU), IA SAS (SK), INAF (IT), LAS (LT), MEC (ES), MPG (DE), NOTSA (INT) , NWO (NL), SER (CH), SRC (SE) and STFC (UK).

Editors: P.T. de Zeeuw & F.J. Molster
ISBN 978-3-923524-62-4



Front cover: The Galaxy Triplet NGC 6769-71 (ESO/VLT)

# Contents

# Preface

Astronomy is the study of everything beyond Earth. It is a science driven by observations, with links to mathematics, physics, chemistry, computer science, geophysics, material science and biology. Astronomy is important for society and culture, and helps attract young people to the physical sciences. The field benefits from and also drives advances in technology. As a result, it is now possible to study objects which are so far away that they are seen at a time when the Universe was only five per cent of its present age, and – perhaps even more astoundingly – to detect and characterize planets orbiting other stars, and to search for evidence of life.

European astronomers have access to a range of observational facilities on the ground and in space. Plans are being made for a next generation of facilities, which would continue to exploit the rapid advances in, e.g., adaptive optics, detector sensitivity, computing capabilities, and in the ability to construct large precision structures, sending probes to Solar System objects, and even bring back samples from some of them. Realizing all the plans and dreams would require substantial investments by national and international funding agencies, with significant long-term commitments for operations. For this reason, funding agencies from a number of European countries established ASTRONET, an ERA-net with financial support from the European Union, to develop a comprehensive strategic plan for European astronomy covering the ambitions of all of astronomy, ground and space, including links with neighbouring fields, to establish the most effective approach towards answering the highest priority scientific questions.

The first step is the development of an integrated Science Vision with strong community involvement, which identifies the key astronomical questions which may be answered in the next twenty years by a combination of observations, simulations, laboratory experiments, interpretation and theory. The next step is to construct a roadmap which defines the required infrastructures and technological developments, leading to an implementation plan. To this end, the ASTRONET Board appointed a Science Vision Working Group and an Infrastructure Roadmap Working Group, both of them with supporting thematic panels.

The Science Vision Working Group identified four key questions where significant advances and breakthroughs can be expected in the coming two decades:

- Do we understand the extremes of the Universe?
- How do galaxies form and evolve?
- What is the origin and evolution of stars and planets?
- How do we (and the Solar System) fit in?

These are amongst the most fundamental questions in science and generate considerable interest in the general public.

The Science Vision Working Group and four supporting panels brought together about 50 scientists with a good distribution of expertise, gender and nationalities. Each of the panels concentrated on one of the key questions, and established the approach, experiment or new facility needed to make progress. Much information already existed in national strategic plans, ESA's Cosmic Vision, and the three ESA-ESO studies. This work led to specific scientific recommendations, which were incorporated in a draft version of the Science Vision, made available to the entire astronomical community in late 2006. The draft was discussed in-depth during a Symposium in Poitiers, Jan 23–25, 2007. Many of the 228 participants from 31 countries provided constructive input, and additional comments were received via a dedicated website (see Contributions). This led to further sharpening of the scientific requirements, an improved balance across the four main areas, and improvements in the text.

This book is the result. Each of the main chapters describes the background of one of the four fundamental questions, identifies the key goals, and then summarizes the experiments or observations needed to make significant progress. Care was taken to describe these in a fairly generic way, focusing on the scientific requirements, and not to identify too closely with specific proposed implementations of missions or facilities, as this is the purview of the infrastructure roadmapping activity that will follow.

Chapter 6 summarizes the main recommendations by thematic area, and distinguishes *essential* facilities or experiments, without which a certain scientific goal simply cannot be achieved, and *complementary* ones, which would go a long way towards answering the question, but may have their main scientific driver elsewhere. Approaches using current facilities (and those under construction) are distinguished from those requiring new developments. Activities needed across the four thematic areas are identified as well. In all cases, the focus is on the most promising avenues for scientific progress, without detailed consideration of cost or technological readiness, which are the subject of the infrastructure roadmapping. Some of the more ambitious facilities may take a significant time to develop, and in some cases may only start to produce a scientific harvest towards the end of the 20 year horizon.

Exploration remains an integral component of the entire field, as many problems require investigation of large numbers of objects, all different. This leads naturally to searches and surveys with modest-sized telescopes, and follow-up by larger facilities, together with a strong programme of theory, numerical simulations and laboratory experiments. Progress relies on a healthy mix of approaches including imaging, spectroscopy and time-series analysis across the entire electromagnetic spectrum, in situ measurements in the Solar System, and use of particles and gravitational waves as additional messengers from celestial objects. Europe has the opportunity to lead the expected scientific harvest, if the Infrastructure Roadmap leads to an effective and timely implementation of the plans outlined in the Science Vision. This is a very exciting prospect.

# 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 past 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 and critical thinking**. 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 the continued well being of science. 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 could 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 science, enters our daily lives directly. This includes 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 and sunsets are red. Yet ignorance of these basic phenomena remains widespread. And while a Solar eclipse can be predicted accurately 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 Solar activity, and the complex effects of the Earth's motion, are 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, and to astronauts on a space mission.

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 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, technically, observationally and theoretically. In spite of a downturn in intake for the natural sciences at university level in general, young students have continued to enter the field at steady level, 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. Public observatories, amateur astronomical societies, and coordinated high school educational projects are all fostering this interest.

Europe 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, in particular through the European Southern Observatory. These have already contributed substantially to advances in science and technology, and have motivated talented 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 (with five named laureates) 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. Their discovery through astronomy now drives 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, extrasolar 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. 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**. Europe's astronomers have access to many optical telescopes in a range of sizes and capabilities, including Solar telescopes, and to a large number of radio telescopes, both single-dish and interferometers. Figure 1.1A illustrates some of these. The largest optical/infrared telescopes are equipped with state-of-the-art instruments, including many that take advantage of progress in adaptive optics. Some are even being linked interferometrically to obtain milli-arcsecond resolution. Numerous large-scale sky surveys are available, and many optical/infrared telescopes in the 2–4 m class now concentrate on challenging wide-field imaging surveys, on obtaining radial velocities for millions of stars in the Galaxy using multi-fiber spectroscopy, or on dedicated programmes to find exoplanets.

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 (Figure 1.1B). Astronomy missions dedicated to specific topics include RXTE, SWIFT, Akari, and CoRoT. SOHO, Ulysses, and Cluster are studying the Sun, its surroundings, and the Earth's magnetosphere. 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, and the Mars rovers Spirit and Opportunity continue to provide stunning science and remarkable images.

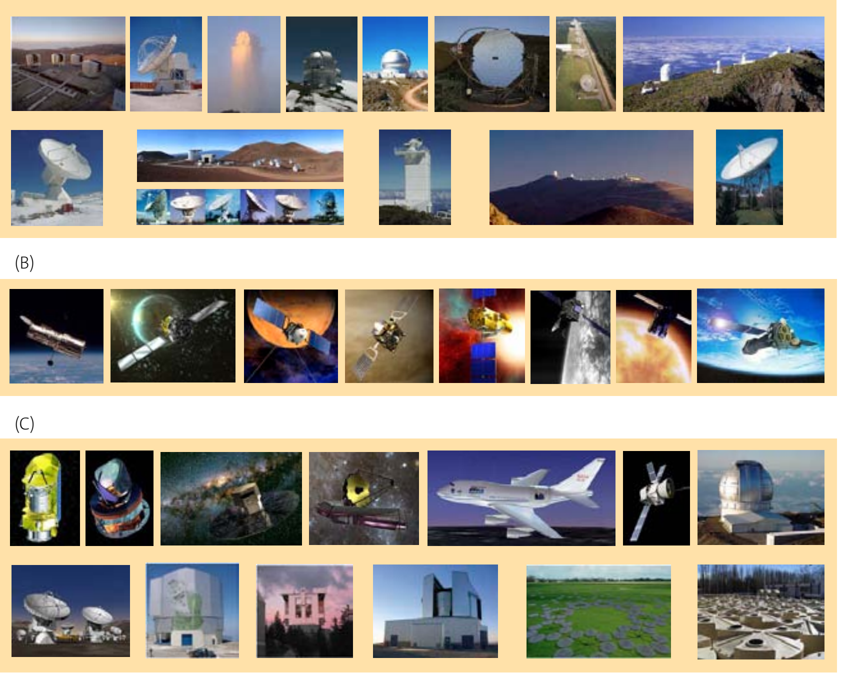
(A)
(B)
(C)


Figure 1.1: A selection of observatories from left to right from top to bottom, ground-based observatories (A): VLT, APEX, GREGOR, Gemini North, Gemini South, Magic, WSRT, Roque De Los Muchachos, IRAM, SMA, Merlin Telescopes, La Silla and Effelsberg; space-based observatories (B): HST, Integral, Mars Express, Venus Express, CoRoT, Rosetta, SOHO and XMM-Newton; observatories under construction (C): Herschel, Planck, Gaia, JWST, Sofia, BepiColombo, GTC, ALMA, VST, LBT, VISTA, LOFAR and Auger .

Figure 1.1C illustrates some of the observational facilities currently under development. These include space missions such as Herschel, Planck, Kepler, Gaia, BepiColombo and the James Webb Space Telescope, as well as the stratospheric observatory SOFIA. On the ground, the GTC and LBT are nearly finished. New survey telescopes are being completed. Various new 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 also see the full power of the 8–10 m class optical/infrared telescopes exploited, with second-generation instruments and interferometric links, and 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 telescopes every three years over the past seventy years (Figure 1.2, 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 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 the optical and the (near-)infrared, eliminating the need for dispersive devices in spectroscopic applications, and resulting in further gains of throughput and sensitivity.

Substantial improvements in capabilities can be expected in the angular resolution of astronomical measurements. Figure 1.2 (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 1.2. 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. This technique was pioneered in radio astronomy, but 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 a highly challenging and somewhat experimental technique, further progress in single-mode optical fibers, integrated optics, lasers and fast control systems is expected to make (sub) milli-arcsecond

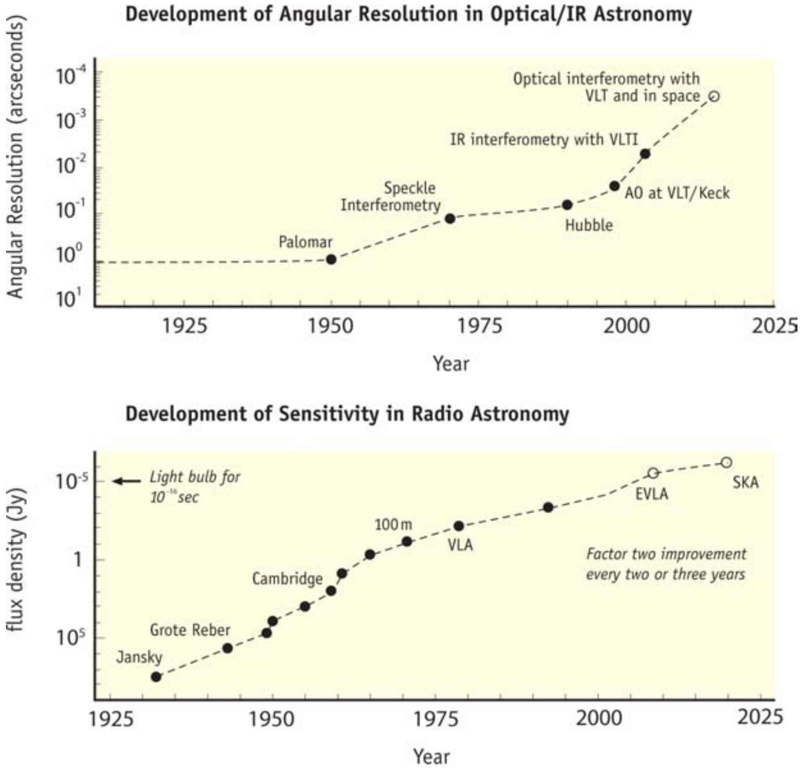



Figure 1.2: Top: Improvement in angular resolution in optical/infrared astronomy. Bottom: Improvement of sensitivity with time in radio astronomy.

imaging interferometry widely applicable in 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. Large integral-field spectrometers and energy-resolving devices, in combination with ever larger imaging detectors, will allow very significant progress in 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 enable a new generation of fundamental time and frequency measurements.

Advances in remote sensing instruments, solar electric and micropropulsion technology, radiation-hardened electronic circuits, digital instrumentation, highbandwidth 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 into the future.

## 1.3 Predicting the future

These technological developments now make it possible to observe planets orbiting other stars, as well as peer deeper than ever into the Universe. This 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 gravitational waves, next generation X-ray telescopes and space missions devoted to characterizing extrasolar planets. Detailed science cases are available for all these missions and facilities.

## 1.4 This document

The plans for astronomy are ambitious. Within Europe, they would require a collective investment of several billion Euros for new instrumentation and associated hardware and operations, spread out over the next two decades. While some funding will be pursued through programmes of the European Union, the bulk of the support will only be accessible from the national funding agencies.

The national funding agencies have proven to be very supportive of astronomy and Solar System exploration over many years, but 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', to be followed by the preparation of a road map for the development of the required infrastructures. Together this will provide the long-term planning directed towards achieving a world-leading position for European astronomy.

This document constitutes the Science Vision, developed by a working group with four supporting panels, comprising in total about 50 scientists (see Contributors). A draft version was distributed widely, and served as the basis for an in-depth discussion during a three-day Symposium in Poitiers in January 2007. Many of the more than 200 participants provided thoughtful and constructive comments. Further input was received via electronic means. The resulting final version serves as input for the development of the infrastructure road map.

In each of the four chapters that follow, the present situation is described, including the expected capabilities of facilities under construction. Then 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 ESA's Cosmic Vision document, and from the ESA–ESO working group reports. Since the ERA-net ASPERA programme 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 summarizes the scientific recommendations. The 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 explain what we see in the sky using exactly the same physical laws that we can measure in the laboratory. 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, such as in the case of gamma-ray bursts, the appropriate physics at work can be highly uncertain. Conversely, astrophysics also offers unique possibilities for probing fundamental physics beyond the level that can be explored in the laboratory. The examples are gravity in the strong-field regime, and particle physics at energies above the TeV scale. From this point of view, astrophysics becomes a prime arena where the frontiers of physics can be advanced. The subject cannot 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 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. To round off 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 remains the ultimate goal. Direct imaging of black hole event horizons will require sub-millimetre and infrared interferometry. Future large X-ray and gamma-ray observatories will also probe these inner regions. The understanding of the nature of supernova explosions and gamma-ray bursts will benefit from the same instruments, as well as from future extremely large optical/infrared telescopes. Cherenkov 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. It is not even sure there has been a beginning; the Universe has perhaps existed forever; what is 'time'? It may seem presumptuous to claim that we have made any progress in replying to those questions. 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 almost always described using the powerful general idea of cosmological inflation. According to this idea, the Universe was set expanding by the 'antigravity' effects of vacuum energy, which is supposed to be very much higher at these early times than it is today. Inflation permits a nearly uniform Universe, with negligibly small spatial curvature. More importantly, it predicts that quantum fluctuations leave behind a spectrum of small density perturbations, which are the seeds for galaxies and anisotropies in the cosmic microwave background (CMB). As a bonus, the theory also predicts relic gravitational waves, and these may be the best way of testing inflation. The amplitude of the gravity waves tells us the energy scale of inflation, and how close it lies to the scale of quantum gravity. By studying both the gravity waves and the density perturbations, we can hope to achieve an understanding of the dynamics of inflation – particularly the means by which the vacuum density was reduced from a high early value to its much lower present level. At present, a range of models exist for this process, but these are built on speculative physics beyond the well-understood Standard Model of particle physics. Inflationary cosmology therefore opens the exciting possibility of doing proper study of the origin of the Universe – and learning new physics in the process.

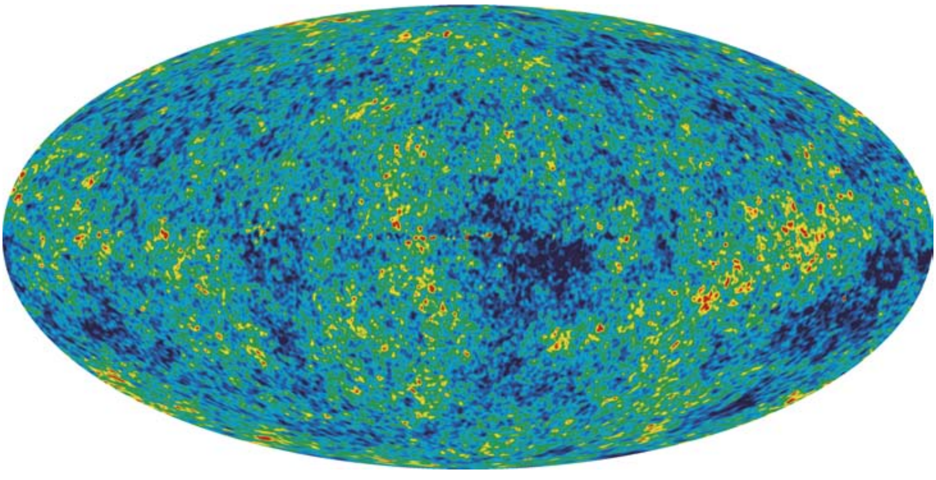


Figure 2.1: A map of the Universe showing the difference in the intensity of the microwave background. The maximum difference in this image is less than $10^{-5}$. The map has been produced with the data of WMAP. [NASA/WMAP Science Team]

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, and much can be inferred from their presence (or absence). In addition to features of inflation, the most widely discussed relics are weakly-interacting dark matter particles, and also so-called topological defects such as cosmic strings, which are structures in primordial scalar fields, analogous to vortex rings in fluids.

### 2.1.2 Key observables

Inflation predicts a very nearly fractal spacetime: One where the deviations from uniformity look the same when viewed on any scale. Technically, this is said to be equivalent to a spectral index $n_s = 1$. But in fact inflation predicts a slight 'tilt' of the spectrum; for the simplest models, this is in the direction of $n_s < 1$, implying greater uniformity on small scales.

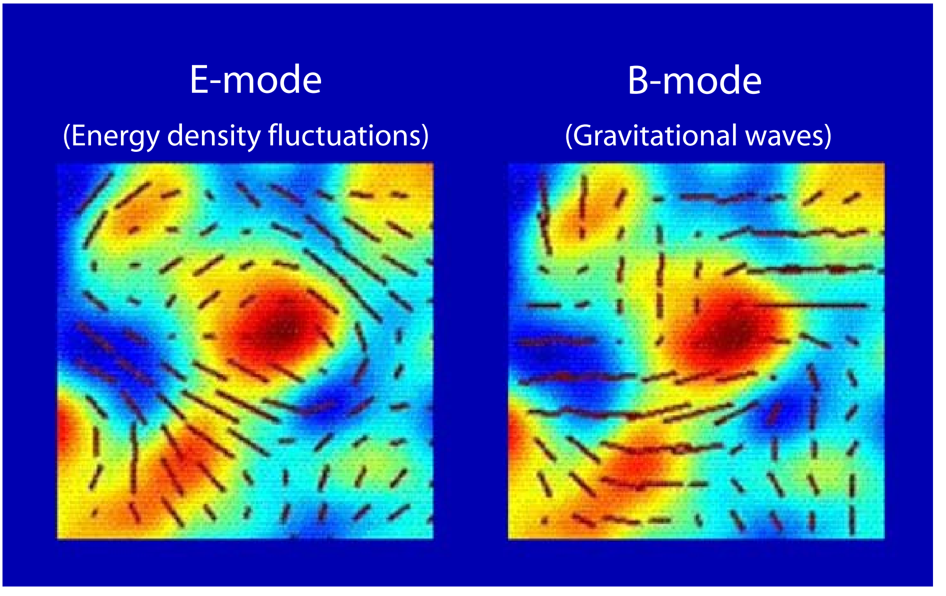



Figure 2.2: An illustration of possible signals in the polarization of the cosmic microwave background. The colour shows the CMB total intensity and the 'sticks' indicate the direction of polarization. The first panel shows scalar ('E-mode') signals, which arise from simple density fluctuations. The 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. Adapted from Seljak & Zaldarriaga (1998).

The data from the WMAP satellite (Figure 2.1) reveal that the spectrum is indeed probably tilted: The fitted value $n_s = 0.95 \pm 0.02$ indicates that the Universe is slightly more uniform on the smallest scales than expected. The data provide 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 (Figure 2.2), 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 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 space interferometer such as the proposed LISA. 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 nano-Kelvin 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 these will be essential in pursuing an understanding of the early Universe.

Direct detection of high-frequency gravitational-wave relics from inflation is predicted to be challenging according to simple models. Larger signatures are possible, however, in models with topological defects. In addition to possible detections with ground-based facilities 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 proposed SKA will have the sensitivity and frequency coverage to discover and time about a thousand millisecond pulsars for this purpose.

Finally, particularly in this area, we have to be continually aware of the possibility of complementary input from laboratory experiments, such a searches for fifth-force deviations from standard gravity. There is also a need to capitalize on developments in supercomputing: Processing of next-generation experimental data, 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 establishing an alternative to General

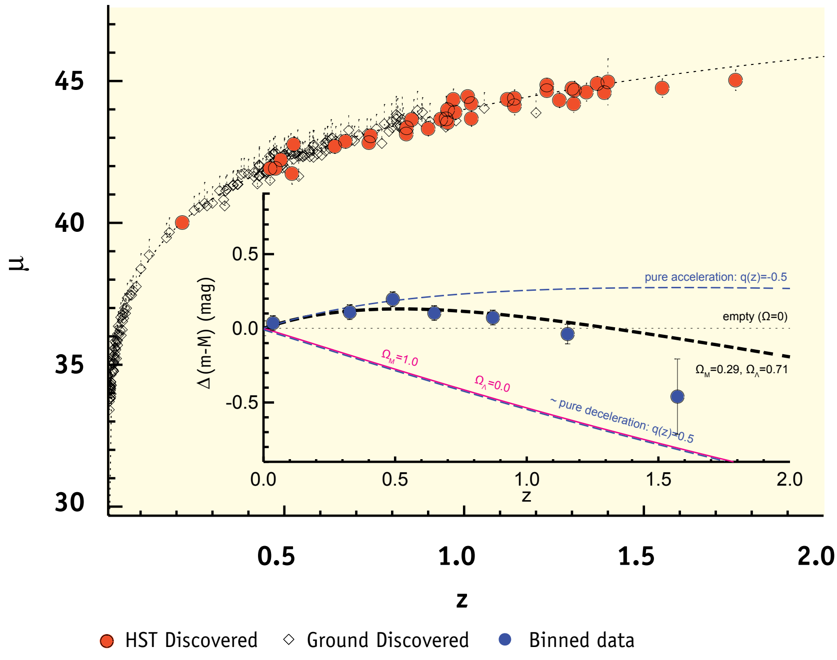



Figure 2.3: Hubble diagram for Type Ia supernovae (§ 2.4) observed with ground-based telescopes and the Hubble Space Telescope. The best fit line has $\Omega_M = 0.29$ and $\Omega_\Lambda = 0.71$ (Riess et al. 2007, ApJ, 659, 98).

Relativity that is physically and cosmologically viable) is a key challenge in cosmology (and fundamental physics).

A combination of recent 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 $\Lambda$. The chief evidence for this conclusion is that the expansion rate is accelerating (from distances to supernovae, see Figure 2.3); that the Universe is flat (from the CMB); and that dark matter cannot provide the critical density (from large-scale structure and galaxy clusters). 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 models allow quintessence to 'track' the energy density of dark matter, but fine tuning is still necessary to obtain the negative pressure required to drive accelerating expansion. New physics possibilities emerge if the quintessence field can interact with dark matter: The changing vacuum energy can then have the effect of inducing time variations in fundamental numbers such as the fine-structure constant, which could be detected in spectroscopy of high-redshift objects. Any interaction between quintessence and dark matter would also amount to the existence of a new long-range force, which would cause an apparent violation of Einstein's equivalence principle.

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 Friedmann–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 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. Detection of plausible particle candidates such as supersymmetric neutralinos at accelerators such as the LHC would be a major advance, although it could not be guaranteed that such particles are sufficiently long-lived to be cosmologically relevant. We consider such possibilities to be of the greatest importance, but beyond the scope of astronomy. For details, one should consult the parallel ASPERA science vision in particle astrophysics (`www.aspera-eu.org`). 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 Cherenkov detectors including H.E.S.S. and MAGIC, as well as by 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 $\Lambda$ 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 $w$ that defines the equation of state of the Universe ($P = w\rho c^2$, where $P$ is the pressure, $\rho$ 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 crucial; any inconsistency between them could indicate a failure of General Relativity.

### 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 progressively dominated by dark matter, so the Universe is expected to be decelerating at redshifts above unity (and it is an important cross-check that this can be seen in the data on supernovae at these distances). 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**. The CMB contains structure down to scales of about 0.1 degrees. ESA's Planck satellite will measure this pattern to effectively perfect precision as far as the temperature pattern is concerned. However, the polarization of the CMB is another story. Here, a post-Planck satellite will be required, in particular to probe the scales larger than a degree, which contain most of the information on inflationary gravitational waves. Such CMB data will be a common basis for many other tests, lifting the degeneracies inherent in these techniques. However, this is only possible if foreground emission from the Milky Way can be understood; to remove foregrounds, deep radio surveys including polarisation need to be performed.

**Gravitational Lensing**. This technique can probe dark energy via the distance-redshift relation, and by its impact on the growth of density fluctuations. Lensing is thus in principle as direct a probe of cosmology as the CMB, although some applications do 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. Possible approaches are to construct a very large CCD imager (or for Europe to become a partner in a venture such as the proposed LSST or Pan-STARRS), or to exploit radio measurements of gravitational image shear and redshifts of neutral hydrogen. Optical lensing studies need a combination of ground-based and space-based data: The latter gives improved and stable image quality for image distortions, and also exploits the low background levels in space to probe the near-infrared, which is required for accurate and reliable redshift estimates.

**Large-Scale Structure and Clusters**. The baryon-oscillation signature of sound waves in the spectrum of large-scale structure can be measured using a 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 should also be possible using H I surveys with the SKA, at least out to redshifts of 1.5. A detailed understanding of the relation between galaxy tracers and dark matter will be required, so a number of independent ways of selecting galaxies is desirable. Constraints on dark energy from such large-scale structure geometrical probes can potentially deliver accuracy comparable to lensing (error on constant $w$ of $\sim 1\%$). Similar precision may also arise from the non-linear part of structure formation: The mass function of clusters and their baryon fraction. This would require $> 100.000$ clusters, best selected from an all-sky X-ray survey.

**Supernovae**. The current supernova surveys will measure a constant equation of state parameter $w$ to around 7 per cent accuracy. 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 (§2.4).

Beyond these 'standard' tests, there are various 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 come from Gaia astrometry and radio pulsar timing. Independent measurements of cosmic acceleration may also be made by LISA if it proves possible to identify some of the galaxies containing massive black hole mergers.

Large-scale effects manifest themselves through modifications of the Friedmann 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 geometrically, via precise mapping of absorbers on the line of sight to quasars. We might also search for varying fundamental constants, either via precision timing with atomic clocks, or via detailed spectroscopic structure of atomic multiplets versus redshift.

## 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. The gravitational radius of an object is just proportional to its mass; this radius is 1.5 km for the Sun. Gravitational astrophysics becomes more exciting for denser objects that are closer in size to their gravitational radius: The radius of a neutron star is about seven gravitational radii 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

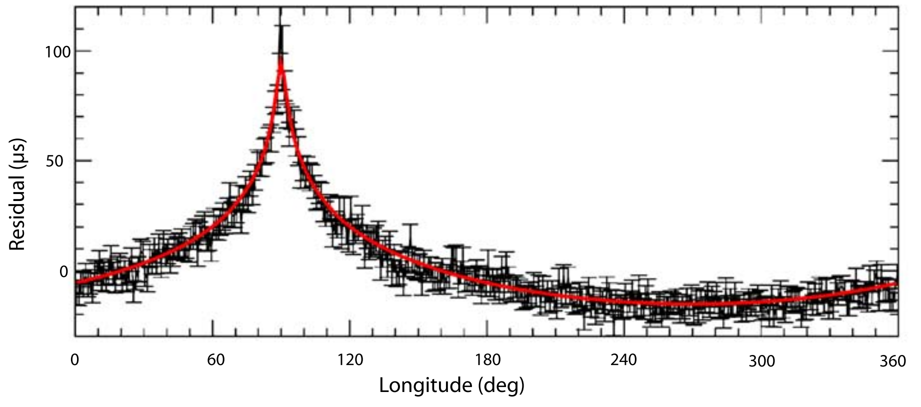



Figure 2.4: The Shapiro time delay measures the relativistic time delay experienced by the pulses from one of the pulsars in the double pulsar system PSR J0737-3039A as they pass through the strong gravitational field of the other neutron star. The red line shows the predicted delay based on General 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, Science 314, 97).

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 this 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 (Figure 2.4), 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 causes 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 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 known stars orbiting the three million solar mass black hole in the Galactic Centre, coincident with the radio source Sgr A*, will probe relativistic gravitational effects down to hundreds of gravitational radii in the next few years (Figure 2.5). The next step is to attain greater sensitivity and resolution, to detect objects orbiting in the strong-field regime. The strong gravity regime can also 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. 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.

It is normally assumed that an accretion flow around a black hole is efficient in the sense that gravitational binding energy is radiated away as the accreted material descends towards the black hole. In this case, the most luminous part of an accretion flow will be the innermost region. Comparison of the light emitted by accreting massive black holes (in 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 spins quite rapidly. The inner luminous parts of the flow are then well within the strong gravity regime. The inner regions are strong X-ray sources and often display rapid variability. 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 of galactic black hole binaries show extreme broadening, indicating that the accretion disc extends down to around two gravitational radii; this would be the innermost stable orbit (larger than the horizon) if 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 cleanly.

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.

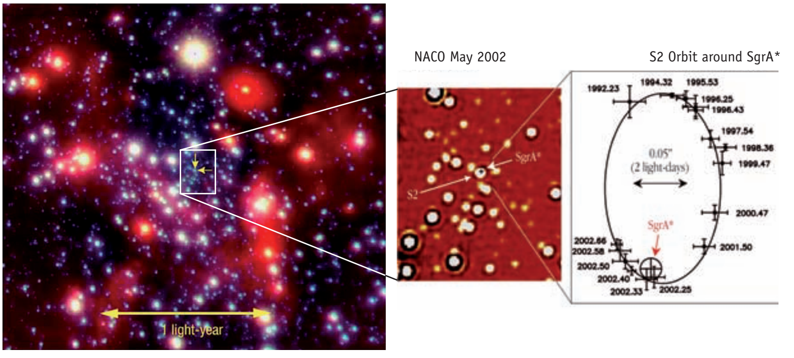



Figure 2.5: The centre of our Galaxy (Sgr A) and the star S2 which orbits it. The measured motions of this and other stars reveal the presence of a black hole of about three million solar masses in the centre of the Galaxy. Based on work by the groups of R. Genzel and A. Ghez. Images courtesy of ESO.

### 2.3.2 Experiments

Event horizon imaging of the Galactic Centre can be achieved by combining individual sub-millimetre telescopes in VLBI-like observations with baselines of $\simeq 4000$ km. 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 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. For massive black hole mergers the frequencies are milli-Hz and a space-borne interferometer is required. Radio pulsar timing can detect gravitational waves in the nano-Hz band.

The electromagnetic emission from the inner regions of black-hole accretion flows is dominated by the X-ray output. It is especially fruitful to be able to study this in the time domain, which implies not only a large-area telescope, but also a facility for wide-area monitoring – identifying X-ray transients and allowing them to be correlated with other observations.

A Galactic census of pulsars with a 1 $km^2$ collecting area radio telescope in a compact configuration would find about 100 relativistic binary pulsars and pulsars orbiting stellar black holes and Sgr A*. This would allow much improved tests of General Relativity.

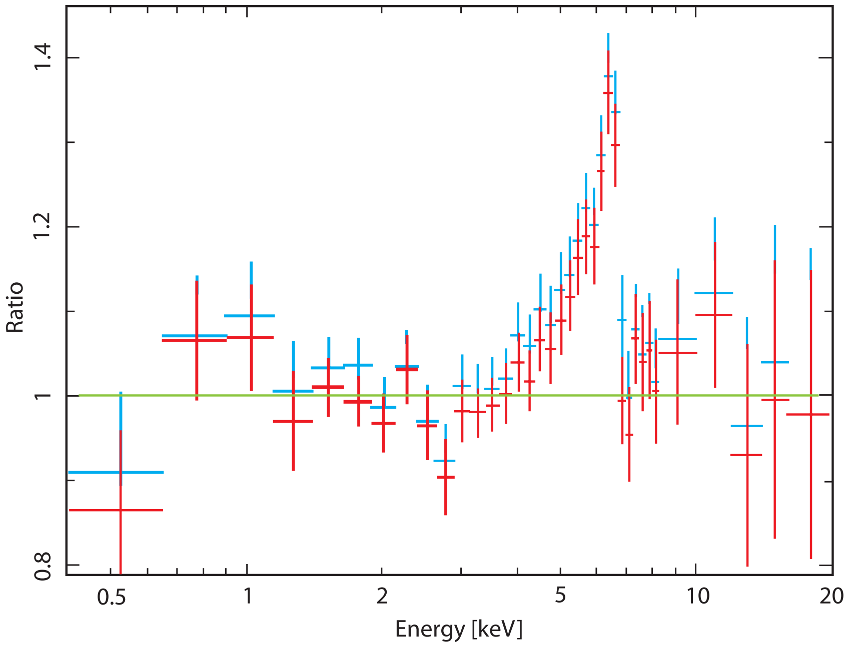



Figure 2.6: XMM-Newton spectrum showing a relativistic Fe K-line around active galactic nuclei. The plot shows the ratio between the actual spectrum and a smooth continuum fit. Image courtesy of MPE and ESA.

The broad X-ray iron lines from active galactic nuclei have low fluxes (Figure 2.6), and measuring the lines on timescales comparable to the light crossing time of the strong gravity region requires collecting area an order of magnitude larger than XMM-Newton, such as envisioned for the proposed missions XEUS and Constellation-X. The lines are brighter in 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 these 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 (Figure 2.7).

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 cosmolo‘gy (§2.2.1).

The only direct observing window into the core collapse event is provided by neutrinos. Their detection 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-

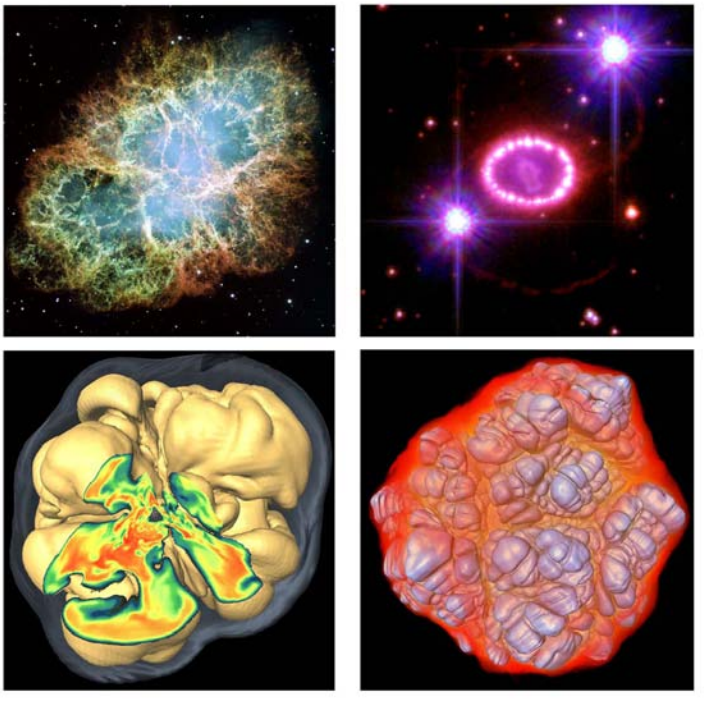


Figure 2.7: Top left: The Crab nebula, a supernova remnant resulting from the historical supernova of 1054 (image courtesy ESO). Top right: HST image of the remnant resulting from SN 1987A in the Large Magellanic Cloud. Bottom: Two three dimensional simulations, left a core collapse supernova and right the resulting supernova of a thermonuclear collapse. Pictures courtesy of Janka (left) and Hillebrandt (right).

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.

### 2.4.2 Key questions

Which stars explode as supernovae? For thermonuclear supernovae, also called Type Ia, the association with white dwarfs is still unsupported by direct evidence. For core-collapse supernovae, we have to date 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 (Figure 2.7) 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?

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

High spatial resolution observations of future explosions, both before and after the event, will be required for a secure identification of 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. A radio telescope with a 1 $km^2$ collecting area is expected to detect 20 000–30 000 pulsars. This will provide excellent statistics of a direct birth product of supernovae and their kick velocities.

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 Cherenkov 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 the observed gamma-rays with the protons accelerated by the internal shock(s). 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 Cherenkov 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 proposed SKA would provide access to long radio wavelengths, which are important for using synchrotron self-absorption as a diagnostic. All-sky monitors, such as LOFAR in the radio, or an orbiting X-ray monitor, will detect the supernova and gamma-ray-burst blast waves as early as possible and follow their evolution. Sensitive X-ray telescopes, such as XEUS, will provide complementary follow-up in the X-rays, offering diagnostics of both line and continuum emission.

Ground-based gravitational wave detectors will, within 6-8 years, be sensitive enough to be able to determine whether short gamma-ray bursts are associated with mergers of neutron stars or of neutron stars with black holes, and to provide statistics on the narrowness of their gamma-ray beams. These detectors may also be sensitive enough to detect gravitational waves from the next supernova in the Galaxy, which would provide unique information on the explosion mechanism and the nuclear physics deep in the core.

The invaluable information on supernovae and gamma-ray bursts provided by new observational facilities will have to be supplemented by complex numerical simulations of the explosions, requiring the largest supercomputers. For the white dwarf explosions the computational requirements are daunting as the thermonuclear 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 the core-collapse simulations, one has to add the neutrino physics and a full description of general relativistic effects. Access 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 galactic nuclei 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 its 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 into a black hole (§ 2.3.1). 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 that can transport matter at high, sometimes relativistic, speed over intergalactic distances (Figure 2.8). 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 lack predictive power. It is clear, however, 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

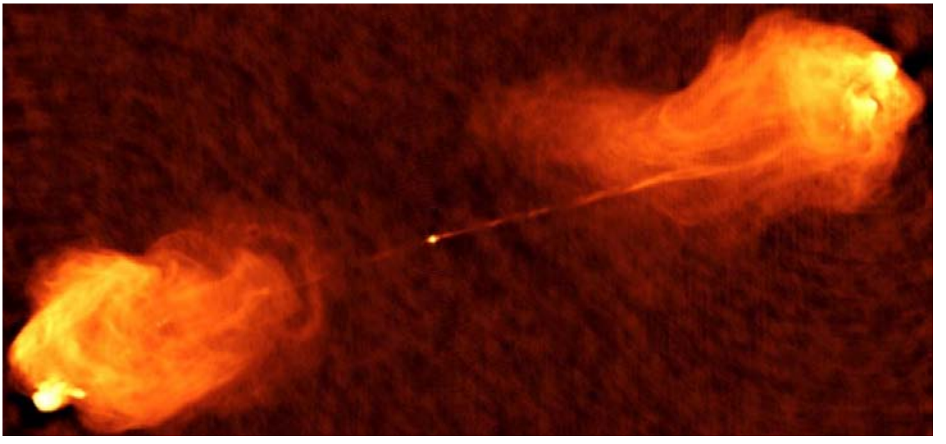


Figure 2.8: Radio image of the jet of the active nucleus of the radio galaxy Cygnus A. Image courtesy of NRAO/AUI.

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 thus accreting black holes and neutron stars.

The big technical challenge in the study of these objects is the physics of multidimensional 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 probe 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: Cosmic rays, gravitational waves and neutrinos. Many experiments to observe these are being constructed or planned. They will provide a truly multi-messenger approach to the study of these extreme and enigmatic objects. Large-scale numerical simulations of collisionless shocks, tenuous plasmas with charge separation, and particle acceleration in these environments will be a key ingredient to understand these phenomena.

The key questions to focus on for the foreseeable future are: How can 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 throughout the entire electromagnetic spectrum, from radio galaxies at the lowest frequencies 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).

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 therefore 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 a large collecting area X-ray satellite, rapidly responding telescopes such as robotic optical telescopes or electronically-steered phased arrays in the radio regime such as LOFAR. 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 submillimetre 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 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 $10^5$ GeV. The observed hard X-rays are consistent with inverse-Compton scattering of these synchrotron photons 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 that 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 an $E^{-3}$ 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 are very distant and therefore very energetic, perhaps gamma-ray bursts or active galactic nuclei. 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 $5 \times 10^{10}$ GeV. Presently there

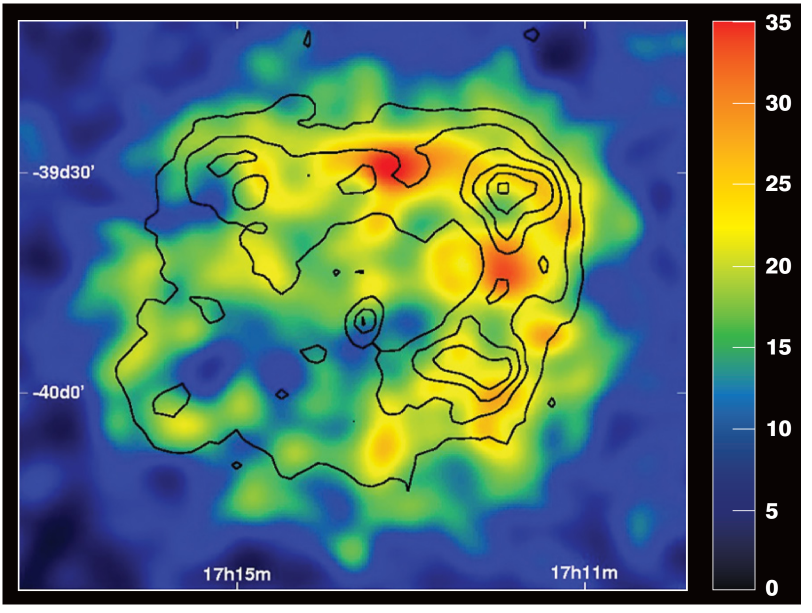



Figure 2.9: TeV gamma-ray image of the supernova remnant RX J1713.7-3946 obtained with the H.E.S.S. telescopes. The superimposed contours show the X-ray surface brightness seen by ASCA in the 1–3 keV range. [H.E.S.S. Collaboration]

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 from the extragalactic cosmic ray sources should also be detectable 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 ultra-high 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. Radio cosmic-ray detection techniques are also capable of achieving the large detection volumes required by these science goals. 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 Cherenkov arrays, such as H.E.S.S. and MAGIC, have demonstrated the power of these instruments for understanding the acceleration of particles in several classes of high-energy objects (Figure 2.9). Next-generation arrays, covering $\sim 10$ GeV to several TeV in energy, should greatly improve the understanding of the nature and acceleration mechanism of the particles.

Direct measurements are being made with balloon- and space-borne experiments (see `www.aspera-eu.org` for details). Medium-scale air shower experiments such as KASCADE-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 is being constructed. 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. At energies above $10^{11}$ GeV, extremely large detector volumes are needed. Novel methods such as radio detection are being investigated, and experiments that 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 are 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 (elements heavier than boron), 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 the formation and evolution of galaxies, and the role of the intergalactic baryons and metals, 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?

## 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.

### 3.1.2 Key questions

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

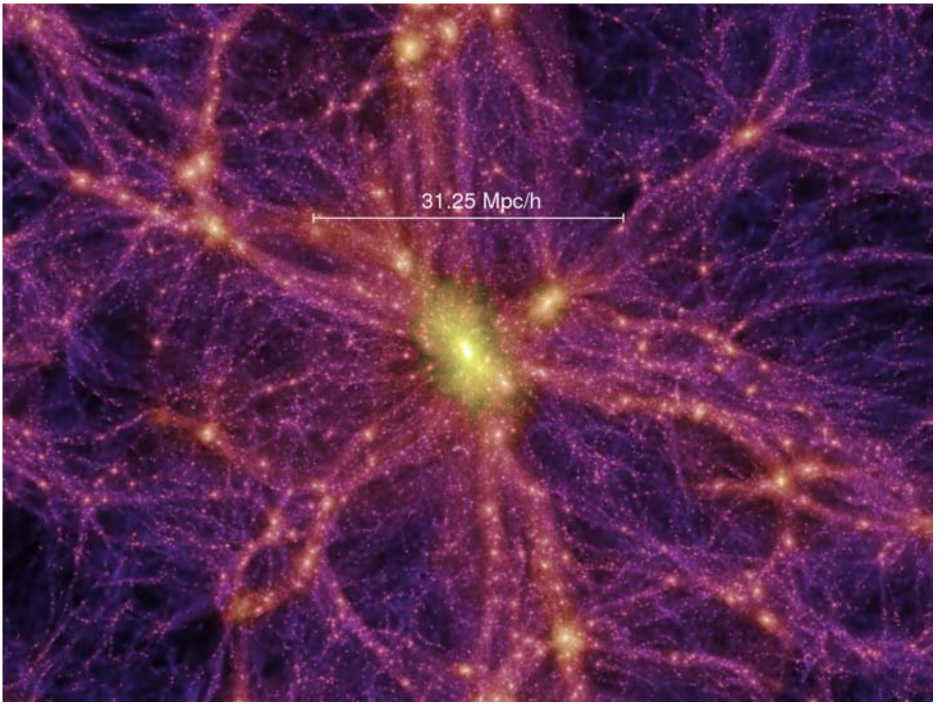



Figure 3.1: Image of a cluster of galaxies as calculated by the largest cosmological N-body simulation carried out to date. It has been used to construct sophisticated semi-analytic models of galaxy formation that cover the complete galaxy population in a representative piece of the Universe down to stellar luminosities. [Image taken from http://www.mpa-garching.mpg.de/galform/millennium/].

imprints of the Dark Ages in the metallicity evolution/pattern of the intergalactic and intracluster medium, in low-redshift Ly-$\alpha$ emitters and in metal-poor halo stars in the Galaxy.

**The neutral Universe**. The most promising approach 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, HI. This line, in principle, allows a superb tracing of the HI in the early Universe, and therefore a reconstruction of the reionization history as governed by the first luminous sources. 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 HI is locked to the kinetic temperature of the gas, as long as the collisional excitation rate of this transition 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-$\alpha$ 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 of 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-$\alpha$ 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 current simulations (like the one of Figure 3.1). 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 (§ 4.1). 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 poorly understood efficiency of dust formation in the ejecta of the first supernovae. 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 a collapsing very massive star en route to the formation of an intermediate mass black hole could result in a gamma-ray burst. Understanding when the first stars formed and whether the very different primordial environment (no metals, no dust, no molecular clouds, 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 (Figure 3.2). 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. This evolution is governed by complex and interlinked physical processes ('feedback'): The ultra-violet 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 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 activity in the near future. The Planck satellite will allow improved measurements of the CMB temperature and polarisation spectra. Sub-millimetre observatories including Atacama Cosmology Telescope (ACT) and ALMA will explore for the first time the sub-arcminute features resulting, amongst other processes, from patchy reionization. High-redshift quasars, radio galaxies and gamma-ray bursts can be used as targets for absorption-line experiments at high-spectral resolution, $R \sim 10^4 - 10^5$ (in the near-infrared with the 8–10 m class telescopes and ELTs; in the radio with E-VLA, ALMA and the proposed SKA), 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 detailed effects of the different feedback types. At least four types of feedback processes occur during the Dark Ages: Radiative, chemical, mechanical and magnetic. The first process, through the emission of ultra-violet photons from the sources, governs the formation/destruction rate of the key gas-cooling species (molecular hydrogen $H_2$), and hence the collapse of the first

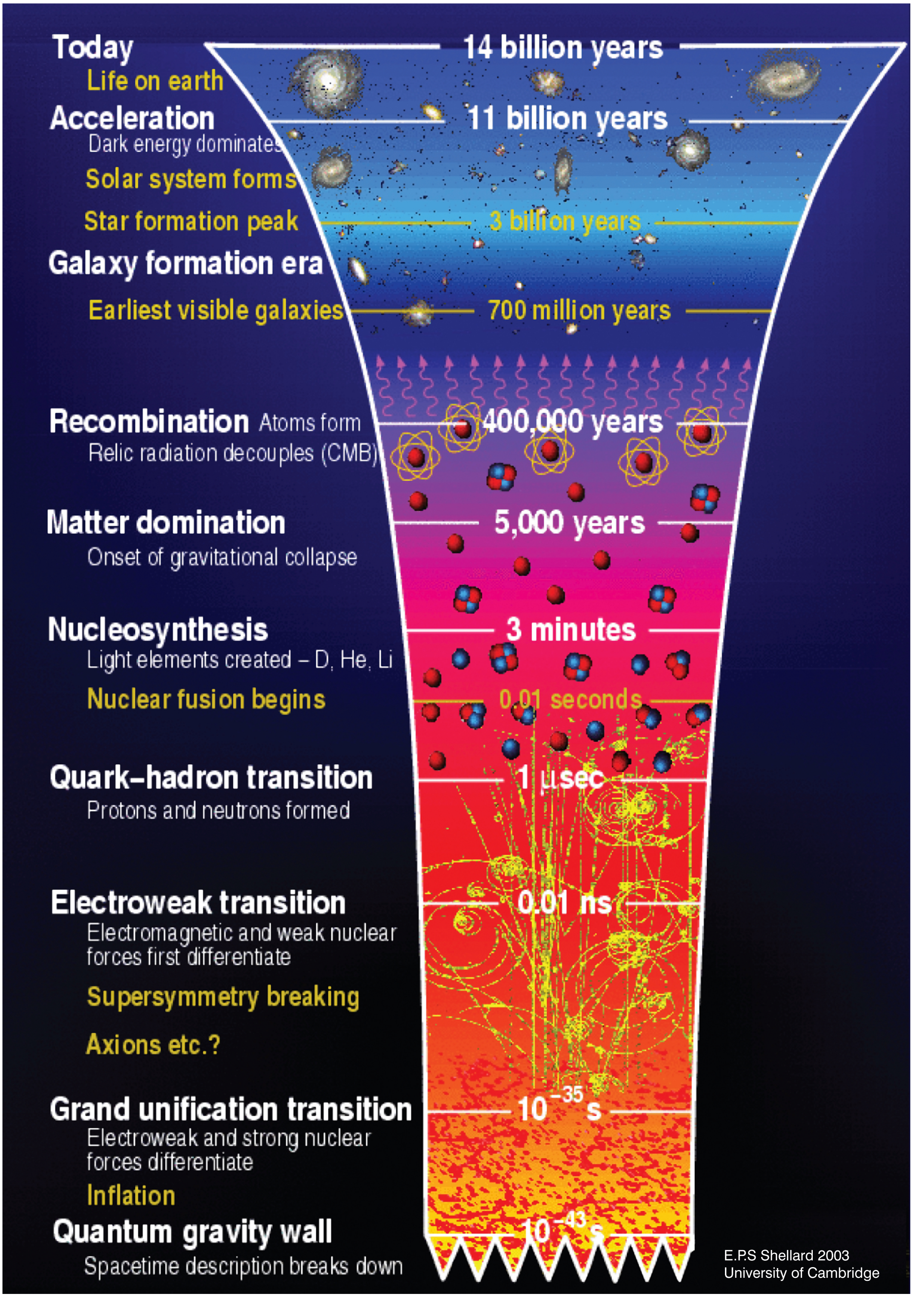



Figure 3.2: The reionization era placed within the evolution of our Universe. Time is going up with the start of the Universe at the bottom and today at the top. Image from Planck: The Scientific Programme (ESA).

(mini-)galaxies. Chemical feedback causes 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 reliable 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 (IMBHs) which could provide a link between stellar evolution and quasar activity. Whether such objects 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, 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.

### 3.1.3 Future strategy/opportunities

Extending 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-$\alpha$ 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 micrometre 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-$\alpha$ 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 a far-infrared/sub-millimetre interferometer in space.

The future of reionization studies relies heavily on the detection of anisotropies on sub-arcmin scales induced by Thomson scattering of the CMB by the ionized gas in massive dark matter halos. This requires detecting relative fluctuations of a few times $10^{-5}$ at high frequencies ($\sim 100$ GHz) ($R \sim 40000$) with 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 large collecting area radio telescopes could be ground-breaking in the study of the Dark Ages via the measurement of the red-shifted H I 21 cm brightness temperature of a large number of independent patches in the sky. In order to succeed it will be necessary to control radio-frequency interference of both cosmic and terrestrial nature, and to develop next-generation receivers. The predicted signatures of feedback 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 require enhanced numerical simulations of both structure formation in large cosmic volumes, and early evolution of protostellar cores with inclusion of detailed physical processes.

Figure 3.3: Detailed images of the galaxy cluster Abel 1689 showing numerous arcs and arclets that are images of 'lensed' background galaxies distorted by the gravitational field of this massive cluster. Picture taken from the HST image archive.

## 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 (§ 2.2). 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 matter.

Galaxy clusters and groups are identified through optical, near-infrared observations either as over-densities in space, and/or from their colours, or through the X-ray emission of the intracluster medium (ICM), i.e. the hot intergalactic gas that accreted into potential wells of the dark matter. Dark matter wells are identified through the distortion of light from background galaxies (see Figure 3.3). About ten thousand galaxy clusters (defined as structures with a mass larger than $10^{14}$ solar masses), and groups (with smaller masses) have been found. Several forthcoming experiments 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 intracluster plasma. Galaxy clusters cast shadows on the microwave sky below 217 GHz (frequency corresponding to the peak emission of the CMB) and shine above. The observed frequency dependence of this effect is the same for all clusters, irrespective of the cluster redshift. This so-called Sunyaev-Zeldovich effect will be easily detectable with the Planck satellite (see Figure 3.4). 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 ICM is strongly affected by merging, star formation and energetic outflows from active galactic nuclei. Conversely, the injection of energy and metals 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 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 early evolution of the cosmic web during the reionization era, together with the growth of ionized regions around protogalaxies, will be revealed by maps of the H I 21 cm brightness temperature on the sub-arcmin scales (regions of negative brightness temperature relative to the mean H I signal). A radio telescope covering a wide wavelength range to observe a range of redshifts and with a sensitivity one order of magnitude larger than now available is necessary for this. At later times (redshifts 2–4), the growth of (dark and luminous) matter 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

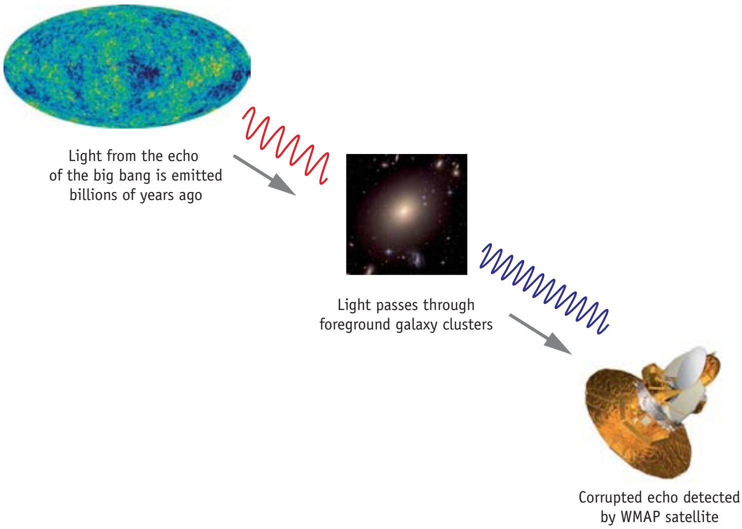



Figure 3.4: Foreground hot electrons distort the microwave background radiation. This is called the Sunyaev-Zeldovich effect and can be detected by instruments that observe the microwave background. This effect is distance independent and will therefore probe all hot electron clouds, in particular those associated with galaxy clusters, between us and the source of the microwave background radiation.

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-$\alpha$ absorbers with an ELT, using high-redshift quasars and compact, luminous galaxies as background targets. This requires observations in the optical at a spectral resolution $R \sim 10^4$ for faint targets with magnitudes down to R(AB) $\sim$ 24.

At redshifts 1–2, the mass distribution of dark matter halos will be given by large-area weak-lensing surveys. This is one of the planned observing projects with the VLT Survey Telescope (KIDS) which will be complemented by near-infrared surveys with VISTA. Similar efforts in the USA include the Dark Energy Survey (DES), Pan-STARRS and the proposed 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.

Clusters of galaxies at high redshift will be discovered with Sunyaev-Zeldovich survey facilities, such as APEX, ACT, SPT, the Planck mission (launch in 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 (a few arcsec) 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 thousand clusters and groups and thereby provide a database for statistical studies of large-scale structure. The continued availability of the X-ray space telescopes XMM-Newton and Chandra is crucial to study the hot ICM and active galactic nuclei in the large number of clusters that will be discovered. High-resolution ultra-violet (such as FUSE or HST) and X-ray spectroscopic capabilities are needed to study the kinematics of the ICM through spectroscopy of extended emission from ionic lines.

The effects of 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 metal 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.

These programmes 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

Elements heavier than boron ('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 metals 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 metals 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 metals 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) intergalactic medium (WHIM), which could contain 40 per cent of the baryons, i.e. about half of the intergalactic gas. Metals 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 hydrogen and metals in galaxies, the ICM, and the general IGM as functions of redshift and environment provide important constraints on cosmological simulations and on the cycling of gas and metals between galaxies and the IGM. 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 HII 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 6, 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 polarization of the CMB. Its end is revealed by the strong increase in ionization level of the IGM around redshift 6. 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 substantially between redshifts 6 and 2.5. Before redshifts of 7, 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 HI 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 and the spin temperature of HI. A major challenge will be to find such very rare, bright sources at redshifts beyond 7. Large-area surveys with LOFAR may uncover a few. High-resolution near-infrared spectroscopy alone of any bright background source will allow an estimate of the evolution of the OI and CIV 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-$\alpha$ 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-$\alpha$ 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$500 kpc for strong OVI absorbers. Probing the absorber-galaxy connection for small column densities of heavy ions will provide an estimate of the overall size of the regions affected by galactic winds, and will require ELT spectra of faint background targets at intermediate resolution and signal-to-noise ratio. In underdense regions of the IGM, the expected column densities of heavy species are very small ($\lesssim 10^{10}$ cm$^{-2}$). Probing metallicity of at most one per cent of the Solar value 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 10 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 per cent of the cosmic baryons are still in the form of warm plasma in the IGM. The missing baryonic matter could reside in a WHIM within diffuse large-scale structures in overdense regions, as suggested by cosmological simulations. The metal 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. Additional hot diffuse gas is in ICM plasmas, galaxy groups and possibly in galactic halos. Moreover, 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 and thus be detectable in the hard X-ray domain.

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 O VI 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 first individual sources detectable at the onset of reionization should be bright, transient explosive events (gamma-ray bursts and pair-instability supernovae). Observing transient sources of near-infrared fluxes of 0.1 to 10 $\mu$Jy 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 a large collecting area (1km$^2$) radio telescope, and intermediate mass black holes will be found by large-aperture X-ray telescopes. The evolution of cosmic reionization will be probed by studying the H I 21 cm forest in high-resolution spectra of very rare, powerful background radio sources. Metallicity will be obtained by combining these observations with ELT near-infrared spectroscopy.

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 ($R \sim 10^5$) and signal-to-noise ($\sim 1000$). Searches at lower redshift can be more easily done with a next generation ultra-violet satellite. 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 in the optical with ELTs should be made at intermediate resolution ($R \sim 10^4$) with multi-integral field units.

A next-generation X-ray observatory operating in the 0.15–1 keV range, with an effective area of several square metres at 1 keV, 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. Detecting the hot phase ($T > 10^6$ K) at higher redshifts will remain a challenge as it requires a high sensitivity down to the 0.15–0.3 keV range. Extending the metal inventory, well known for ICM plasmas, to galaxy groups and halos of starburst or massive galaxies will be achieved by deep X-ray resolved imaging coupled to medium-low resolution spectroscopy of extended emission from ionic lines.

Advances in the problem of missing baryons in the local Universe will continue after the next servicing mission of the HST (2008, installation of COS) and with next-generation ultra-violet satellite and future large-aperture X-ray telescope. 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-$\alpha$ 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 neighbouring 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 basic 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 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 covered by the Sloan Digital Sky Survey, the largest galaxy redshift survey carried out to date. In spite of

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.

these advances, however, the theory contains unsolved problems. The growth of the gaseous and stellar components of galaxies is not related in a simple way to the build up 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 (Figure 3.5).

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 galaxies with high rates of star formation, the energy in these explosions is sufficient to heat the gas and drive many of the metals synthesized by the supernovae out of the galaxy. These metals 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 may play a significant role in 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 or trigger 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 the 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 Universe, 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. 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. We also need to understand how this evolution changes as functions of a galaxy's type, mass, and environment.

### 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 content 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 (Figure 3.6) will shift out of the ALMA bands, but the lines would be accessible to a large centimetre-wavelength array such as the proposed SKA.

About half the cold gas in nearby galaxies is in atomic form. A $1 km^2$ radio observatory will enable measurement of H I masses for Milky-Way mass galaxies out to redshifts of 2.5. At redshifts below one, a 1000km baseline radio facility will establish the rate at which (atomic) 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. In local galaxies the star formation rate is tightly coupled to the surface density of cold gas, and the maps of molecular and atomic gas produced by ALMA and SKA will test whether such relations extend to galaxies observed at earlier cosmic epochs. 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.

These observations of the distant Universe will need to be supported by in-depth studies of local galaxies, with sensitivities and resolutions extending down to individual star-forming clouds. Such investigations are crucial for testing and calibrating the diagnostics of the molecular gas, atomic gas, dust, cosmic rays, magnetic fields, and star formation rates of galaxies, and for probing the physical processes that couple the star formation to the interstellar medium (§ 4.1).

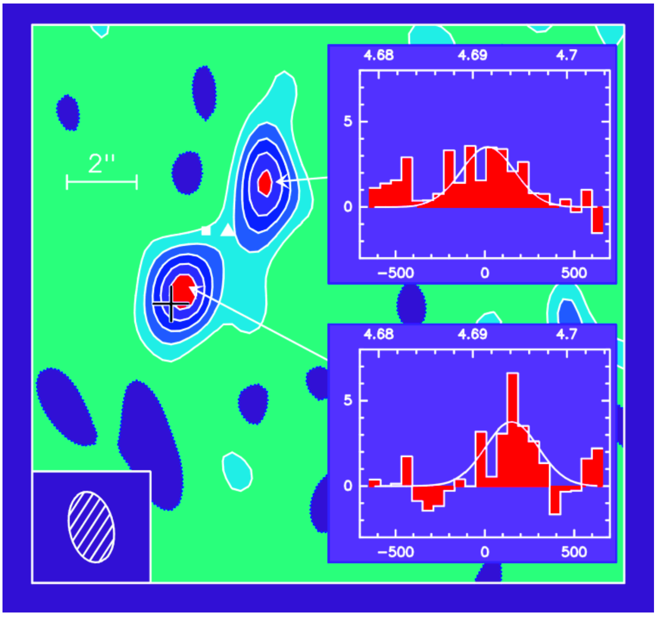



Figure 3.6: Dust and CO J=5-4 emission from the high redshift galaxy BR1202: A young merger at z = 4.7. Image taken with the IRAM interferometer.

Some information on star formation distributions in distant galaxies is already available, either from observations of the redshifted ultra-violet continuum from HST or maps of H$\alpha$ emission from deep imaging and integral-field spectroscopy in the near-infrared 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 bandpasses of these instruments, and sources are lost in confusion noise. Herschel and JWST 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 new generation of infrared facility. A larger aperture telescope would provide higher sensitivity and reduced confusion noise for imaging of faint distant sources, and enable infrared spectroscopy of high-redshift objects. The latter is critical for separating the contributions of active galactic nuclei from star formation. Such capabilities can be provided by a cold far-infrared space observatory, as discussed in ESA's Cosmic Vision. Observations from Antarctica could also address important aspects of these problems.

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.

**Dynamical evolution and dark matter in galaxies at all redshifts**. 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 (down to the 0.01 arcsec scale in the H band) and spectroscopy ($R \sim 5000$) in the near-infrared is needed to understand the internal structure of these distant galaxies. It will allow us to measure the distribution of metals, dust and stars across the galaxies, the 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.

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 active 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. They also will help to constrain the evolution of the dark matter distributions in the galaxies.

The most effective observational probe of this evolution is integral-field spectroscopy in the visible and infrared. This requires both high sensitivity and exquisite spatial resolution. Such work is just beginning, using integral field instruments in combination with adaptive optics on 8–10 m class telescopes (Figure 3.7). However these applications are limited to the very brightest and largest distant galaxies, and a comprehensive investigation will require a 30 m class ELT equipped with multi integral-field instruments.

Another important observational probe of dynamical evolution are the motions of individual stars in the nearest galaxies. Such observations can be used to estimate merger rates and quantify the effect of the environment. Substructures such as counter-rotating cores and tidal features provide direct evidence of a dynamical interaction, but it is impossible at this point in time to quantify how

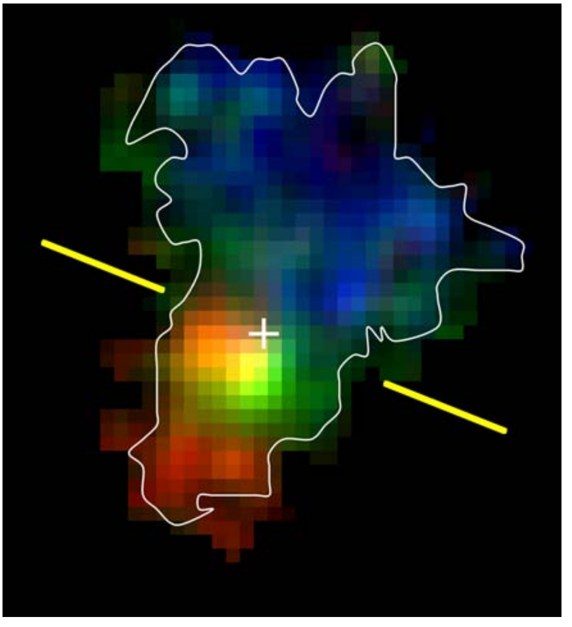


Figure 3.7: A composite velocity map of H$\alpha$ emission (red, green, blue: Spanning a total of $\sim$450 km s$^{-1}$) in the galaxy BzK155043 (z = 2.4). These VLT-SINFONI observations achieved an angular resolution of 0.15 arcsecond. The galaxy appears to be a disc and rotates at 230 km s$^{-1}$ about the yellow axis, which is centred on the nucleus of the galaxy (white cross) [courtesy ESO].

common these features really are. A systematic census of galaxies of all classes in different environments and at various redshifts will yield direct measurements of the merger rates, and hence allow us to quantify the impact of mergers and interactions on galaxy evolution.

With an ELT multi-object spectrograph in the optical-infrared range (8000–10000 Å), the motions of millions of giant stars in nearby galaxies sampling the full Hubble sequence lie within reach. These kinematic studies will shed light not only on the merger history of a galaxy, but also allow a measurement of its gravitational potential, including its total mass, its density profile and shape of its dark-matter halo (see also § 3.5).

**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 diameter one will be able to image in the optical and near-infrared 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 Ca II triplet (at 860 nm) allows for a basic metallicity measurement, 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 giants in M31, and perhaps as far as Centaurus A.

**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, stellar populations, 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. Many of these massive galaxies will be heavily obscured by dust, and observations in the infrared and sub-millimetre will be needed to identify the objects, and measure their redshifts and properties. As discussed earlier this work will begin with Herschel, JWST, and ALMA, but ultimately will require a large-aperture infrared telescope.

## 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. The first evidence for extended dark halos surrounding galaxies came from observations of the Local Group. Surveys of stars in the Galactic halo have revealed fossil evidence for hierarchical accretion, and have led to the discovery of new galaxies such as Sagittarius, which is currently being disrupted by the tidal field of the Milky Way's dark halo.

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 (see Figure 3.8). 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 fossil record of the formation process of the Galaxy can be retrieved in two complementary ways: Through the orbital dynamics of its constituent stars and through the chemical element abundances of these stars. 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, measurement, and modelling of these substructures will provide answers to key questions: How many mergers has our Galaxy experienced? When did these mergers take place? What were the properties of the Galactic building blocks, and how to do they compare to those of the present-day galaxies in the Galactic neighbourhood?

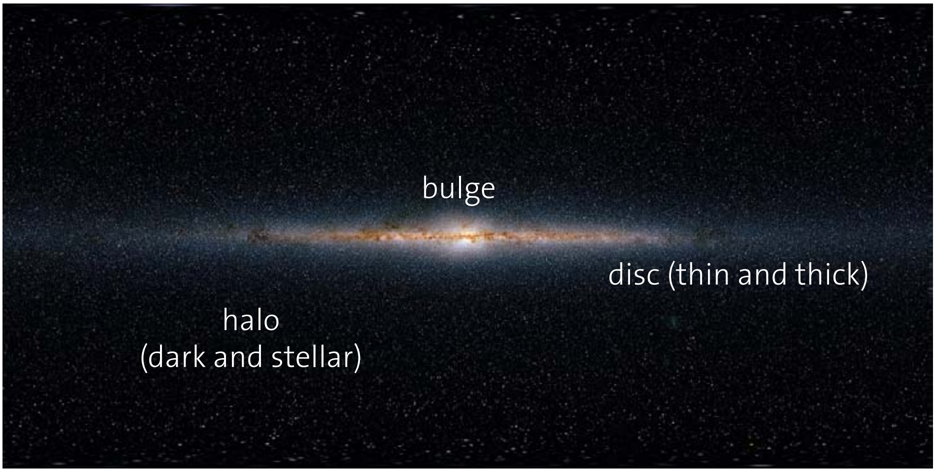



Figure 3.8: 2MASS all-sky image (blue, green, red are resp. 1.2, 1.6 and 2.2 microns) showing the central bulge and disc of the Milky Way Galaxy. The Galactic components are not only distinct morphologically as seen in this picture, but also in their characteristic ages, chemical abundances and kinematical distributions. This must be a consequence of their different formation epochs and mechanisms, and suggests they have retained different signatures of the assembly history of our Galaxy.

The chemical abundance patterns in the atmospheres of low-mass stars provide strong constraints on the history of star formation, of the initial mass function, and of the assembly of our Galaxy, especially at very early times. Large samples of stars belonging to all of the stellar components of the Galaxy are needed in order to find the rare, pristine (very metal-poor) stars, which must date back to the earliest episodes of star formation in the Universe. 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 quite similar in most respects to an elliptical galaxy. At a distance of only 8 kpc, 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. The formation timescale of the bulge, its chemical enrichment history, and its relation to the galactic bar and disc are largely unknown.

The Galactic disc offers an environment where the processes of star formation, nucleosynthesis, chemical evolution, and galaxy building are ongoing. Newly formed elements produced by supernovae, planetary nebulae, and other stars, are returned to the interstellar medium out of which subsequent generations of stars are formed. The disc of the Milky Way offers a unique laboratory for characterizing and understanding these complex physical processes.

**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 yield insights into how the dark matter is distributed, but the observations are generally confined to the inner regions of galaxies, where luminous matter makes its largest contribution. Kinematical measurements of distant (halo) tracers are strongly preferred, because they provide 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 its time evolution, because the streams are formed by stars on parallel orbits. Different insights are provided by the satellites of our Galaxy, which as 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 density profiles).

### 3.5.3 Future experiments

Disentangling the surviving fossils of our Galaxy's assembly process requires a vast catalogue of stellar properties, including ages and chemical abundances, spatial distributions (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 particularly good example; it contains the most metal-poor stars known, and it is believed to be the natural repository of debris from disrupted galaxies. However fewer than one per cent of the stars belong to this component. Any criterion that would favour the selection of halo stars will introduce biases which are impossible to correct a posteriori. Furthermore, although positional information may be sufficient to discover tidal streams in the outer halo, it is insufficient to disentangle the fossils in the inner part of our Galaxy (which contains most of the stars). There the dynamical timescales are so short that full phase-space information (and possibly chemical abundances) will be needed to disentangle the assembly history.

Ongoing surveys, such as SDSS, UKIDSS, the European Galactic Plane Surveys (EGAPS), and the VISTA and Pan-STARRS facilities under construction will provide high-quality photometry of billions of stars and inventories of non-stellar objects. A deep wide-field survey telescope such as the proposed LSST would provide high-precision photometry of stars extending to much fainter limits, and provide a unique resource for Galactic structure studies.

A major milestone will be reached with ESA's Gaia mission, which will measure very accurate positions, distances and motions of a billion stars all with spectrophotometry. Its long list of science goals includes fundamental calibrations of the stellar luminosity functions and initial mass functions (§ 4.1), 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. Gaia may eventually be complemented with a near-infrared astrometric mission (e.g., the proposed Japanese JASMINE mission), which would extend these measurements to the Galactic bulge.

Gaia will not only provide a representative census of stars throughout the Galaxy, but it will also make unique observations of millions of stars in Local Group Galaxies, especially in the Magellanic Clouds, in a dozen dwarf galaxies and in M31. Besides a clear discrimination between field stars and Local Group stars and a precise mapping of tidal tails and inter-galactic features, Gaia will provide unique clues to the dynamics of the Magellanic Clouds - Galaxy interactions and, more globally, orbits for a significant sample of the Local Group galaxies in a region large enough to provide direct constraints on the mass distribution and dynamical history of the Local Group to about 200 kpc. With a sufficient number of bright stellar tracers, Gaia will also measure the distortions of the dark halo structures and determine their density and extent.

It is crucial to supplement the Gaia dataset with dedicated ground-based spectroscopic programmes, in order to obtain the radial velocity and detailed chemical abundances for fainter stars. This work will be greatly aided by the construction of wide-field multi-object spectrographs with both low-resolution and high-resolution capabilities for 8 m telescopes, along the lines of the proposed WFMOS. Such a facility would offer key advantages including greater multiplexing and wavelength coverage together with the much larger field of view. This provides enormous leverage in Galactic studies by providing sufficiently large and diverse samples of stars to separate the complex dependencies between dynamical and chemical populations of stars that reflect in turn the complex formation histories of the Galaxy.

These surveys will provide a large number of tracers with precise kinematics allowing accurate measurements of the mass distribution of our Galaxy and the Local Group. 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.

Extending these large-scale spectroscopic surveys to other galaxies in the Local Group will allow us to investigate the formation and evolutionary history of our nearest neighbours. Current facilities limit this work to the brighter more evolved objects. Access to an ELT is crucial to enable this work to be extended to fainter limits in the Local Group, moderate spectral resolution giving access to, for example, red giants and black hole binary stars in M31, while at higher resolution it will be possible to observe Solar-type stars in the Magellanic Clouds.

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. Theoretical work is also needed to better understand the physical processes that cycle chemical elements through the interstellar medium into subsequent generations of stars.

# 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 drives 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 intricate 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 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 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 astronomy in 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 the 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 fields, gas dynamics including turbulence, 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 influence of stellar populations on galactic structure; 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, 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 need 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 (Figure 4.1). 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

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.

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 mass 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, due to the remarkable progress in infrared and millimetre astronomy at high angular resolution.

**Formation and evolution 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 one 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.

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 companion 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.

**The initial mass function for stars.** Our Sun is an average, inconspicuous star, one among billions in our Galaxy. But when looking at all the stars within a distance 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.

The distribution of stellar masses at birth (the initial mass function; IMF), is the key output of the star formation process. Surveys show that the shape of the 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.

### 4.1.2 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 fields in the interstellar medium, in particular in molecular clouds about to form stars, are still to be clarified theoretically and constrained observationally.

Conversely the feedback effect of newly formed stars on the parent environment and the forming stars is also critical 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 starbursts depleted in low-mass stars as a result of such processes? The subsequent dynamical evolution of the stellar clusters is also a key process to understand as clusters are the likely birthsites of most stellar populations in galaxies. 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.

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 (§ 4.4). In dense star clusters, where the encounters between stars are relatively frequent, close passages may disrupt the circumstellar 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 circumstellar 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 others are being planned (e.g., the 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 can be imaged with the proposed 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 Galactic magnetic field can be made on the scales of molecular clouds. On the 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 molecular lines. The results will constrain simulations of the role of the magnetic field in 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 $10^5$ 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 adaptive optics. 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 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 eight 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-arcseconds which have moderate magnitude differences between the components.

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. Furthermore, Gaia will bring a third dimension to star forming regions by measuring the distances of many stars observable in obscured regions.

The uniformity of the IMF need to be tested 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.

## 4.2 Do we understand stellar structure and evolution?

### 4.2.1 Background

The basic understanding of stellar structure and evolution is well-established through nearly a century of extensive model computations and comparisons with observed photometric and spectroscopic properties of stars. At a superficial level the models and the observations appear to agree in, e.g., the colour-magnitude diagrams of stellar clusters. Yet this apparent success hides a great deal of uncertainty in the modelling, which in a fundamental way affects our understanding of stellar evolution and its use in other areas of astrophysics.

After the initial stages of star formation (§ 4.1) and the first ignition of stable hydrogen fusion, stellar evolution proceeds through a sequence of nuclear 'burning' stages, involving successively heavier elements as the core temperature increases. The longest phase in this process is on the main sequence, during core hydrogen burning. For low-mass stars the nuclear reactions end with the production of carbon or oxygen after which cooling processes, involving neutrino emission, prevent further temperature increase; after a period of strong mass loss the end product is a white dwarf, a very compact star which subsequently cools over a period of billions of years. For stars of initial mass higher than around 10 solar masses, nuclear reactions continue until the formation of iron-group elements, at the maximum of nuclear binding energy, after which no further energy can be generated by nuclear fusion. The resulting core collapses, to nuclear densities, with a dramatic release of gravitational potential energy which leads to a supernova explosion (§ 2.4).

Since the evolution is controlled by the gradual conversion of lighter elements into heavier, it is greatly affected by processes that mix the products of the nuclear reactions; these include convection, penetration beyond convection zones and circulation and instabilities induced by rotation. Such processes can also bring products of the reactions to the stellar surface, changing its composition and providing an observable signature of processes deep in the stellar interior. Similarly, settling changes the initial stellar composition. An important example is mixing and settling of lithium which may reduce the observed lithium abundance relative to the primordial value resulting from Big Bang nucleosynthesis, a very important cosmological diagnostic. Advances in the modelling of this process and observation of its result have recently explained a former discrepancy between the observed abundances and the cosmological predictions.

Massive stars are very luminous and hence have a short lifetime. They dominate star-forming regions through effects of mass loss and radiation, most dramatically, of course, in supernova explosions. Their strong mass loss also substantially affects their evolution and plays an important role in the life cycle of matter in the Universe (§ 4.3). The loss of the outer layers of the star apparently plays an important role in determining whether the supernova explosion ending the life of the star leads to the generation of a gamma-ray burst (§ 2.4).

A substantial fraction of stars (about 60%) reside in binary systems. This leads to complex evolution in the cases where the stars are sufficiently close to interact in some evolution stages. The result can be mass transfer from one component to the other, as well as strong mass loss from the system with a consequent loss of angular momentum and shrinkage of the orbit, resulting in a bewildering zoo of end products of the evolution, including white dwarfs in close orbits. Mass transfer onto a compact star often takes place through an accretion disc, leading to emission of X-rays. Of great importance are systems where a white dwarf accretes material either from a companion star or merges with another white dwarf. During this process it can reach the Chandrasekhar mass limit and explode via a thermonuclear runaway producing a Type Ia supernova (§ 2.4). Owing to their well-defined luminosities such thermonuclear explosions serve as important standard candles in cosmology (§ 2.2).

This description of stellar properties is based to a large extent on theoretical modelling and hence requires observational support. Photometric and spectroscopic observations provide information about the stellar surface temperature and composition and, if the distance is known from astrometry, the luminosity, although in all cases subject to uncertainties in the observational calibration or stellar atmosphere models. Additional valuable information is available for stars in clusters which may be assumed to share a common age and initial composition. In the relatively rare cases of well-observed binary stars stellar masses and, for eclipsing binaries, radii, can be determined with substantial accuracy. However, such observations provide no direct information about the internal properties of the star; these must be inferred from modelling.

Asteroseismology, based on observations of stellar pulsations, allows essentially direct investigations of stellar interiors. The oscillation frequencies can be determined with exquisite accuracy and provide measures of the structure and dynamics of the stellar interior (Figure 4.2). With a sufficient number of observed modes, as available in many types of stars, specific information can be extracted which provides a direct test of the models of the stellar interior. Pulsating stars are found in essentially all parts of the Hertzsprung-Russell (HR) diagram (Figure 4.3), and hence asteroseismology can test and extend our understanding of all phases of stellar evolution.

Surface abundances and isotope ratios derived for stellar atmospheres, winds and ejecta provide strong constraints on mixing and nucleosynthetic processes in stellar interiors. For the most massive stars, spectroscopy may be the only way to extract this information, as spectra provide a crucial means of deriving many essential parameters such as temperature, surface gravity, mass-loss rate

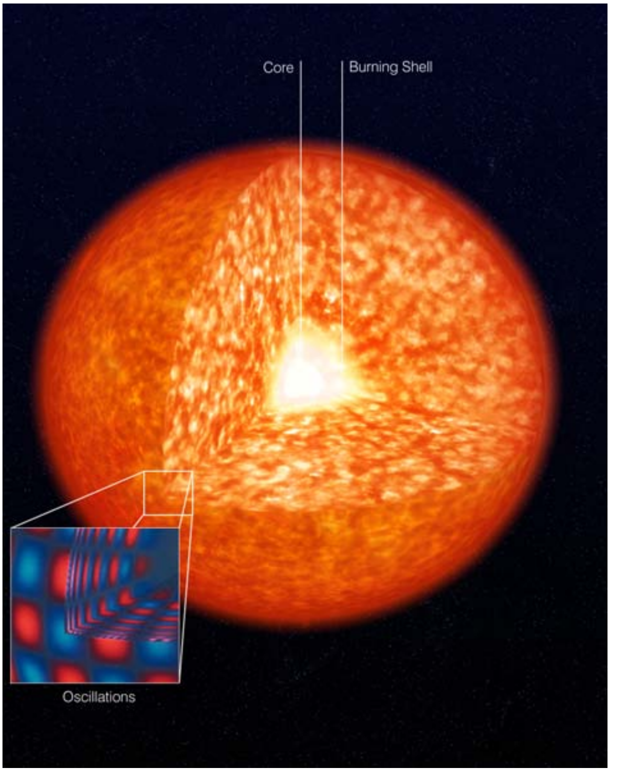



Figure 4.2: Oscillations generated inside a star will travel outwards and will become visible as parts of the surface moving up and down. The mode, frequency and intensity of these oscillations give valuable information about the inside of the star. It can be compared to the information about the inside of the Earth that is obtained from the propagation of the seismic waves below the crust. Credit: ESO and KU Leuven.

and magnetic field strength. The ability to obtain spectra for spatially resolved stars and winds other than the Sun would revolutionize stellar atmosphere theory and lead to enormous advances in our understanding of stellar properties.

Many stars show magnetic activity similar to Solar activity, including chromospheres and coronae, observable in the ultra-violet and in X-rays. By studying such stars we can improve our understanding of the causes for the activity and its variation with time, assisting our understanding of Solar activity and perhaps our ability to predict it. Helioseismology has provided detailed information about the Solar internal structure and rotation and beginning insight into other types of flows and interior magnetic properties. Such results are crucial for understanding the origin of Solar magnetic activity, and ultimately the possible variations in the effects of the activity on the Earth (§ 5.3).

### 4.2.2 Key questions

The main uncertainties in stellar modelling concern the hydrodynamic phenomena in the interiors. Convection, and penetration of motion beyond convectively unstable regions, remain a serious problem, affecting the internal composition structure and the later evolutionary stages, as well as the abundances inside and at the surface, of which the last one is an observable. 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. The size of the mixed 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. Mixing processes in the interiors of massive stars have dramatic effects on their subsequent evolution, including supernova explosions. 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. Further complications must be taken into account in the modelling of binary systems. A crucial question for cosmology is the extent to which Type Ia supernovae are reliable standard candles (§ 2.4).

Extrasolar planets appear to be preferentially associated with stars with relatively high surface abundances of heavy elements. It is not known, however, 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 and hence provides important input to the study of planetary-system formation (§ 4.4).

The structure and dynamics of stellar atmospheres remain uncertain, with important consequences for the interpretation of photometric and spectroscopic observations, particularly with regards to stellar abundances. A related question concerns mass loss in both massive main-sequence and red-giant stars, of key importance to stellar evolution and the enrichment of interstellar matter.

Our understanding of stellar activity is also very limited. It is believed that dynamo effects involving interaction between rotation and convection controls the generation and variation of magnetic fields, but the details remain highly uncertain. A better understanding of these phenomena would certainly improve our understanding, and perhaps our prospects for predicting, the Solar activity cycle. Particularly important in this context would be determination of rotation in the vicinity of stellar analogues to the Solar tachocline, the transition region at the base of the convective envelope which likely plays a dominant role in the generation of the magnetic field.

### 4.2.3 Strategy for the future

Modelling of all aspects of stellar evolution needs refinement. Progress has been made on the modelling of near-surface convection for selected stars, but little is known about the properties of convection and other mixing processes in stellar interiors. Large-scale hydrodynamical simulations, coupled with theoretical insight, will play a major role in these developments, and the results must be included in general calculations of stellar evolution. Modelling of binary evolution represents particular problems and suitable diagnostics must be found to test the results of such calculations. An interesting example is the possibility of identifying binary white-dwarf systems through both high-precision astrometry and gravitational-wave observations, leading to stringent constraints on the properties of these systems.

The physical properties of stellar matter, such as opacity, the equation of state and nuclear reaction parameters, need refinement; this requires laboratory experiments as well as theoretical calculations. The results will be tested through detailed observations of abundance distributions or direct inference of stellar structure through asteroseismology, thus leading to a fruitful interplay between basic physics and astrophysics.

Accurate observational determination of stellar global parameters is crucial. In particular, luminosity and mass will be provided by high accuracy astrometry. Many results have already been obtained from the astrometric data provided by the Hipparcos satellite (fine structure of the HR diagram, age of local halo stars – solving for the first time the discrepancy between the age of the oldest objects in the Milky Way and the expansion age of the Universe). Its successor, Gaia, will map one billion stars in six dimensions and provide parallel astrophysical diagnostic from photometric and spectrometric measurements. It will provide accurate global parameters covering all types and populations, reaching the rarest and the most rapidly evolving objects, including the most massive stars. Gaia will provide a very accurate position of the zero-age main sequence in the HR diagram, thereby contributing to the improvement of the determination of the helium content; it will provide membership, extremely accurate distance (to one per cent) and photometry of most galactic open clusters, providing a dramatic improvement in quantifying all effects of metallicity and helium content; it will provide membership, distance and photometry of all globular clusters within 10 kpc, thereby giving an excellent definition of the turn-off regions which are clues to age determination. The simultaneous availability of seismological measurements and absolute magnitudes will also allow to separate the effects of rotation and overshooting on the position of a star in the HR diagram.

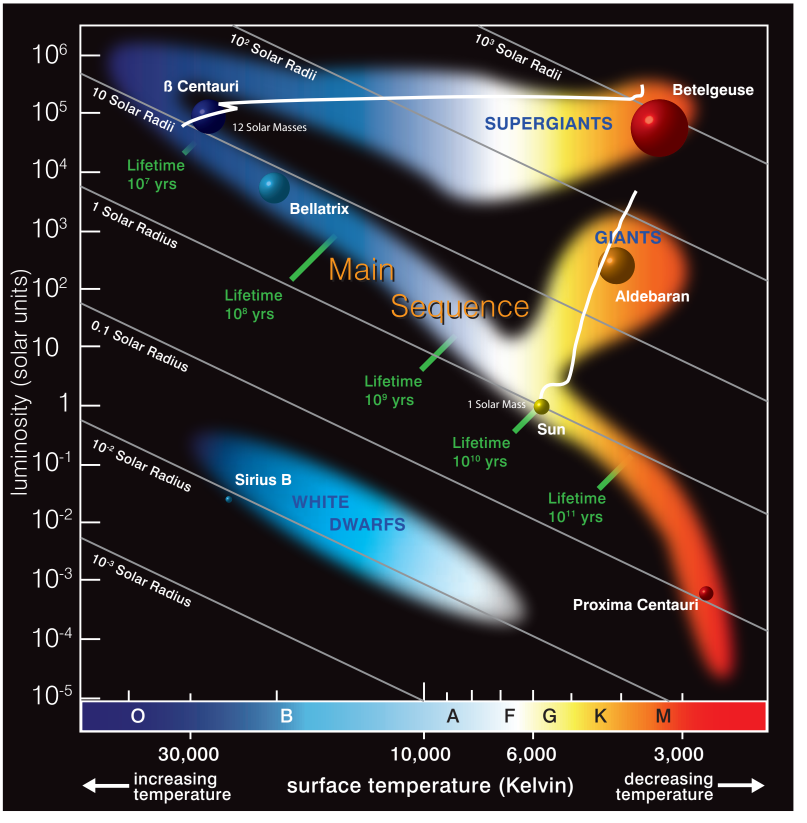



Figure 4.3: The Hertzsprung-Russell diagram showing temperature versus absolute luminosity of stars. Evolutionary tracks have been given for a star of 1 and 12 solar masses (white lines). The position of several types of stars have been indicated. The colours indicate the colour of the stars (adapted from The Internet Encyclopedia of Science).

It will be crucial to carry out ground-based high resolution spectroscopic follow-up of Gaia observations for a careful selection of stars to obtain precise stellar surface compositions and learn about the transport processes in stellar interiors. With the increasing availability of very large telescopes with high-resolution spectrographs such studies can be extended to fainter stars, including also stars in external galaxies, broadening the observational base for understanding these processes. Extension of such observations to the ultra-violet would obviously increase the range of elements and stellar types that can be investigated.

Stellar activity must be investigated through observations in the ultra-violet and X-rays, as well as with high-resolution spectroscopy to measure directly the stellar magnetic fields. Long-term monitoring of many stars is essential to find evidence for activity cycles, corresponding to the Solar 11-year sunspot cycle.

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 stars near the main sequence. The WIRE and MOST satellites are providing excellent data on stars with relatively large amplitudes. Within the next few years the CoRoT and Kepler missions will provide detailed asteroseismic investigations of a substantial number of stars, including Solar-like pulsators, through very precise space-based photometry. Extension of such studies to even larger numbers of stars would be extremely valuable, particularly for relatively nearby and bright stars, for which accurate supplementary observations are possible. Furthermore, experience from helioseismology shows that full utilization of the potential of asteroseismology will require observations extending over several months or even years of carefully selected stars, at the highest possible signal-to-noise level. This will require a network of telescopes dedicated to Doppler-velocity observations of stellar pulsations, since the intrinsic stellar 'noise' is far less serious for Doppler observations than for intensity observations. Such a network is technically entirely feasible and should be pursued as a high priority. In addition to asteroseismology it would also provide very valuable data on extrasolar planets.

The coming years will show increasingly reliable asteroseismic determination of stellar ages from observations of Solar-like stars, of great value to investigations of the evolution of the Galaxy (§ 3.5). In addition, information will be obtained on the extent of convective envelopes and cores, and on the importance of other mixing processes which affect the composition and hence the oscillation frequencies. Furthermore, 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. Such observations will provide stringent tests of the increasingly realistic modelling that will be carried out in parallel.

Stellar observations with sub-milliarcsec resolution, with large ground- or space-based optical, infrared and radio interferometers, will allow direct imaging of stellar surfaces and hence investigations of stellar surface convection and other surface features. In particular, space interferometry in the ultra-violet can provide detailed observations of surface magnetic activity. Interferometric observations are also required to meet the long-term goal of carrying out asteroseismology with sufficiently high spatial resolution to study modes of relatively high spherical-harmonic degree. This will allow analysis of the internal dynamics of Solar-like stars, including the generation of stellar magnetic fields through a rotation-convection dynamo.

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

### 4.3.1 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 metals 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**. The interstellar medium is an impressive chemical reactor. 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 the 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 occur with greatly enhanced abundances in the so-called 'hot cores' associated with low- and high-mass star formation (Figure 4.4). Even larger polycyclic aromatic hydrocarbons (PAHs) are seen as well. Finding a link between these species and the carriers of the diffuse interstellar bands – a century-old spectroscopic mystery – would transform our understanding of cosmic dust.

Some of this material will be incorporated into protoplanetary discs where ultraviolet, X-ray and thermal processing further modifies the composition. 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 amorphous 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, annealing and transport. Thus, mapping dust features in discs provides a direct probe of disc evolution.

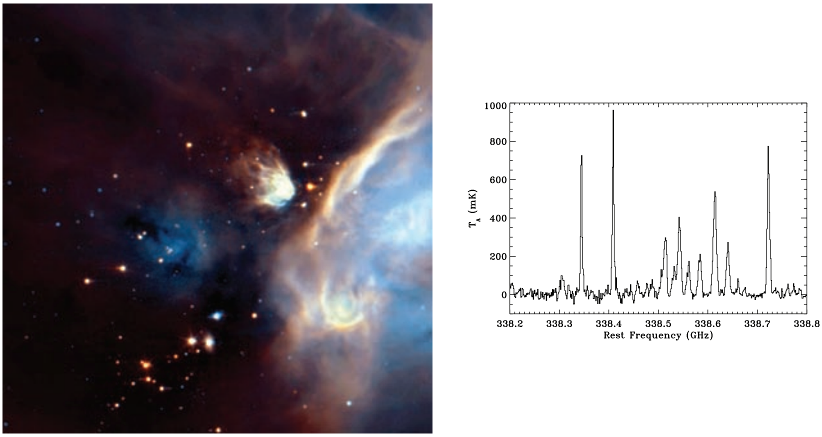



Figure 4.4: Detection of hot methanol in the young protostar IRAS 16293-2422.

**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 (AGB) phase. This is characterized by a poorly-understood thermally pulsing phase during which the star ejects matter at an increasing 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. As a result, these low-mass stars feed back the light elements (He, C, N, O) into the interstellar medium where they are then available for future generations of stars. Understanding the evolution from spherical red giants to the complex morphologies 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. Simulations indicate that winds of hot massive stars may be extensively clumped on small scales, due to the intrinsic instability of the driving line-force. This could imply that the currently accepted mass-loss rates may need revision by a large factor, which would have dramatic consequences for evolutionary calculations and predictions throughout the upper Hertzsprung–Russell diagram, and affect all astrophysical fields where massive stars play a decisive role. Furthermore, 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.

### 4.3.2 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 sub-millimetre 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 AGB track in the Hertzsprung–Russell Diagram, expelled during mass loss episodes, and the dependence of mass loss on time, stellar mass and metallicity (§ 4.2.2). The final stages of the evolution of massive stars remain surrounded by many uncertainties (§ 2.4). The interaction of mass-loss, rotation and magnetic fields, and their impact on the angular momentum distribution and hence the evolution of massive stars are all critical issues. The rate at which mass is lost, and how it depends on metallicity, is uncertain, in particular for Wolf–Rayet stars. In some scenarios, the catastrophic mass-loss events occurring during short outbursts are dominant. These outbursts characterize the Luminous Blue Variable (LBV) phase of massive star evolution and their mechanism is a mystery. Studying these short-lived rare evolutionary phases is complicated by the small number of LBVs which are within easy reach of current facilities. The role of binarity and multiplicity in understanding massive star populations also needs to be addressed in a systematic manner.

### 4.3.3 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 smaller than 0.1 arcsecond corresponding to less than 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 physical and chemical properties of molecules and solids and of the physical processes (e.g., photo dissociation, photo ionization and grain coagulation) and chemical reaction rates in the interstellar medium, by means of theoretical calculations and laboratory experiments. Frequencies of many molecules in the ALMA wavelength range are still lacking and the Herschel THz range is largely unexplored spectroscopically. Also, basic collisional rate coefficients to analyse ALMA and Herschel data are highly incomplete, and laboratory studies of PAHs and gas-solid processes, important to interpret JWST data, are 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**. The emergence of products of internal nucleosynthesis at the stellar surface during the AGB phase allows through spectroscopy improvement of our understanding of the internal energy sources, and in particular to study the effects of the largely unknown internal mixing processes on nucleosynthesis. With larger telescopes such studies become also possible in other galaxies, hence enabling the study of these processes at different metallicities.

The advent of powerful instruments for infrared (Herschel, JWST) and millimetre (ALMA) observations will allow detailed probing of late AGB star winds.

Spatially resolved spectral observations of the molecules and dust is necessary to unravel the mass loss mechanisms which determine the final fate of low- and intermediate-mass stars and to further characterize the processes in these outflows which lead to the formation of cosmic dust. The infrared and millimetre wavelength observations will need to be complemented by observations at other wavelengths (X-ray, ultra-violet and centimetric) and at very high angular resolution as provided by optical-infrared and radio interferometry.

Resolving the close circumstellar environments of young massive stars, LBVs and Wolf–Rayet stars in the Milky Way and nearby galaxies will require access to an ELT with adaptive optics, while discovering more of these rare objects will require narrow-band galactic plane surveys. Interferometry and spectropolarimetry will be required to probe the extended winds of supergiants.

## 4.4 How do planetary systems form and evolve?

### 4.4.1 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 cores of planets (Figure 4.5).

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 represents the debris of the planetary formation pro-

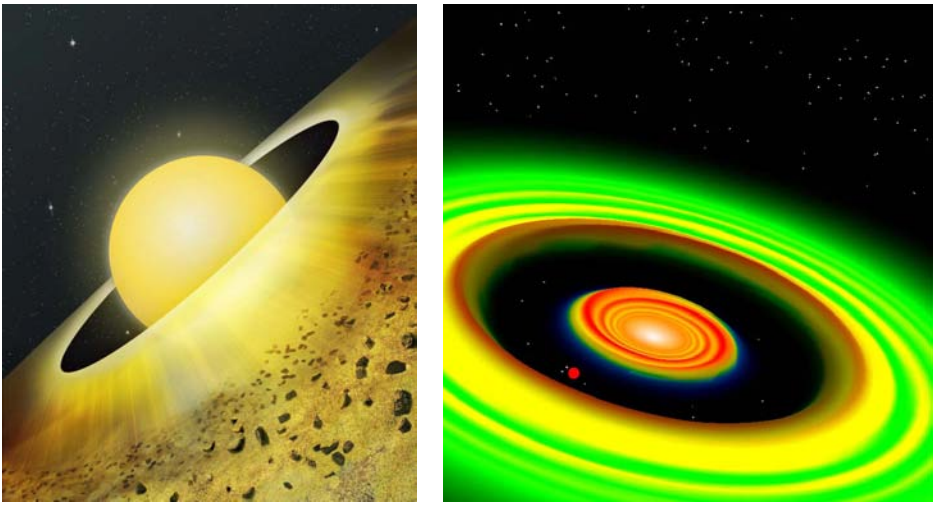


Figure 4.5: Dust evolution and planet formation in circumstellar discs. Left: Artistic view of the formation of pebbles in circumstellar discs as suggested from millimetre wave observations of the TW Hydrae system. Right: Simulation of the formation of a gap in a disc around a young star due to the gravitational effect of a newly formed giant planet. Image credit: Bill Saxton, NRAO/AUI/NSF (left); Geoff Bryden (right).

cess, for this reason these are called 'debris discs'. The study of debris discs is a powerful indirect probe of the planetary system they host.

### 4.4.2 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 very difficult to observe, except in small and warm regions 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 also 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 the disc-star-jet system. This is partly due to the lack 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 and the interaction with magnetic fields).

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.

### 4.4.3 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 or in laboratory experiments 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 (§ 5.5). The mechanism of giant planet formation via gas accretion onto the core in an evolved disc is favoured by the fact that the present composition of giant planets gas is depleted in volatiles (H, He) compared to the Solar composition, and by the existence of a large rocky core. However, 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). Observations 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, ELT, SKA, and a far-infrared interferometer in space).

## 4.5 What is the diversity of planetary systems in the Galaxy?

### 4.5.1 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 has steadily increased each year. The vast majority of these discoveries have been accomplished by high precision Doppler surveys on samples of more than a thousand nearby stars. This is illustrated in Figure 4.6 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

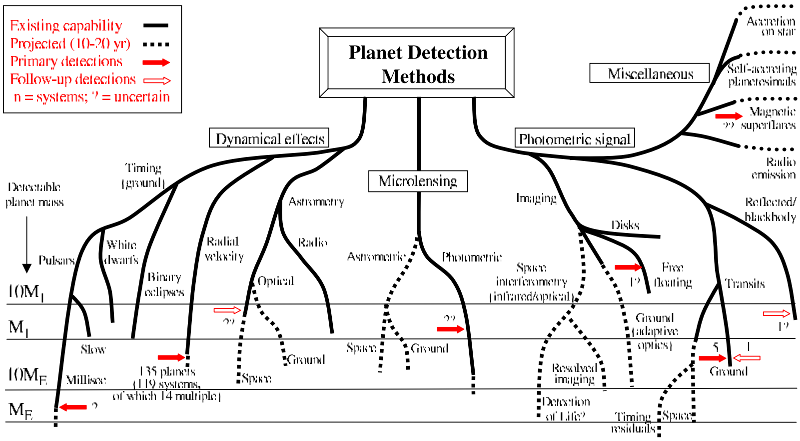



Figure 4.6: Detection methods for extrasolar planets. The lower extent of the lines indicates, roughly, the detectable masses that are in principle within reach of present measurements (solid lines), and those that might be expected within the next 10-20 years (dashed). The (logarithmic) mass scale is 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. Question marks indicate uncertain or unconfirmed detections. The figure takes no account of the numbers of planets that may be detectable by each method. Figure adapted from from ESA-ESO report nr 1: Extra-Solar Planets (eds. Perryman et al).

striking feature is the presence of giant planets on very short period 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 with 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 distance 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 typical 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 primordial nebula but its precise link to planet formation mechanisms is a matter of debate (§ 4.2.2).

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 detections of secondary eclipses on HD 209458 and TrES-1 with Spitzer have revealed the flux emerging from these planets themselves.

## 4.5.2 Key questions

Searches for extrasolar planets have revealed a large diversity of 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 is an almost unexplored topic.

The first earth-like planet has recently been discovered around a low mass star. 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 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 with an ELT or a space interferometer. This may then allow a future spectroscopic characterization for the search for extrasolar life (§ 4.6).

### 4.5.3 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, Kepler and Gaia satellites will provide an enormous enhancement in our capabilities of detecting planets via transits, 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. It is crucial to carry out ground-based follow-up and/or complementary observations to the data provided by Corot, Kepler, and Gaia, in order to confirm the planet detection by other methods and to obtain complementary information on the planet or planetary system, and also to fully characterise the host star.

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. Similar instruments attached to much larger telescopes should extend the survey to more distant Solar-type stars and also to the coolest nearby stars (exploiting the near-infrared). The outcome of this effort is likely 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 sufficient). 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 extremely 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 ground-based telescope 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 a sufficiently sensitive mid-infrared interferometer in space.

In both cases it is essential to achieve very high contrast imaging at the level of $10^8 - 10^9$. For ground-based telescopes this requires major developments in the field of extreme adaptive optics, coronagraphy and differential imaging, as this level is nearly 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.

## 4.6 Is there evidence for Life on exoplanets?

### 4.6.1 Background

The first direct detection of the atmosphere of a giant hot planet orbiting a star outside the Solar System was performed by HST when the planet passed in front of its parent star (HD209458), 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 prerequisite 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. Although we could reasonably expect to characterize gaseous giants with the instrumentation already planned 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 main-sequence lifetime of the star. 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 depends mainly on the stellar luminosity and age, but also on the planetary atmosphere and on possible internal heat sources. The presence of life can be inferred by the detection of life-related compounds like $O_2$ and $O_3$; all these molecules are in principle observable in the ultra-violet or infrared wavebands, but are difficult to detect.

### 4.6.2 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 two decades, extrasolar planetary systems search and characterisation (like the study of infrared spectral signatures of selected exo-planet atmospheres with JWST) programmes together with technological developments for innovative instrumentation concepts should be carried out in the future with this as the 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 with the most likely 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, on space-based infrared interferometers capable of measuring specific biosignatures. The serendipitous detection of radio signals from extraterrestrial civilizations within distances of up to tens or even hundreds of parsecs from the Sun will also be a possibility with the next generation of sensitive ground-based radio observatories.

# Chapter 5

# How do we fit in?

The Solar System is our unique vantage point for exploration of the Universe. The Sun, heliosphere and Solar System bodies – planets, satellites, asteroids and comets – play a critical role in unravelling the secrets of stellar physics, planetary system formation, and fundamental astrophysical processes. We can study our star and its planetary system in exquisite detail, but, this is just a snapshot in time; the Sun's life stretches across billions of years. To understand the past and future of the Solar System, we must compare it to other stars and their planetary systems. Key questions include:

- *What can the Solar System teach us about astrophysical processes?* Processes such as magnetic reconnection, plasma heating and acceleration, as well as shocks and turbulence occur throughout the Universe. These phenomena can be studied on all spatial and temporal scales in our own Solar System by remote sensing, and sometimes even in situ. Furthermore, probing the Solar interior by helioseismology is a litmus test for all theoretical models of stellar internal 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 across this 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?* Variations in Solar activity can affect the terrestrial environment and endangers human activities on various time scales. We need to understand and predict the disturbances of the space environment which are linked to the Solar output.
- *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 benefited 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 requires a coordinated effort in developing new capabilities to enable a system-level look at the Solar System. This effort needs to combine theory, simulations, observations, laboratory experiments, and in situ exploration.

## 5.1 What can the Solar System teach us about astrophysical processes?

### 5.1.1 Background

Due to their proximity, the Sun, Solar wind and Earth environment provide unique opportunities to test astrophysical processes including magnetic field generation, energy conversion and particle acceleration. Solar and space plasma research has contributed significantly to fundamental physical problems in atomic physics, plasma physics, nuclear fusion, and elementary particle physics. These fields act as catalysts for cross disciplinary studies where laboratory experiments and modelling contribute to the understanding of phenomena which are ubiquitous in the Universe, but can only be observed at high resolution at the Sun, in the Solar wind, or in planetary magnetospheres and ionospheres.

We cannot see magnetic fields directly, so for most astrophysical phenomena we determine field structure and evolution through measurements of the emitting plasmas trapped within the fields. In the Solar System we can study magnetic fields remotely at high resolution or, for the Solar wind and magnetosphere, we can observe in situ complex magnetic configurations which are also present in astrophysical phenomena ranging from exoplanets to high-energy jets.

Magnetohydrodynamics (MHD) – the mathematical description of a magnetised fluid – has proved an excellent starting point for understanding Solar System plasmas. Many current problems (e.g., energy conversion, boundary layers) require studies at small scales, where MHD is invalid and must be replaced by

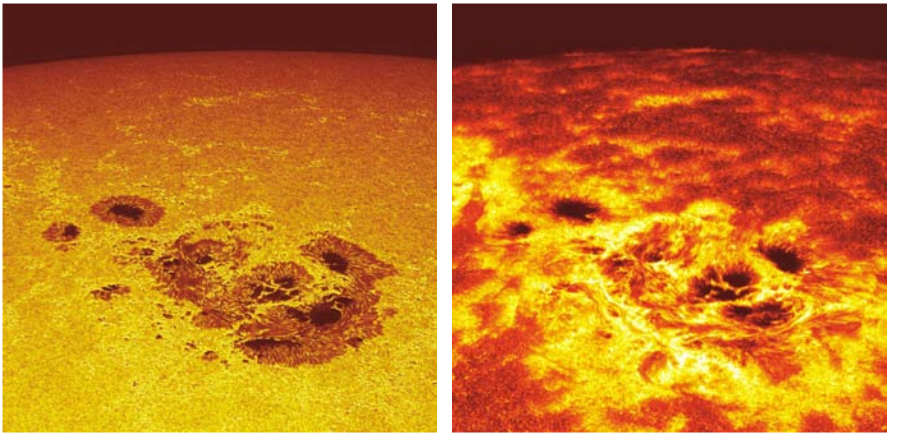


Figure 5.1: A sunspot region observed in November 2003. The left picture shows the lowest visible layer of the Solar atmosphere, the photosphere, in the light of CH molecules, and the right picture shows the 1000 km higher layer, the chromosphere, in the Ca II H line of once-ionized calcium. In the photosphere, 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.

a plasma kinetic description. Local thermodynamic equilibrium may not be a valid physical concept in many cases, especially the collisionless plasmas of the Solar wind and planetary magnetospheres. There is significant potential to apply these techniques at the level of heliospheric phenomena as these are increasingly resolved by ground and space-based telescopes. However, application to other stars is more problematic as it requires much greater resolution. Many stellar phenomena, e.g., starspots, activity cycles, stellar coronae and certain types of winds, use the Solar paradigm to describe their likely nature, yet they manifest themselves only by a variation of total flux in various spectral regions and time domains. Observing these at high resolution has become possible with the commissioning of multi-meter-class telescopes and adaptive optics, space-borne X-ray, extreme ultra-violet and visible observations, advanced field and particle instruments and high time-resolved radio data (Figure 5.1).

The complex interaction of magnetic fields and plasmas produces an array of Solar System phenomena and reveals the sites of basic plasma processes. Space-borne ultra-violet instrumentation, especially on board SOHO and TRACE, allowed the detailed study of transient, small-scale events occurring throughout the Solar disc (Figure 5.2). These include explosive events, blinkers, nano-flares and heating events, characterised by jets and brightenings. These events are at the limits of current resolution; but their effect may be crucial to the heating of the corona and the acceleration of the Solar wind. Larger, active phenomena such as flares and mass ejecta clearly reveal evidence for reconnection, acceleration and shocks. Their behaviour is key to understanding energetic particles, relativistic electron beams, magnetic activity and mass ejection propagation, all of which are fundamental to phenomena elsewhere in the Universe.

The magnetic field structures the environment of all objects embedded in Solar wind plasma, such as planets and comets. The presence of an internal magnetic field, such as the Earth's magnetic field, is especially important. Above the neutral atmosphere, the terrestrial environment is an ionized and increasingly collisionless medium, embedded in a magnetic field of planetary origin, which is a relatively easy accessible region to observe in situ plasma processes. The consequences of the interaction between plasmas and fields, such as auroras, radio emissions, particle acceleration by shock waves, transport and turbulence, are observed inside the magnetospheric cavity, and are key to understanding the dissipation of energy in the Earth's environment. The terrestrial case is a prototype of the interaction of a star with its planets, and gives direct evidence of plasma processes at work in less accessible contexts.

Collisionless plasma shocks are some of the most spectacular and energetic phenomena in the Universe. Generated by supernovae, stellar winds, or the rapid motion of neutron stars, they have important effects such as plasma heating, particle deflection and particle acceleration to very high energies. The interaction of the fast-moving Solar wind with the Earth's magnetosphere results in a bow shock with Alfvén Mach numbers up to 20, comparable to other astrophysical objects. In situ measurements by the Cluster mission reveal spatial and temporal variations at different physical scales, which represent a fundamental part of the phenomenon. It is crucial to study these variations simultaneously on electron, ion and fluid scales to sample the proper scale of each process, and to progress in the understanding of the full range of shock phenomena.

When two magnetized plasmas flow against each other, a boundary forms between them. As long as the magnetic field remains frozen in, this boundary should remain perfectly tight, each plasma confined on its side, without any reconnection between the field lines of each. This configuration exists in the terrestrial magnetopause, the boundary between the Solar wind plasma and the magnetospheric plasma, and at the heliopause, where the region of Solar influence interacts with the interstellar winds. Spacecraft measurements have shown that a small fraction of the Solar wind penetrates into the magnetosphere via reconnection events. The four Cluster spacecraft have allowed us for the first time to disentangle the complex temporal and spatial variations of these phenomena. The THEMIS and future MMS missions, supported by an appropriate array of remote sensing instruments on the ground, will allow further progress in this understanding. Finally, in neutral media, turbulence is at the origin of many (if not all) mechanisms which control the dynamics: 'anomalous' viscosity, resistivity and heat conductivity. Such processes are even more important than in

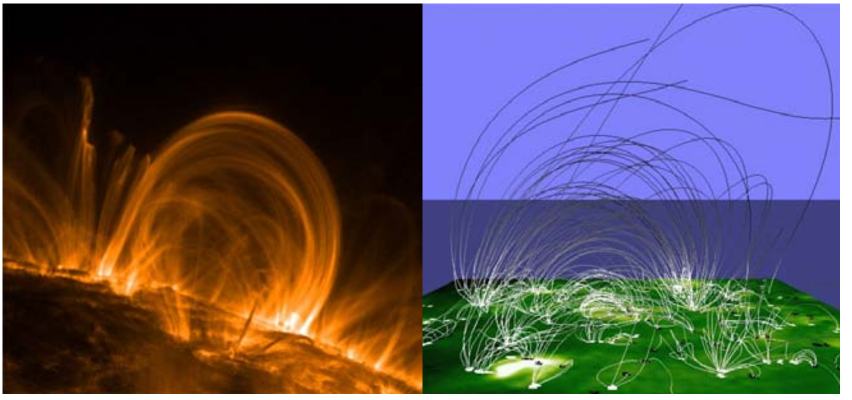


Figure 5.2: The complex magnetic fields in the Sun's atmosphere are well illustrated by this image of (left) an active region on the limb, taken in extreme ultra-violet using the NASA TRACE spacecraft, revealing million Kelvin plasmas trapped in magnetic loops and (right) the so-called magnetic carpet, the dynamic tangled magnetic fields in the Solar atmosphere driven by small-scale emergence of fields and photospheric motion, providing the complex environment for many transient phenomena [Courtesy Stanford SOI-MDI team and NASA/TRACE team].

space plasmas because they are 'collisionless'; the anomalous aspects serve to not only accelerate 'normal' transport mechanisms, but replace them. Turbulence is closely related to magnetic reconnection, which can be viewed as a way to dissipate magnetic energy into kinetic energy, and demands the existence of small-scale, localised processes.

The Sun is of paramount importance for astronomy 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.

### 5.1.2 Experiments

Magnetic field emerges at the Sun's surface at a variety of scales: from tiny flux tubes smaller than the spatial resolution of current telescopes to large, complex active regions. Most flux is produced by a global dynamo mechanism, deriving its energy from Solar rotation and the large-scale velocity field in the Solar convection zone, responsible for the overall magnetic field and its variation through the 11-year activity cycle (Figure 5.3). Understanding the Solar dynamo is essential for understanding stellar magnetic cycles and the origin of Solar variability. Theories exist which explain a global dynamo in terms of differential rotation within the convection zone, but current knowledge about the internal state of the Sun is too limited to discriminate between competing models or constrain them in a meaningful way. Helioseismology has established that differential rotation occurs throughout the convection zone, the rotation rate is larger at the equator than at the poles, while the Solar interior below rotates like a solid body. 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. The relationship between rotation rate, differential rotation, and activity is not known. With regard to the dynamo, it is not clear which processes dominate and to what extent several processes occur concurrently. Progress depends on our ability to develop detailed MHD models of entire stars and on measuring velocity fields with a precision much less than one meter per second in the Solar convection zone using helioseismology.

Some Solar magnetic flux may be caused by induction effects driven by small-scale velocity fields near the Solar surface ('local dynamo'). Although small-scale magnetic fields are found in the photosphere, a detailed study of their origin and evolution requires much higher spatial resolution and light collecting power than available today. The tangled, highly-dynamic magnetic fields of the low Solar atmosphere are known as the magnetic carpet (Figure 5.2); this environment is the seat of a wide range of basic astrophysical processes. Progress in these areas requires ground-based telescopes with apertures of several meters to resolve scales of order 10 km in the photosphere at visible wavelengths with sufficiently high sensitivity to do spectro-polarimetry with a precision of one part in ten thousand, assisted by high-order multi-conjugate adaptive optics.

It is essential that we can study the wide range of Solar physical processes by making significant progress in our ability to observe transient phenomena in the Sun's atmosphere down to the smallest scales. This requires a combination of next-generation sub-arcsecond imaging and spectroscopy, with temporal capabilities down to seconds, in extreme ultra-violet and X-ray wavelengths, to cover all relevant energies. This will need spatial, temporal and spectral resolutions significantly better than available today.

A thorough understanding of the structure and evolution of magnetic fields in the corona and chromosphere is beyond our current capability. Only photospheric magnetic fields can be accurately measured and mapped currently. Spectro-polarimetric diagnostics are mostly based on visible and infrared lines which to a large part originate in the photosphere. Very recently, Hanle-effect based diagnostics as well as several infrared spectral lines formed above the photosphere have been used to provide some quantitative information on chromospheric magnetic fields. The most promising magnetic diagnostic at higher altitudes appears to be Hanle effect measurements in, e.g., the Lyman hydrogen series, or ionized OVI, NV, and CIV lines known to be excited by radiation from the Solar disc and, therefore, should be polarized by scattering.

A major goal is the routine mapping of the magnetic fields in the chromosphere and corona, utilising the forward scattering polarisation signals of the (extreme) ultra-violet lines when observing at or close to the Solar disc centre. This can be achieved by spectro-polarimetric observations in selected (extreme) ultra-violet lines with appropriate spatial and temporal scales, of order 1 arcsec and seconds. In active regions the spectrum and circular polarisation of bremsstrahlung and gyroresonance emission at centimetre wavelengths is a tool that can also provide routine measurements of the coronal magnetic field through spectral imaging observations with a synthesis array such as FASR. These developments provide exciting pointers towards routine coronal magnetic mapping.

Our view of the Sun from the ecliptic plane has always restricted our understanding of the polar regions. The magnetic, thermal and velocity structure at the Solar poles is unknown, which leaves a considerable fraction of the Solar surface and its role in the generation and processing of magnetic field unexplored.

We also require a thorough understanding of how the Solar atmosphere extends into the heliosphere; how is the Solar wind generated and how does it evolve? The Ulysses mission enabled plasma and field studies of the heliosphere from a polar orbit with perihelion and aphelion 1.4 and 5.4 AU; it does not view the Sun directly but has made important discoveries about the Solar wind structure. The next step requires a combination of high resolution remote sensing (imaging and spectroscopy) and in situ instruments in out-of-ecliptic orbits (a minimum of 30 degrees – in particular to enable efficient flow measurements through Doppler analyses) and close encounter platforms with extended periods within 0.25 AU. Ground-based instruments, especially at radio waves, will complement these high-latitude and encounter strategies. In the sub-millimetre regime the Sun has an unexploited, yet rich potential of diagnostics in a wavelength range of nearly three orders of magnitude, which will be exploited by ALMA.

Multiple-spacecraft missions such as Cluster have demonstrated the capability to probe three-dimensional features at the scale of the satellite separation. Future missions such as MMS will allow exploring plasma turbulence down to dissipation scales and thus to test models applicable to a broad range of astrophysical situations. To better understand collisionless shocks, magnetic reconnection and plasma turbulence requires an in situ mission capable of probing simultaneously the major scales of the physical processes in the magnetosphere and Solar wind and should involve a fleet of spacecraft, e.g., forming three embedded tetrahedrons. Such techniques should also be a major target for other, more remote environments such as Jupiter's magnetosphere.

At the edges of what we regard as space, the terrestrial magnetosphere, ionosphere and upper atmosphere act as filters which respond to Solar forcing on a variety of temporal and spatial scales. They are also the ultimate regions where energy generated within the Sun is dissipated. This dissipation is very complex, involving many branching pathways and processes. Unravelling these

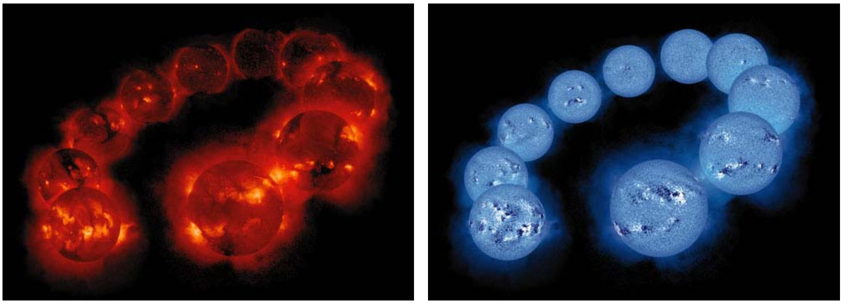


Figure 5.3: Sunspots and the corresponding magnetic fields follow an 11-year cycle, the sunspot cycle. The composite image on the left shows ten magnetic maps of the Sun approximately one year apart, from one maximum to the next. As the cycle fades, the large regions disappear. As the next cycle of activity 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).

processes and the ways in which they change, for example in response to Solar variability, provides key information for understanding the development of our own and other planetary systems. The operation of ground-based facilities, such as the SuperDARN network and the EISCAT radars (and their next-generation successors), capable of measuring the state and the transport of the terrestrial plasma environment and its coupling to the upper atmosphere provides a vital underpinning to the magnetospheric spacecraft missions described earlier, and is also essential to studying the response of the geospace system to Solar energy and its dissipation into the terrestrial environment.

Theoretical modelling of mechanisms such as magnetic turbulence, reconnection, and energy dissipation is an essential tool to support observations, often requiring the use of large-scale numerical simulations. Three-dimensional simulations based on the Vlasov-Maxwell equations are currently only possible at the level of small-scale dynamics. Coupling to large scales and interfacing between codes based on different physical models constitutes a major challenge.

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.

## 5.2 What drives Solar variability on all scales?

### 5.2.1 Background

When looked at from a distance, in visible light, the Sun is a remarkably stable star. Its luminosity varies by less than one part in a thousand over time scales of weeks, and by less than one part in ten thousand over decades, presumably one of the conditions for a stable climate on Earth. The picture changes when one takes a close-up view of the Solar atmosphere and its emissions across different regions of the electromagnetic spectrum. The Solar output is modulated by the magnetic field at the surface, causing variations in the emitted radiation, most prominently in the X-ray, ultra-violet and radio spectral regime, over time-scales ranging from less than a second to decades, and, in visible light, even centuries.

Even the so-called 'quiet Sun', loosely defined as regions of relatively little magnetic complexity and activity in the corona is not as stable as previously thought. The corona had long been treated as an equilibrium fluid, but recent observations suggest a more complicated picture. Noticeable differences have been inferred from imaging spectroscopy, showing that electron temperatures in coronal holes hardly attain 1 million K, compared to heavy ions where temperatures of 100 millions of K have been reported. Non-Maxwellian features are frequently observed in electron populations in the interplanetary space. The question of where these features originate is open, but calls for a non-MHD view of coronal heating and Solar wind acceleration. Before understanding Solar variability, therefore, we must also address the 'constant' Sun.

The interaction of magnetic fields and plasmas is the main process causing Solar variability. In the lowest layers of the Solar atmosphere, magnetic fields manifest themselves most prominently via active regions, the seats of the dark sunspot groups, whose latitudes vary with the activity cycle. Concentrated magnetic fields have profound effects on convection, causing a variety of small-scale structures, 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 on time scales of weeks, caused by a decrease of luminosity when a sunspot group transverses the visible disc. However, there is a positive correlation of Solar luminosity and activity; the Sun is on average brighter during Solar maximum, which is counterintuitive. The source for the excess brightness is not fully established, but appears to be connected with an increase of small-scale magnetic structure ('plages'), during Solar 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.

Different Solar regions have different elemental compositions. Element abundance can be measured in situ, so, the study of composition is a powerful tool for identifying Solar wind source. At Solar minimum, when large coronal holes cover most of the high latitude regions, measurements reveal different abundances in the high speed wind, compared to those measured in slow wind from low latitudes. Elements with a low first ionization potential are a factor 3-4 more abundant than in fast wind, where abundances are nearly photospheric. Because elemental composition is determined in the lower atmosphere, fast and slow winds are likely generated in different regions. We do not know which features are sources of the fast wind, nor do we have an accepted theoretical explanation for this first ionization potential effect. Recent observations do suggest that the Solar wind is accelerated at lower altitudes than previously thought.

The extended Solar atmosphere is a system of extremes, with a range of ten orders of magnitude in particle density from the photosphere to the edge of the heliosphere, and including the coolest parts of the Sun – the so-called temperature minimum at a few thousand Kelvin – and the hottest – the corona at millions of Kelvin during flares. The magnetic field pervades the entire atmosphere and drives most 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 X-ray, extreme ultra-violet and radio waves, and, variations can be measured in the Solar wind, in situ. Energetic events such as flares increase the radio and X-ray flux by orders of magnitude over time scales of seconds to minutes.

### 5.2.2 Experiments

We note that most requirements here are effectively satisfied by the experiments described in § 5.1.2.

The Sun shows variability on all time scales and that fact highlights the need for temporal measurements of Solar processes by developing a more homogeneous framework for time scales. This should be expanded to include long-term trends in space weather, e.g., geomagnetic activity, Solar energetic particle events and planetary space environments. On the practical level this shows the importance to acquire, archive, document and disseminate long-term time-series data. It is difficult to overstate the importance of good documentation as a tool for enabling new science from long-term time series (Figure 5.3).

While the 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 and the transient events in the upper atmosphere is only possible with ultra-violet and X-ray telescopes in space and with radio telescopes on the ground. In situ measurements of the composition and magnetic field of the Solar wind and its relation to transient events at the Sun require instruments on spacecraft. Non thermal particle populations are ubiquitous in the Solar atmosphere, including flares and minor activity in the corona, and quiet-time populations in the Solar wind. They are key to understand energy conversion and transport, but their origins are not understood. This calls for closer cooperation between remote sensing and in situ diagnostics.

Magnetoconvection, the interaction of the convective velocity field with the magnetic field in a plasma, is strongest in the photosphere, where the magnetic, thermal and kinetic energy densities are similar. This determines the appearance of sunspots and pores. The thermal and three-dimensional structure of a sunspot is to a large degree unknown. Together with the high conductivity of the plasma, magnetoconvection results in a highly intermittent magnetic field distribution, which has a tendency to concentrate at convective downflows. 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 the interior into the upper atmosphere is connected with the topology change, which is the likely source of energy for coronal heating and energetic events. How is energy transferred? Several wave types have been proposed that are thought to travel through magnetic flux tubes. However, the exact characteristics of such waves, and the wave-particle interactions, the mechanisms by which wave energy is transferred to the magnetic field, and 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. This implies adaptive-optics-assisted ground-based telescopes with apertures of several meters to achieve the required spatial resolution and light collecting power. 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. Recognising the range of scales over which Solar variability occurs, from the smallest transient phenomena through to the large scale transient phenomena such as flares, it is clear that a combination of next-generation spectroscopic and imaging instrumentation is needed, especially in the (extreme) ultra-violet range. Our current interpretation of Solar transient phenomena is often based on plasma diagnostic information combined with projected magnetic information based on photospheric magnetic measurement. Thus, routine magnetic measurements in the Solar atmosphere are a 'holy grail' of the Solar physics community and this requires the (extreme) ultra-violet spectro-polarimetry also detailed above.

There are many questions regarding Solar magnetic fields. How much flux is present on the Sun, how is it distributed and how does the distribution vary on different time scales? How does flux emerge and how is it removed? Answering such questions requires long-term synoptic observations with a network of medium-sized synoptic telescopes continuously observing and providing vector magnetograms with a resolution of better than an 1 arcsec, and high cadence.

## 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 (Figure 5.4). The goal is to understand physical mechanisms which originate on the Sun and which influence 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.

The most essential research areas to support the development of space weather are: (i) understanding Solar brightness variations at all wavelengths, in particular the ultra-violet that influences the chemical composition of the terrestrial atmosphere, (ii) understanding particle acceleration processes in the Solar atmosphere, in planetary radiation belts and in the auroral zones of the magnetised planets, (iii) modelling propagation of coronal mass ejections (and their embedded magnetic fields) from the Sun to the Earth, (iv) understanding and modelling the dynamics of planetary magnetospheres – most importantly gaining a deep understanding of how substorms work, and (v) modelling the behaviour of planetary upper atmospheres (both neutral and plasma) and their response to space weather. All of these areas require progress in plasma astrophysics, going beyond MHD to use hybrid and kinetic models of plasmas and developing models of the coupled magnetosphere-ionosphere-atmosphere system which can be validated and constrained by high quality observational data.

Many of these topics involve short-term events, i.e., transient phenomena in the Solar atmosphere and their impacts in the heliosphere,including flares and coronal mass ejections. However, it is possible that Solar variability has a noticeable influence on the Earth's climate on 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,

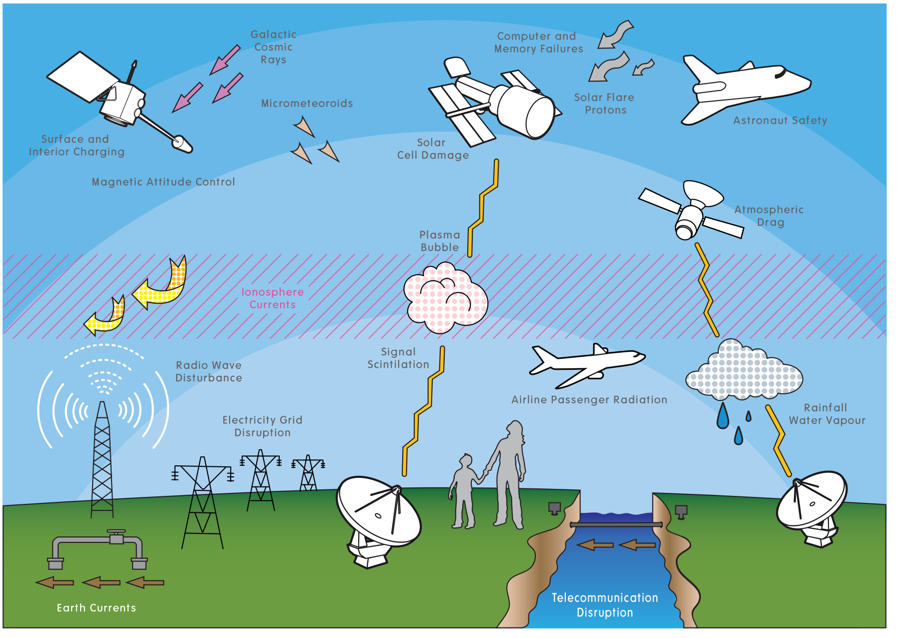



Figure 5.4: Solar eruptions such as flares and coronal mass ejections, triggered by changes in the magnetic field in the Solar atmosphere, accelerate charged particles and magnetic fields towards Earth, creating space weather. This cartoon summarizes some of the resulting hazards. Recreation of picture from Lou Lanzerotti, Bell Laboratories.

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 to long-period variations which modulate the activity cycle maxima, and may result in extended minima like the 17th century Maunder minimum when sunspots were absent for some 60 years. This period coincided with the 'small ice-age' when extremely cold winters occurred in Europe. A detailed understanding of the Solar dynamo is needed to understand longer-term Solar activity.

### 5.3.2 Experiments

In order to protect mankind's investments and space activities, space weather prediction needs to be developed and improved. We must the configuration of

Solar magnetic fields and changes in those fields, leading to early recognition of configurations which could lead to 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/coronal magnetic field. Suitably equipped monitoring missions with coronagraph and extreme ultra-violet instruments are essential.

A limitation is the fact that only one Solar hemisphere can be observed directly from near-Earth. The far-side magnetic field is currently not accessible at all. Furthermore, Earth-directed activity appears in projection against the Solar disc and is difficult to detect and to characterise. This situation will improve once the two recently launched STEREO spacecrafts move far enough away from Earth to give a stereographic view of the Solar environment oriented towards Earth, permitting a much better characterisation of earthward moving ejecta.

Another emerging approach is Solar surface far-side imaging using helioseismology. This method has passed proof-of-principle tests and has potential for space weather forecasting at least for basic conditions two weeks ahead of time. A Solar far-side imaging mission would be ideal. These observations need to be complemented by comprehensive in situ and remote sensing measurements of interplanetary space, of conditions in the Earth's magnetosphere, radiation belts and ionosphere, consistent with those detailed in § 5.1.2.

Finally, a coronal and heliospheric imaging system above the Solar poles would enable a complete view of the streamer-belts and, thus, a view of the complete coronal mass ejection picture, not possible from the ecliptic plane. This would also provide a more complete study of stellar mass loss through discrete ejecta.

Synoptic data from upcoming missions, e.g., the proposed Solar Dynamics Observatory, as well as ground-based facilities will provide useful input to space weather studies. The same applies to the detailed high-resolution investigations with the recently launched Japanese Hinode mission.

## 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 extrasolar planet (51 Peg *b*) in 1995, and the additional

200 extrasolar 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 (§ 4.5). 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 (§ 4.4).

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 that 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 extrasolar 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

Figure 5.5: 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. [ESA]

past (Figure 5.5). 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.

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.

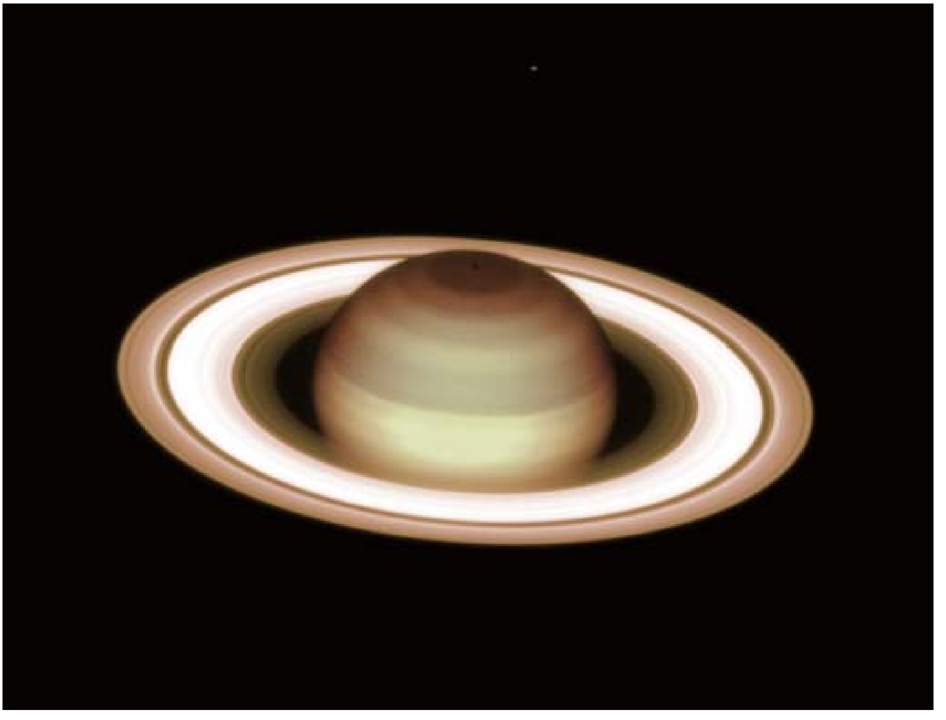


Figure 5.6: With the use of adaptive optics, ground-based diffraction-limited images can now be obtained in the near-infrared range. This image of Saturn was obtained at the VLT using the NAOS-CONICA system. The angular size of the Saturn disc is about 15 arcsec. The image quality is sufficient for the ring systems to be separated. Bands and zones are also clearly resolved on Saturn's disc. [ESO]

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 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.

**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 (Figure 5.9). 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. Finally, the giant planets have ring systems (Figure 5.6). Analysing the ring structure, including the role of shepherd moons, is important for understanding their origin and evolution.

**Small bodies and extraterrestrial matter**. Our understanding of cometary nuclei has advanced enormously during the 20 years since the exploration of Halley's comet (Figure 5.7). At that time even the verification that the solid nucleus did exist was considered a major advance. By now we have obtained 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 sam-

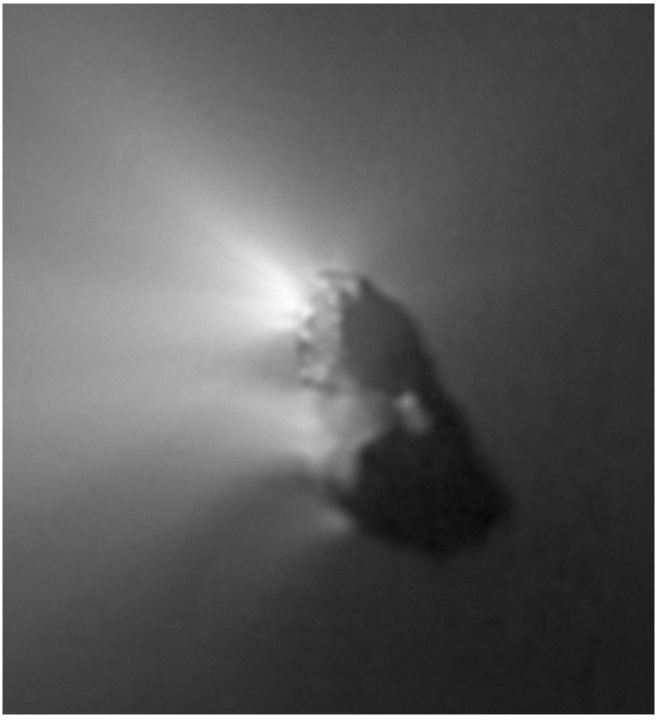


Figure 5.7: The nucleus of comet Halley from a distance of about 500 km. The picture was taken on March 13, 1986, with the Halley Multi-Colour Camera on board the ESA satellite Giotto, and it was the first detailed image of a cometary nucleus. The image distinctly shows an irregular and inhomogeneous 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.

ple 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 $\sim 0.5$ 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 captures from the Oort Cloud are 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 Hayabusa mission to the 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.

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. Interplanetary dust particles have been collected around Earth orbit for decades. 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

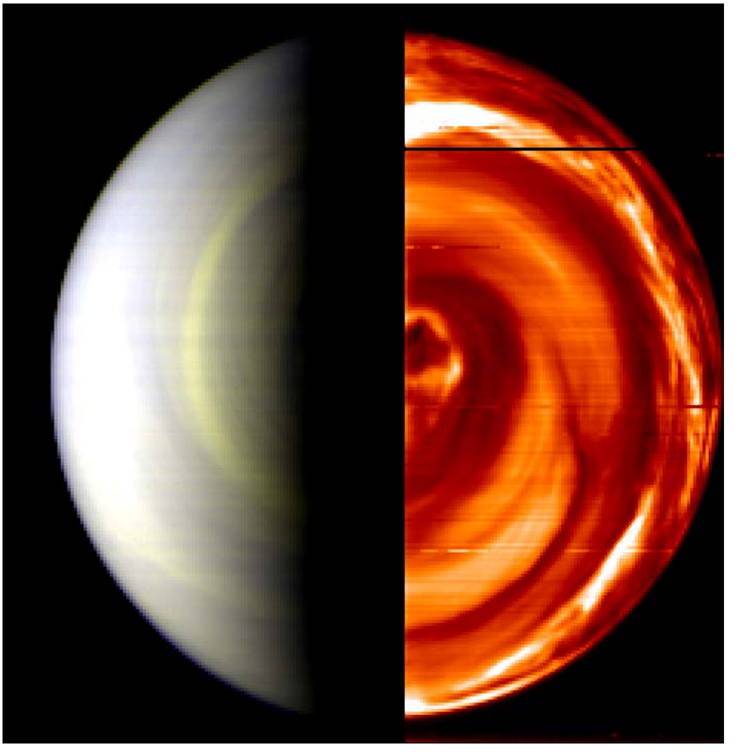


Figure 5.8: The south pole of Venus observed by VIRTIS on board Venus Express in June 2006 one day after the spacecraft entered the Venus Insertion Orbit. The upper image (in false colours) is taken in visible light, while the lower image is taken in the near-infrared range, at 1.7 micron. At this wavelength, 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.

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. Finally, radioactive debris have been deposited on Earth by nearby supernovae, as found in oceanfloor sediments. Such terrestrial archaeology supplements meteoritic studies, in exploring the nearby surroundings of the Solar System at earlier times.

### 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? What is the origin of Mercury's intrinsic magnetic field? 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. In the past few years, Europe contributed significantly to the exploration of the Solar System with Mars Express, Venus Express (Figure 5.8) and Smart 1. Future missions include ExoMars and Bepi Colombo.

The next steps of Mars exploration should include networks of small stations with geophysical 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 chemical nature of the condensates, the nature of the giant planets' complex dynamical systems, the structure of their magnetospheres, and their interaction with the Solar wind.

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 ac-

Figure 5.9: 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.

quisition 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? Which outer satellites could host liquid water below their surface, and exhibit cryovolcanism? These questions are relevant for the possible emergence of life in the outer Solar System.

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 (Figure 5.9), should be followed by space missions towards Jupiter and its system (in particular Europa), Saturn (including a probe) and its satellites (including 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 accuracy, based on knowledge of their dynamical origin and the biases involved in the search programmes. 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. The fate of medium-sized objects (less than about 100 meters) in the terrestrial atmosphere is not well known and is closely related to the internal structure of the bodies. 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. Europe has a long tradition of international networks for observing bright fireballs providing orbital information for recovered meteorites. 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 a 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 USA search programmes, an international programme will be required to secure follow-up observations and analyse 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, 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 took place already in the early stages of Solar System history. In particular, comets have probably partly contributed to the terrestrial oceans and could possibly have brought prebiotic molecules.

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 by ESA's Rosetta mission to comet Churiumov–Gerasimenko. Additional information will come from millimetre and sub-millimetre 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, USA 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 describe the main science goals in the four thematic areas, and identify key technologies and facilities that are needed to achieve these in the next two decades. The main scientific goals for each main area and required facilities are summarized here in no particular order. The latter have been divided into two classes: *Essential*, implying that the scientific goal cannot be reached without this facility, and *complementary*, where the facility will provide very beneficial information to reach the scientific goal, and without it there would be delays, difficulties and stronger requirements on other facilities. The list is further divided in current facilities which include those under construction and are fully funded, and future facilities. The description of future facilities is fairly generic, but the list includes projects which are in the detailed design phase.

## 6.1 Cross disciplinary requirements

There are several items that appear in most if not all of the main themes. Investments in those areas are therefore a high priority for all of astronomy.

- **Theory and simulations:** Astronomy has evolved from a following science (applying fundamental results from other fields), to a leading science (astronomical discoveries and interpretations inspire other fields). In order to maintain this position, and to remain able to interpret and guide future observations, continued investments have to be made into the development of theory and simulations.
- **Computing resources:** Substantial high-performance computing resources will be mandatory, not only for processing and analysis of the extensive

observational data, but also for the theoretical calculations and simulations including detailed physical processes and feedback mechanisms. The combination of these two aspects is crucial for careful comparison of observational datasets and theoretical predictions.

- **Astronomical data management and the Virtual Observatory:** With the ongoing expansion of instrumental capabilities across the electromagnetic spectrum, the importance of archival datasets and the combination of information from multiple datasets is growing rapidly. This trend will accelerate dramatically in the future, with the commissioning of dedicated survey facilities. The scientific output of these facilities will be maximized by making adequate provision for systematic archiving of the data, with common standards that allow for interaccessibility of datasets and effective scientific exploitation of the data. The first steps in this direction are being made by individual data centres and through the European Virtual Observatory, which in turn is part of an International Virtual Observatory effort.
- **Laboratory astrophysics:** Obtaining ground-truth is difficult when observing objects that are lightyears away. Experiments in a laboratory on Earth have therefore always played a crucial role in the interpretation and verification of the data. The need for this continues to increase. With the improving spectral resolution and sensitivity, more materials, in the solid and gas phase, will become distinguishable by their atomic and molecular transitions, and need therefore to be measured. New simulations and theoretical calculations of processes and reaction schemes remain to be verified in the laboratory. Finally, for sample return missions, there will be a strong demand for the laboratory analysis of these samples.

## 6.2 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 composition, and its formation. The specific science goals reflect this mixture:

1. Measure the evolution of the dark-energy density with cosmological epoch, to search for deviations from a cosmological constant;
2. Test for a consistent picture of dark matter and dark energy using independent and complementary probes, thus either verifying General Relativity or establishing the need for a replacement theory;
3. Measure the polarization of the cosmic microwave background at ten-degree scales, to search for the signature of relic gravitational waves;
4. Directly detect astrophysically-generated gravitational waves to measure strong-gravity effects, in particular arising from black-hole coalescence;

5. Make direct studies of regions near the event horizon of supermassive black holes in galactic nuclei, to test strong gravity and to understand how large-scale relativistic jets are launched;
6. Understand the astrophysics of compact objects and their progenitors, particularly the functioning of the supernova explosion and gamma-ray burst mechanisms;
7. 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 current and 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 for independently cross-checking the reality of some of the more subtle phenomena at issue. Many of the current facilities will have long lifetimes in their existing forms, and have well-motivated plans for upgrades. Except where there is inevitably a fixed lifetime (e.g., a cryogenic satellite), or where a future facility is a clear replacement, it should be assumed that current facilities will continue to be required.

**Current facilities**

- Planck will be essential for goals 1 & 2 and complementary for goal 3 since cosmic microwave background (CMB) information underlies most of statistical cosmology. Planck will measure the unpolarized properties of the CMB effectively to the ultimate limits imposed by cosmic variance;
- Direct dark-matter detection experiments and indirect search for dark matter via high energy gamma-rays are essential for goal 2, and have the potential to prove that it is a supersymmetric relic particle;
- SWIFT, XMM and INTEGRAL are providing essential data for goal 5 & 6 allowing to address the study of neutron stars and black holes in binary systems and to clarify the progenitors and the physics behind GRBs;
- Ground-based gravitational wave detectors will be essential for goal 4. Low-mass black-hole mergers generate signals in the kHz bands that can be accessed from the ground. Existing detectors may lack the sensitivity for carrying out these astronomical measurements now, but are needed as a basis for more powerful facilities;
- The 8–10 m class optical telescopes (including the VLT, GTC, LBT and Gemini), as well as interferometers, like the VLTI are essential for goal 5. Astrometric studies of orbits near the centre of our Galaxy are the best evidence for a nearby massive black hole, and this case should strengthen with time and increased sensitivity due to upgrades of instruments;

- H.E.S.S. will be essential for goals 5 & 7 in order to study the production of very high gamma-rays;
- The Pierre Auger observatory will be essential for goal 7. To determine the origin of high-energy cosmic rays will require huge air shower arrays and air fluorescence detectors;
- LOFAR will be complementary for goal 7. Synchrotron radiation from cosmic-ray shower debris can help pin down the energy of the initial event.

**Future facilities**

- The planned SKA will be essential for goals 1 & 2. A sufficiently sensitive radio telescope can obtain H I redshifts and image shapes for a large fraction of all galaxies out to redshift one, allowing gravitational lensing and large-scale structure studies of cosmological geometry. The main requirement is a large field of view, so that the whole sky can be imaged within a few years. Improved tests of General Relativity can also be expected from the large sample of new pulsars expected from such a facility, which will be complementary for goal 5;
- An X-ray survey satellite will be essential for goals 1, 2 & 6. Large-area X-ray imaging is important both for the detection of clusters of galaxies, and also for monitoring to detect transient explosive sources;
- A wide-field imaging telescope in orbit will be essential for goals 1 & 2. Combining photometric redshifts with sub-arcsecond resolution imaging enables the principal tests of dark energy: Gravitational lensing, baryon oscillations, cluster evolution, and new supernova samples. Ground-based colours should be allied with space precision imaging and near-IR photometry. Wide-field spectroscopy from the ground is also needed to calibrate the photometric redshifts;
- An Extremely Large Telescope (ELT) will be essential for goals 2 & 6. Fundamental cosmology will require spectroscopic classification and testing of systematics of supernovae at extreme redshifts, and also through hyper-accurate quasar spectroscopy, which can limit variation of the fundamental constants and measure directly the acceleration of the Universe. The understanding of thermonuclear and core-collapse supernovae and their connection to gamma-ray bursts, calls for detailed spectroscopic observations of the debris in the optical and infrared. At the highest redshifts, this will only be possible with an ELT. Both these applications require a mirror substantially enhanced from the current 8–10 m standard, to 30–40 m;
- A Cherenkov Telescope Array will be essential for goals 2 & 7. The physics potential is well defined by the impressive results from todays TeV gamma-ray telescopes. Detailed understanding will be gained of the acceleration of relativistic particles in supernova remnants and active galactic nuclei, and dark-matter annihilations may also be detected;

- A cosmic microwave background polarization satellite will be essential for goal 3. Beyond Planck, primordial gravitational waves from inflation can be detected from CMB polarization;
- LISA will be essential for goal 4, which requires detection of low-frequency gravitational waves that cannot be probed with ground-based experiments. It will be complementary for goal 3;
- A large collecting area X-ray observatory will be essential for goals 5 & 6. Emission processes close to the inner edge of black hole accretion flows can be probed by sensitive X-ray measurements;
- Very long-baseline (4000 km) interferometry at sub-millimetre wavelengths will be essential for goal 5. In combination with the large collecting area of ALMA, this technique offers the combination of resolution and dust-piercing wavelength to image near to event-horizon accretion flows within the Milky Way;
- A large volume ($km^3$) neutrino telescope will be complementary for goals 2, 6 & 7. The detection of high-energy neutrinos (in particular from the Galactic Centre) is a sensitive probe of dark-matter annihilation, and of the sites of cosmic ray acceleration. Smaller neutrino detectors (megaton-class water Cherenkov detectors or 50-kiloton-class liquid scintillation detectors) will be complementary for goal 7 in the event of a nearby core-collapse supernova.

## 6.3 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 the following important goals can be identified:

1. Map the growth of matter density fluctuations in the early Universe, both during and after the Dark Ages;
2. 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;
3. Determine the evolution of the galaxy cluster mass function and constrain the equation of state of the dark energy;
4. 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;
5. Measure the metallicity of the warm-hot phase of the intergalactic medium in the local Universe and solve the missing baryons problem;

6. Measure the build up of gas, dust, stars, metals, magnetic fields, masses of galaxies and thus the evolution of the Hubble sequence with cosmic time and the connection between black hole and galaxy growth;
7. Obtain a comprehensive census of the orbits, ages, and compositions of stars in our own Galaxy and the nearest resolved galaxies, aiming to produce a complete history of their early formation and subsequent evolution.

This work will draw on current facilities (including those under construction) but will require new facilities covering the wavelength range from the low-frequency radio to the gamma-rays.

**Current facilities**

- LOFAR will be essential for goal 1, and complementary for goals 2 & 6. It will reveal the first phases of galaxy formation via the observation of high redshifted (luminous) radio galaxies;
- Planck will be essential for goal 1, and complementary for goal 3 by tracing the earliest density fluctuations via the Sunyaev-Zeldovich effect and the polarization of the CMB;
- JWST will be essential for goal 2, and will be complementary for goals 4 & 6. Its infrared capabilities will allow the study of the first ionizing objects, dust production and evolution over cosmic timescales, (early) star formation and many other distant objects;
- XMM–Newton remains essential for goal 3, and complementary for goal 6. 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 often in combination with observations in optical, radio and sub-millimetre domains;
- The 8–10 m class optical telescopes are essential for goals 4 & 6, and will be complementary for goal 3, as the instrument to study high-redshift quasars and gamma-ray bursts and the absorption lines in their light;
- ALMA will be essential for goal 6, and complementary for goals 1, 2, & 3. ALMA will provide important information on the reionization, the evolution of dust and molecules at high redshift, the earliest epochs of structure formation, dust-obscured star formation and nuclear activity;
- HST (after the next servicing mission) will be essential for goal 5, and will be complementary for goal 6, by giving invaluable information about baryons in the intergalactic medium and (until JWST is operational) about the star formation in distant galaxies;
- Gaia will be essential for goal 7 by measuring very accurate positions, distances and motions of a billion stars all with spectrophotometry providing in effect a stellar census of our Galaxy;

- Herschel will be complementary for goal 6, and will study molecules and dust at intermediate and high redshift.

**Future facilities**

- An ELT with adaptive optics and high-resolution imagers will be essential for goal 6, and complementary for goal 2. Spectrometers ($R \sim 5 \times 10^4$), and highly multiplexed (near-)infrared spectrographs (including multiple integral-field instrumentation) will be essential for goal 1, and complementary for goals 2, 4 & 6. It 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;
- SKA with large surface area and long baselines will be essential for goal 1, and complementary for goals 3 & 4 (for which combined observations with an ELT are required) & 6. The large surface area and large frequency coverage will be essential for goal 6, and complementary for goals 1 & 2. 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 sub-millimetre to centimetre work that will be opened up by ALMA and LOFAR;
- It will be essential to continue to have the capability to detect gamma-ray bursts, which provide a unique way of probing the early Universe. Follow-up observations are necessary for goals 2 & 4;
- An X-ray space mission with moderate-resolution spectroscopy ($R \sim 1000$) will be essential for goals 2 & 5, and complementary for goals 3 & 4. This mission should address key problems relating to the intergalactic medium, missing baryons, black hole evolution, and galaxy assembly;
- A 4–8 m ultraviolet space telescope will be essential for goal 4, and complementary for goals 2, 5 & 6. Such a facility 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;
- A 4–8 m cooled infrared telescope for spectroscopy will be essential for goal 6. It will trace dust-obscured galaxy formation, star formation, and black hole formation and growth back to the reionization epoch;
- LISA will be complementary for goal 2 by providing unique information about the (first) mergers of black holes;
- A wide-field optical-infrared imaging telescope capable to measure weak gravitational distorsion of faint objects, will be complementary for goals

3, 6 & 7, and should provide firm evidence of the earliest cosmic star formation and constrain dark matter and dark energy;

- A far-infrared space interferometer will be complementary for goals 3 & 6 in connection with ALMA. It would allow observations of $H_2$ molecules at high redshift (dust-obscured and shock-heated regions).
- A wide-field, highly-multiplexed, low and high resolution spectroscopic survey telescope will be complementary for goals 6 & 7, it will be necessary to disentangle the relations between masses, ages, morphologies, and environments in order to understand the formation of galaxies.

## 6.4 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 determining our place in the Universe. To address these challenges several important goals can be identified:

1. Determine the initial physical conditions of star formation, including the evolution of molecular clouds, and the subsequent development of structures in general, and the formation and mass distributions of single, binary or multiple stellar systems and stellar clusters;
2. Unveil the mysteries of stellar structure and evolution, also probing stellar interiors;
3. Understand the life cycle of matter from the interstellar medium to the processing in stars and back into the diffuse medium during the last stages of stellar evolution;
4. 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;
5. Explore the diversity of exo-planets in a wide mass range from giants to Earth-like, to characterise the population of planetary systems in relation with the characteristics of their host stars;
6. 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, optical/infrared interferometry on the ground and in space, very large effective area and angular resolution millimetre and radio interferometers, transits, microlensing, radial velocity and astrometric planet searches.

**Current facilities**

- The 8–10 m class optical telescopes are essential for goals 1, 2, 3 & 5, and complementary for goals 4 & 6, and need to be fully exploited and upgraded;
- The HST and current generation of X-ray space observatories are essential for goals 1 and 2, and complementary for goals 3 & 4. It will be an asset to keep these facilities at least until a next generation of facilities will become available. Alone and in combination with simultaneous observations in the optical, radio and sub-millimetre domain, they provide important information of all phases of stellar evolution and activity;
- Herschel and JWST will be essential for goals 1, 3 & 4, and will be complementary for goals 2 & 5. High angular resolution and sensitivity in the near- to far-infrared, as provided by space missions provide essential data not accessible from the ground;
- ALMA and its future upgrades will be essential for goals 1, 3 & 4, and will be complementary for goals 2 & 5. High angular resolution and high sensitivity millimetre continuum and line spectroscopy will be key to understand the physical and chemical evolution of dust and gas in the early phases of planet formation;
- Large millimetre single dish telescopes equipped with focal plane arrays for line and continuum observations are essential for goals 1 & 3, and complementary for goal 5, because of their ability to rapidly survey large areas of the sky and find sources for detailed studies of star- and disc formation and evolution with ALMA and the proposed SKA;
- Gaia will be essential for goals 2 & 5, and complementary for goals 1 & 6. Higher accuracy astrometric measurements from ground and space will extend the detection capability and resolve the degeneracy in mass and orbital inclination, quantify the relative orbital inclination of multiple exoplanetary systems and address fundamental topics in stellar structure and stellar populations in the Galaxy.

**Future facilities**

- Near- and mid-infrared imaging and spectroscopy at high spatial resolution and sensitivity provided by an Extremely Large Telescope with high-performance adaptive optics will be essential for goals 1, 2, 3, 4, 5 & 6;
- A next generation of ultraviolet and X-ray missions will be essential for goals 1, 2 & 3, and be complementary for goals 4 & 5. They will provide higher sensitivity, spectral and spatial resolution to study the gas and dust around stars;
- SKA with long baselines and full frequency coverage will be essential for goals 1, 3, & 4, and complementary for goals 2 & 6. It will allow to resolve individual obscured cores of molecular clouds, trace the largest molecules and image the magnetic fields;
- The availability of a both high-contrast and high-resolution sensitive infrared interferometer in space will be essential in order to answer goals 1, 4 & 6, and they will be complementary for goals 2, 3 & 5;
- High-precision photometry (better than 0.01 millimag) and long term monitoring from space in conjunction with a ground-based dedicated network of moderate aperture (2m) robotic telescopes will be essential for goals 2, 5 & 6, and complementary for goal 3;
- The provision of high spectral resolution in the near-to-far infrared has to be included in planning for a fully adaptive ELT or interferometry from space, as the next step in the study of gas dynamics, solid state chemistry and molecular hydrogen 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: This capability will be essential for goals 3 & 6, and will be complementary for goals 1 & 4;
- The next generation of high precision radial velocity monitoring instruments will be essential for goals 5 & 6, and complementary for goal 3; With a velocity resolution better than 0.1 m/sec it allows the detection of earth like planets in habitable zones around solar type stars.

## 6.5 How do we fit in?

The Solar System is uniquely accessible to in situ and detailed remote sensing measurements. To better understand the formation, evolution, and detailed properties of the Solar System, and to understand its impacts on human activities, the following goals have been identified:

1. Utilise the vicinity of Solar System plasmas, in (i) the Sun, (ii) the heliosphere and (iii) planetary environments, to develop a detailed understanding of physical processes which apply to astrophysical phenomena;

2. Develop a unified picture of the Sun and the heliosphere including the planetary environments, including a systems-level view of energy flow from the Sun to the Earth;
3. Understand the underlying mechanisms for Solar variability and transient activity, the subsequent variability in the heliosphere and the resulting impacts on the Earth and other planetary environments;
4. Understand the role of turbulence and magnetic fields in the evolution of the primordial nebula, the mechanism of particle growth, and the elemental and isotopic ratios in this nebula, and in Solar System bodies;
5. 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;
6. 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;
7. 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.

Solar System research requires a wide range of complementing facilities which cover the full electromagnetic spectrum and probe a variety of space environments and planetary surfaces. These facilities are combined into groups of similar characteristics in this list.

**Current facilities**

- The space solar observatories SOHO, STEREO, Hinode, ACE, Cluster, SDO and Ulysses are essential for goals 1–3 until next-generation Solar imaging and spectroscopy and in situ missions are available;
- Ground-based metre-class solar telescopes such as THEMIS, SST and GREGOR remain essential for goal 1(i) and complementary for goals 2-3, until a large aperture solar telescope becomes operational;
- Genesis and Rosetta provide measurements of comets and minor bodies in the Solar System, essential for goals 4 & 5;
- ALMA and LOFAR will provide the capability to do solar observations with a spatial resolution comparable to optical facilities, and will be complementary for goals 1& 2. Similar high resolution observations of planets, comets and trans-Neptunian objects are essential for goal 5;
- Mars Express, Mars Odyssey, Mars Reconnaissance Orbiter, Mars Exploration Rovers, ExoMars, Venus Express, Cassini, and Bepi-Colombo provide in situ measurements of planets which are essential for goals 6 & 7;

- Monitoring of many stars with high-resolution multi-object spectrographs at 4–8m class telescopes is complementary for goal 1.

**Future facilities**

- A solar high-latitude, multiple Solar-encounter mission, climbing to at least 30 degrees out of the ecliptic and approaching to within 0.25 AU of the Sun, including a combination of high-resolution imaging and spectroscopy and in situ instruments is essential for goals 1(i), 1(ii), 2 and 3. It will establish the links between the Sun and its heliosphere as well as the polar regions of the Sun and the three dimentional nature of a star;
- Radio spectral imaging at centimetre to metre wavelengths, is essential for measuring magnetic fields in the corona, to identify sites of particle acceleration and to track travelling disturbances through the corona (goals 1(i), 1(ii) and 2);
- A large-aperture (3–5 m) ground-based solar telescope with adaptive optics and integral-field spectro-polarimeters, with a precision of one part in $10^4$, to resolve scales of order 10 km in the photosphere, to observe astrophysical processes at their intrinsic scales, and thereby observe the interaction of magnetic fields and plasma motions in the Solar atmosphere, is essential for goal 1(i) and complementary for goals 2 and 3;
- A medium-aperture (1–2 m) (extreme-)ultraviolet satellite facility with X-ray capabilities, incorporating sub-arcsecond resolution imaging and spectroscopy, cadences down to seconds and wavelength selections appropriate to the temperature range of the Solar atmosphere – up to relativistic electrons – including, for the first time, (extreme-)ultraviolet magnetic mapping of the Solar transition region and corona, to study fundamental Solar processes that cannot be studied from the ground, is essential for goal 1(i) and complementary for goals 2 and 3;
- An in situ mission probing simultaneously the three major scales of the physical processes in the magnetosphere and Solar wind involving a fleet of spacecraft forming three embedded tetrahedrons. Essential for goals 1(ii), 1(iii) and 2;
- A high quality network of ground-based radars, such as SuperDARN and the next-generation EISCAT facility, providing both global context for the magnetospheric missions and specific detailed measurements in support of conjugate studies. Essential for goal 1(iii) and 2;
- A combination of magnetically conjugate ground-based and space-borne instrumentation to investigate the interchange of plasma populations in the Earth's magnetic environment is essential for goal 3;
- Future exploration of minor bodies in the Solar System, in particular near-Earth asteroids, is essential for goal 4 & 5;

- Space missions to the outer Solar System to explore the Jovian system, in particular Europa, and the Saturnian system, in particular Titan and Enceladus are essential for goals 6 & 7;
- A Mars sample return mission as follow-up to recent orbiter, lander and rover missions, coupled with laboratory infrastructure for sample analysis, is essential for goals 6 & 7. Future exploration of other terrestrial planets is also essential for goal 6;
- JWST and an ELT will allow imaging spectroscopy of planetary atmospheres and surfaces, comets, and trans-Neptunian objects and are essential for goals 6 & 7;
- Participation in an international network of ground-based, synoptic instruments that monitor continuously the full-disc Solar spectral irradiance and magnetic and velocity fields is complementary for goals 1(i), 2 & 3.

# 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 A | Royal Observatory Edinburgh | UK |
| Claes Fransson | co-chair A | Stockholm Observatory | Sweden |
| Jacqueline Bergeron | chair B | Institut d'Astrophysique | France |
| Rob Kennicutt | co-chair B | Institute of Astronomy Cambridge | UK |
| Leonardo Testi | chair C | Osservatorio Astrofisico di Arcetri | Italy |
| Rafael Rebolo | co-chair C | Instituto de Astrofísica de Canarias | Spain |
| Oskar von der Luhe | chair D | Kiepenheuer-Institut für Sonnenphysik | Germany |
| Therese Encrenaz | co-chair 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 | Roadmap lead | Liverpool John Moores University | UK |
| Frank Molster | | NWO | Netherlands |

**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 | | ESO | Germany |
| Subir Sarkar | | Oxford University | UK |
| Peter Schneider | | Universität Bonn | Germany |
| Ralph 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 |
| Richard Harrison | co-chair | Rutherford Appleton Laboratory | UK |
| Willy Benz | | Universität Bern | Switzerland |
| Michele Dougherty | | Imperial College London | UK |
| Artie Hatzes | | Thüringer 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 F. Aharonian, E. Amata, L. Amati, P. Andre, T. Appourchaux, D. Baade, M. Barstow, R. Beck, C. Benn, D. Bomans, S. Borgani, J. Borovička, F. Boulanger, E. Branchini, K. Butler, P. Caraveo, C. Cesarsky, K. Chyzy, P. Crowther, H. Dannerbauer, C. De Breuck, R. Diehl, N. Dimarcq, L. Drury, P. Dufton, A. Efstathiou, F. Eisenhauer, S. Ettori, C. Evans, H. Falcke, L. Foschini, A. Franceschini, M. Gerin, R. Gilmozzi, V. Godunova, A. Gomez de Castro, J. Greiner, M. Haberreiter, W.-R. Hamann, L. Hanlon, M. Hapgood, A. Herrero, I. Howarth, J. Hron, G. Kanbach, A. Kaufer, J. Knödlseder, R. Konstantinova-Antova, M. Kontizas, A. de Koter, M. Kramer, T. Krichbaum, N. Langer, L. Lara, M. Lehnert, D. Lennon, P. Lilje, A. Lobanov, N. Markova, D. de Martino, K. Menten, E. Meurs, S. Molendi, G. Monnet, F. Najarro, F. Nicastro, Å. Nordlund, J. Palous, F. Pasian, L. Pasquini, A. Pauldrag, J.-C. Pecker, E. Pian, L. Piro, N. Przybilla, J. Puls, T. Reiprich, I. Robson, S. Schindler, W. Schmutz, K. Schuster, A. Shukurov, R. Siebenmorgen, S. Smartt, C. Snodgrass, M. Thompson, C. Trundle, M. Urbanik, H. Völk, N. Walton, G. Wolk, and all the participants of the ASTRONET Symposium in Poitiers, January 23–25, 2007.

# 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/`

ACE: The Advanced Composition Explorer is a satellite launched to measure and compare the composition of several samples of matter, including the Solar corona, the Solar wind, and other interplanetary particle populations, the local interstellar medium (ISM), and galactic matter.
`http://www.srl.caltech.edu/ACE/`

ACT: The Atacama Cosmology Telescope is designed 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 saw recently first light. ACT will be located on Cerro Toco in the Atacama Desert of the Chilean Andes at an alititude of 5100 m.
`http://www.physics.princeton.edu/act/`

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

AGB: Asymptotic Giant Branch phase of stellar evolution.

Akari: Akari (Previously known as ASTRO-F or IRIS - InfraRed Imaging Surveyor) is an all sky infrared (1.7 to 180 micron) survey. It was launched in May 2006 and run out of helium in August 2007. It will continue its survey at the shorter wavelengths.
`http://www.ir.isas.jaxa.jp/ASTRO-F/`

ALMA: The Atacama Large Millimeter/submillimeter Array is a Joint Europe–North America–East Asia project to build a synthesis telescope of up to 64 12-m and 12 7-m antennas that will operate at millimetre and sub-millimetre wavelengths at the 5000 m Chajnantor site in Northern Chile.
`http://www.eso.org/projects/alma/`

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

ANTARES: Astronomy with a Neutrino Telescope and Abyss environmental RESearch; a 30 million cubic metre 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 sub-millimetre 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`

ASCA: The Advanced Satellite for Cosmology and Astrophysics (formerly named Astro-D) was Japan's fourth cosmic X-ray astronomy mission. ASCA carried four large-area X-ray telescopes. It operated from 1993 until 2000.
`http://heasarc.gsfc.nasa.gov/docs/asca/asca2.html`

ASPERA: ASPERA is an FP6 ERA-net programme which started in July 2006, and comprises 16 national funding agencies in Europe together responsible for funding astroparticle physics research.
`http://www.aspera-eu.org/`

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`

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

Bepi-Colombo: An ESA mission in cooperation with Japan, which will explore Mercury. It is expected to be launched in 2013.
`http://sci.esa.int/home/bepicolombo/`

Cassini-Huygens: NASA-ESA-ASI mission to explore Saturn and its moons. The Huygens capsule made a successful descent and landing on the surface of Titan in January 2005. Its mission will last until at least 2008.
`http://saturn.jpl.nasa.gov/home/index.cfm`

CCAT: Cornell Caltech Atacama Telescope is a proposed large sub-millimetre telescope in the high Andes of northern Chile, designed 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: The imaging X-ray space observatory Chandra (formerly AXAF) was launched in 1998. 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/

Cherenkov Telescope Array: 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).
http://www.mpi-hd.mpg.de/hfm/CTA/CTA_home.html

Cluster: The four spacecrafts of Cluster are to study the small-scale structures of the magnetosphere and its environment in three dimensions. To achieve this, the four identical spacecrafts will fly in a tetrahedral configuration. The separation distances between the spacecraft will be varied between 600 km and 20.000 km, according to the key scientific regions.
http://sci.esa.int/cluster/

CMB: Cosmic Microwave Background.

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/

CoRoT: The Convection, Rotation and planetary Transits satellite is a 27cm space telescope which is sensitive to tiny variations of the light from stars. It will be used to detect extrasolar planets via micro eclipsing of their parent star and to probe the inner structures of stars via stellar seismology.
http://smsc.cnes.fr/COROT/

COS: The Cosmic Origins Spectrograph is a fourth-generation instrument to be installed on the Hubble Space Telescope (HST) during the 2008 servicing mission. COS is designed to perform high sensitivity, moderate- and low-resolution spectroscopy of astronomical objects in the 1150-3200 Å wavelength range.
`http://www.stsci.edu/hst/cos`

Deep Impact: In July of 2005, the Deep Impact spacecraft released a small, 370 kg Impactor directly into the path of comet Tempel 1 with a speed of about 10.2 km/sec in order to know more about the inside of a comet. The resulting crater and its corresponding ejecta have been observed by the spacecraft and many other observatoria in the world and in space.
`http://deepimpact.jpl.nasa.gov/home/index.html`

DES: The Dark Energy Survey is a high precision multi-bandpass wide-area survey instrument, designed to produce photometric redshifts from $0.2 < z < 1.3$. DES will be mounted at the Blanco telescope at Cerro Tololo Inter-American Observatory. The survey data will cover 5000 sq-degrees, with 4000 sq-degrees overlapping the Sunyaev-Zeldovich CMB survey being conducted by the SPT. DES is expected to start its observations in 2009.
`http://www.darkenergysurvey.org/`

EGAPS: European Galactic Plane Surveys is a confederation of mainly European astronomers, who pursue new advanced wide-area, high spatial resolution, broad and narrow band, deep optical and infrared surveys of the Milky Way.
`http://www.egaps.org`

EISCAT: The European Incoherent SCATter facility consists of three incoherent scatter radar systems, at 224 MHz, 931 MHz, and 500 MHz, in Northern Scandinavia used to study the interaction between the Sun and the Earth as revealed by disturbances in the ionosphere and magnetosphere
`http://ion.le.ac.uk/eiscat/eiscat.html`

ELT: The Extremely Large Telescope will be an optical/infrared 30-50 m class telescope. Design and instrument suites are presently under study.
`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.
`http://cordis.europa.eu/coordination/era-net.htm`

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: The European Southern Observatory is the intergovernmental European Organisation for Astronomical Research in the Southern Hemisphere. On behalf of its thirteen member states ESO operates a suite of ground-based astronomical telescopes located at Cerro La Silla, on Cerro Paranal and, in the future, on Chajnantor in Chile.
http://www.eso.org/

E-VLA: The Expanded Very Large Array project will provide a radio telescope of unprecedented sensitivity, resolution, and imaging capability by modernizing and extending the existing Very Large Array.
http://www.aoc.nrao.edu/evla/

ExoMars: ExoMars is an ESA-led Mars exploration mission currently under development. It is a robotic mission that will consist of a Mars orbiter, a descent module and a Mars rover. The approximately 40 kg exobiology payload will be a lightweight drilling system, a sampling and handling device, and a set of scientific instruments to search for signs of past or present life. The launch is expected in 2013.
http://www.esa.int/SPECIALS/Aurora/SEM1NVZKQAD_0.html

FASR: The Frequency Agile Solar Radiotelescope is a concept for a ground-based synthesis imaging radiotelescope designed specifically for observing the Sun.
http://www.ovsa.njit.edu/fasr/

FLAMES: The Fiber Large Array Multi-Element Spectrograph is a 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-colour 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/

Galileo: The Galileo spacecraft measured Jupiter's atmosphere with a descent probe and conducted long-term observations of the Jovian system from orbit. It found evidence of subsurface saltwater on Europa, Ganymede and Callisto and revealed the intensity of volcanic activity on Io.
`http://galileo.jpl.nasa.gov/`

GBT: The Green Bank Telescope is a movable 100 metre 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 will be launched early 2008.
`http://glast.gsfc.nasa.gov/`

GREGOR: GREGOR is a 1.5 metre solar telescope on Tenerife.
`http://gregor.kis.uni-freiburg.de/`

GTC: The Gran Telescopio CANARIAS, is a segmented 10.4 m telescope that is being installed on the Roque de los Muchachos Observatory (La Palma, Canary Islands, Spain). This Spanish lead project is further supported by Mexico and the University of Florida. First light with a partially finished mirror was achieved in July 2007.
`http://www.gtc.iac.es/home.html`

Hayabusa: Hayabusa (also known as MUSES-C) is a Japanese asteroid sample return mission. Launched in 2003, it arrived at asteroid Itokawa September 2005 and will bring back some asteroid material to Earth in 2010.
`http://www.jaxa.jp/projects/sat/muses_c/index_e.html`

Herschel: ESA's Fourth Cornerstone Mission is a Far-Infrared and Submillimetre Telescope (formerly called FIRST). The satellite, to be launched in 2008, will observe the sky in the wavelength range from 60 to 670 microns. It will act as the higher frequency complement of ALMA.
`http://sci.esa.int/home/herschel/index.cfm`

H.E.S.S.: 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. H.E.S.S. is located near the Gamsberg in Namibia.
`http://www.mpi-hd.mpg.de/hfm/HESS/HESS.html`

HINODE: Hinode, formerly known as Solar-B is a Japanese satellite equipped with three solar telescopes: The solar optical telescope for the observation of Solar magnetic fields, the X-ray telescope and the EUV imaging spectrometer to study the heating mechanism and dynamics of the active Solar corona.
`http://solar-b.nao.ac.jp/index_e.shtml`

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 against its colour. 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 metres. 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 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: Intracluster Medium

IGM: Inter Galactic Medium

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

Integral: The INTErnational Gamma-Ray Astrophysics Laboratory is a gamma-ray observatory in space. Its instruments provide spectroscopy and imaging of gamma-ray emissions in the energy range of 15 keV to 10 MeV. It was launched in 2002. Integral is an ESA mission in cooperation with Russia and the USA.
http://sci.esa.int/integral

IRAM: The Institut de Radioastronomie Millimétrique operates two major millimetre observatories: A 30 meter diameter telescope on Pico Veleta in the Sierra Nevada (Southern Spain), and an array of six 15 meter diameter telescopes on the Plateau de Bure in the French Alps.
http://www.iram.fr/

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 $\mu$arcsec at 15.5 mag at 900 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 three 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 $28\,\mu$m, with a 6.5 m mirror. It is a collaboration between NASA, ESA and CSA. JWST is expected to be launched in 2013.
http://sci.esa.int/jwst/

KASCADE-Grande: KASCADE-Grande is an extensive air shower experiment array to study the cosmic ray primary composition and the hadronic interactions in the energy range $10^{16} - 10^{18}$eV. The experiment is situated on site of the Forschungszentrum Karlsruhe. It measures simultaneously the electromagnetic, muonic and hadronic componenets of extensive air showers of cosmic rays.
http://www-ik.fzk.de/KASCADE_home.html

Kepler: The NASA Kepler mission is designed to detect (terrestrial) exo-planets by the partial obscuration (about 1/10000) of their host star.
http://kepler.nasa.gov

KIDS: The Kilo-Degree Survey is a large (1500 square degrees in four bands) survey to be performed with the VIT 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

Kuiper Belt: The Kuiper Belt is a region of the Solar System extending from the orbit of Neptune (at 30 AU) to approximately 55 AU from the Sun. It consists mainly of small bodies (remnants from the Solar System's formation) and at least one dwarf planet Pluto.

LBT: The Large Binocular Telescope is an optical/infrared telescope that utilizes two 8.4 m mirrors. Due to its binocular arrangement it will have a resolving power equivalent to a 22.8m telescope. It is a collaboration between INAF, Max-Planck-Geschellschaft, University of Arizona, Ohio State University and the University of Notre Dame.
`http://medusa.as.arizona.edu/lbto/`

LBV: Luminous Blue Variable stars

LHC: Large Hadron 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 2008.
`http://lhc.web.cern.ch/lhc/`

LISA: Laser Interferometer Space Antenna is a planned ESA/NASA interferometric gravitational wave detector. It consists of three free-flying spacecrafts that form a triangle 5 million km across, orbiting the Sun trailing the Earth by 20 degrees. They will detect gravitational waves from thousands of galactic binaries, and likely from merging supermassive black holes.
`sci.esa.int/lisa/`

LMT: The Large Millimeter Telescope is a 50 m diameter single-dish telescope optimized for astronomical observations at millimetre 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 USA and aims for first light in 2008.
`http://www.lmtgtm.org/`

LOFAR: The Low Frequency ARray is a high resolution and high sensitivity radio interferometer working at frequencies between 30 MHz and several hundred MHz; Stations will be located in the Netherlands and Germany. Discussions with various international partners in Germany, UK, France, Sweden, Italy and Poland have been started to extend the number of stations.
`http://www.lofar.org/`

LSST: The Large Synoptic Survey Telescope is a proposed ground-based 8.4m, 10 square-degree-field telescope that will provide digital imaging of astronomical objects across the entire sky. The LSST will cover the sky every three nights, opening a movie-like window on objects that change or move on rapid timescales: Supernovae, potentially hazardous near-Earth asteroids, and Kuiper Belt objects. The distortions in the shape of billions of galaxies will be measured to provide multiple tests of Dark Matter and Dark Energy.
`http://www.lsst.org/lsst_home.shtml`

MAGIC: The Major Atmospheric Gamma Imaging Cherenkov telescope is an imaging atmospheric Cherenkov telescope that is located at La Palma and has started measuring since late 2004.
`http://www.mppmu.mpg.de/`

Mars Exploration Rovers: The NASA Mars Exploration Rovers (Spirit and Opportunity) study since January 2004 the surface of Mars in situ. Their primary scientific goal is to search for and characterize a wide range of rocks and soils that hold clues to past water activity on Mars. The spacecraft are targeted to sites on opposite sides of Mars that appear to have been affected by liquid water in the past. They both have several instruments onboard. `http://marsrovers.jpl.nasa.gov/`

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`

Mars Odyssey: The NASA Mars Odyssey mission, launched in 2001 and arrived at Mars in 2002, maps the chemical elements and minerals on the surface of Mars, looks for water in the shallow subsurface, and analyses the radiation environment to determine its potential effects on human health. `http://marsprogram.jpl.nasa.gov/odyssey/`

Mars Reconnaissance Orbiter: Launced in August 2005, the Mars Reconnaissance Orbiter, is looking for evidence of water on Mars. Furthermore it will act as a relay station for other Mars (surface) missions. `http://mars.jpl.nasa.gov/mro/`

Messenger: The Mercury Surface Space Environment Geochemistry and Ranging spacecraft from NASA carries 7 instruments to study Mercury remotely. It has been launched in 2004, and will start orbiting Mercury in 2011. `http://messenger.jhuapl.edu/`

MHD: Magneto-Hydro-Dynamics

MMS: The Magnetospheric Multiscale mission is a NASA Solar-Terrestrial Probe mission comprising four identically instrumented spacecraft that will use Earth's magnetosphere as a laboratory to study the microphysics of three fundamental plasma processes: Magnetic reconnection, energetic particle acceleration, and turbulence. `http://mms.space.swri.edu/`

MOND: Modified Newtonian Dynamics, is a modification of the usual Newtonian force law at small accelerations 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/`

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

Oort cloud: The Oort cloud is an immense spherical cloud of comets surrounding the planetary system and extending approximately 3 light years from the Sun.

Opportunity: See Mars Exploration Rovers.

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

PAHs: Polycyclic Aromatic Hydrocarbons

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/`

Pierre Auger Observatory: The Pierre 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 and is in the final stages of construction already starting to collect data.
`http://www.auger.org/`

Pioneer 10: Pioneer 10 was the first spacecraft to travel through the asteroid belt and reach the outer Solar System. During its Jupiter encounter in 1973, Pioneer 10 imaged the planet and its moons, and took measurements of Jupiter's magnetosphere, radiation belts, magnetic field, atmosphere, and interior. These measurements of the intense radiation environment near Jupiter were crucial in designing the Voyager and Galileo spacecraft.
`http://www.nasa.gov/centers/ames/missions/archive/pioneer.html`

Planck: Space mission designed to image the anisotropies of the Cosmic Background Radiation Field over the whole sky, with unprecedented sensitivity and angular resolution. ESA plans to launch Planck 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 programme. 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 in situ data.
`http://www.estec.esa.nl/spdwww/rosetta/html/index.html`

RXTE: The Rossi X-ray Timing Explorer is an X-ray satellite launched by NASA in 1995. It observes and monitors the Universe with photon energies ranging from 2 to 200 keV.
`http://heasarc.gsfc.nasa.gov/docs/xte/XTE.html`

SDO: The Solar Dynamics Observatory is a satellite that will study the Solar activity and how space weather results from that activity. It will collect data about the interior of the Sun, the Sun's magnetic field, the hot plasma of the Solar corona, and the irradiance that creates the ionospheres of the planets.
`http://sdo.gsfc.nasa.gov/`

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 $km^2$, a wide frequency coverage, large field of view, high spatial resolution and a high time resolution.
`http://www.skatelescope.org/`

SMA: The Submillimeter array is an interferometric telescope at the submillimeter wavelengths, and consists of 8 movable antennas. The maximum distance between the antennas is 500 metre. It is sensitive for wavelength from 0.3 to 1.7 millimetre. It is located at the summit of Mauna Kea.
`http://smawww.harvard.edu/`

SOFIA: The Stratospheric Observatory for Infrared Astronomy is an airborne telescope built in a Boeing 747SP. It will do observations from 0.3 microns up to 1.6mm (although the focus is on the infrared and sub-millimetre observations) while in the stratosphere. The project is supported by DLR and NASA.
`http://www.sofia.usra.edu/`

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/`

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/`

Spirit: See Mars Exploration Rovers.

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 large-area (sub-)millimetre wave surveys of faint, low contrast emission, as required to map primary and secondary anisotropies in the cosmic microwave background. The SPT had first light in February 2007.
`http://spt.uchicago.edu/`

SST: Swedish 1 m Solar Telescope. Vacuum refractor on La Palma, and first solar telescope to reach 0.1 arcsecond resolution.
`http://www.solarphysics.kva.se`

Stardust: The Stardust spacecraft visited comet Wild 2 in 2004 and brought cometary material back to Earth in January 2006. The analysis of the samples is ongoing. Additionally, the Stardust spacecraft brought back samples of interstellar dust. These samples have not yet been analysed.
`http://stardust.jpl.nasa.gov/home/index.html`

STEREO: Solar TErrestrial RElations Observatory is a two-year mission, launched in 2006, 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/`

SuperDarn: The Super Dual Auroral Radar Network is an international radar network for studying the Earth's upper atmosphere, ionosphere, and connection into space.
`http://superdarn.jhuapl.edu/`

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

SVWG: Science Vision Working Group

SWIFT: NASA mission aimed at gamma ray burst studies, launched in November 2004 and is expected to operate for 7 years.
`http://swift.gsfc.nasa.gov/`

SZ-Array: The Sunyaev-Zeldovich Array is a radio telescope whose purpose is to search for clusters of galaxies in the Universe using the Sunyaev-Zeldovich Effect. Information gained from the SZA will be used to derive several cosmological results such as the matter density of the Universe and the evolution of clusters over time.
`http://astro.uchicago.edu/sza/index.html`

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 2007.
`http://www.telescopearray.org/`

THEMIS: The NASA mission Time History of Events and Macroscale Interactions during Substorms launched in February 2007 is a 2-year mission consisting of 5 identical probes that will study the violent colourful eruptions of the aurora that occur during substorms in the Earth's magnetosphere.
`www.nasa.gov/mission_pages/themis/index.html`

TRACE: The Transition Region and Coronal Experiment is a NASA space telescope designed to investigate the connections between fine-scale magnetic fields and the associated plasma structures on the Sun by providing high resolution images and observation of the Solar photosphere and transition region to the corona.
`http://trace.lmsal.com/`

TUNKA: A cosmic air shower detector that makes use of Cerenkov light emission. It consists of 25 wide angle integral detectors. The detectors are deployed in a square of 340x340 $m^2$ in the Tunka Valley, Siberia, Russia.
`http://dbserv.sinp.msu.ru/tunka/`

UKIDSS: The UKIRT Infrared Deep Sky Survey is a near-infrared sky survey, the successor to 2MASS. UKIDSS began in May 2005 and will survey 7500 square degrees of the Northern sky, extending over both high and low Galactic latitudes, in JHK to K=18.3.
`http://www.ukidss.org/`

UKIRT: The United Kingdom Infrared Telescope is a 3.8 m telescope dedicated solely to infrared astronomy, UKIRT is sited in Hawaii near the summit of Mauna Kea.
`http://www.jach.hawaii.edu/UKIRT/`

Ulysses: The joint ESA-NASA Ulysses deep-space mission will make the first-ever measurements of the unexplored region of space above the Sun's poles. The studies encompass the heliospheric magnetic field, heliospheric radio and plasma waves, the Solar wind plasma including its minor heavy ion constituents, Solar and interplanetary energetic particles, galactic cosmic rays and the anomalous cosmic ray component. Other investigations include cosmic dust, interstellar neutral gas, Solar x-rays and cosmic gamma-ray bursts.
`http://helio.estec.esa.nl/Ulysses/`

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/

VISTA: Visible and Infrared Survey Telescope for Astronomy. 4 m telescope to be completed at ESO's Paranal Observatory in 2007. 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 in which signals collected simultaneously by radio telescopes at different locations in the world are correlated afterwards. This effectively allows them to operate as one giant radio telescope with extremely high spatial resolution. There are VLBI networks in Europe (e-VLBI), Japan and the USA which can be combined to a global network. Since 1997 these networks can be combined with the space-based antenna HALCA, resulting in micro-arcsecond resolution.

VLT: ESO's Very Large Telescope, comprising four 8.2 m diameter telescopes on Cerro Paranal in Northern Chile. They operate in the optical and infrared. The first light was in 1998, and all four telescopes were completed by 2000. They are equipped with a full set of instruments and a second generation is under construction.
http://www.eso.org/observing/vlt/

VLTI: The Very Large Telescope Interferometer, which combines the light of up to four unit 8.2 m telescopes of the VLT and/or the up to four auxilary telescopes at a special VLTI focus, to obtain diffraction-limited imaging at milliarcsecond resolution in the near- and mid-infrared. This is a unique feature of the VLT.
http://www.hq.eso.org/projects/vlti/

Voyager 1 & 2: Launched in 1977 and built for a 5 years lifetime, the primary mission was the exploration of Jupiter and Saturn. After making a string of discoveries there the mission was extended. Voyager 2 went on to explore Uranus and Neptune, and is still the only spacecraft to have visited those outer planets. The adventurers' current mission, the Voyager Interstellar Mission (VIM), will explore the outermost edge of the Sun's domain, and beyond. Voyager 1 is farther from Earth than any other human-made object and speeding outward at more than 17 kilometres per second.
http://voyager.jpl.nasa.gov/

WHIM: Warm-Hot Intergalactic Medium

WIRE: The Wide Field Infrared Explorer, was originally meant to study galaxy evolution at high redshift. Unfortunately, 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 acts 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, launched in 2001, is a NASA Explorer mission measuring the temperature of the cosmic background radiation over the full sky with unprecedented accuracy. The satellite will operate until 2009.
`http://map.gsfc.nasa.gov/`

WSRT: The Westerbork Synthesis Radio Telescope consists of 14 radio telescopes, ten are fixed and four can be moved on a rail. The maximum distance is about 3km. They are sensitive for radio waves between 120 MHz and 8,3 GHz. The WSRT is located in the Netherlands.
`http://www.astron.nl/p/observing.htm`

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`

XMM-Newton: The X-ray Multi-mirror spectroscopy Mission is 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-Newton is the high-throughput spectroscopic complement of NASA's Chandra.
`http://sci.esa.int/home/xmm-newton/`

European astronomy has now regained full worldwide competitiveness, being at the forefront in many domains. To strengthen this position and to extend it to all astronomical domains and across Europe, a group of European funding agencies set up the ASTRONET programme, supported by the European Commission, with the goal of establishing a comprehensive long-term plan for the development of European astronomy.

The first stepping stone is the Science Vision for European Astronomy presented in this book. It is the result of intense work by thematic panels, with detailed feedback from the community at large through an open Symposium in which 228 scientists from 31 countries participated.

The Vision analyses the key scientific questions that future research should address and the observing capabilities that are needed to achieve these goals in the next twenty years. The Vision also stresses the need for investments in theory and numerical simulations, high-performance computing resources, efficient astronomical data archiving, as well as in laboratory astrophysics.

This two-year-long effort is being followed by the building of a prioritised roadmap for the observational facilities needed to implement the Vision. This will ensure that Europe fully contributes to humankind's ever deeper understanding of the wonders of our Universe.

ASTRONET
www.astronet-eu.org