We are in the midst of a brilliant era of cosmic exploration. From the discovery of pulsars and quasars in the 1960’s, to the detection of other planetary systems and possible acceleration of the universe’s expansion in the late 1990’s, the pace of astronomical discovery continues to grow. This unprecedented golden age is being driven by the advent of new technologies and powerful, innovative research tools. Telescopes on the ground and in space complement one another and gather signals from across the entire electromagnetic spectrum, while the interpretation and ultimate understanding of these observations is being fostered by increasingly sophisticated theoretical calculations and computer simulations.

This chapter outlines some of the major themes of astronomical research that will dominate the first decades of the 21st century. Emphasis is placed on areas in which Canadians have often played significant international roles. The major scientific theme that will run through all of astronomy in the new century is, in essence, the quest for the ‘Origins of Structure’ in the universe. This quest addresses the formation of planets, stars, and galaxies, as well as the geometry and fate of our universe.

2.1 The Formation of Planetary Systems

How did our own solar system form? How common are planetary systems around other stars, and do any of them resemble ours? And is there life elsewhere? These are among the oldest of human questions. The hypothesis that the Earth orbits the Sun was first proposed nearly 2300 years ago by the remarkable Greek mathematician, Aristarchus of Samos. In the 16th century, Copernicus re-introduced the correct physical picture of the solar system as a set of planets orbiting the Sun, but did not extend the same concept to other stars. Two and a half centuries later, Immanuel Kant introduced the idea that the Sun and planets could have formed by condensing out of a great rotating disk of gas and dust. Today, we know that the gaseous disks that Kant could only speculate about are in fact frequently found around most, if not all, young stars forming in the Milky Way. Within this last decade of the twentieth century, we
have finally begun to find other planets around nearby stars, opening the door to a new era whose impact upon human history may ultimately be more significant than the discovery of the New World. The recent detection of planets around other stars is based on finding the small cyclic motions due to the gravitational pull of a companion planet. The biggest effects are produced by the most massive planets: for example, in our solar system, Jupiter (by far our largest planet, at 318 times the mass of the Earth) causes the Sun to wobble in a 12-year cycle at a speed of 13 metres per second. To detect such tiny orbital motions of other stars that might have their own Jovian-type gas-giant planets, astronomers must be able to measure stellar velocities to an accuracy of just 3 metres per second, the speed of a fast walker! Moreover, these velocity changes must be monitored over many years for their cyclic nature to reveal itself.

The spectroscopic techniques and instruments capable of such high precision measurements were pioneered by a Canadian team over a decade ago. Building on these techniques, several other international groups have now surveyed over 300 nearby solar-type stars; the first discoveries of Jovian-size planets around a few such stars were announced in 1995. Twenty such planets have now been found, most of them with measured masses between 0.5 and 5 times the mass of Jupiter. A major surprise about these new planets, however, is that many of them are much closer to their parent suns than the Earth is to our Sun. Most astonishing of all is the discovery this year of the first planetary system: three Jovian-mass planets have been detected to be orbiting the star Upsilon Andromedae. We now have the first concrete evidence that entire families of planets revolve around stars like our Sun!
A host of new and compelling questions flow from these discoveries. Why are giant planets found so close to their central stars? Jupiter itself probably formed out in the gas-rich suburbs of the primordial solar system and never moved very far in or out from its place of birth. For other stars, their “Jupiters” might have migrated inwards from their birthplaces through the tidal interaction between the newly formed massive planets and the material in their surrounding, remnant protoplanetary disks. Does the fate of the much smaller terrestrial planets in one of these systems depend critically upon how common such processes might be? Planet searching will go into high gear in the coming years, and there is little doubt that, by the end of the next decade, we will have a considerably larger sample of detections to draw from. The theory of planet formation, aided by highly sophisticated computer simulations, will grow into one of the most lively subjects in astrophysics.

A much more ambitious goal of planetary research will be to find Earth-like – and beyond that, life-bearing — planets. This task will require entirely new instruments both on the ground and in space. Terrestrial-sized planets are the lightweights, and so exert gravitational effects upon their stars too tiny to be detected with current instruments. It may become possible to “see” them directly, by the use of still-experimental imaging techniques that block out most of the interfering light of the star itself. This will be an incredibly difficult challenge, but its fundamental importance will almost certainly drive the development of more ambitious technologies later in the 21st century. Ultimately, as the search for terrestrial planets becomes feasible, it may become possible to search for traces of carbon dioxide, water, and oxygen in their atmospheres, as signatures of biological activity.

Along with these fully formed planets, astronomers have now found examples of stars so young that they still have around them the raw material from which their planets will form. These protostellar disks of gas and dust typically extend out to 100 - 1000 AU from their central young stars (where 1 AU or “astronomical unit” is the distance from the Earth to the Sun; about 150 million kilometres). The evidence suggests that stars themselves build up by accreting gas out of these disks, while the dusty residue is used to build the planets, asteroids, and comets. Thus, planet formation and star formation processes are inseparably linked. Both take place in placental, cold dusty disks of gas, and we need to understand the physical state of these gaseous disks if we are to unravel the complete story of planetary systems. To refine this basic picture, we need to learn about the motion and chemical state of the gas and dust in...
real protoplanetary disks. Where do planets form in these disks? How are the gas and dust coupled and how do they evolve chemically? Exactly how long does it take for a planet to form? Such questions can be tackled only if many more examples are studied, and at much higher spatial detail than it has yet been possible to do. The dust in protostellar disks emits radiation typically at infrared wavelengths of a few microns, while the cold molecular gas in the disk is seen at the longer millimetre wavelengths; thus two very different types of telescopes and instruments must be used to gain a complete picture of the disk structure and evolution. A few protoplanetary disks have now been detected, but even with our best current telescopes, they cannot be resolved to better than several tens of AU. Over the next decade, a major goal will be to gain views of protostellar disks at the scale of one AU or even less.

Within our own solar system, we will also learn considerably more from the actual relics of planet formation -- the comets. Comets are small objects which come in from the remote outer parts of the solar system and put on a brief, spectacular show while they pass close to the Sun. Built essentially like large, dirty snowballs, they are the frozen leftover debris of planet formation and thus can potentially tell us a great deal about the early conditions of the primordial solar nebula. Comets are believed to reside in two distinct regions: the first, called the Oort Cloud, is a vast spherical region of space extending from 10,000 to 50,000 AU from the Sun. Billions of comets reside there which are believed to have been flung outward from the inner regions of the solar system at an early era by the massive Jovian planets. The second population of comets is much closer in, residing within the so-called Kuiper Belt, a zone which extends roughly from the 30 AU orbit of Neptune out to 50 AU or more. This “Trans-Neptunian” region is a topic of great current observational and theoretical interest. For example, “planets” such as Pluto and the other large Kuiper Belt objects are actually more akin to the comets in their composition and orbital properties. Canadian astronomers have contributed significantly to the search for Kuiper Belt objects and to studying the way they interact dynamically.
with the inner parts of the planetary system. Better knowledge of how comets are dislodged from their normal locations to plunge into the inner regions of the solar system is, of course, vital to understanding how frequently the Earth will undergo major, and damaging, collisions with them. The cometary collision with the Earth that may have resulted in the extinction of dinosaurs 65 million years ago, released as much energy as 10 million of the world’s largest hydrogen bombs, or the equivalent of about 100 million megatons of TNT! The recent spectacular impact of the comet Shoemaker-Levy 9 with Jupiter in 1994 is a graphic reminder that such destructive impacts are still possible in our solar system.

2.2 The Lives of the Stars

Our Sun is the nearest star. Along with the planets, it is about 4.6 billion years old, and will continue to look much the way it does now for another 5 billion years. Unique among all stars, we can study both its surface and its interior in exquisite and challenging detail. Its hot outer atmosphere is in a perpetually dynamic state as strong magnetic fields channel the hot gases along giant loops and prominences. Sometimes, the magnetic fields change abruptly and release energetic particles into space, potentially to interact with the Earth’s magnetosphere and upper atmosphere in a myriad of ways. All studies of the solar atmosphere to date suggest that solar disturbances are linked to a powerful magnetic dynamo which is in turn driven by the upwelling of hot gas in the deeper layers of the Sun.

Stepping outward from our planetary system, we pass into the stellar realm. Whereas interplanetary distances are a few AU’s, interstellar distances are vastly larger. They are commonly measured in parsecs, where one parsec is equal to about 200,000 AU or about three lightyears. Some stars undergo activity which is hundreds or thousands of times more powerful than that on the Sun. New arrays of radio telescopes now planned by the international community will finally allow astronomers to sample and image tens of thousands of stars comprising their full range of size and activity. At optical wavelengths, sensitive detectors and new ways to process the detected light will probe deeper into the atmospheres of both quiescent and active stars to determine their physical conditions and state of motion.

Stars are ultimately powered by nuclear fusion “reactors” that run deep in their cores. At temperatures of tens of millions of degrees and pressures that are a hundred billion times greater than that exerted by the Earth’s atmosphere upon us, the lighter atomic nuclei at the centre of our Sun, principally hydrogen and helium, are fused into heavier ones. This liberates tremendous amounts of energy. All the other elements found in nature, such as carbon, oxygen, silicon, and iron, are built up this way in successive generations of stars. Nature’s most elusive known particles, the neutrinos, are also released as a by-product of these fusion reactions. The panorama of stellar evolution, that encompasses how a
star burns its nuclear fuel to exhaustion and then either dies away quietly or explodes violently, represents one of the most well developed fields of astrophysics. This subject requires the input from many areas of physics, as well as the use of sophisticated, numerical techniques and modern, high-speed computers.

Our theories and models of stellar structure will be tested during the coming decades to a level almost unimaginable 15 years ago. Although the Sun’s deep interior cannot be directly seen, the rates of nuclear reactions there can be directly measured by detecting the flux of all types of neutrinos that travel to us from the Sun’s center. The measured flux of one type of neutrino is a factor of two smaller than predicted by our best computer models of the Sun. The solution to this well-known puzzle may lie in the ability of the neutrinos to change their type. New international neutrino observatories, including the Sudbury Neutrino Observatory (SNO) with its unique detector consisting of 1 kiloton of Canada’s reserve of heavy water, as well as the Super Kamiokande observatory in Japan with its 22 kilotons of ordinary water, have been designed to resolve this problem. They will thereby make critically important contributions to both astrophysics and particle physics.

A second remarkable way of deducing the interior conditions within stars is through the new technique of stellar seismology – in essence, by analyzing how stars vibrate. In the same way that geologists can use the aftershocks and tremors from earthquakes to deduce the interior structure of the Earth, the way that the Sun “rings” in response to the disturbances created by its own boisterous inner convective motions can tell us a great deal about its internal structure. Astronomers measure the tiny stellar oscillations using highly sophisticated spectroscopic observations from ground-based optical telescopes, and increasingly, from space-based observatories. The coming decades will allow us to extend these studies to other stars of very different sizes, masses, and evolutionary states.

The evolution and fate of a star depends upon its mass and composition. The smallest stars that can become hot enough in their interiors to ignite nuclear fusion are about 0.08 times the mass of the Sun, while the biggest known are nearly 100 times the Sun’s mass. The low-mass stars live many billions of years, die out rather quietly, and do very little to their surroundings. High-mass stars, by contrast, live short but extravagant lives as they burn their fuel at prodigious rates for only a few million years before finally exploding as supernovae. These supernovae inject huge amounts of energy back into the gaseous interstellar medium, along with the heavy-element products of the nuclear fusion chain. Thus the massive stars, although they are small in number compared with the low-mass ones, are the ones which drive the grand cycle of stellar history in our Galaxy, the Milky Way. From interstellar gas clouds, stars are born and begin their active nuclear burning lifetimes; they evolve, eject their processed material back into space, and the material mixes with other interstellar gas to create a new generation of stars. The powerful outflows of gas from massive stars also exert strong effects on the heating of the interstellar gas and the rate at which it can form new stars. It is also no
exaggeration to say that all living things upon the Earth, and probably other planets, are the children of the stars. The elements heavier than hydrogen and helium, out of which we are made, were created in the interiors of earlier stars.

The supernova explosion, which is the endpoint for stars more massive than about 8 times the mass of the Sun, is the most spectacular stellar phenomenon. For a brief period of days or weeks, the expanding supernova shell can be brighter than the combined light from billions of normal stars, making it a beacon bright enough to be seen even in a remote galaxy. The famous supernova SN1987A, first discovered by a Canadian astronomer, was the visible result of the destruction of a 15-Solar-mass star in our nearby dwarf companion galaxy, the Large Magellanic Cloud. It has been the most closely studied supernova in history and has played a crucial role in allowing astrophysicists to compare the predictions of models with the real thing. A particularly important result was the confirmation that the supernova detonation event, which is set off by the sudden collapse of the dense core of the star, should release a huge flux of energetic neutrinos. Neutrino bursts were received from SN1987A at detectors in Japan and the USA just before the optical light wave arrived. The next comparable supernova explosion in our galaxy, or its neighbours, will be readily detected by their successors at SNO and Super-Kamiokande.

A central neutron star or black hole may be left behind as the dead remnant of the original star’s core. A neutron star has a mass similar to that of the Sun, yet may be no more than 20 km in radius. Such an object can be denser than even the nucleus of an atom, and a full understanding of such a compressed state of matter is still being explored. But it is the black hole which represents the ultimate condition of collapsed matter. If the entire core of the star collapses to within a critical radius (equal to about 3 kilometers for one solar mass of material), it becomes completely sealed off to outside view: not even light can escape from it, and outside material such as gas from neighboring stars can be swallowed up in it but cannot come back out. Nothing can prevent such a collapsing star from reaching a state known as a singularity at its centre. Such an entity is governed by both quantum mechanics and general relativity for which there is, as yet, no unified physical theory. By carefully measuring the orbital motions of stars in binary star systems that appear to have massive, but unseen companions, astronomers now have superb evidence for the existence of at least 8 stellar mass black holes in our galaxy. What appeared to physicists 70 years ago to be an inconceivable consequence of Einstein’s general theory of relativity, has in the late 1990’s become a near certainty; black holes exist. We will surely need to know how common they are in the Galaxy and to learn how they are formed.

Neutron stars may also represent the key to understanding the remarkable phenomenon of gamma-ray bursts that have been observed by satellites since the early 1970’s. These are short-lived bursts of extremely high energy radiation, visible only from satellite gamma-ray detectors, but which occur everywhere around

This striking HST image of supernova SN1987A in the Large Magellanic Cloud shows bright rings of glowing gas many light-years across, encircling the site of the explosion.

Image by Christopher Burrows (ESA/STScI) and NASA
the sky. Very recently, the “afterglow” from a few of these bursts has been seen at longer wavelengths (X-ray, optical, and radio), and the evidence now suggests that these are occurring in galaxies at “cosmological” distances in the early universe. The huge amounts of energy that pour out of such bursts could be released during the rare collision and merger of two neutron stars, with the end result being a black hole. The potential exists to use these unique objects, not just to investigate stellar collapse to a black hole, but to probe conditions within galaxies at much earlier times.

2.3 Star Formation

How do stars actually form? Once a star has emerged from the raw material of the interstellar medium, we can predict in impressive detail how it will evolve and end. But its very first stage is still poorly understood: it is an extremely challenging problem on both observational and theoretical grounds, and a fully detailed theory of star formation has not yet been constructed. Yet it is crucial to further progress in understanding the complete history of the Galaxy. Star formation sets the stage for the process of planetary formation, and the feedback between actively forming stars and their large-scale surroundings, through stellar winds and supernovae, affects the formation and evolution of an entire galaxy. No true progress can be made in any of these fields without a deep understanding of how star formation occurs, here and now.

A particularly important observed quantity that a successful theory must explain is called the Initial Mass Function (IMF), which is the relative number of stars formed at a given mass. The IMF prescribes the numbers of high-mass stars (and thus the rate at which the composition of the Galaxy’s raw material can cycle through supernovae), as well as the numbers of the so-called “brown dwarfs” – stars whose mass falls below the 0.08-Solar-mass limit for hydrogen fusion. Brown dwarfs are ultra-faint, making them exceedingly hard to detect; but if large numbers of them exist, they could make up a significant fraction of the ubiquitous “dark matter” that dominates the mass of the galaxies. The new generations of optical and infrared telescopes, such as Gemini and NGST, will play major roles in the search for brown dwarfs and the study of star-forming nebulae all over the Milky Way.

Active star formation takes place in the disk of our own Milky Way Galaxy inside special types of dense nebulae called molecular clouds. These clouds were discovered serendipitously in the mid 1960’s, by astronomers using radio telescopes tuned to the millimeter wavelengths at which many kinds of molecules in space emit radiation. A molecular cloud in the Milky Way may contain up to several million solar masses in cold gas, over a region of space tens of parsecs across. The temperatures within these clouds are normally 10 - 20 degrees above absolute zero (or about minus 250 degrees Celsius) and the gas is mostly
in the form of hydrogen molecules, with trace amounts of more easily detected material such as carbon monoxide (CO). It is primarily the millimetre-wavelength emission of CO molecules that allows us to map and study these clouds. Astronomers have, in the last few years, mapped molecular clouds in other nearby galaxies and have even found CO emission from galaxies that are only a few billion years old!

The new tools for observing at infrared and millimetre wavelengths that began to be developed over the last 20 years have revolutionized the observations of star forming regions. The optical light from stars still embedded in their natal molecular clouds is blocked because it is absorbed by dust that is mixed in with the cool, dense, molecular gas. However, infrared light from young stars and heated dust, as well as the millimetre radiation from gas molecules, can propagate almost unhindered out of these dense dusty realms. The enormous amount that we will learn about star formation in greater detail provide one of the strongest motivations for investing in the more advanced infrared and millimetre telescopes that will come on line in the next decade.

Molecular clouds would quickly undergo gravitational collapse without some means of internally supporting their huge weight. They are too cold to be supported simply by the internal pressure of hot gas, as stars are; instead, they rely upon the pressure that is provided by the magnetic fields that are woven through them. The actual conversion of clumps of gas into stars takes place within smaller, high pressure regions known as molecular cloud cores. New, high-detail infrared images of molecular cloud cores also make it clear that stars seldom form in isolation, but emerge as members of stellar groups and clusters. Star formation is a highly gregarious process, for reasons that still need to be discovered.

Many regions within our Milky Way galaxy can be found where dense amounts of gas and dust have collected together. These nebulae are the sites of new star formation. In this nebula, called RCW38, a cluster of stars at the center is lighting up the surrounding filaments of gas and dust from which it formed.

Image from Very Large Telescope (VLT) at the European Southern Observatory (ESO), Chile. Copyright ESO

A “ridge” of molecular gas many light years long within the Orion nebula, as seen at a wavelength of 450 microns. Two bright knots near the center indicate the densest parts of the ridge where individual stars are beginning to form. This image is taken with the SCUBA instrument on the JCMT telescope.

Courtesy Dr. Doug Johnstone, CITA
Finally, individual stars within such clusters may form through the gravitational collapse of individual molecular cloud cores; regions of gas much less than a parsec in size. As a slowly rotating core collapses inward, it rotates ever more rapidly to conserve the total amount of “spin” (angular momentum) held by the original extended cloud. Exactly the same type of process occurs as spinning figure skaters pull in their arms in order to spin even faster. The increasing spin of the collapsing gas prevents it from reaching the central protostar directly; instead it forms a rotating disk. Astronomers are making a concerted effort to detect such collapse motions, and to observe the process of disk formation and evolution. Such research programs require the extremely powerful millimetre wave and infrared telescopes that are being planned for the coming decade.

One of the most spectacular and beautiful aspects of star formation is the pervasive appearance of violent outflows and jets from the vicinity of young stellar objects. No theory anticipated the discovery, in 1980, of highly energetic molecular outflows that accompany the birth of all stars. An outflow is now recognized to consist of gas that is swept up by the pressure of an underlying, highly collimated jet. While most molecular outflows move at tens of kilometres per second, jets have speeds in the range 200 – 400 kilometres per second and may be several parsecs long. Jets and outflows carry off much of the energy and angular momentum that is released as the accretion disk spirals into the central protostar. Outflows and jets are closely linked with the presence of accretion disks around young stars. They are also among the first visible sign-posts of the star formation. Sorting out the currently contending theories for the way the jets are created and how they are linked to the formation of stars is a major challenge that will require high-detail observations at optical, infrared, and millimetre wavelengths, capable of probing to within an AU of the central protostar. Investigating the characteristics of accretion disks, jets, outflows, and the entire sequence of steps in star formation will provide one of the major scientific themes of our Long Range Plan for the next decade.

2.4 Galaxy Formation and Evolution

The outward step from stars within our own Galaxy, to the realm of the galaxies themselves, is a huge one. Distances between these immense “cities of stars” are measured typically in millions of parsecs, between which are unimaginably vast stretches of nearly empty space. The Milky Way is a collection of perhaps a hundred billion stars along with several billion solar masses worth of interstellar gas, all held together by their mutual gravity. Galaxies are the great arenas in which stars, planets, and life itself play out their history.
Some galaxies are much larger than the Milky Way; but most are much smaller. Galaxies vary considerably in their mass and take on several different forms; from the small “irregulars” that have quite chaotic and unorganized structures and contain lots of gas, to the “disk galaxies” with their elegant spiral arms rotating like majestic pinwheels around a bright nucleus, to the “ellipticals” that are structurally simple, enormous collections of stars. In the first few decades of the 20th century, it was commonly thought that large galaxies were “island universes” living passively in ever growing isolation as they are carried farther apart by the cosmic expansion. Today, little is left of that simple picture. Most galaxies are found gathered together in groups and clusters; the sparsest of these may contain as few as a dozen individual galaxies, while the largest hold thousands. Galaxy clusters and superclusters define vast filaments and sheets spanning hundreds of millions of parsecs that are separated by even larger voids. How did this proliferation of structure arise? What characteristics of their raw gaseous material and dark matter control the form of galaxies at any time and their distribution in space?

Galaxies as a whole should be viewed as dynamic entities whose masses, gas content, and stellar populations evolve with time. Small galaxies that are gas-rich can completely transform their appearance during sporadic “starbursts” in which much of the gas condenses quickly into stars. The huge numbers of supernova explosions that accompany such bursts are sufficient to blow much of the remaining gas out of these tiny galaxies and scatter their heavy elements into the highly diffuse, intergalactic medium in which galaxies themselves are immersed. Larger galaxies such as the Milky Way can grow by accretion, as occasional small satellite galaxies fall into them. The largest galaxies of all, the giant ellipticals at the centres of large clusters of galaxies, have been found by satellite X-ray imaging to be surrounded by huge, tenuous halos of gas heated to millions of degrees; the total amount of mass in this gas in some cases adds up to considerably more than the mass in all the stars of all the galaxies in the cluster.

In large groups of galaxies, the “intracluster space” between the galaxies is often filled with hot gas that is detectable at X-ray wavelengths. This hot, tenuous material can exceed the mass in all the individual galaxies, and may represent gas left over from the early formation stages, mixed in with some ejected from within the galaxies by supernova explosions. This ROSAT X-ray picture of the Coma cluster of galaxies shows a vast, 6 million light year wide cloud of this high-temperature gas.

The Large Magellanic Cloud, a satellite of our own Milky Way, is about 160,000 light years away but is easily visible to the naked eye in the night sky of the southern hemisphere. Patches of active star formation (light red) are scattered all over the older main body of this small galaxy.

This elegant pinwheel form is the spiral galaxy NGC1232, in the southern sky. Our own Milky Way galaxy may look like this, with bright, young stars and gas clouds lining up along the spiral arm segments.

The Large Magellanic Cloud, a satellite of our own Milky Way, is about 160,000 light years away but is easily visible to the naked eye in the night sky of the southern hemisphere. Patches of active star formation (light red) are scattered all over the older main body of this small galaxy.
Much of the gas may be unused material left over from the early cosmos, while some of it is material ejected from within the galaxies by supernovae.

Many large galaxies have supermassive black holes at their centres that are from millions to billions of times more massive than the Sun. Astronomers can measure the masses of these “monsters” in the hearts of galaxies by studying the orbital motions of luminous, radio-emitting, blobs of gas as well as those of luminous massive stars, around the centres of galactic nuclei. The presence of supermassive holes may be the only viable way of explaining the energy liberated at the centres of super-active galaxies. The nuclei of these galaxies generate up to 100,000 times more energy than the total liberated by all of their stars, in a region that is much smaller than a parsec in size (such a region can easily accommodate even a billion solar mass black hole, which is only 20 AU in size). The engine that is almost certainly responsible for this copious output involves the accretion of gas through a disk that orbits a monster black hole. The efficiency of such an accretion process onto a central rotating black hole rivals that of nuclear reactions, making the accretion of gas onto black holes nature’s most efficient form of energy conversion. By feeding massive black holes with sufficient gas (about a mass of our Sun per year in the most extreme cases), the nuclei of galaxies are transformed into the active high-energy states called Seyfert galaxies or even quasars, the most luminous types of galaxies. The universe currently harbours relatively few quasars in comparison with the large numbers of quasars that were active billions of years ago as galaxies were being assembled. This is another sign that galaxies evolve significantly with time and that the monsters lurking at their centres are now being relatively starved for gas.

One of the greatest adventures of astrophysics over the next two decades will be to understand the long-vanished first epoch of galaxy formation. It is almost certain that the most active epoch of galaxy formation happened in the first 20 percent (3 billion years) of the universe’s history and that the present-day level of activity, fascinating though it is, can be only a shadow of its former self. How do we gain direct evidence of this — that is, how do we catch galaxies in the act of formation?

The direct route is to take advantage of the finite speed of light. The farther away that a galaxy is, the longer it has taken for its light to reach us. Thus, we view it as it was, eons ago. This “lookback time” gives us, in essence, a time-machine that allows us to view faint, distant galaxies as they were only a few billion years after the Big Bang. The new tools deployed during the past decade, such as the Hubble Space Telescope (HST) and the giant 8 metre class telescopes on the ground, have just started to give us the ability to bring such remote targets within reach. These extraordinary new observations are beginning to show that the further back in time we look, the patchier the distribution of galaxies seems to become. The scene at these early times is dominated more and more by smaller, more gaseous pieces, rather than a few large, fully formed galaxies of stars.
These results, preliminary though they are, give considerable support to theoretical interpretations which predict that galaxies build up by a “bottom-up” process wherein small systems merge to form larger ones. As we look even further back to a time before galaxies had converted their gas into stars, we must also employ radio telescopes that allow us to make high resolution studies of the state of protogalactic gas. The new generations of world observatories will play pivotal roles in these studies.

A complementary approach is to deduce the histories of galaxies through our increasing ability to resolve and study the individual stars within other galaxies. The new telescopes on the ground and in space now allow us to see relatively nearby galaxies (ones only a few million parsecs away) in considerable detail — star-forming gas clouds, clusters of stars, and even the individual stars themselves. The advantage of studying nearby galaxies is that their constituents can be measured in immensely more detail than cosmologically remote objects. In a remarkably direct sense, what we do by probing the makeup of nearby galaxies resembles archaeology on a cosmic scale; the “layers” of stars formed at different times within the galaxy can be either simple or complex depending on their individual histories. The very oldest parts of these galaxies, and the objects most interesting for cosmology, are the globular star clusters that are tightly clustered groups of thousands of old stars (their study has long been a Canadian specialty). The ages of the globular clusters are in the range from 12 to 14 billion years, according to the best combinations of stellar evolution theory and observational data that we have at present. Because these ancient objects are present in every large galaxy, they provide what seems to be a common thread in the story of galaxy formation, and give us a key link between stellar astrophysics and cosmology.

Galaxies formed in the early universe by transforming their raw material — the hydrogen and helium gas emerging from the Big Bang, along with the ubiquitous, non-luminous “dark matter” — into the forms we observe today. Thus an understanding of galaxy formation flows seamlessly from some of the basic ideas of cosmology. The current large-scale picture that we have of galaxy formation is that the original, tiny ripples in the density of matter that emerged out of the Big Bang (see below) should grow with time. Eventually, these regions become dense enough to become distinct entities that separate out of the general expansion. The gas that collected into the resulting dark matter “wells” cooled to form the clouds that were probably the nurseries for the first stars that were born in our Universe. The radiation from these first stellar objects lit up the Universe as galaxies began to assemble and grow. Theoretical modeling has advanced stride for stride with the new observations, and will continue to do so as it is driven onward by computing technology. N-body simulations involving millions and even billions of simulated “particles” will acquire
such a level of detail that deep comparisons will be possible between our models and the rich data that the next decade’s complementary observatories will be amassing. Realistic simulations must include dark matter, stars, and gas. An essential point brought home by these numerical experiments is that the dynamic range of the formation process is enormous: it involves the clustering of the gas and dark matter on all scales from molecular clouds within the protogalaxies up to whole groups and superclusters of galaxies on scales a million times larger. Galaxy formation is, therefore, at once a local and a global process, and our theoretical and computational techniques will need to become ever more sophisticated in order to deal with this.

2.5 Cosmology

Humans have always wondered: What is the universe made of? How did it begin, and how will it end? These questions are simple, yet profound. Our current understanding of the structure of the universe on its ultimate large scales follows from the remarkable theory of general relativity, published by Albert Einstein in 1915. It sets out a view of gravity and space that has since become one of the most exhaustively tested theories in all of physics. Its basic tenet is that matter curves the very structure of space and time, much the way that a ball lying on a taut rubber membrane stretches its surface. At the same time, curved space tells an object how to move as it travels through the curves and valleys.

On the very largest scales that we can observe, now up to billions of parsecs, galaxies and clusters of galaxies appear to be uniformly spread throughout the universe. Einstein showed that a universe containing such a homogenous
distribution of matter cannot simply be static; it must either expand or contract with time. In the 1920’s, Vesto Slipher and Edwin Hubble used measurements of the distances and motions of galaxies around us to reveal that, in fact, the galaxies recede from each other: they had discovered the expansion of the Universe, now called Hubble’s Law. The rate of expansion that is deduced from this – the average velocity of the galaxy divided by its distance — is known as the Hubble constant, $H_0$. This concept can be visualized by imagining galaxies to be like raisins embedded in a loaf of baking bread. As the dough rises, it carries each raisin increasingly farther away from every other raisin. Just like the dough in this example, it is the expansion of space-time itself that carries galaxies increasingly farther apart from one another.

An immediate consequence of Hubble’s Law is that the size of the universe changes with time. It must have been much more “compact” in the past, originating in a “Big Bang”. Huge efforts have gone into measuring the value of the Hubble constant $H_0$ because, according to the simplest form of the Big Bang model, it gives us a measure of the age of the universe. Current measurements of $H_0$ translate into an “expansion age” of 10 to 13 billion years for standard cosmological models. Depending upon the exact value, there could be a major outstanding problem in reconciling the age of the universe and the ages of its oldest stellar systems, the globular clusters. If the universe is undergoing accelerated expansion, then the best current models suggest that it could be 15 billion years old.

Two other consequences of the Big Bang model have been beautifully confirmed. Just as hot gas cools when it expands, so too does the universe. The expected fossil radiation from the era before the galaxies existed, when the universe was nothing but a mixture of dark matter and hot, dense, and opaque gas, was detected in 1965. This is the Cosmic Microwave Background Radiation (CMBR) that fills all of space, the visible relic of this long-vanished early epoch. It was emitted when the universe was cool enough to become transparent to this radiation and for the radiation to decouple from the matter; when it was less than a million years old. Today, the CMBR appears as cold microwave radiation whose temperature is just 2.735 degrees above absolute zero. The tiny ripples in it (no larger than a few parts per million) detected by the COBE satellite mission a decade ago may represent the seeds of regions where clusters of galaxies would eventually grow. Astronomers are making concerted efforts both in ground and space-based observatories to measure these tiny fluctuations down to much smaller angular scales, because these contain the direct record of how structures such as the galaxies evolved from a nearly smooth pre-galactic universe. The coming decade will see the launch of several important new missions whose task will be to map this microwave background radiation on much finer scales than COBE achieved.

A third triumph of the Big Bang picture came from the realization, half a century ago, that there must have been a time in the early universe when the material was hot and dense enough to resemble the interior of a star. During this
brief epoch, nuclear fusion reactions would have occurred, starting with the basic subatomic particles (neutrons and protons) that were present then. This era of element production happened when the universe was less than three minutes old and had cooled to a temperature of a billion degrees. The ongoing expansion quickly cooled the gas below the point where fusion could continue, but the end result of this brief phase of nucleosynthesis turned the material into a mixture of very simple elements: primarily hydrogen, some helium, and trace amounts of lithium and deuterium. The detailed calculations from the model are found to correspond superbly with the proportions of these elements that are found in stars. Thus the light elements in today’s universe are the remnants of the Big Bang; all the heavier ones are the later result of nuclear fusion within the stars themselves.

But despite the resounding successes of the standard Big Bang model, a number of major unresolved issues remain. Whether or not the universe will continue to expand forever must still be determined. The answer depends in part upon making good measurements of the average density of matter in cosmological space. According to general relativity, the geometry (overall curvature) of the universe depends upon this average mass density: if it is too low, the mutual gravitational attraction of all the galaxies and dark matter will be too weak to bring the eternal expansion to a halt. If the density is, however, higher than a certain threshold value, then the universe will expand to a maximum size and then recollapse to a “Big Crunch”. The ratio of the average density of all kinds of matter in the universe to this threshold value is called Ω₀; if Ω₀ is less than 1, then the universe will expand forever.
Many of the major observational programs in cosmology are designed with the goal of measuring this elusive density ratio. Since dark matter makes up most of the mass of the universe, astronomers have devised sophisticated techniques to measure both its luminous and dark matter content. The presence of dark matter is determined by its gravitational influence upon luminous matter. Thus, the orbits of stars within disk galaxies are largely determined by the distribution of dark matter within them. The same can be said about the motions of galaxies within rich galactic clusters. Finally, astronomers can also map out the distribution of dark matter in galaxies and clusters of galaxies by looking at their ‘gravitational lensing’ of yet more distant objects. This important technique is based upon the prediction of Einstein’s theory of general relativity that light, too, follows a curved path as it nears and passes concentrations of matter. This wide variety of observations, some performed by Canadian groups using the Canada-France-Hawaii telescope, tend to give a value of $\Omega_0$ of at most 0.3, well below the threshold density. However, this result does not explain another major fact, namely, the observed smoothness of the CMBR radiation across the sky.

In the standard Big Bang model, different regions of the universe could have begun expanding at different moments and therefore would have very different temperatures today. Somehow the universe must have been homogenized at early times. While the standard picture outlined above has no explanation for this, a new model developed in the 1980’s called “inflationary cosmology” does. In this picture, our visible universe originated from a finite region which, when it was much less than the size of a proton, underwent an era of tremendously rapid (inflationary!) expansion in its first, unimaginably small, fraction of a second. This initially tiny region subsequently settled into the normal expansion state described by the standard Big Bang model. Though this approach explains the smoothness of the CMBR, it also predicts that the density ratio $\Omega_0$ should be almost exactly 1.0, which is at odds with many of the current painstakingly accumulated measurements discussed above.

A new breakthrough in observational cosmology that has the potential to address these puzzling difficulties is the putative discovery that the universe may actually be accelerating its rate of expansion. The surprising faintness of supernovae in very distant galaxies suggests that they are actually farther away than the standard model of the Hubble expansion would predict, which implies perhaps that the expansion rate of the Universe is accelerating. This could be driven by the “vacuum energy” that Einstein introduced into his general relativity equations, though he later considered it his “greatest mistake” after Hubble’s discovery of the expansion. This “cosmological constant” measures the amount of the putative vacuum energy density in...
space. The coming decades will see concerted efforts by astronomers to survey the universe even more deeply to test this most remarkable idea.

On the threshold of a new century, modern astronomy and astrophysics are poised to attack many of the most challenging and profound questions in all of science. The scientific goals of elucidating the origins of planets, to those of stars and galaxies, and outwards to the very structure of the universe itself, will require the construction of a series of complementary new observatories that will project far beyond the capabilities of our current facilities and require the development of many new technologies. We turn next to survey the achievements and capabilities of Canada’s current complement of observatories and their role in the coming years.