1

Introduction

The discovery of the expanding universe in the first half of the twentieth century is one of the most exciting exploits in the history of astronomy; it is the main theme of this book.

An intense and occasionally even fierce controversy divided the astronomical world at the beginning of the twentieth century: does the Milky Way represent the whole Universe, or is our Cosmos composed of an enormous or even infinite number of island universes, each of them similar to our own Milky Way? The question was definitely settled in 1925 in favour of island universes, a concept that had been proposed by Immanuel Kant in 1755. Yet, even before the debate was closed, that grand silent Universe had become the playground of theoreticians, who tried to fathom its structure with their abstract mathematics.

Modern cosmology began in 1917 when Albert Einstein published his 'Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie' (Cosmological considerations on the general theory of relativity). In his theory, space and time form a unity: *spacetime*. They are no longer absolute and independent concepts, the structure of spacetime is intrinsically related to the material and energetic content of the Universe: spacetime is structured by gravitation.

More than two thousand years of astronomical observations showed the Universe to be stable and practically immutable in space as well as in time. Thus, when Einstein merged relativity with cosmology, he quite naturally wanted to accommodate a static universe that did not change with time. Yet, gravitation makes matter condense. With the aim of counteracting that force and maintaining the stability of the Universe, Einstein introduced his famous cosmological term, now generally called the cosmological constant. The resulting solution provided bold answers to age-old questions: the Universe is finite, we know its size, and we know how much matter it contains.

However, the idea of a static universe was soon challenged. In 1922, the Russian Alexander Friedmann showed that the most natural solution of Einstein's fundamental equations of General Relativity is a dynamic universe: it may expand or contract.

Redshifts play an important role in cosmology. A prism splits white light into a continuous range of colours from red to blue, as seen in the rainbow. A range of wavelengths defines the portion of light corresponding to a given colour: short wavelengths for blue, longer ones for red. Stars and galaxies consist of chemical elements like hydrogen, oxygen, carbon and iron. When these elements radiate or absorb light they produce in their spectra characteristic patterns of lines at fixed wavelengths. Often the spectra of stars show that the whole set of these patterns is shifted away from the wavelengths expected from our laboratories. The best known explanation for such wavelength shifts is the Doppler shift: if a radiating object is moving towards us, its spectrum is shifted to the blue; if it is moving away from us it is shifted to the red. Since 1912, Vesto Slipher had observed that spectra of spiral nebulae showed very large redshifts.

The first scientific text – backed by theoretical and observational evidence – that explicitly advocated an expanding universe in the sense we see it today was published by Lemaître in June 1927. He did not know about Friedmann. When he theoretically re-discovered the dynamical universe, he combined that theory with observations, and showed that we live in an expanding universe. His principal observational proof came from spiral nebulae, for which he combined Slipher's redshifts with distances, published by Hubble in 1926.

In the autumn of 1927, Lemaître showed his discovery to Einstein. But Einstein would not hear about a dynamically evolving world. In 1922, he had already shrugged off Friedmann's discovery, and he told Lemaître that his model of the expanding universe seemed mathematically correct, but it was physically abominable. Only in 1931, after Eddington and de Sitter had enthusiastically welcomed Lemaître's publication of 1927 as a true breakthrough for cosmology, and Hubble together with Humason had in 1929 observationally confirmed Lemaître's velocity–distance relationship and its implication that galaxies move away from us, did Einstein finally accept the new concept.

Scientific discoveries are never the sole merit of one individual. They have to be seen in a historical perspective. Yet, all the milestones are tied to outstanding individuals. This is no contradiction. Masterstrokes of individuals become possible when a favourable territory has been prepared and a propitious environment provides indispensable support. This was as valid for Newton as for Einstein. The story we have to tell shows that even geniuses can go astray or fail to recognise the obvious solution to a burning problem.

After an introductory, very condensed summary of Ptolemy's *Almagest*, we start our journey through history in the Middle Ages with the revival of astronomy in the West. We concentrate on developments that are relevant for the eventual discovery of the expanding universe. The cultural revival owed much to close contact with Islamic culture, which had preserved and further developed the classical Greek scientific tradition. These contacts were intensified after the re-conquest of Toledo by the Spaniards in 1085. The translation of classical Greek and Islamic authors from Arabic into Latin in the twelfth and thirteenth centuries gave an enormous boost to Western culture.

We begin our history with Johannes de Sacrobosco. He belonged to a group of scholars who took advantage of the Islamic heritage to create a European astronomy. From him we learn how the scientific community of the thirteenth century saw the Universe. Then, in Cusanus, we meet an eminent philosopher of the early Renaissance, whose ideas about the Heavens fundamentally diverged from traditional scholastic thinking. The following generations – including Nicolaus Copernicus, Giordano Bruno, Johannes Kepler and Galileo Galilei – belong to the 'heroic' set of the sixteenth and early seventeenth century who fought to replace the geocentric by the heliocentric view of the world. In that period we also find the first Western description of a galaxy.

Essential ingredients of modern astronomy entered in 1644 with Descartes' concept of a universe in perpetual evolution, and Newton's formulation of the law of gravitation in 1687. Nebulae became a subject of speculation and investigation with Halley's report to the Royal Society in 1716; Kant first proposed the concept of galaxies in 1755. A new era of observational cosmology began with the giant telescopes of Herschel and Rosse, and the spectroscopy of Huggins. After a few decades of calm, photography greatly helped to revive interest in nebulae, and the debate as to whether some nebulae represented island universes was resumed. In 1925, Kant's hypothesis had definitely won the day. By this time, modern cosmology, which had started with publications by Einstein and de Sitter in 1917, was already on its way.

Our principal aim is to bring together the crucial stepping-stones, through which the concept of an expanding universe and the first suggestions of the Big Bang and vacuum energy merged into a coherent picture in the period from 1917 to approximately 1934. Although Slipher's redshifts were an early intriguing observational stimulus, the cosmological debate was largely theoretical initially. The most important participants were de Sitter, Eddington, Einstein, Friedmann, Lanczos, Lemaître, Robertson, Tolman and Weyl. Observational contributors who were directly involved were Hubble, Humason, Lundmark, Slipher, Strömberg and Wirtz.

On our journey we try to cover a large field in a very limited number of pages. For nearly every person, fact or idea mentioned, there exist detailed investigations, scattered in books or specialised publications. To mention them all, and the shades they add, would result in an encyclopaedic volume, which is not our aim. Those interested in historical details will profit from scanning through the volumes of the *Journal for the History of Astronomy*. We give, however, full references to all the sources that helped us to reconstruct the discovery of the expanding universe. Most of them are the original publications that constitute the scientific backbone of the evolving cosmology of the early twentieth century. A mathematical appendix is given for those who would like to have the rudiments of the theoretical arguments immediately available; it cannot, of course, replace a proper introduction to General Relativity.

When citing sources written in German or French we have translated them directly into English if the original is generally easily accessible. Otherwise we give the original text followed by our translation.

Cambridge University Press
978-0-521-51484-2 - Discovering the Expanding Universe
Harry Nussbaumer and Lydia Bieri
Excerpt
More information

2

Cosmological concepts at the end of the Middle Ages

Greek antiquity turned astronomy into a mathematical science. Ptolemy of Alexandria (*c*. AD 100–170) expounded its essence in the *Almagest* (see Toomer 1998). In that work, the philosophical roots (Plato, Aristotle) were merged with the observational and mathematical foundations (Eudoxus, Callipus, Aristarch, Hipparchos). It presents the Universe as a spherical system built around the motionless Earth. The philosophers demanded the motions of the heavenly bodies to be circular and of constant speed. If they appeared to be different, they had to be traced back to the prescribed ideal motions. The planets, which comprised the Moon as well as the Sun, moved in spherical shells, placed concentrically around the Earth. Ordered according to increasing distance the shells contained: the Moon, Mercury, Venus, the Sun, Mars, Jupiter and Saturn. Immediately adjacent to Saturn followed the shell of the stars. It was surrounded by the invisible *primum mobile*, the shell that provided the impetus to keep the system going.

The *Almagest* contains the theory of that structure, as well as detailed instructions on how to calculate the apparent motions of the planets as seen against the background of the stars, 1022 of which were catalogued. With the decay of the Roman Empire, Greek scientific culture, including the *Almagest*, was lost in the western part of Christendom. It was, however, absorbed into Islamic scientific culture.

After the long dormant phase of the Early Middle Ages, when in the Western Christian world practically all classical astronomical knowledge had been forgotten, contact with Islamic culture, particularly after the re-conquest of Toledo in 1085, gave a new impetus to European science. The translation of the *Almagest* from Arabic into Latin provided the techniques to calculate the paths of the planets. The basic cosmological concept, however, was much influenced by St Thomas Aquinas (1227–1274). He worked on a synthesis of Aristotelian philosophy and Christian teaching. There was, of course, a permanent discussion about how much biblical teaching had to be taken literally. A fundamental difference existed between the Aristotelian eternities of the astral world as opposed to the biblical creation of all things within one week. Nevertheless, a semi-official consensus on astronomical truth emerged.

Philosophy in the second half of the sixteenth and the first half of the seventeenth century had to come to terms with Copernicus and his shift from a geocentric to a heliocentric system. But hardly had that been accepted, when in 1644 Descartes widened the finite spherical cosmos into a universe with no bounds and no centre. At least as important, he changed the biblically created, forever immutable heavens into a dynamical universe, where the properties of matter, and not the unpredictable whim of God, were the creative forces that formed stars and planets. One generation later, Newton's *Principia* (first published 1687) gave scientists the tools to explain and predict the paths of planets and the forces that acted between the stars.

Because of its ability to measure time and geographical positions with great accuracy, astronomy in the seventeenth century attracted considerable financial support from kings and governments. It also captured the curiosity of the general public. In the second half of the seventeenth century, if we believe Molière, astronomy became a very popular pastime of Parisian upper-class ladies. *Les femmes savantes* was first performed in 1672, one year before distances in the Solar System were determined to within about 10% of their actual values by French and English astronomers. In that comedy the landlord complains that his wife, his sister, and his daughter had emptied the top floor of his house in order to install an observatory to find out what was happening on the Moon, instead of keeping the house in order and preparing him a decent meal. If the three ladies were seasoned observers, they might even have spotted Andromeda; the first Western report about that galaxy had been published by Marius in 1614. But nebulae had to wait for the eighteenth century to become objects of serious astronomical research.

The spherically closed universe of antiquity and the Middle Ages

The astronomical model of the late Middle Ages was basically the antique geocentric system. The *Tractatus de Sphaera* by John Holywood (1195?–1256) was probably the most important astronomical textbook, and became part of the standard medieval university curriculum. It was copied, printed and reprinted up to and beyond the times of Copernicus. The author was better known by his latinised identity of Johannes de Sacrobosco. *De Sphaera* was written around 1230. Although it was based on the astronomy of Ptolemy and his Arabic commentators, it gave a rather elementary account of spherical astronomy, with none of the sophistications of Ptolemy, such as a proper treatment of epicycles. Epicycles were frequently included as additions by later commentators. The last important edition was by the Jesuit Christoph Clavius, and was first published in Rome in 1570. The large number of further editions, even in the seventeenth century, bears witness to the importance of this textbook of Ptolemaic astronomy long after the publication of Copernicus' *De Revolutionibus* in 1543.

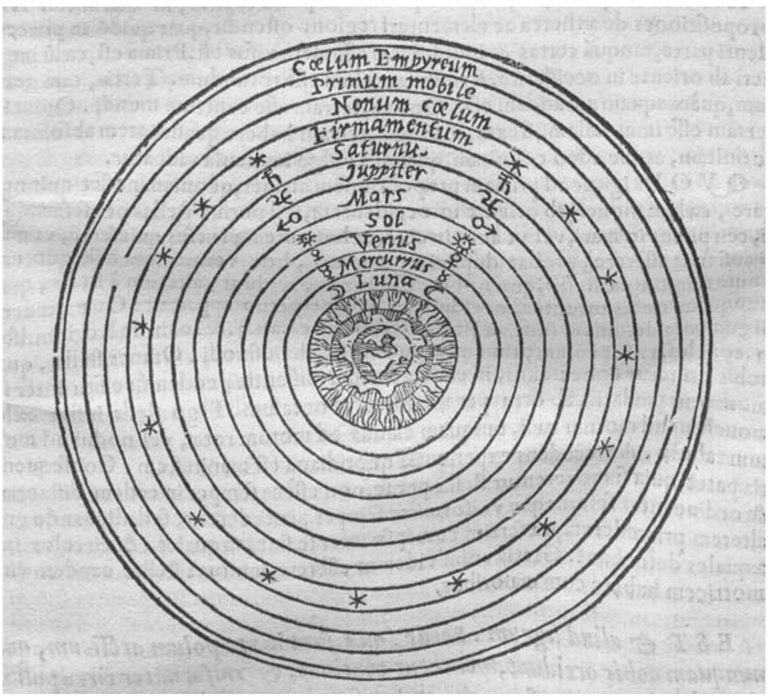

Fig. 2.1 Astronomical cosmos of the Middle Ages. The Earth is in the centre. It is surrounded by water, above water – air, above air – fire. Further still are the guiding spheres of the Moon, the Sun, the planets and the stars. Depending on the author or commentator, there may be several spheres beyond the fixed stars: they take care of additional periodic motions, such as precession, deemed important by the respective author. Often there is just one additional sphere: the *primum mobile*, the unmoved mover, the cause that sets the Universe in motion. In this image, we find beyond the fixed stars (here called *Firmamentum*), the Ninth heaven, the *Primum Mobile* and the *Coelum Empyreum*, the abode of God. (*Sacrobosco in Sphaeram*. Christoph Clavius, Rome, 1581. ETH-Bibliothek, Zurich, Sammlung Alte Drucke.)

The *Almagest* explained the motions of the heavenly bodies, and gave instructions how to calculate them. The most obvious motion is the daily rotation of the sky around the axis through the north and south poles; the celestial equator divides the northern and southern hemispheres. The celestial equator is really an extension of the Earth's equator. We know this apparent motion is due to the daily rotation of the Earth. Because of our annual path around the Sun we see the Sun moving annually on a great circle across the sky. This circle is called the ecliptic. Because of the tilt of the Earth's rotational axis against the plane of the ecliptic, the ecliptic is tilted against the celestial equator by 23.4°. The ecliptic cuts the celestial equator at the two equinoxes, where the Sun crosses the celestial equator in the spring and autumn. The planets move in planes that have only small inclinations against the Earth's orbital plane, with a maximum of 7° for Mercury. The projections of the planetary orbits therefore lie in a small band centred on the ecliptic: the Zodiac.

Observation shows that we are dealing with periodic motions. The daily period is due to the Earth's rotation, the annual period originates in the annual orbit of the Earth around the Sun, and there are the periods due to the planetary orbits around the Sun. There is an additional long-term period of 25 780 years due to the precession of the Earth's axis. Since Newton, we have known that the gravitational forces of the Sun and the Moon acting on the Earth's equatorial bulge cause the precession. Hipparchos had discovered this slow precessional shift of the equinoxes along the ecliptic of approximately 1.4° per hundred years in approximately 130 BC.

For us it is fairly easy to visualise these motions and their effects on the apparent paths of the celestial bodies. However, observers in antiquity and the Middle Ages had to explain apparent motions within the model of an immobile Earth in the centre of a rotating universe. The *Almagest* gave this explanation and the book of Sacrobosco repeated it, but on a very much simplified level. *De Sphaera* was not simply copied. Many different editors augmented it with comments. In this way each generation enriched the cultural heritage; but we have here neither the time nor the space, nor the detailed knowledge to do justice to all the individuals who contributed to it. Instead, we can summarise a few of Sacrobosco's essential points; our source is Thorndike's translation of 1949.

Sacrobosco states that the Universe is divided into two, the ethereal and the elementary regions, i.e. heaven and Earth. The elementary region, existing subject to continual alteration, is divided into four. The Earth is placed at the centre in the middle of all, above which is water, above water is air, above air fire, which is pure and not torpid and reaches the sphere of the Moon. Revolving around the elementary region with continuous circular motion is the ethereal,

Cambridge University Press
978-0-521-51484-2 - Discovering the Expanding Universe
Harry Nussbaumer and Lydia Bieri
Excerpt
More information

which is lucid and immune from all variation in its immutable fifth essence. The ethereal is divided into the following: the ninth sphere, which is called the 'first moved' or the *primum mobile*; the sphere of the fixed stars, which is named the 'firmament'; and the seven spheres of the seven planets: the Moon, Mercury, Venus, the Sun, Mars, Jupiter and Saturn. Each of these encloses its inferior spherically.

Two movements have to be distinguished. One is of the outermost heaven. Then there is another movement, oblique to this and in the opposite direction, of the inferior spheres on their axes, distant from the former by 23 degrees. But the first movement carries all the others with it in its rush about the Earth once within a day and night, although they strive against it, as in the case of the eighth sphere, one degree in a hundred years. This second movement is divided through the middle by the Zodiac, under which each of the seven planets has its own sphere, in which it is borne by its own motion, contrary to the movement of the sky, and completes it in varying lengths of time – in the case of Saturn in 30 years, Jupiter in 12 years, Mars in 2 years, the Sun in 365 days and 6 hours, Venus and Mercury in about the same time and the Moon in 27 days and 8 hours.

In chapter two, Sacrobosco comes back to the two movements. 'Be it understood that the first movement means the movement of the primum mobile, that is, of the ninth sphere or last heaven' and 'The second movement is of the firmament and planets'. Thus, the ninth sphere not only provided the fundamental motion for the lower spheres but was also the background for the slow precession of the eighth sphere.

Some astronomers thought precession to be variable (the variability was called trepidation). But, according to Platonic–Aristotelian dogma, all observed motions had to be composed of regular motions, so it was thought necessary to introduce an additional sphere for trepidation. Depending on whether trepidation was explicitly taken into account, the number of shells varied. Thus, in some editions of Sacrobosco we find nine spheres, whereas in others we might find ten or eleven. The eleventh would then be a purely theological sphere: the Empyreum, the immovable place for God and his elect – cosmology in those days was no less complicated than it is today.

There are two fundamental points we should keep in mind about the cosmological concept of the Late Middle Ages: (1) The Earth is a sphere and is located in the middle of the firmament, where it remains immobile. Contrary to what one still frequently reads, people of the Middle Ages knew very well that the Earth has a spherical shape, and that was never a serious theological problem. (2) There is an important division between the earthly sphere in continual alteration and the heavenly immutable confines. It corresponded to the religious imagery of those times, as embodied in Dante's *Divine Comedy*: the corrupt Earth with Hell at its centre and its ever-consuming fire for the damned, whereas the saved ones are with God in the highest sphere where everlasting perfection is found.

Whilst the *Almagest*, or Sacrobosco's tuned-down version, *De Sphaera*, provided the technical side of cosmology, Genesis delivered the information on the origin. On the first day God created heaven and Earth, and separated light from darkness. On the second day the *firmament* is created. On the fourth day, lights are created to separate day and night. He made two big ones, the bigger one to rule the day, the smaller one to rule the night, and stars He made as well, and placed the lights on the firmament.

The Cosmos also figured prominently in the imagery of the mystics, as shown in a vision of Hildegard von Bingen who lived from 1098 to 1179. Hildegard, abbess of Rupertsberg, was a German mystic whose written work was about religion, medicine, music, and also cosmology. Her treatise *Causae et curae*, which deals with health and illnesses and other aspects of the human condition, begins with a theological–cosmological background of our Universe. She tells about the creation of the world, the construction of the Solar System and the Cosmos, and of the elements that make up the Universe. In the pictorial representation of her visions we see a spherical universe that holds in its centre the spherical Earth, but the true centre is the human being. The Universe is built around Man, the apex of creation.

This was the cosmic order as accepted by practically everyone. Yet, at the beginning of the Renaissance this order was fundamentally questioned. In the middle of the fifteenth century, Cusanus proposed a dramatically different alternative, but had little response. Then, in the middle of the sixteenth century, Copernicus suggested a much less dramatic correction. Finally, in the middle of the seventeenth century, Descartes, supported by Tycho Brahe's and Galileo's observations, reshuffled and enlarged the Copernican Sun-centred universe fundamentally, and this time irrevocably.

Cusanus and his universe without centre or boundary

Cosmology deals with the Universe in its totality. As we have just seen, up to the end of the Middle Ages this consisted of the Earth-centred Solar System, with the stars in a spherical shell placed just beyond Saturn. A change of paradigm was announced in the philosophy of Cardinal Nikolaus von Kues (1401–1464) or, latinised, Nicolaus de Cusa, but better known as Cusanus.

Cusanus lived at a crucial moment in the history of the Church. Under the threat of Turkish invasion of the Byzantine Empire and Constantinople itself, an attempt was made to reunite the Greek and the Roman Church. Constantinople hoped for military assistance, whereas the Roman Catholic Church hoped to