Chapter 1

A brief history

1.1 The pre-Copernican view

Pre-Copernican theories in which the Earth was the center of the universe have long lost the attention of scientists. This is not only because such theories have been superseded since the Copernican Revolution, but also because in such hypotheses the origin of the Earth, Sun and planets is inextricably bound up with the origin of the universe. The Earth could hardly be younger than the rest of the universe if it occupied the central position. We are now aware that the solar system is less than one-third of the age of the observable universe. This makes it no longer necessary, as was the case with the authors of the Book of Genesis, to seek a common origin for Earth, Moon, Sun and stars. Most of this progress has been made by the discovery of new facts, not by theories. Galileo's observations, like those of Darwin, have done more to give us a correct view of the world than most of the theorising about it over the centuries.

1.1.1 The Greeks

The Babylonian and Greek astronomers observed the strange motion of the planets against the fixed positions of the stars [1]. In this manner, they became aware that there were two classes of heavenly objects in addition to the Sun and the Moon. It is curious that although the ancient astronomers devoted much study to the movements of the planets, they did not spend much time considering the origin of the solar system. The whole question of origins seems to have been mostly the province of philosophers.

Some astronomers, however, took up the challenge. Among them was Anaxagoras (*c.* 500–428 BC) who considered that the Moon was a stone. He thought that the Sun was a red-hot mass of iron bigger than Peloponnesus, the southern region of Greece that is about the size of Sicily. This idea that the Sun might be made of iron was based on a reasonable interpretation of the available evidence. An iron meteorite had fallen about 467 BC in ancient Thrace [2] and Anaxagoras concluded that the visitor had come from the Sun. He was banished from Athens because his views about the composition of the Sun and the Moon were considered to be heretical. Little of his work has survived, but apparently he pictured the Earth at the center of a sort of large cosmic whirlpool. In this he anticipated the notions of Descartes in the 16th century, demonstrating the truism that few ideas are truly original.

The great trio of Greek philosophers Socrates, Plato and Aristotle, whose ideas have formed the basis for Western culture, were mostly concerned with questions of purpose. They distinguished carefully between the Earth, with its obvious imperfections, and the heavens, which they held to be unchanging. Four elements – earth, air, fire and water – sufficed to make up the Earth. The heavenly bodies in contrast were composed of shining crystal, a perfect fifth element, or quintessence. The Moon was also made of this. The dark patches that could easily be seen on the face of the Moon were thought to be the reflections in this perfect mirror from the mountains and oceans on the Earth.

The doctrine of Socrates (*c.* 470–399 BC) held that the heavens were perfect, in obvious contrast to the Earth. This left no room for any changes or evolution and so did little to encourage scientific investigation. Plato (*c.* 428–347 BC) concerned himself with the motions of the planets rather than their origin. He did suppose, however, that the Earth was moving. In his scheme, the heavenly bodies were supposed to move in perfect circles. This problem of perfectly circular orbits continued to haunt astronomers as late as Copernicus, nearly two millennia later, until Kepler finally broke the spell. Aristotle (384–322 BC), the third member of the trio, also thought that the heavens were permanent and thus not subject to the earthly laws of physics as he perceived them. His views, wedded to the concept of a providential Old Testament God who designed all for our well-being, were to dominate Western culture for two thousand years.

A refreshing contrast to these views was proposed by Aristarchus of Samos who lived around 250 BC. He placed the Sun at the center of the solar system, and included the Earth with the rest of the planets. He realized that the Earth was small in relation to the Sun. Aristarchus appears to be the first person who suggested that the Earth both rotates and revolves around the Sun. This idea was not forgotten, but lay around until revived by Copernicus eighteen centuries later. It is fitting that a prominent crater on the Moon is named for Aristarchus.

Epicurus (341–270 BC), who was a strong critic of the views of Aristotle, did not give the heavens any special or separate status. He supposed that the heavenly bodies formed by random collisions of atoms, whose existence had been proposed by Democritus (*c.* 470–400 BC) 150 years earlier. The Epicurean School rejected divine explanations and believed in physical causes. Unfortunately it did not encourage investigations into natural phenomena, so that no scientific advances resulted. Epicurean philosophy was mostly concerned with freedom and happiness and was very popular. It survived until replaced by Christianity in the 4th century AD. Our best surviving statement of the physical theory of Epicurus comes from the Roman poet and philosopher Lucretius (96?–55 BC). In his long poem *De rerum natura (On the Nature of Things)* he adopted many of the ideas of Epicurus. He encouraged a materialistic outlook, discouraged superstition and paid little attention to astrology, which was popular then as now.

Among others deserving a special mention, Eratosthenes of Cyrene (276–195? BC) who was Director of the Great Library (Mouseion) at Alexandria, correctly calculated the radius of the Earth. His answer to this classical problem was within about one percent of the modern value.

1.1.2 Ptolemy

In the 2nd century AD, Ptolemy compiled a summary of Greek astronomical thought and data in his book, the *Almagest*. It was a triumph of the use of geometry in understanding the solar system. This work was the definitive work on astronomy until the end of the Middle Ages and so remained the acceptable explanation for over a millennium. Like Lucretius, very little is known of his life. His birth and death dates are unknown, although the Arab sources recorded that he lived for 78 years. Ptolemy remains an obscure figure. He seems to have been endowed with bad judgment, as he rejected both the Sun-centered solar system of Aristarchus and the essentially correct value for the size of the Earth that Eratosthenes had worked out. Both decisions set back the progress of scientific knowledge for the next 1500 years. Perhaps Ptolemy's major achievement [3] was to salvage the star catalogue of Hipparchus, the greatest of the ancient observational astronomers, who had worked in the second century BC. His catalogue listed 850 stars arranged in six orders of apparent brightness, more or less in line with modern concepts.

Like his Greek predecessors, Ptolemy felt that the imperfect Earth could not be given a place among the heavenly bodies, that were composed of shining crystal in their cosmologies. The system devised by Ptolemy placed the Earth at the center of the universe. The motions of the planets followed extremely complicated paths. Despite its theoretical defects, it was a practical success and remained in use up to the late Middle Ages. However, many of its problems had been long understood by sceptical observers. One of these was Alfonso X (The Wise), King of Castille, (1221–1284 AD) who is commemorated by having one of the larger craters on the Moon named in his honor.

> *"Alfonso was one of the first sovereigns who encouraged the revival of astronomy in Europe. This science can reckon but few such zealous protectors but he was ill seconded by the astronomers whom he had assembled at a considerable expense and*

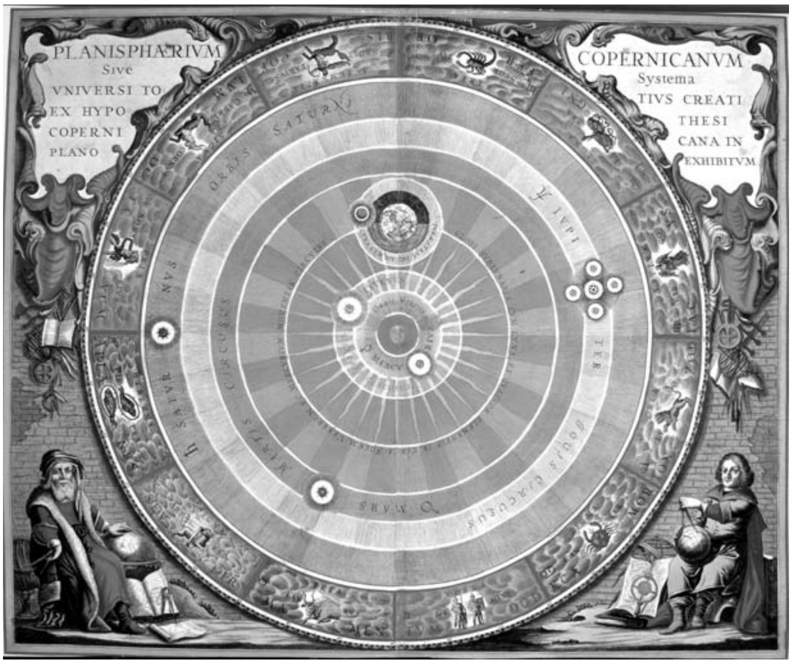

Fig. 1.1 The Copernican System, as illustrated by Andres Cellarius in 1661.

the tables which they published did not answer to the great cost they had occasioned. Endowed with a correct judgement, Alfonso was shocked at the confusion of the circles, in which the celestial bodies were supposed to move. He felt that the expedients employed by nature ought to be more simple. 'If the Deity' said he, 'had asked my advice, these things would have been better arranged'" [4].

Despite such opinions, scientific knowledge in Europe by the 14th century was less advanced than in Greece and Alexandria in the second and third centuries BC. The level of mathematics was about that which the Babylonians had achieved two millennia before.

1.2 The Copernican Revolution

The Copernican Revolution is usually dated at 1543, the year of the publication of the great work of Nicolaus Copernicus (1473–1543) *De revolutionibus orbium coelestium, libri VI (On the Revolutions of the Celestial Spheres, volume VI)*. The model of Ptolemy had placed the Earth at the center of the universe. It was clear to casual observers that the Earth was flat and that the Sun, Moon, planets and stars all revolved around it. Furthermore, the Ptolemaic System, for all its complexity, worked well enough for practical matters, including navigation. Columbus used it. Minor problems were accommodated by complicated adjustments until a complex array of epicycles and the like, to which Alfonso had objected, encrusted the whole scheme.

Copernicus, however, placed the Sun at the center (Fig. 1.1). Why did he do this? He had no evidence from observations to support his notion. One can do little more than speculate 450 years later, but he seems to have viewed the Sun-centered system as more intellectually and aesthetically satisfying than the Earth-centered model of Ptolemy. Clearly one places the light in the center of the room. Copernicus did not refer to the ideas of Aristarchus of Samos who had proposed a Sun-centered system eighteen centuries earlier [5].

But the new idea did not arise in a vacuum, any more than did Darwin's theory of evolution. Along with Alfonso, other thinkers in the Middle Ages, of whom Nicolas of Cusa (1401–1464) and Regiomontanus (1436–1476) were examples, had laid the intellectual framework for dismantling the old system.

However, the new scheme of Copernicus did not work as well as that of Ptolemy for practical applications. The planets remained in circular orbits, so Copernicus still had to use even more epicycles than Ptolemy to account for their motions.

1.2.1 Tycho Brahe and Kepler

The next significant step in understanding the solar system was taken by Tycho Brahe (1546–1601). His chief accomplishment was the precise pre-telescopic measurement of planetary positions. He was also concerned about the problems with the complicated system of Ptolemy and produced yet another model in which the Sun and the Moon indeed went around the Earth, as everyone could see. However, in his system the other planets rotated around the Sun. This compromise cosmology was popular, as it appealed to common sense observations and did not conflict with the scriptures. Variations survived until late in the 17th century, finally vanishing as the motions of the planets became widely understood.

Tycho was so unpopular with the other residents of his island of Hven that they demolished his splendid observatory when he lost Danish royal favor and had to leave in 1597 [6]. After various wanderings, he finally arrived with his data in Prague where he was appointed Imperial Mathematician. Here chance plays its role. Another refugee arrived in Prague in 1600. Johannes Kepler (1571–1630) had been banished from Graz in Austria, a victim of religious persecution. He became Tycho's assistant and succeeded him as Imperial Mathematician when Tycho died suddenly in 1601. Kepler thus inherited or perhaps just took ("usurped" was his word) the boxes that contained Tycho's monumental observations. These data formed the basis for Kepler's basic discoveries of the laws of planetary motion [7]. Kepler's great contribution was to discover that the orbits were elliptical and he became an advocate of the Copernican System.

However, like many other scientists, he was mainly concerned with other matters so that, as Jaki [8] has commented, "the three major gems in Kepler's works on astronomy lay in a vast field of errors, of irrelevant data, of mystical fantasies, and of useless speculations". Despite such distractions and with a vast amount of labor, Kepler was able to fit the orbits of the planets into spheres based on the five "perfect" geometrical solids: cube, tetrahedron, octahedron, icosahedron and dodecahedron. These are the only solids bounded by identical faces and were thus considered "perfect". Kepler considered that he had answered a fundamental question why were there only six planets with five intervals between them. Kepler's view was that this cosmic limit was imposed because of the small number of "perfect" solid forms. However, the planetary orbits, on the basis of Kepler's own laws, turned out to be elliptical, not circular and so his elaborate geometrical system fell into ruin.

1.2.2 Clockwork systems

The development in the late Middle Ages of mechanisms for driving clocks naturally suggested that similar processes might be responsible for the well-ordered motions of the planets. Clocks had become more sophisticated as clockwork became perfected and often included astronomical models as well as historical and religious displays. One of the earliest was constructed by Richard of Wallingford in 1320 at St Albans. Another famous example is the great clock at Strasbourg, dating from 1364. Such mechanical marvels led to the idea that perhaps the universe was some kind of giant clockwork. As a clock requires a builder, this suggested that the universe had been created by a master craftsman.

Once the solar system had been constructed by an omnipotent clockmaker and the system was set running, it could continue to operate under the laws of physics. Such ideas went back to Nicolas of Oresme (1330?–1382), a bishop who had conceived of God as the master clockmaker. Kepler was an enthusiastic supporter, suggesting that perhaps magnetism was the driving force, just as falling weights drove earthly clocks.

The clockwork idea was also consistent with the Bible. Archbishop Ussher (1581–1656) calculated that the creation of the world (including the universe) had occurred in 4004 BC on Sunday, October 23 at 9.00 a.m. This date, although now derided, was carefully calculated from the available biblical record. What it represents is essentially that of recorded history. The significance of this date, if correct, was that the universe had not had much time to evolve and everything must have been created in the beginning, more or less as it now appeared.

1.2.3 Galileo

The Copernican Revolution did not resemble those of more modern times. Fifty years after the publication of his system by Copernicus, little had changed. His ideas had disturbed neither the public nor the church. What was needed was some crucial observation to decide between Copernicus and Ptolemy. This came, as is common in scientific progress, with a technical advance. The telescope had been invented about 1600 by Hans Lippershey, a Dutch spectacle-maker, apparently by accident. When the news reached Italy, the Senate of Venice asked Galilei Galileo (1564–1642), a skilled maker of instruments, to make some. He was the son of a lute player and composer, but had decided not to follow his father's career. We are still living with the consequences of that decision. It was not of course the intent of the Venetian state to upset the accepted view of the world. Their reasons were more practical. Telescopes would obviously be useful for an empire based on sea power [9].

Galileo's observations are famous. The Milky Way was composed of stars, and so maybe the universe was infinite. The Moon was not a smooth mirror after all, but rough like the Earth and so perhaps made of the same material. Venus showed phases like the Moon, including a full face, informing Galileo that Venus was passing behind the Sun. Another critical observation that led to the collapse of the Ptolemaic System came when Galileo discovered in 1610 that four satellites were rotating around Jupiter. Copernicus was right after all.

1.2.4 Descartes and Newton

René Descartes (1596–1650) next took up the challenge of the origin of the solar system [10]. His view of the world was a completely mechanical one. He postulated that there was no basic difference between the forces driving a clock, the solar system or living matter. He proposed that the universe contained many circular eddies. Like a whirlpool, matter accumulated in the center of the vortex to form the Sun. Coarser particles were captured to form the planets. Satellites formed in secondary whirlpools surrounding the planets.

By the time that Isaac Newton (1647–1727) appeared, the Copernican System had long dominated thought. Newton's work was the culmination of the work of Copernicus, Kepler and Galileo. Newton was impressed by the tidy nature of the solar system. He was irritated by the qualitative notions of Descartes and showed that the complexity of the solar system could be dealt with by exact physical laws. The planets were securely tucked into their orbits and the space between was apparently clean. Newton followed the popular belief, at least publicly, that the world had been created essentially in its present form only a few thousand years before, in accordance with the biblical timescale that Archbishop Ussher had calculated. This left no time for the system to evolve from a more primitive state as Descartes had imagined. Thus it required a Creator, who had set the system up much as we now see it and who had ordered each planet to move in its particular orbit.

The success of Newtonian mechanics reinforced the notion that the solar system was some type of celestial clockwork. This theme of a celestial clockmaker came to dominate thinking about the solar system in the 17th and 18th centuries. These ideas bore fruit in the construction of mechanical models of the solar system. Models of the solar system date back to antiquity [11].

The 18th-century models were named orreries after the 4th Earl of Orrery, Charles Boyle (1676–1731) [12]. These instruments became very popular. There is a fine example in the Meteorite Hall of the Natural History Museum in Vienna of a "Kopernikanische Planetenmaschine" made in 1761 for the Austrian Emperor. When Louis XV (1710–1774) constructed a new wing at Versailles, an orrery was placed in the central room, in contrast to the chapel which forms the center of the old wing. This was in keeping with the philosophy of the Age of Enlightenment.

Newton, however, noted that there were small variations in planetary orbits. Jupiter's orbit appeared to be shrinking and Saturn's expanding. Unable to solve this problem, Newton concluded that God had to intervene from time to time to make periodic repairs or adjustments to keep the system in balance, in effect winding up the clockwork. This led to complaints by Leibnitz (1646–1716), that Newton was guilty of heresy by supposing that God had created something less than perfect. Given supreme power, the construction of a well-ordered planetary system should not be beyond the powers of a competent clockmaker. Surely God would not have constructed an imperfect system and would have had enough foresight to create perpetual motion, rather than acting as a maintenance man, who had to wind up the clock and make fine adjustments to the planetary orbits.

1.2.5 Kant

A little later, the great philosopher, Immanuel Kant (1724–1804) considered the problems of the solar system [13]. He produced a correct explanation for the Milky Way, proposing that it was an edge-on view of a disk of stars. He showed remarkable foresight by suggesting that the fuzzy lentil-shaped nebulae were distant island universes similar to the Milky Way, a cosmological leap in understanding that was not substantiated until the second decade of this century, nearly 200 years later. These essentially correct insights perhaps explain why his concepts for planetary origins are usually acclaimed rather than critically examined.

In the words of one critic, "Kant took but a glance and not a thorough look at the staggering problem of the origin of planetary systems. His explanation of it contained more nebulous statements at crucial junctures than there was nebulosity in the rudimentary form of the solar system . . . contrary to the stereotype accounts of Kant's planetary theory, he did not compare the rudimentary form of the solar system to a nebulous agglomeration of matter" [14]. Kant believed that the operations of Newtonian mechanics could not of themselves produce the regularities of the solar system without divine guidance, and that "the material universe . . . has no freedom to deviate from this perfect plan" [15].

Kant's model for the origin of the solar system was based heavily on an analogy with the galaxies. It began with a chaotic distribution of particles with slight density variations, which would accrete material and grow with time. The material was assumed to be rotating and to develop into flattened rotating disks. The Sun formed at the center, and the planets formed at secondary condensations within the disk. He postulated the existence of many additional planets outside the orbit of Saturn, although these possessed large eccentricities, with a gradual transition to the comets. He assumed that density fell off with distance, and accounted for the anomalously low density of the Sun at the center by deriving the material in the Sun from beyond Saturn.

Unlike Laplace, who proposed that the ring of Saturn (no divisions were known at the time) was solid, Kant correctly thought it to be composed of a multitude of small particles. He supposed, with less prescience, that all the planets, including the Earth, were initially surrounded by a ring, meanwhile populating the planets with beings whose intelligence increased with distance from the Sun. In his model *"our own Newton would not outrank a monkey on Saturn"* [16].

It seems clear that the many contradictions in Kant's hypothesis do not accord with the general popular acclaim that it has received. Perhaps this is due to his eminence as a philosopher. It is a commentary on the inherent difficulties of accounting for the solar system that one of the foremost thinkers of the Enlightenment should have failed to produce an internally consistent explanation. His model is often linked with the hypothesis of Laplace, to which we now turn.

Fig. 1.2 Pierre-Simon, marquis de Laplace (1749–1827).

1.3 Laplace and his followers

We can date modern thinking about the origin of the solar system from the appearance in 1796 of the *System of the World* by Pierre-Simon, marquis de Laplace (1749–1827) (Fig. 1.2) [17]. Laplace was impressed, as Newton had been earlier, with the regularities in the solar system as it was known in the late 18th century. The planets all lay in a plane. They all moved in the same anticlockwise sense around the Sun in nearly circular orbits. The satellites revolved around their parent planets in the same direction

Laplace ignored the inconvenient fact that at least two satellites of Uranus, discovered by Herschel (1738–1822) in 1787 were orbiting in a plane perpendicular to the rest of the solar system. Although "the devil is in the details" is a useful truism, there comes a point where details can usefully be ignored. The wisdom comes in knowing what is important and what to ignore [18]. This regular arrangement led Laplace to the concept that the system had arisen far in the past from a primitive rotating cloud, the "solar nebula". This idea has survived. This was in contrast to the ideas of Newton, who had believed, at least in his written works, that the solar system had been created in its present form only a few thousand years earlier.

Laplace, however, was an inhabitant of the Age of Enlightenment. Born into what we would now call a middle-class farming family, he had, unlike Lavoisier, survived the French Revolution and was a distinguished member of the French scientific establishment at the beginning of the 19th century. He was able to show that the apparent variations in the orbits of the planets were self-correcting and so God was not needed to adjust the system.

Laplace gave a copy of his book to Napoleon to whom he had taught mathematics when the Emperor had been an artillery cadet. Bonaparte, seeing no mention of God, presumably the designer of the system, asked Laplace about this omission. Laplace, having solved the problem that had concerned Newton, made his famous reply that he had *"no need for that hypothesis"* [19].

A watershed had been crossed. Now the solar system could be considered as having arisen by the operation of natural processes from a primitive beginning, rather than being created perfect in the instant. This marks the beginning of modern attempts to understand how the Sun and the planets came into being [20].

During the 19th and the first half of the 20th century theories of solar system origin were advanced. Although they showed much diversity, they can be classified into three categories [21].

- *Tidal theories*, in which the formation of the planets occurred from material extracted from the Sun or a passing star after these bodies had formed.
- *Accretion theories*, in which material was captured by the Sun from interstellar space.
- *Nebular theories*, in which the planets formed directly either concurrently or consecutively from the same nebula as the Sun.

1.4 Tidal theories

Tidal theories are now out of favor, although they were very popular in the past. The initial idea seems to be due to Buffon [22] who proposed that a cometary collision with the Sun ejected a disk of material. The masses of comets, then unknown, were thought to be about 0.1 solar masses. When the true masses of comets became established, the theory languished until the comet was replaced with a collision with a passing star [23]. Other proposals involved a head-on collision [24] or a collision between two nebulae, the true nature of nebulae as galaxies not being established at that time [25].

The Chamberlin–Moulton hypothesis attempted to overcome the angular momentum difficulties with the Laplacian models. In these, the Sun should have an equatorial rotation exceeding 400 km/s, instead of the observed 2 km/s. The Chamberlin–Moulton solution proposed that the approach of another star would cause an increase in solar activity, resulting in the ejection of solar material [26]. Now this theory is chiefly of interest because these clouds of ejected material condensed, forming what were termed planetesimals, from which the planets were accreted. The term planetesimals has survived and occurs frequently in current models, although in a different context.

Further attempts to deal with the thorny problem of angular momentum, which had caused the eclipse of the Laplace nebular hypothesis, led to the further development of tidal theories. These used the tidal action of passing stars to draw out a cigar-shaped filament from the Sun [27]. Although these theories were very popular, particularly in the public domain, and explained many features of the solar system, various objections were raised [28, 29].

Although the tidal theories had been erected to account for the angular momentum of the planets, difficulties still emerged in attempting to overcome this particular problem. It was soon shown that the