Chapter 2
From Newton to Einstein: The Discovery of Laws of Motion and Gravity

The Law of Gravity

The three-body problem is Sir Isaac Newton's problem. It is not only an academic problem, but a serious problem, affecting the whole existence of mankind. Nobody before Newton had thought of it, but as soon as it dawned on Newton, he started to work on it feverishly, even though it gave him a headache (Fig. 2.1).

So what is the problem? Newton had discovered that celestial bodies attract each other by gravity. Previously it had been thought that celestial bodies revolve around their central bodies so that a single body can orchestrate the motions of the others. But Newton realized that everything depends on everything else in the universe, and that there is really no stable configuration of the Earth and the celestial bodies. The question was if any long term stability is possible at all.

The bodies of most immediate interest to us are the Earth on which we live, the Moon which goes around the Earth raising its waters along its path, and the Sun which similarly pulls the Earth. And of course the Sun and the Moon attract each other[1] also even though it does not have such easily detectable consequences as the ocean tides on the Earth. How do these three attractions balance each other? Could it be that they compound one day in such a way that the Moon crashes on the Earth, ending the lives of most creatures here, including our own? Nothing certain can be said about this until the three-body problem is solved.

[1] You may think that the Earth exerts greater attraction to the Moon than the distant Sun. Actually, the reverse is the case—Sun's gravitational influence on the Moon is more than double the influence of the Earth. So the question appears: Why doesn't the Moon fall into the Sun?

We should remark on the beautiful coincidence of sizes and distances: diameter of the Moon is approximately 1/400 of the Sun's, the Moon's mean distance is 1/389 of the Sun's. Thus the Moon and Sun are nearly same size as seen from the Earth. Because of the ellipticity of the Moon's and Earth's orbits, during the solar eclipse the Moon can be seen a little bigger than the Sun (in this case we observe a total eclipse), or a little smaller (in this case we see an annular eclipse). Total eclipses occur slightly less often than annular.

M. Valtonen et al., *The Three-body Problem from Pythagoras to Hawking*,
DOI 10.1007/978-3-319-22726-9_2

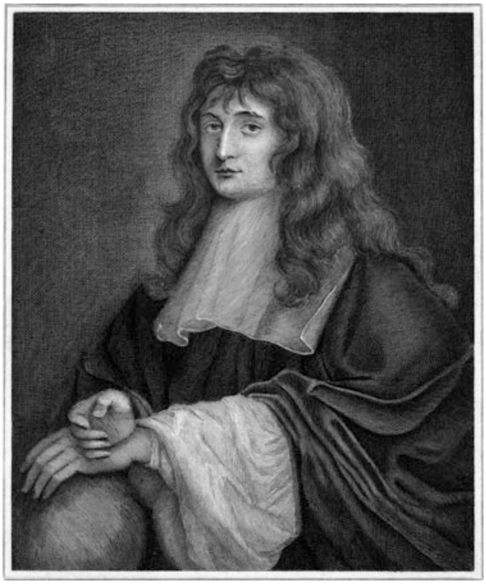

Fig. 2.1 Portrait of Isaac Newton when B.A. at Trinity College, 1677 Engraved by Burnet Reading (1749/50-1838) from a head painted by Sir Peter Lely (1618–1680), published 1799 (Credit: Institute of Astronomy, University of Cambridge)

The same story applies to the planets. Each planet circles the Sun in its orbit because the gravity between the Sun and the planets forces them to do so. But in addition, the planets attract each other. The planet on the inside lane below the Earth's orbit, as seen from the Sun, is Venus. Venus is almost as big as the Earth, and gives a mighty tug to the Earth each time it passes by. Venus moves faster than the Earth (Kepler's third Law), and thus these celestial encounters are quite frequent. Could it be that the Venus' tugs would gradually build up to the point where the Earth is thrown off its regular orbit, and made to plunge to the Sun and burn up, or made to escape far from the Sun where the Earth and everything on its surface would freeze permanently without the heat of the Sun's rays? We have to solve the three-body problem of the Earth, Venus and the Sun to see if it will happen one day.

For Newton this was an urgent problem. We now know that the planetary system has survived for 4.5 billion years without major changes. Therefore, being optimistic, we can say that nothing too bad will happen for the next 4.5 billion years, either. But Newton relied on the biblical time scale, the calculation based on generations from the Bible since the creation of the universe. It gives the age of the universe as small as 6000 years. The survival of the system over only 6000 years does not give us much confidence that the stability would persist even to the next human generation.

Such were the thoughts of Newton when he embarked on this scientific endeavor, one of the most important in the history of mankind.

If we could report that Newton solved the problem, the story would be short. Rather, the story has had twists and turns and it has branched into many other areas of research of equal importance. This is the story that we begin here.

According to Newton's biographer William Stukeley, the idea of the law of gravity came to him when he one day was sitting under an apple tree of his country home at Woolsthorpe, Linconshire.

> After dinner, the weather being warm, we went into the garden and drank thea, under the shade of some apple trees. . .he told me, he was just in the same situation, as when formerly, the notion of gravitation came into his mind. It was occasion'd by the fall of an apple, as he sat in contemplative mood. Why should that apple always descend perpendicularly to the ground, thought he to himself. . .

This may be true, but there is more to it than Newton cared to remember at his old age. In fact, one could say that the idea was in the air, so to speak. Perhaps it did require an apple to remind Newton of what he had learnt as a student of Trinity College at Cambridge during the previous years.

The reason why Newton was at his country home at all was that a plague had forced the closing of the University of Cambridge soon after his graduation. Newton had more than a year to spend leisurely at Woolsthorpe and to think about marvelous things he had learnt from his teachers and from books.

The teacher who had impressed Newton above all was Isaac Barrow (1630–1677), professor of mathematics. He himself had an illustrious career, including a 3-year study tour in central Europe. There Barrow had met the leading mathematicians of the time, including the last student of Galileo Galilei, Vincenzo Viviani (1622–1703). Thus was the wisdom of one of the giants of modern science transmitted to another through first hand student-teacher relations.

Galileo Galilei, as well as Simon Stevin, had studied the effect of gravity near the surface of the Earth, and determined that the speed of fall of all bodies changes uniformly and in the same way, independent of the weight of the body. As we mentioned previously, he is said to have dropped bodies of different densities, but of the same surface area, from the leaning tower of Pisa to determine that they all reach the ground at the same time if they are released together. Thus Galileo knew the law of gravity for the Earth. He did not think it universal enough to be applied to planetary motions.

What was needed were new tools of mathematics to deal with a constant change, such as occurs in the fall of bodies, or in the motion of planets when they constantly deviate from the straight line orbit, i.e., fall toward the Sun. At every interval of time, the speed changes by a constant amount. In order to calculate the orbit, one has to add up the progress of the body from one interval to the next. In reality, the body does not jump along like a rabbit, but moves smoothly. To describe the smoothness mathematically, we have to make the intervals of time very small, and to make their number correspondingly large. In fact, in the limit we have to consider letting the interval go to zero while their number goes to infinity. This is exactly the same process that Archimedes used in the squaring of the circle almost 2000 years earlier!

This is where Barrow had made his mark. It was left to Newton to develop the methods of calculation further. However, he was not alone in this pursuit. Around the same time the German mathematician Gottfried Leibniz (1646–1716) had developed similar methods. From today's point of view, Leibniz's techniques, with further work by the Swiss Bernoulli brothers Jacob and Johann, have survived better. However, at Newton's time serious priority issues developed, mixed with nationalistic feelings. Could a German from Leipzig be cleverer than an Englishman from Cambridge? No, if you asked the English!

Newton did not publicize his findings readily, and this lead to other priority issues as well. Newton solved all such issues, at least to his own satisfaction, in his memoirs where he declared that all his major discoveries were made when he was forced to spend his days at Woolsthorpe. Newton described how he passed his time there, making major discoveries in mathematics, optics and astronomy one after another. He said:

> All this was in the two plague years of 1665–1666. For in those days I was in the prime of my age for invention & minded Mathematicks & Philosophy more then at any time since.

The other priority dispute that bothered Newton greatly was related to the explanation of the orbital motion of planets. It was already well accepted in scientific circles that the Sun is the central body of our system. Planets go around it each in their own orbit. We have learnt above how Nicolaus Copernicus had suggested this more than 100 years earlier, followed by Johannes Kepler's discovery of elliptical orbits, and the laws of motion in these orbits. It had occurred to several people that an inverse square law of central force was needed to balance the centrifugal tendency of planets to continue along the tangent, rather than to curve around. This was good and obvious for circular orbits. What about the elongated orbits of the planets? The elongation, or eccentricity as it is called, is not much but still significantly different from zero (i.e., the orbit is not a circle).

Robert Hooke (1635–1703) as the newly elected secretary of Royal Society tried to engage Newton to a correspondence on this problem. He explained what his understanding of the question was and then asked how the inverse square central force could lead to eccentric orbits. Could Newton with his superior mathematical skills solve the problem?

The correspondence continued for a while but then stopped abruptly in 1680. It has been surmised that it was exactly at this time that Newton had found the answer but did not want to share the credit for the discovery. What Hooke had explained to him was probably known to Newton already so that there was no need to give him credit. Whatever the reason was, it was annoying to Hooke that he could not get further with Newton.

Then in 1684 Hooke decided to send a younger member of the Society, Edmond Halley (1656–1742), to Cambridge to meet Newton. According to Abraham DeMoivre:

> After conversing some time, Halley asked Newton 'what he thought the Curve would be that would be described by the Planets supposing the force of attraction towards the Sun to be reciprocal to the square of their distance from it.' S^r Isaac replied immediately that it would be an Ellipsis, the Doctor struck with joy & amazement asked him how he knew it, whereupon D^r Halley asked him for his calculation without any further delay, S^r Isaac Looked among his papers but could not find it, but he promised him to renew it, & then to send it him…

Fig. 2.2 René Descartes (*left*, credit: Wikimedia Commons) and Edmond Halley (*right*, Engraving by John Faber of a portrait by Thomas Murray. Credit: Institute of Astronomy of University of Cambridge)

Newton kept his promise and sent the calculation to Halley in November. Halley returned to Cambridge and persuaded Newton to make the discovery public and write it more fully for the Register of Royal Society. In this way Newton's priority to the discovery was guaranteed. The final written version of the discovery of the law of gravity arrived at London at the end of 1684 or early 1685, and is known as *De Motu*, from key words contained in a longer title.

Now Halley persuaded Newton to continue writing and have his whole new system of mechanics published. Halley agreed to cover part of the expenses of this work. Newton's model for his major work was *Principia Philosophiae* from 1644 by the Frenchman René Descartes (1596–1650) (Fig. 2.2), a leading philosopher and physicist from the Netherlands. However, Newton did not quite arrive to this goal but was satisfied with a less grand title called *Philosophiae Naturalis Principia Mathematica,* i.e., he felt that he had been able to solve only the mathematical part of the problem. Descartes had an explanation for the reason behind gravity, while Newton did not in 1687 when the work appeared.

Descartes' explanation was that a central body creates space whorls in its surroundings, and other bodies are carried along these whorls. The Sun creates the major whorl and carries the planets, while planets create secondary whorls that carry the Moons. Intuitively this sounded plausible and was easy to understand. Descartes' theory was popular even after Newton's *Principia* had been published. *Principia* required mathematical skills which were not common among the scientists of his generation.

Descartes is remembered especially for uniting algebra and geometry. Up to then geometry was regarded the primary science and algebra only an auxiliary method. Descartes brought them to equal status. Much of the way of writing algebraic equations today comes from Descartes. Newton as well as Gottfried Leibniz were able to take off from there, and to develop what is now called differential calculus. Among the early founders of this field of mathematics was Isaac Barrow, Newton's teacher. This form of calculus was an essential tool for Newton's physics.

Newton's Failure

Newton's *Principia* gave the laws of motion and the mathematical form of the law of gravity. However, in the concluding *General Scholium* he had to admit that he had not discovered the essence of gravity, the reason why the celestial bodies attract each other the way they do. It is enough, he concluded, that gravity and its laws lead us to describe "all the motions of the celestial bodies and of the seas" and much else. In the science today this would be viewed as quite a satisfactory state of affairs.

However, in Newton's time the situation was regarded less so. The Dutchman Christiaan Huygens (1629–1695), one of the leading physicists of the time, commented that he also might have produced many, if not all, of Newton's mathematical results, but he was inhibited by the lack of physical intuition of this approach. The current view then was that actions need intermediaries to be transmitted, somewhat like the set of gears of a clock which converts the force of the spring to the motions of the clock hands. In the next decades Newton thought hard what there might be between the Sun and the planets that transmits the force between the celestial bodies, but he could think of none that would not at the same time cause friction in the planetary motion. Thus he turned to the ancients for their views on the matter.

Newton observed: The power of gravity must proceed from a cause that penetrates to the very centres of the Sun and the planets, without suffering the least diminution of its force; that it operates not according to the quantity of the surfaces of the particles upon which it acts (as mechanical causes would do), but according to the quantity of the solid matter which they contain. Only a spiritual cause could penetrate adequately without constituting a frictional drag on the motions of planets and comets and producing slowing down of their motions that was contrary to the phenomena. Newton equated this spiritual cause with God which is omnipresent in literal sense.

Already Pythagoras and Plato had claimed that the world was divinely ordered and maintained by a "soul of the world composed of harmonic numbers". The supervisory harmonic function was due to "the providence and the preservation of God". Newton called attention to the Stoics, one school of followers of Plato, to have had the right ideas. They held that there exists dualism in nature, a distinction between spirit and matter. Newton observed that only spirit could penetrate to the very centres of bodies without acting on the surfaces of bodies. In the context of the Stoic thought, the all-pervasive spirit was the *active principle*, penetrating and binding the *passive principle* of matter. Sometimes the Stoics are given credit for the discovery of the concept of gravity, and a germ of the theory could thus be traced as far back as to Plato.

These concepts were actively discussed in European intellectual circles from the classical period up to Newton. Thus it is quite understandable that Newton accepted the divinity of the active principle. He said that the Divine mind will form and reform the parts of the universe, and this is what we call gravitation. For us today, more important than Newton's explanation of the nature of gravity, is the mathematical theory of how it can be handled in practical calculations.

Newton's Physics

One of the most significant concepts in the *Principia* was universal gravity. Of course, gravity holds us to the ground on the Earth. Something forces the distant Moon to circle the Earth. Is it the same force? Huygens had found earlier that the acceleration of an object in a circular path toward the center is its orbital speed squared divided by the radius of the path. The orbital speed of the Moon and the size of its orbit can be determined from astronomical observations; thus we know how big the central acceleration at the Moon must be to keep it in its orbit.

To establish that the universal law of gravity follows the inverse square law, Newton compared the acceleration towards the center of the Earth at its surface with the acceleration caused by the Earth further away, at the distance of 60 Earth radii which is the distance of the Moon. The gravitational acceleration should be lowered by 60 squared = 3600 when we go from the Earth to the orbit of the Moon, where it should equal the circular acceleration of the Moon toward the Earth. Newton carried out the comparison by using the value of the radius of the Earth and was able to confirm the inverse square law. Because of the greatly reduced acceleration, the Moon falls in 1 min as far as an apple (on earth) falls in 1 s.

Newton summarized his research on motion in three laws of mechanics. The first rule was adopted from Galileo and was also used by Descartes:

> I. Every body continues in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed upon it.

Under the influence of an external force the state of motion changes, in other words, the body experiences acceleration. In his second law, Newton concluded that

> II. The change of motion is proportional to the motive force impressed and inversely proportional to the mass of the body; and is made in the direction of the right line in which that force is impressed.

We may state this more briefly:

force = *mass* times *acceleration*,

or using mathematical symbols, $F = ma$. The law of reaction, Newton's third law, completes the basic rules of mechanics:

III. To every action there is always opposed an equal reaction; or, the mutual actions of two bodies upon each other are always equal, and directed to contrary parts.

In other words, for the force exerted by one body (an "agent") on another, the other exerts an equal and opposite force on the "agent."

Newton could thus write down the mass dependence of the law of gravity. According to Newton's second law, the force has to be proportional to the mass of the affected body. For example, the force by which the Earth pulls the Moon must be proportional to mass of the Moon. But from Newton's third law, consider the case from the point of view of the Moon. The force by which the Moon pulls the Earth must be equal and opposite, and also proportional to the mass of the Earth. Thus in all, the gravitational attraction between two bodies has to be proportional to both masses, i.e., to the *product of the masses* of the two bodies as well as being *inversely proportional to the square of the distance* between them (using mathematical symbols: $F=GmM/r^2$, where m and M are the two masses, and r is the distance between them. G is a constant number to be determined by experiments).

The Best of all Possible Worlds

In the eighteenth century Newton's new principles gained wide acceptance, and new ways of looking at them were developed. One important development was initiated in 1744 by the French mathematician Pierre Louis Maupertuis (1698–1759) (Fig. 2.3) who pointed out that the motion between two points, say A and B, may be derived from a general principle called the *principle of least action*. The principle was quantified by Leonhard Euler and Joseph-Louis Lagrange

Fig. 2.3 Pierre Louis Maupertuis (*left*, Wikimedia Commons), and Leonhard Euler (*right*, painting by Joseph Frédéric Auguste Darbès. Credit: Musée d'art et d'histoire, Genève)

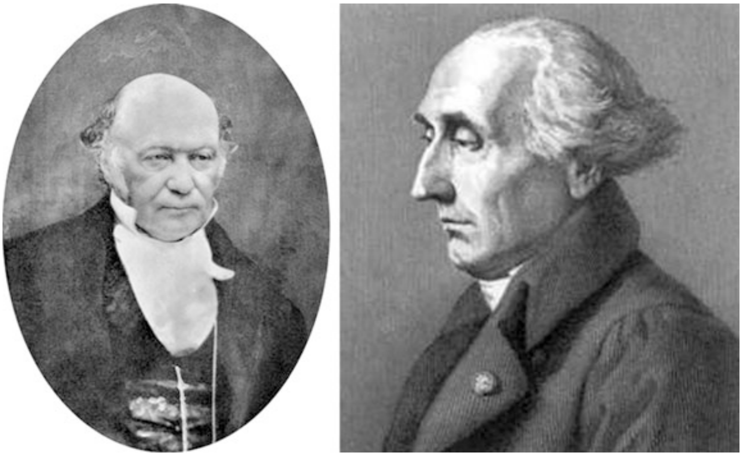

Fig. 2.4 William Rowan Hamilton (*left*, Wikimedia Commons) and Joseph Louis Lagrange (*right*, Wikimedia Commons)

and later by the Irishman William Rowan Hamilton (1805–1865) (Fig. 2.4) in 1834. What it says is that among all possible routes between points A and B nature chooses the one which minimizes a quantity called “action”. In Hamilton’s formalism, which is nowadays exclusively used, the action is the product of energy and time: thus nature tends to save both of these quantities.

Let us take an example to explain the principle of least action. Our task is to drive a car from a certain address in a city to another address, say from home to work, and to do it most cheaply. Our time as well as fuel cost money. The best route is not necessarily the shortest route, since it may have more traffic lights, and add to the cost of time. A more round-about route using high ways part of the way may turn out to be better even when the fuel consumption is greater. We may install a little computer in the car to make a calculation of the cost for all possible routes, and then let a GPS system guide us through this route.

In this example we still have a choice not to follow the GPS but take other matters into consideration. In a physical system, the principle of least action says that we have to take the optimum route, whether we like it or not. That is, we are dealing with a law of nature.

Thus it appears that the moving body sniffs out all possibilities before choosing how best to get from point A to point B. In classical mechanics this was only an alternative way to calculate an orbit. One might view the old way of using Newton’s equations at each point along the way as the standard way. Then one does not need to know where the end point B is until the journey is finished. In using the principle of least action, the initial and final positions determine how the motion must happen.

Which one of the two principles is then correct, the least action principle of sampling all possible orbits and choosing “the best”, or going strictly under the influence of forces at each point locally? Before the birth of *quantum mechanics* we did not have an answer. In 1932 the English physicist Paul Dirac (1902–1984) pointed out that in fact the sampling of all orbits is what moving bodies do. It is not so obvious for macroscopic bodies, as the distance unit for the sampling interval is related to the Planck's constant, and thus only orbits very close to each other need to be considered. The constant was first introduced by the German physicist Max Planck (1858–1947); it is the natural unit of measuring action, and a very small quantity for everyday use. The ideas were further developed by American Richard Feynman (1918–1988), and it is now well accepted that the sniffing of the "best" orbit is what actually takes place in nature.[2]

The way the sniffing is carried out mathematically is by calculating a quantity called *Lagrangian*. It is the difference of the kinetic energy due to speed of the body, and the potential energy, due to the elevation of the body. For example, a ball raised to the top of a hill has no kinetic energy while at rest, but it has potential energy. When the ball rolls down, the potential energy decreases at the same rate as the kinetic energy increases as a result of the ball gaining speed. The exact expression for the kinetic energy is ½ times mass times the speed squared. (A student at an exam on mechanics says that the kinetic energy is directly proportional to the mass and the square of the speed, and inversely proportional to deuce. Not quite full marks for the answer.) In other physical systems the Lagrangian may be defined in other ways, but it is important that the Lagrangian may be calculated at every point in space and at every moment of time. We may compound it for all possible orbits between points A and B, and get the action for each orbit. Then we only have to take the orbit with the smallest action and there we are: we know exactly how the body is going to move from point A to point B!

At least formally, this is much simpler than Newton's laws of motion. Newton's three laws are replaced by a single rule. The principle of least action is sometimes expressed by saying that our world is the best of all possible worlds.

The actual discovery of the quantity called Lagrangian, and the derivation of the laws of motion using the least action principle, is due to William Rowan Hamilton. He was born and lived in Dublin. He studied at Trinity College, where his talents were recognized early on by John Brinkley, the Royal Astronomer of Ireland,

---

[2] Quantum mechanics was a new formulation of mechanics which was developed to explain phenomena at the atomic level where Planck's constant is the suitable measure of action. The development was initiated by the Dane Niels Bohr (1885–1962), who created a model for an atom where light electrons orbit around heavy atomic nuclei. The model had resemblance to the Solar System with the heavy Sun at the center and light planets going around it. However, it soon became obvious that Newton's rules did not apply. The new rules were generated in a rush around 1925 by Germans Werner Heisenberg (1901–1976) and Max Born (1882–1970) as well as Austrian Erwin Schrödinger (1887–1961) and others. One of the essential features of this model is that bodies are represented by probability waves. These waves can interfere with each other, sometimes amplifying each other, sometimes cancelling each other. Where the wave is strong, there the particle is likely to be. The statement where the body actually is loses significance, except when its position is actually measured. Thus a particle going from point A to point B is not located anywhere in particular until it comes to B where it is found by a detector

himself a former pupil of Maskelyne, the Astronomer Royal of England. Brinkley stated: "This young man, I do not say will be, but is, the first mathematician of his age". No wonder then that Hamilton was appointed professor of astronomy in 1827, even prior to his graduation. This allowed him to take residence at Dunsink Observatory where he resided the rest of his life. Sounds dull, but scientifically, it was a highly original and very productive life.

Let There Be Light

What is the basic reason behind the inverse square law of gravity that makes the force of gravity weaken to ¼ when the distance is doubled, to 1/16 when the distance is four times greater and so forth? Actually it has to do with the same reason why you need 16 tiles to cover a square of 4 tiles wide in your garden rather than four tiles; that is, the surface area goes like the side squared. This became obvious, at least to the Englishman Michael Faraday (1791–1867), while studying electric and magnetic forces in the early part of nineteenth century. The French Charles Augustin Coulomb (1736–1806) demonstrated that the electrical force is a stronger version of Newton's force, in the sense that it is also an inverse square law like in Newton's law of gravity. The Dane Hans Christian Ørsted (1777–1851) showed that electricity and magnetism, which so far has been considered totally different, had a connection. In France, Andre-Marie Ampère (1775–1836) was quick to elaborate a theory of how this might be.

Faraday's major insight was his new interpretation of how a force is transmitted between bodies. He saw lines of force penetrating through space. For Faraday the concept of lines of force came naturally from experiments with magnets. When he sprinkled needlelike iron filings on a piece of paper lying on a bar magnet, he found that the filings lined up in very definite directions, depending their position relative to the magnet. He thought that the poles of the magnet are connected by magnetic lines, and that these lines are visualized with the help of the iron needles which line up parallel to the lines of force. For Faraday these magnetic lines were real, even though invisible.

How does this lead to the inverse square law? The strength of the force may be counted from the number of lines of force that pass through a given "tile". If we cover a sphere surrounding the source of the lines with tiles, the number of tiles increases as the surface area, which is proportional to the square of the radius. Since the number of lines does not increase, the number going through one tile goes down as the inverse square of the radius.

Because this concept was not mathematical, however, it was rejected by most scientists. Two important exceptions were Faraday's younger associates, William Thomson, also known as Lord Kelvin (1824–1907) and James Clerk Maxwell (1831–1897). They found a way to describe the lines of force mathematically.

Faraday believed that gravity could be treated similarly. Rather than saying that a planet knows by some strange reason how it has to move around the Sun, Faraday introduced a *gravitational field* which guides the planet in its orbit. The Sun generates a field in its vicinity, and planets and other celestial bodies feel the field and behave accordingly. Similarly, a charged body generates an *electric field* in its surroundings. Another charged body recognizes the field and reacts to it. Also there is a *magnetic field* associated with magnets.

Maxwell combined the separate laws of electromagnetism discovered by Coulomb, Ampère and Faraday into what is known as Maxwell's equations, treating electricity and magnetism together as a single phenomenon; electromagnetism. From Maxwell's equations one could deduce that vibrating electric and magnetic fields proceed through space with a high speed which Maxwell calculated. The value was so close to the measured velocity of light that Maxwell wrote:

> The agreement of the results seems to show that light and magnetism are affections of the same substance, and that light is an electromagnetic disturbance propagated through the field according to electromagnetic laws.

Thus light is made of electric and magnetic fields which oscillate perpendicular to the direction of propagation. In a remarkable experiment in 1887 the German physicist Heinrich Hertz (1857–1894) tested Maxwell's hypothesis of *electromagnetic waves*. He was able to produce and detect another form of electromagnetic radiation, radio waves.

The principle of least action may be now extended to electromagnetism. There the Lagrangian is built up from magnetic field and electric field at each point in space, by taking the difference of the energies in the two kinds of fields. This formulation may be generalized to a full description of all electromagnetic phenomena.

The fact that the speed of light is finite was first realized by the Danish astronomer Ole Rømer (1644–1710) while working at Paris observatory in 1676. He studied the motions of the four major moons of Jupiter, especially the innermost Io, and noticed that sometimes it comes from behind Jupiter too early, sometimes too late, depending on the season. Taking Kepler's laws to be valid, such bad timing was impossible. He realized that the messages of the appearances, carried by light, take some time to come from Jupiter to us. The time is longest when the Earth and Jupiter are on the opposite sides of the Sun, and shortest when the Earth is on Jupiter's side.

This first measurement of the speed of light, which we now know to have been rather correct, did not gain approval of the scientific community. When Rømer returned to Denmark, he was not invited to an academic position that the major discovery would have deserved, but he ended up earning his living in other ways, finally as the Copenhagen harbor master. It remained to the Englishman James Bradley (1693–1762) to acquire the definite proof of the finite speed, from the way it affects the positions of "fixed stars" in the sky over the year. The result was communicated to the Royal Society in January 1729 by Edmond Halley, his former teacher and patron. The apparent motion of "fixed" stars in the sky over the year was also the final proof of the annual motion of Earth around the Sun, and thus of the Copernican world view.

The Rise of Relativity

In Newton's universe there is a universal time which in principle is measured by synchronized clocks in every point of the three dimensional space. For a long time there was no need to question this principle. The trouble came in 1887 when American physicists Albert Michelson (1852–1931) and Edward Morley (1838–1923) tried to measure the motion of the Earth through space by studying from which direction the light comes at highest speed. Indeed, one would expect that the light should appear to come fastest from the direction toward which we are heading. This is based on everyday experience when moving through air. But it is not what experiments told: light travels with the same speed independent of the motion of the measuring apparatus.

In order to understand this fundamental result, it was necessary to introduce an entirely new way of thinking of the nature of space and time. It was done by the Swiss/German physicist Albert Einstein (1879–1955). Einstein was born in Germany, and received his early education in Munich. He excelled in Mathematics and Physics where he raced well ahead of his fellow students through self-study. For example, at the age of 12 he discovered his own proof of the Pythagoras' famous theorem. When he was 15 the Einstein family moved to Italy to start a company dealing with electric machines, and Albert worked in the enterprise. In his free time he continued his self-studies. However, in order to gain admission to higher education, he needed a high school diploma which he passed in Aarau, Switzerland, after spending one academic year there. He had generally top marks, except for French which was his weak subject. Thereafter he studied in the Technical University of Zürich.

At the age of 21, in 1900, he graduated, and would have liked to continue academic studies. However, his teachers thought him lazy and unsuitable for further academic work. Besides, and this may have been the crux of the matter, he did not show the physics professor Heinrich Weber proper respect, and addressed him Herr Weber rather than Herr Professor Weber, as was customary. After 2 years of temporary employment, Einstein finally became a technical officer of the patent office of Bern. It turned out to be a suitable job even though not academic, and while working there, Einstein finished his PhD in the University of Zürich. The supervisor Alfred Kleiner's main comment on the final thesis was that it was too short. Einstein added a sentence, and Kleiner let it pass.

There was nothing in Einstein's early career to anticipate the 1905 miracle: three articles in the esteemed journal *Annalen der Physik* which made Einstein perhaps the most famous scientist of the century.[3] The articles dealt with Brownian motion, "light gas", and Special Relativity. The first article gave crucial arguments in favor of matter consisting of atoms, a fact by no means generally accepted at the time. The second article gave a new interpretation of the nature of light and the third,

[3] The story of the 1905 miracle is told in many places, e.g., in Walter Isaacson "Einstein. His Life and the Universe." (Simon & Schuster, New York 2007).

most famous, article discussed in a novel way the concepts of space and time, which among other things, later led to a prediction of the enormous reserves of energy hidden in matter. In this research he followed the footsteps of Henri Poincaré and the Dutch physicist Hendrik Lorentz (1853–1928) who had not quite been able to put the pieces of the puzzle together.

Einstein's research was not unnoticed, but it took a while before it became general knowledge among professionals. Einstein was appointed a docent in the University of Bern in 1908, but his university career started properly a year later when he became associate professor at the University of Zürich. He moved to Prague in 1911. The time in Prague was significant to Einstein's career since he learnt there new mathematical methods with the help of his assistant Georg Pick. These were necessary for his next great step forward in physics.

Only a year later Einstein returned to Switzerland, to his *alma mater* in Zürich where he started developing the General Theory of Relativity together with Marcel Grossmann. This is a new theory of gravity which improves on the Newton's theory. By this time Einstein had become famous and was invited in 1914 to become the head of the Physics Department at the Kaiser Wilhelm Institute in Berlin and a member of the Prussian Academy. Here he published the foundations of the General Relativity in 1916. During the solar eclipse of 1919, British delegations organized by Arthur Eddington (1882–1944) observed the bending of light predicted by Einstein, thus making his theory a serious rival of Newton's theory.

Because of Hitler's rise in power in Germany, Einstein had to migrate to United States. In 1934 Einstein settled at Princeton, New Jersey, living there for the rest of his life, working to unify electromagnetism and gravity under a single theoretical framework. He did not succeed, and neither have others.

In his Special Theory of Relativity, Einstein accepted Michelson and Morley's observation that the speed of light is a constant, c, independent of the state of motion of the observer. He did not ask why, but rather addressed what consequences derive from this odd fact. What are space and time? The constant speed of light does not make sense in everyday life. The speed of light can be the same for everybody only if space and time are linked in a way which nobody had anticipated.

The entanglement of space and time means that we live in a special kind of *four-dimensional* world. The nature of time is different from the three spatial dimensions (length, width, height), and not only because we measure time by a clock, while distance is measured by a ruler. Hermann Minkowski, Einstein's former mathematics professor, explained this view in 1908 as follows:

> From henceforth, space by itself, and time by itself, have vanished into the merest shadows and only a kind of blend of the two exists in its own right.

The constant c already appears in Maxwell's equations. It is curious that the first relativistic theory was Maxwell's electromagnetism which was constructed before relativity theory itself! When he invented his famous equations, Maxwell did not know that they were hiding a treasure, the theory of relativity.

Soon after sending out his three fundamental works in 1905, Einstein realized that the relativity theory leads to a very special and unexpected connection between mass and energy. The theory suggests that all matter has hidden energy by the amount

Energy = *mass* times *speed of light* squared.

(Using mathematical symbols: $E = mc^2$). Since the speed of light is a big number, this formula implies that even a small bit of matter contains a huge amount of energy. If one gram of matter could be turned totally into energy, it would provide the same amount of energy that is liberated by burning 10,000 barrels of oil. The enormous power of nuclear energy is based on liberating a small fraction of the mass of the atomic nucleus into energy. The first controlled liberation of nuclear energy was achieved by the Italian physicist Enrico Fermi and his group in Chicago in 1942; since then the nuclear power has become one of the main resources available to mankind. This is also how the energy of the Sun is produced.

Geometry and Gravity

Newton's view of space was based on Euclidean geometry. Even in the Special Theory of Relativity the spatial part of the 4-dimensional space-time is Euclidean. Euclid, who worked in Alexandria around 300 BC, developed a system of geometry which is still part of our mathematics curriculum. It was based on five "obviously true" axioms, out of which a rich collection of 465 theorems were derived, the essential knowledge of geometry. Among the five axioms, the most widely discussed is the last axiom which states that

> through a given point in a plane one can draw one and only one line parallel to a given line in this same plane.

Euclid and many of his followers had misgivings about this *Parallel Postulate*. Though it seems intuitively true, there was no way of confirming it experimentally. In practice we are always dealing with limited segments of straight lines, and cannot observe the whole of the line. But perhaps it could be inferred from the other four axioms? In fact, for two millennia mathematicians tried to demonstrate that the fifth postulate is implied by the others. All these attempts failed.

Not until the nineteenth century did it become clear that the fifth axiom can be replaced, ending up with other systems where geometric relations are different from what we are used to. Among the many possibilities, there are two interesting cases: *hyperbolic geometry* which was discovered independently by German Carl Friedrich Gauss (1777–1855) (Fig. 2.5), Russian Nikolai Lobachevski (1792–1856) and Hungarian János Bolyai (1802–1860), and *spherical geometry*, invented by the German student of Gauss, Georg Riemann (1826–1866). Besides the Euclidean *flat geometry*, these two are the only possible descriptions of space which is homogeneous and isotropic, that is, where all places and directions are equivalent.

Gauss invented the term "Non-Euclidean Geometry" but did not publish anything on the subject since "he was most unwilling to get involved in something that would put him under criticism". In a private letter of 1824 Gauss wrote:

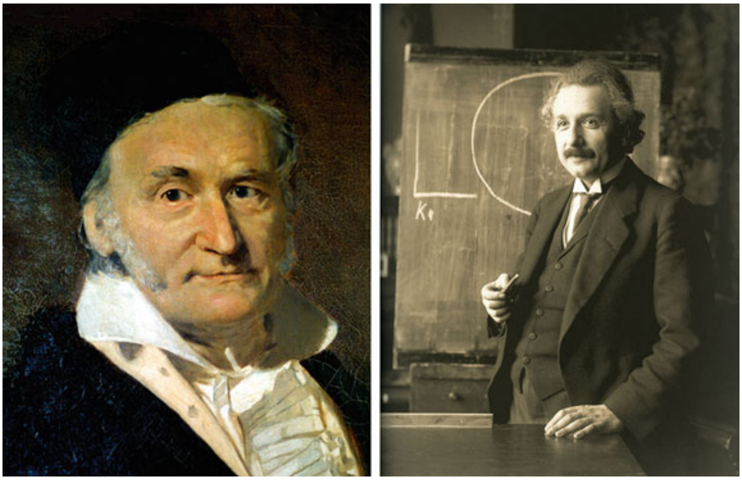

Fig. 2.5 Carl Friedrich Gauss (*left*, credit: Wikimedia Commons) and Albert Einstein (*right*, Wikimedia Commons)

> The assumption that (in a triangle) the sum of the three angles is less than 180° leads to a curious geometry, quite different from ours, but thoroughly consistent, which I have developed to my entire satisfaction.

Riemann developed the mathematical methods which are required for calculations in the non-Euclidean geometry. The work was continued by Enrico Betti (1823–1892), Gregorio Ricci-Curbastro (1853–1925) and Tullio Levi-Civita (1873–1941) in Italy. Albert Einstein learnt the necessary mathematical methods from the latter for his theory connecting geometry and gravity. This field of mathematics was not in university curricula, and it took a while even for Einstein to learn about it.

The common way to visualize non-Euclidean geometry is to use surfaces as examples. It is difficult to imagine a four dimensional space, not to mention what its curvature might mean. Our brains are not used to tackle such problems; thus it is best to limit ourselves in looking at two dimensional surfaces. A creature who does not possess any third dimension off the surface, and who does not even understand what the third dimension would mean, can carry out geometrical measurements on the surface to find out the overall geometry. One may draw a triangle, and measure the sum of the internal angles. If the result is more than 180°, that determines right away that the creature lives on a spherical surface. Alternatively, he may draw a circle and measure it. If the ratio of the circumference to the diameter of the circle is less than π, the creature would know that he lives inside spherical geometry (Fig. 2.6).

A flat surface is an example of a two dimensional Euclidean space. The laws of Euclidian geometry which we are familiar with, are valid in this and only in this geometry: the sum of the internal angles of a triangle is exactly 180°, the ratio of the

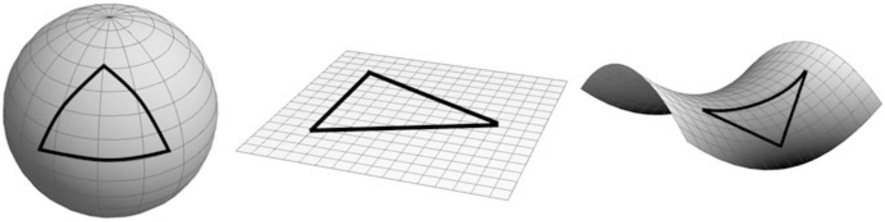

Fig. 2.6 Spaces with different geometry. Sum of the internal angles of triangles is (*left to right*): more than 180 degrees (positive curvature), 180 degrees exactly (zero curvature) and less than 180 degrees (negative curvature)

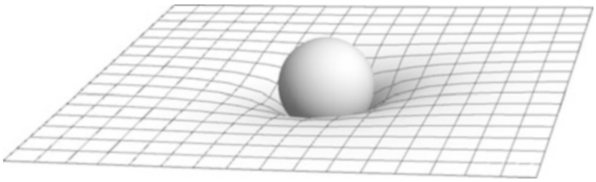

Fig. 2.7 Matter makes the surrounding space curve

circumference (s) of a circle to its radius (r) is exactly 2π ($s = 2\pi r$), and one can draw one and only one straight line through a point, parallel to another straight line.

According to Einstein, the root cause of gravity is curvature of space-time. Matter makes the surrounding space curve, and bodies react to this curvature in such a way that there appears to be a gravitational attraction.

From the known geometry of space-time, it is possible to calculate the orbit of a body that is not influenced by anything else besides gravity. Einstein did not view gravity as a force. In a flat space-time force free motion happens on a straight line, but in a curved space-time the force free motion can create practically closed orbits. Take a planet circling around the Sun. It moves forward as straight as possible, but because the Sun has curved the space-time, the orbit becomes an ellipse. We may illustrate this with a stretched horizontal sheet of rubber ("flat space"). An iron ball placed in the middle of the sheet causes a dip in the surface. Now roll a marble ball along the sheet. With a push in the right direction, you may get the light marble to roll around the big ball, perhaps in an elliptic orbit. It appears as if there is a central force pulling the marble, when in fact the closed orbit arises from the form of the surface (Fig. 2.7).

In case of the motion of planets around the Sun both Newton's theory and Einstein's theory give practically the same result. The most important difference arises with Mercury. The long axis of Mercury's orbit precesses slowly due to the influences of other planets. But Einstein's theory gives an extra precession by 43" per century as compared with Newton's theory. In fact, this little bit extra had already been observed and it was an unsolved problem at the time!

The explanation of the motion of Mercury was the first success of Einstein's new gravitation theory. Another consequence is the bending of light rays when they pass close to the surface of the Sun. Because of it, stars appear to shift away from their usual places in the sky when the Sun is in the foreground close to them. Normally we don't see the Sun and the stars at the same time, but during a solar eclipse it is possible. When the shift of stars by the expected amount was detected during the solar eclipse of 1919, it was hailed as victory for Einstein.

The crucial observation was carried out by Arthur Eddington, the director of the Cambridge observatory. As conscientious objector he did not take part in active war service in World War I, and barely managed to avoid going to prison by organizing his now famous solar eclipse observation trip. He had become aware of Einstein's General Relativity, and obtained whatever publications he could, taking into consideration that Eddington was not supposed to read "enemy literature". In spite of these limitations, he became one of the leading experts in the field, and it was natural for him to organize the expedition to verify the theory. There was also strong opposition to his activity, as he was seen taking the side of the German Einstein against the English Newton. At a Royal Society meeting, a Fellow who was well acquainted with General Relativity but was skeptical about it, wanted to dampen Eddington's enthusiasm. He challenged Eddington to tell, if he considered himself to be one of only three men who understood the theory (the others being obviously Einstein and the Fellow himself). As Eddington did not respond, he prompted Eddington not to be so shy and answer. Eddington replied: "Oh no! I was just wondering who the third one might be!"

In General Relativity we are dealing with space-time rather than space and time separately; therefore the action is calculated not only over the flight time of the orbit but over the 4-dimensional space-time. The quantity which is compounded over the space-time, going from point A to point B, is called the scalar curvature (also called Ricci scalar after Ricci-Curbastro). Using it as a Lagrangian one may derive Einstein's law of gravity, in place of Newton's law of gravity. This is how the space-time curvature enters mathematically into the law of gravity. The scalar curvature is number 2 divided by the square of the radius R of a sphere ($2/R^2$), if we are describing two-dimensional spherical geometry, in three dimensions 2 is replaced by 6 (3 times (3–1) times (3–2) = 6, the rule goes). The action formalism was noted parallel with Einstein by the German physicist David Hilbert (1862–1943) in 1915, and it makes the statement of the law of gravity extremely simple: the bodies move in such a way as to make the compound of the scalar curvature as small as possible.

The discovery of the least action principle for gravity has an interesting history. Einstein proposed this principle first with Marcel Grossmann in 1914, but it was not until November 4 in 1915 that he first derived the law of gravity from this principle. However, even this formulation was not quite correct, and on November 25 he presented the revised gravity equations. In the meantime, on November 20, David Hilbert had sent an article for publication which not only had the correct gravity law but also its derivation from the least action principle. Did Hilbert then discover General Relativity rather than Einstein? Hilbert never made this claim, but the action is called Einstein-Hilbert action in recognition of Hilbert in the development of the theory, while the theory on the whole is ascribed to Einstein. Actually, it is not known whether it was Hilbert or Einstein who wrote the correct equations of General Relativity first, as Hilbert modified his article in preparation for printing in December, and the modifications have not survived for the historians of science to assess the situation.

In 1917 Einstein made a new addition to Einstein-Hilbert action. He realized that a constant number could be added to the action without violating the principles of General Relativity. The constant has the effect of producing a repulsive force between bodies in the universe. It is important to note that this repulsion first appears only in the scale of the universe, as a repulsion between galaxies. He introduced the constant in order to create a model for a boundless, static and eternal universe. The idea was appealing, but it was soon contradicted by the discovery of the expansion of the universe. At that point Einstein lost interest in the repulsive force. However, it has made a comeback in recent times. Today this repulsion goes by the name of *dark energy*. We will discuss it in Chap. 6 when we talk about galaxies.

After all the developments over more than 300 years since the publication of *Principia,* the Newton's law of gravitation and the laws of motion are practically enough for most astronomical computations. It is true that General Relativity is more accurate, and has to be included from time to time, but even then it is usually done as a small correction to Newton's laws. At the atomic level we encounter different problems; at the limit of large bodies the atomic level theories converge to Newton's laws.

However, we notice that Newton's major philosophical problems have now been answered. We have an intermediary for gravitational attraction, the space itself which has physical properties just as Riemann had speculated. It is not a problem how Newton's force can enter inside physical bodies. Neither Divine spirit of Newton nor active principle of Stoics is needed!

Springer
http://www.springer.com/978-3-319-22725-2

The Three-body Problem from Pythagoras to Hawking
Valtonen, M.; Anosova, J.; Kholshevnikov, K.; Mylläri, A.;
Orlov, V.; Tanikawa, K.
2016, XI, 173 p. 96 illus., 70 illus. in color., Softcover
ISBN: 978-3-319-22725-2