6 The gravitational mechanics of the Earth

In Cambridge Encyclopedia of Earth Sciences, Chapter 6, 1982.

Physicists have deduced that there are four types of force that operate between material bodies: the electromagnetic force, the gravitational force and the strong and weak nuclear forces. The second of these, gravity, is the most important in astronomy and the Earth sciences. It is well known that Isaac Newton (1643-1727) gave a correct mathematical description of the laws of gravity while he was in his early twenties. Newton’s law of gravitation states that, for two point masses, the force attracting them is proportional to the product of their masses, and inversely proportional to the square of the distance of their separation. Although Newton’s law has been superseded by Einstein’s general theory of relativity it is completely adequate for all practical purposes in geophysics and much of astrophysics.

In order to write down Newton’s law $F$ can represent the force of attraction, $m_1$ and $m_2$, the pair of point masses and $r$ the distance between them. The law is then written thus: $F = G m_1 m_2 / r^2$. The quantity $G$, the gravitational constant, is found experimentally and is difficult to measure with high precision. In the international metric system its value is $6.670 \times 10^{-11}$ N m$^2$/kg$^2$. This small value for the constant reflects the fact that gravity is an intrinsically weak force. In everyday life we are only aware of the gravitational attraction of the Earth itself, which has a large mass ($M = 5.98 \times 10^{24}$ kg); we experience this force as ‘weight’. Geophysics and astrophysics deal with very large masses—planets, moons, stars and galaxies—as well as enormous distances. Under such conditions it is gravity that is the dominant natural force.

The law of gravity is central to the study of the physics of the Earth as it determines the general shape of a planet, influences many tectonic processes that shape the planetary surface and controls the orbital and rotational motions. Much of the discussion in this chapter will deal with the far-reaching consequences of this simple law.

Gravity and the solar system

When we consider the orbits in the solar system, such as the Earth’s path round the Sun, we have to take into account the gravitational attractions of many masses, not just the two that appear in Newton’s law. The total attraction experienced by any one planet is found by taking the vector sum of the individual attractions. This was once a formidable computational problem which had to be laboriously worked out, but it is now relatively easy to solve by computer. Knowledge of the planetary orbits has been greatly increased through space missions and accurate high-speed computing.

Orbital and spin motions

As Copernicus (1473-1543) suggested, the Earth moves around the Sun in an elliptical orbit, the orientation of which is fixed in space. When the motion is averaged over a long time interval the mean orbital plane defines a useful frame of reference for describing the orbital motions of the Moon and other planets, as well as for describing the orientation of the rotation axes of planets and satellites in the solar system: this plane is called the ecliptic. Most solar system bodies do not move far out of this plane. The Earth’s instantaneous rotation or spin is about an axis that is inclined to the ecliptic by about $66^\circ$, an angle that remains more or less constant. The Earth’s equator is therefore inclined at an equal angle to the ecliptic and this inclination, known as the obliquity of the equator, is responsible for the annual seasonal motion of the Sun’s path in the sky as seen from the Earth.

The Moon moves about the Earth, also in a slightly eccentric orbit, in a plane that maintains an almost constant inclination, about $5^\circ$, to the ecliptic. Because of the obliquity, the inclination of the lunar orbit upon the equator varies periodically with the period of the lunar month, about twenty-eight days, the time it takes for the Moon to make one revolution about the Earth. The values of the constants in these motions partly reflect the conditions at the time of formation of the planets: for example the length of the year and the eccentricity of the Earth’s orbit. Other ‘constants’, such as the period of the lunar orbit and the length of the Earth’s day, are consequences of the subsequent dynamic evolution of the solar system.

The motions of the Earth–Moon–Sun system cannot be fully treated by discussing only the attractions between two bodies at a time, although this simplification does explain the dominant characteristics of the motions. Thus the elliptical motion of the Earth about the Sun is principally due to the interaction of the gravity fields of these two bodies, while the motion of the Moon about the Earth is, in the first approximation, a consequence of the gravitational interactions between these two bodies only. In a complete discussion the problem must be treated as one of mutual gravitational interactions between the three bodies. Thus, because of the Sun’s attraction of the Moon, the latter’s motion will oscillate in orientation and size by small amounts about a mean orbit that is almost elliptical.

If the three bodies, Sun, Moon and Earth, could be considered as point masses or as spherically symmetric bodies the resulting orbital and spin motions could be readily evaluated from Newton’s laws. But the Earth and the Moon are not perfectly spherical and this causes considerable complexity in solving the equations governing the motion of the Earth and Moon.
Motions of the spin axes
The most significant departure from the spherical shape of the Earth is its flattening at the poles. Our planet's shape can be better approximated by an oblate spheroid, symmetrical about the spin axis. The flattening, which is a direct consequence of the planet's rotation, can be described in several ways but, most simply, it means that the polar diameter is about 43 km shorter than the equatorial diameter. This departure from radial symmetry causes the Sun and Moon to exert additional forces (torques) on the Earth, so inducing shifts in the position of the rotational axis in space. These shifts are the precession and nutation discussed in the following section. Neither is the Moon a strictly radially symmetrical body and the Earth's and Sun's attraction will similarly cause its spin axis to precess in space. This is referred to as the lunar libration.

The rotation of the Earth
Precession and nutation
Viewed by an observer on the Earth and near the north (or south) pole, stars appear to trace out concentric circles whose centre defines the celestial north (or south) pole, the extension of the Earth's rotational axis. The celestial north pole currently lies close to the star Polaris. Careful observations over many years have revealed that the position of this celestial pole changes relative to the stars, as was noted by Hipparchus about 120 BC. The rotation axis is observed slowly to trace out a cone, with a half-angle of 23½°, about the pole of the ecliptic. It takes about 26,000 years to go full circle around the heavens. In about 3000 BC the pole star was Alpha Draconis; Alpha Cephei will be near the pole in AD 7500. This steady motion of the rotation axis in space is termed the precession of the Earth.
As noted above, the major axis of the oblate Earth is inclined to the ecliptic. In consequence the net gravitational force on the Earth due to the Sun does not pass through the centre of mass of the Earth. This results in a torque being exerted about the centre. The torque attempts to draw the equator into the plane of the ecliptic but the spinning Earth
resists this. Instead the torque achieves a motion of the spin axis about the pole of the ecliptic.

The Moon acts on the Earth in a similar way and the observed precession is the sum of the polar and lunar torques plus a rather minor contribution arising from the other planets. The Earth's orbit about the Sun is somewhat eccentric and twice a year the Sun passes over the equator where it is aligned with the Earth's bulge; consequently the solar torque varies periodically, as does the Moon's torque. The net result is that the secular precessional motion of the rotational axis is perturbed by small oscillations or 'nodding' motions called the forced nutations. The principal nutation term has a period of nineteen years, and the size of the nodding motion is nine seconds of arc. This arises from a nineteen-year periodicity in the inclination of the Moon's orbit. It was first detected and explained by the astronomer James Bradley (1692-1762) in 1747.

Although the precessional and nutational motions are mainly of astronomical interest, they are also of some geological consequence in that their amplitudes depend on the oblateness of the Earth and therefore provide some information on the internal structure of the Earth. This is discussed below.

Polar motion

If an observer at either pole recorded photographically the small concentric circles traced out by the circumpolar stars, with time the centre of the circles would be seen to be slightly displaced on the photographic plate, and this centre would itself trace out a small circle over about a one-year period. This reflects a motion of the rotational axis relative to an Earth-fixed reference frame and is referred to as polar motion. It is quite distinct from precession and nutation. If the motion is observed from space the Earth as a whole appears to wobble about its axis. This motion was predicted by L. Euler (1707-83) in 1765, but it was not observed until the end of the nineteenth century.

The motion of the pole is deduced from careful measurements of the positions of the stars. Figure 6.2 illustrates the pole path as observed over a two-year period—a meandering motion that is typical of the polar path observed over the past eighty years. An analysis of such observations indicates that the motion is mainly made up of two periodic oscillations, one with a fourteen-month period, the other with a twelve-month period, both with amplitudes of the order of 0.1 arc second. Hence the wobble of the rotational axis at the north and south poles, relative to the Earth's surface, is merely a few metres, and precise astronomical measurements are needed to observe it. Yet these observations yield considerable information on the Earth, which will now be discussed.

To understand the fourteen-month oscillation the Earth may be regarded as a rigid oblate spheroid as in the previous discussion of the precession and nutation. If the rotational axis of a rigid body is initially aligned with its principal axis of maximum inertia then that body will rotate uniformly with the two axes always remaining parallel; however, if for any reason the two axes are tilted relative to each other, Euler's theory predicts that the rotational axis appears to wobble about the principal axis. For the Earth this theory predicts a 305-day wobble. But the observed period is not 305 but 430 days, as first noted by S. C. Chandler in 1892. This increase is explained by the Earth not responding to the rotation as a rigid body. The motion must be extended to the rotation of a deformable Earth in which the mantle's elastic deformation must be considered as well as movements in the oceans and in the liquid core. The theory is relatively complicated but one aspect of the Earth's non-rigidity can be mentioned here.

Consider that the Earth has a liquid core filling a spherical cavity. If the fluid had no viscosity then there would be no frictional forces linking motions in the core to the mantle, and vice versa. The outer mantle would be able to move relative to the inner core without disturbing the fluid motions within. If this simplification described the real situation then it would reduce the theoretical period of the wobble by about thirty days. The elasticity of the mantle on the other hand increases the period by about 120-130 days and the oceans lengthen it by a further thirty days, so that the computed period is now close to the observed 430 days. From this discussion it can be seen that observations of polar motion assist the deduction of the Earth's structure.

The observed wobble, now referred to as the Chandler wobble, being a free oscillation, should ultimately be damped since no free oscillation will persist indefinitely in physics. Yet the observations suggest that the oscillation has persisted for nearly a century, albeit with considerable fluctuation in amplitude and phase, and this suggests that there is some mechanism exciting this motion. One mechanism (there are several) is the excitation of the wobble by large earthquakes. This
6.3: Motion of the instantaneous rotational axis \( \omega \) about the principal axis. Until time \( t_1 \), when the rotational axis is at \( P_1 \), the pole has rotated about the principal axis at \( A \). At time \( t_1 \), an earthquake occurs and modifies the mass distribution such that the principal axis is now at \( B \). The pole now moves about this new position until the next earthquake occurs at time \( t_2 \) when the pole is at \( P_2 \). The principal axis then jumps to \( C \).

is illustrated in Figure 6.3. Suppose that while the rotational axis wobbles about the principal axis at \( A \), a large earthquake occurs at time \( t_1 \). This seismic event, if large enough, changes the mass distribution within the Earth and shifts the position of the principal axis from \( A \) to \( B \). The rotational axis now moves about this new position of the principal axis leaving behind it a kink in the pole path. At some later time \( t_2 \), a second earthquake occurs and a second kink forms. In this way the Chandler wobble can be maintained as long as sufficiently large earthquakes occur at irregular but frequent intervals. This points to the main problem with this model; for while some earthquakes appear to be sufficiently large to shift the principal axis by the requisite amount, there do not appear to have been enough of them to maintain the wobble throughout the past hundred or so years. It is more likely to be a combination of the catastrophic seismic shock and slower deformations preceding or following the main shock that displaces the principal axis.

The other, annual, oscillation in the polar motion is a consequence of the Earth's mass distribution being periodically modified by seasonal redistribution of mass within and between the atmosphere, oceans and surface and ground water.

In addition to these two periodic motions the pole path observations exhibit a secular drift in a direction roughly along a meridian 20° west of Greenwich and at a rate of about 0.002–0.003 arc second per year. This is apparently a consequence of an exchange of water mass between the polar ice sheets and the oceans and of the Earth's response to the changing load on its surface.

Changes in the length of day

The third aspect of the Earth's rotation concerns the angular velocity about the instantaneous rotational axis. Astronomers observe the times of transit of a star across their meridian using precise atomic clocks to establish the time scale, and this provides a measure of the length of the day. These observations indicate that the interval between successive transits varies perceptibly, and that the Earth is usually either ahead or behind the time kept by the clocks. Typically the length of the day fluctuates by 1 part in \( 10^6 \), equivalent to about \( 10^{-5} \) second, a small but observable amount. Observations of the length of day have been made regularly since about 1820, but only with the introduction

6.4: A schematic representation of the forces acting on the Earth that will perturb the rotational motion.
6.5: Proportional changes in the length of day as observed on three different time scales. (a) From 1820 to 1970. Only fluctuations that persist for about ten years or longer are evident. (b) From 1962 to 1977. Seasonal variations are clearly evident. (c) Upon removal of the seasonal effects, highly irregular variations are seen. Some of these may represent 'noise' in the astronomical data while some are due to a rapid exchange of angular momentum between the Earth and atmosphere.

Of atomic time in about 1955 have the observations been sufficiently precise to establish a comprehensive picture of these changes. Prior to this time only changes that persisted for five to ten years or longer could be seen but now changes on a week-to-week basis are detectable.

Figure 6.5 illustrates some results. In Figure 6.5(a) only the variations occurring on a 'decade' time scale are indicated and most notable is the sudden decrease in spin velocity from about 1870 to 1900, during which period the length of day increased by nearly 10 milliseconds. This is followed by a reversal in trend from 1900 to about 1930 and a second period of deceleration continuing into the 1970s. The origin of these changes has remained obscure but only the core contains sufficient mass and mobility to explain these fluctuations. The most likely explanation is that they are caused by interactions at the core–mantle boundary between the magnetic field (in and moving with the core fluid) and the electrically conducting lower mantle. This results in a variation in the extent to which the core moves with the mantle and in a concomitant change in the mantle spin. The theory is complex and not readily verifiable since the magnetic field responsible is largely shielded from the observer by the low electrical conductivity of the upper mantle.

Of the higher frequency fluctuations observed clearly since 1955 the annual and semi-annual behaviour is mainly caused by an exchange of angular momentum in the east–west circulation of the atmosphere with the angular momentum of the solid Earth. In simple terms, as the westerly winds speed up so the Earth slows down, and vice versa. In addition to these seasonal terms many irregular changes in rotation rate occur in the frequency range of about 0.3 cycle per year to perhaps as high as 10–20 cycles per year and these are also largely due to irregular fluctuations in the atmospheric circulation.

The astronomical observations of the Earth’s rotation have provided very fascinating data whose interpretations impinge upon many aspects of the Earth sciences. Briefly, any force that exerts a torque on the Earth’s crust, or that results in a redistribution of mass within the Earth, is a candidate for perturbing the Earth’s rotation. There is a wide range of phenomena that do this, including the secular tidal torques, to be discussed below, mantle convection, fluctuations in the magnetic field, relative motions in the core, oceans and atmosphere, the direct attraction of the Sun and Moon and the concomitant tidal deformations.

The motion of the Moon

Orbital motion

The complex problem of determining the Moon’s orbital motion occupied many of the great mathematicians of the eighteenth century.
including Euler, Clairaut, D'Alembert, Lagrange and Laplace. To the intrinsic interest of the problem was added the desire to test Newton's theory. A further incentive was the substantial financial rewards that governments and scientific societies offered to those mathematicians who could provide the precise tables of the Moon's motion essential for navigation at sea.

The simplest approximation to the motion of the Moon is an ellipse that remains fixed in size, shape and orientation. This ellipse has a mean radius of about 380,000 km and an eccentricity of about 0.055, and is inclined by about 5° to the ecliptic. The orbital period of the Moon moving along this simplified ellipse would be about twenty-seven days for motion observed relative to the background stars. This is referred to as the sidereal month. During one sidereal month the Sun moves eastwards by nearly 30° and the time between successive full Moons—the synodic month—which occurs when the three bodies are aligned, is in consequence longer than the sidereal month by nearly two days.

The simple motion assumed above is much perturbed by the gravitational attraction of the Sun and to a lesser degree by the Earth's departure from a point mass as well as by the attractions of the other planets. The main consequence of the solar attraction is to precess the orbit in its plane, such that the line of apsides—the line joining the point at which the Moon is nearest to the Earth (perigee) to the point at which the Moon is furthest away (apogee)—rotates in the plane of the orbit with a period of about 8.9 years (Figure 6.6). A second consequence is that the orbital plane itself precesses about the ecliptic, such that the line of nodes—the intersection of the orbital plane and the ecliptic—makes one revolution in about 18.7 years, in a direction opposite to the Moon's orbital motion. Superimposed upon these secular motions of the orbit are a number of periodic perturbations due to the variation of the solar attraction with the continual changing Earth–Moon–Sun geometry.

**Librations**

The Moon always shows the same face to the Earth. However, this represents only an average state, for although the lunar spin velocity is constant, the Moon's orbital velocity along the ecliptic path about the Earth is not. Thus there are some times when the spin is behind its orbital motion and other times when it is ahead; at such times it becomes possible to see a little beyond the limits of the average disc directed towards the Earth. Furthermore, since the inclination of the Moon's orbit on the ecliptic is 5.2° and the inclination of the Moon's equator on the ecliptic is about 1.5°, it becomes possible to see about 6.7° of latitude beyond the two poles. Together these circumstances are referred to as the optical librations. They permitted nearly 60 per
cent of the Moon’s surface to be photographed and investigated from the Earth long before the lunar orbiter programme provided nearly complete photographic coverage. The librations have also permitted the lunar topography to be estimated for the limb regions, observations that continue to provide control on lunar mapping.

Of geophysical interest are the much smaller physical librations, analogous to the Earth’s precession and nutation. These are a consequence of the Earth exerting a torque on the asymmetrical mass distribution of the Moon. Their magnitudes do not exceed a small fraction of a degree, and while their existence was predicted by Newton they were not observed until 1839. The amplitudes of these oscillations depend on the distribution of mass within the Moon, and in common with the precession observations of the Earth their observation provides some information on the shape and mass distribution of the planet. The Moon’s spin being only about one twenty-seventh of that of the Earth, the lunar polar flattening is much less than that of the Earth.

The Moon is subject to a solid tidal deformation in the same way that the Earth is tidally deformed (see below) but with the difference that the lunar tidal bulge is predominantly and permanently oriented towards the Earth. Thus the Moon possesses an equatorial bulge and the Moon’s shape can be described approximately as a triaxial ellipsoid with the major axis directed towards the Earth.

Tides

Ocean tides

Perhaps a more familiar consequence of the Moon’s gravitational attraction than precession and nutation are the Earth’s tides. This phenomenon is perhaps most readily understood by viewing the Earth as a rigid sphere covered by an ocean layer of uniform depth. The Moon’s gravitational attraction on the Earth is slightly greater on the side of the Earth towards it than elsewhere and this causes a small bulge in the ocean layer that is directed towards the Moon. On the opposite side of the Earth, away from the Moon, the gravitational attraction is a minimum; however, the lunar attraction on the solid Earth exceeds that on the water since the former is closer to the Moon. The solid Earth is pulled more towards the Moon than is the water which actually appears to be pulled away from the Earth as a second bulge. During the daily rotation of the Earth underneath the Moon, the tidal bulge moves with the Moon and therefore around the surface of the Earth with one bulge always facing the Moon and the second bulge directly opposite. The time between successive transits of the Moon across the observer’s meridian is about twenty-five hours. At any point on the Earth’s surface the successive lunar tides pass at intervals of about twelve and a half hours.

Likewise, the gravitational attraction of the Sun also raises two tidal bulges of about half the amplitude of the lunar tide and these travel around the Earth in nearly twenty-four hours. When the Sun and Moon are aligned the two tides reinforce each other, producing the spring tides; this occurs every two weeks at full or new Moon. When the Sun and Moon are separated in longitude by 90° the two tides partly cancel and the combined tide has a minimum amplitude. These are the neap tides. The actual pattern of tidal periodicities is much more complicated than this since the intensity and direction of the force of gravity varies with the Moon’s orbital motion along its inclined and eccentric path. An analysis of a long series of tidal observations reveals a large number of oscillations whose periods cluster around twelve and twenty-four hours as well as a number of much longer periods up to 18.6 years.

Observed tidal patterns

The actual spatial pattern is complicated by the ocean–land distribution, by the variable depth of the ocean and by frictional forces along the sea floor. Figure 6.7 illustrates the tidal pattern of the principal semidiurnal lunar tide around the British Isles, where the tides are particularly complex. The lines of equal phase, called cotidal lines, join

![Image of tidal patterns](image)

6.7: The principal lunar semidiurnal tide around the British Isles. The co-range (black) lines show the occurrence of equal amplitude in centimetres. The equal phase (red) lines give the phase of the tide relative to the lunar transit across the Greenwich meridian. The arrows denote the sense of rotation at the amphidromes (points of zero tidal amplitude).
points at which high water occurs at the same time. Of note are the regions where the tidal amplitudes are zero. These are the amphidromic points about which tidal currents rotate without there being a change in amplitude. One such point occurs in the North Sea between England and Denmark and the tidal currents rotate about it in an anticlockwise manner. Thus along the western boundary of the North Sea the maximum tidal amplitude occurs progressively later from the Shetlands to the Strait of Dover. Figure 6.8 illustrates a global representation of the same tide. These global representations do not illustrate the detail of the local tidal patterns along the coasts where the geometry of the coast line and of the sea floor can result in tidal amplitudes exceeding 12 m, as in the Bay of Fundy, Canada.

The ocean tides and the associated tidal currents are of importance in navigation in coastal and shallow seas. They also influence the coastal landform by forming sand bars or barrier islands that are subsequently stabilized by vegetation and may become a permanent feature of the local geography.

Tidal energy
One frequently mentioned use of the ocean tide is as a source of energy, by using the tidal currents to drive turbines. The total energy stored in one cycle of the ocean tide is of the order of $10^{18}$ joules. This is comparable with the geothermal loss of heat but four orders of magnitude less than the solar energy received by the Earth during the same period. For much of the oceans the energy density is extremely low and only in some coastal areas, where the tidal currents are sufficiently enhanced by local conditions, may the extraction of this energy be worthwhile (see also Chapter 9).

Solid tides
Sensitive instruments capable of measuring small changes in the local gravity or small tilts of the Earth's crust indicate that the solid Earth is also subject to tides, implying that the Earth does not respond as a rigid body to the gravitational attraction of the Sun and Moon. Typically, an observer on the Earth's surface will move up and down in

6.8: The global lunar semidiurnal tide. The co-range or equal amplitude lines are black and the cotidal or equal phase lines (in lunar hours) are red.
twelve hours by as much as 500 mm. At the same time gravity on the surface will change fractionally. This is the solid or elastic tide. Its amplitude is a function not only of the magnitude of the tide-raising force but also of the elastic properties of the Earth. The Earth’s response to this force is not instantaneous since the planet is not a pure elastic body and the tidal bulge lags behind the applied force by a small angle that is unlikely to exceed a degree. The observations of the amplitude and lag of the tidal response are of considerable geophysical interest in that they reflect elastic and anelastic properties of a periodically stressed planet at relatively low frequencies. As such they are complementary to seismic studies of the planet’s response to the high frequency waves excited by earthquakes and to rotation studies which measure the response at much lower frequencies.

The most reliable observations of the tide are obtained with precise gravimeters located on the Earth’s surface. During the tidal cycle gravity varies because of the direct attraction by the Sun or Moon, the variable distance of the tidally deformed surface from the centre of mass and the redistribution of mass inside the body. The combined effect of these last two factors is about 1.5 per cent of the direct attraction. The gravimeter cannot distinguish between the solid tide and the attraction of nearby ocean water, a contribution that will also vary periodically because of the ocean tide. A simpler remedy is to observe the solid tides far from coastlines. Figure 6.9 gives an example of a record taken at a station in central Australia.

**Tidal friction**

If, for the moment, we consider only the solid tide, the bulge will be aligned with the Earth–Moon axis if the tidal response is that of an elastic body. But in a more realistic model the deformation is subject to frictional dissipation and the response is somewhat delayed. During the ‘delay’ time the Moon will have moved through a small angle along its orbit. Thus the bulge appears to be ahead of the Moon, as shown in Figure 6.10. The lunar attraction on the nearest misaligned bulge exceeds that on the farside bulge and a torque is exerted which does not vanish when averaged over an orbital period of the Moon. The consequence of this torque is a change in the Earth’s angular momentum or, equivalently, a decrease in the spin of the Earth such that the length of day is increasing, at present by about 0.001 seconds in a hundred years. At the same time the bulge exerts an equal but opposite torque on the Moon, slows the Moon down and the torque exerted by the bulge on the Moon slows the Earth down as well. This secular change is of the order of a few centimetres per year. Both changes are small but when integrated over longer time intervals the consequences become very significant. For example, after about two thousand years the Earth is misorientated relative to stars by some 10–15° in longitude and the position at which a solar eclipse is observed is displaced by this amount from a position computed on the basis of a uniform rotation of the Earth. Thus the theory of the Earth’s tidal acceleration can be tested by predicting, on the assumption of uniform rotation, places and times of eclipses or of other astronomical configurations and comparing them with observations of these events as recorded in the literature and history of older civilizations. Conversely, once the acceleration has been established such comparisons can be used for dating purposes.

Over longer time periods the consequence of the small but persistent tidal acceleration becomes even more dramatic. Four hundred million years ago the length of day was about twenty-two hours and the year consisted of about four hundred days. Curiously enough this can be tested against the records of tidal and diurnal cycles contained in the fossils of certain corals and bivalves. The growth of these organisms is controlled by the daylight and tidal cycles which influence the biological processes that deposit the thin incremental layers of calcium carbonate in their skeletons. The available results confirm that the tidal acceleration has continued over at least the last 500 Ma. The fossil remains of one of these early ‘astronomers’ is seen in Figure 6.11.

---

**Figure 6.9** Gravity variations in central Australia due to the solid tides over a nine-day period in October 1978.

**Figure 6.10** Because of the lunar attraction the elastic Earth deforms into an ellipsoid whose major axis is aligned with the Moon (top). If there is a delay Δt in the response due to the Earth not being a perfectly elastic body, the bulge will have been rotated through a small angle ΔωΔt. The Moon will also have moved through a much smaller angle nΔt and the bulge will no longer be aligned with the Earth–Moon axis. The torque exerted by the bulge on the Moon slows the Earth down and the torque exerted by the bulge on the Moon slows the Moon down in its orbit.
Gravity

The shape of the Earth

Gravity varies from place to place on the Earth's surface, reflecting the asymmetrical distribution of mass in the Earth's crust and mantle. A major departure from symmetry is the Earth's oblateness. Because of diurnal rotation any element of mass within the Earth is subject to the gravitational attraction by the remainder of the planet and by the centrifugal force acting in a direction perpendicular to the rotational axis. This latter force is at a maximum at the equator and acts in the opposite direction to the within-planet attraction. At the poles the centrifugal force vanishes. If the Earth responded as a fluid to these latitude-dependent forces its equilibrium form would be an oblate ellipsoid. The observed flattening is close to that expected for a fluid body, suggesting that the Earth responds essentially as a fluid to forces that act over very long time periods.

A more precise description of the shape necessitates the introduction of the term 'geoid', a level surface that closely approximates the ocean surface. Because of the anomalous mass distribution within the Earth the geoid departs from the ideal fluid ellipsoid, in some places rising above it and in other places falling below it by up to 100 m (Figure 6.12). When scientists talk about the shape of the Earth they usually mean the geometric form of the geoid. Only at sea is the geoid directly accessible and its form can be determined from gravity observations taken at sea level. On land the geoid is deduced from gravity observations taken at the surface.

6.11: The epitheca of a Devonian fossil coral. The epitheca forms part of the coral skeleton that records the growth rhythms in the form of a fine structure of ridges. This example illustrates thirteen bands, each consisting of about thirty ridges. These are interpreted as daily growth lines modulated by a monthly influence in the growth.

6.12: The global geoid as deduced from an analysis of satellite orbit perturbations and from surface gravity. The heights (m) relate to the best-fitting ellipsoid. Negative regions, where the geoid lies below this ellipsoid, are red.
Gravity anomalies

Gravity can be measured with a pendulum apparatus since the period of a pendulum depends only on its length and on gravity. Gravity is therefore essentially a measurement of a time interval and of length—the period of the pendulum and its effective length. Thus, if the pendulum length is kept constant, differences in gravity will be reflected in differences in the period. Gravity can also be measured using gravimeters, in which a weight is suspended from a coiled spring whose length changes in proportion to a change in gravity. Whereas pendulum instruments can measure absolute gravity, gravimeters can measure only changes in gravity, but the latter are more sensitive to small changes as well as simpler and faster to operate. Most gravity measurements are now made using gravimeters that measure gravity relative to a base station at which absolute gravity has been measured. Pendulum measurements typically have an accuracy of one part in $10^6$, while gravimeters operated on land may be three or four orders of magnitude more precise. A major development in the past two decades has been the measurement of gravity at sea with one part in $10^6$ precision, despite the disturbances caused by the accelerations of the ship. Thus gravity measurements are now available for many regions of the world's oceans as well as on land. In recent years absolute gravity has also been measured with great precision using 'free-fall' instruments in which the time interval required for a mass to fall through a known distance in a vacuum is measured. With these instruments such precision has been attained as to raise the possibility that secular and long-period changes in absolute gravity can now be measured.

Gravity measurements taken on the Earth's surface vary from place to place because of the Earth's oblateness, because of the variable distance from the centre of mass and because of an asymmetrical mass distribution within the planet. For geological and geophysical studies only the last of these effects is of real interest and it is convenient to correct the observed gravity for the first two factors (the free-air and Bouguer corrections, see page 106).

The difference between a corrected gravity measurement and the theoretical value for an idealized ellipsoidal Earth is termed the gravity anomaly. Gravity anomalies of a short wavelength or extending over a small area reflect near-surface density anomalies whereas anomalies persisting over large areas generally reflect mantle anomalies. A distinction that is a consequence of the attenuation of the force of gravity with distance. Figure 6.13 illustrates several examples.

The small- and intermediate-sized anomalies are readily surveyed with gravimeters but the regional and global anomalies are more difficult to determine in this way, not only because the ground survey approach is time-consuming but also because the long wavelength fluctuations are difficult to separate from the shorter wavelengths and the instrumental drifts that make gravimeters unsuitable for measuring the long wavelength anomalies, particularly at sea. However, satellites have provided a much more direct way of measuring the global and regional gravity field.

Measurement by satellite

An artificial satellite in an orbit about and close to the Earth would follow an elliptical orbit like that described for the Moon if the Earth were a radially symmetrical sphere. The better approximation of the Earth as an oblate spheroid requires that the Earth exert a torque on the satellite that tries to rotate the orbital plane into the equator. But because the angular momentum of the satellite motion is conserved the orbital plane precesses about the Earth's symmetry axis. Hence the effect of the Earth's flattening is primarily to cause the satellite's orbit to precess, both about the rotational axis and within the orbital plane, at a rate that depends on the exact form of the flattening. Typically the rates of precession are a few degrees per day and are readily deduced from a series of observations of the satellite's position from tracking stations on Earth.

Precise observations of an artificial satellite's motion show that the orbit is periodically perturbed from the precessing elliptical motion due
to the departures in the mass distribution from the oblate spheroid approximation. As the satellite passes over a density anomaly it is accelerated in its orbit, and while the density anomalies may be relatively small the cumulative accelerations of the satellite during successive passes become measurable.

The study of the Earth's gravity field using satellites involves celestial mechanics, satellite geodesy and geophysics. The problem of celestial mechanics is to describe the motion of the satellite not only because of the Earth's gravity but also because of a variety of other forces. The satellite is attracted by the Sun and Moon. Furthermore, the tidal deformations of the Earth result in a time-dependent attraction by the Earth on the satellite. Other forces of a non-gravitational origin include the drag forces experienced by the satellite moving in the tenuous atmosphere that exists even at altitudes well above 2000 km, and the force that radiation emitted by the Sun exerts on the satellite.

The problem of satellite geodesy is to compare the motion described by the equations of celestial mechanics with observations of the satellite's position and to deduce the various parameters—gravitational or other—entering into the celestial mechanics theory, describing the anomalous gravity field or defining the atmospheric drag coefficient or the intensity of the solar radiation. The problem for geophysicists is the interpretation of these results, to which we return below.

Tracking methods

Observations of satellite positions can be made in several ways by optical or electronic methods. A much used technique has been to photograph a satellite at night against a star background while the satellite is still illuminated by the Sun. These observations can give satellite positions accurate to about 10 m. Another much used method is the Doppler tracking of a satellite that emits a continuous radio signal at a constant frequency. Because of the motion of the source relative to the observer, the received signal is shifted to a lower frequency by an amount that varies during the spacecraft's passage over the tracking station. The comparison of the observed frequency with the standard frequency contains information on the satellite's motion relative to the observer and on its position in its orbit with a precision that ranges from about 10 m to 1 or 2 m.

During the past decade one of the most precise tracking methods to be developed has been the determination of the distance to the satellite by measuring the travel time of a very short laser pulse transmitted from the tracking station, reflected by an array of reflectors and received back at the station. Accuracies of a few centimetres are now possible in this way, which means that spacecraft accelerations due to quite small forces can be determined with considerable precision.

The satellite's motion is about the Earth's centre of mass and a logical choice for the origin of a terrestrial reference system is about this point. The positions of the tracking stations in this geocentric system will generally not be known with an accuracy that is compatible with that of the observations and it becomes necessary to determine simultaneously the orbital and force-field parameters and the
station positions. The latter can be determined in several ways.

A simple geometric procedure exists in which the satellite is used as a target that can be observed simultaneously from a number of stations. Assume, for example, that laser range measurements are made simultaneously from three stations whose relative positions are known from conventional terrestrial geodetic measurements. These observations determine the satellite's location as the intersection of three spheres with radii equal to the observed ranges. A simultaneous observation from a fourth station, of unknown position, fixes this station on a sphere centred at the satellite (Figure 6.14). Repeating the observations when the satellite is at different points in its orbit fixes the position of the fourth station relative to the other three. It is possible to measure relative station positions over continental and intercontinental distances with accuracies that approach those of the measurements themselves. This leads to the possibility that plate tectonic motions can be measured directly rather than inferred from less direct geophysical evidence.

While this geometric approach is simple in concept and does not require information on the forces that act on the satellite, the method is not widely used because the condition that the satellite be visible from many stations is restrictive. Instead, a procedure is usually adopted in which the station positions and the unknown parameters quantifying the forces are both determined simultaneously. Once these parameters have been determined with a satisfactory precision the motion of the spacecraft can be regarded as known and predictable and any additional observations can be used to determine the station's position relative to the orbit. This procedure is widely adopted, particularly when rapid position determination is required, as for navigation using the Doppler tracking of satellites.

Satellites used in these studies are varied. Almost any satellite can be photographed, whereas laser ranging requires that the satellite be fitted with cube-corner reflectors which reflect the incident laser pulse back to the station. Doppler observations require an on-board continuous transmitter of a constant frequency radio wave. To achieve high accuracies in describing the orbital motions, it is important that the air drag and radiation pressure forces are reduced. The simplest way of ensuring this is to have spherical satellites, free of protruding solar panels and antennae, with uniform reflectance properties and made of dense material. Geodetic requirements led to the design of two very dense, spherical satellites which are covered with cube-corner reflectors and are tracked entirely with lasers. The first satellite (Figure 6.15) was launched by the European Space Agency in 1975 into an orbit at an altitude of about 800 km. This satellite is used mainly in the determination of the Earth's gravity field and in the study of other forces on the satellite, principally those due to the periodic tidal deformations of the Earth. The second satellite, launched as part of the US space programme, is considerably larger and heavier and has been placed in a much higher orbit at an altitude of about 5000 km. Being further from the Earth, it is less perturbed by the Earth's gravity field than are lower spacecraft and this makes the satellite most useful for the precise determination of positions of points on the Earth and for measuring continental drift.

For gravity studies the lower the satellite the more sensitive it is to the Earth's anomalous density structure and the more useful it is, geophysically speaking. However, the air drag force also becomes more important and the drag perturbations begin to dominate the gravity field perturbations. Also, the lower the satellite, the more difficult it becomes to track it regularly from the ground. Together these factors limit the heights of geodetically useful satellites to about 700–800 km, meaning that only gravity anomalies of an extent exceeding about 2000 km can be detected using the methods outlined above.

An important development has been the use of radar altimeters to measure directly the height of the satellite over the oceans. The ability to track satellites from the ground with high precision means that the orbits are now well known and that it is possible accurately to compute the satellite's position within a reference framework whose origin lies at the centre of mass of the Earth. If the height of the satellite is measured using an on-board radar then the position of the reflecting
surface can be determined relative to the orbit. Thus, for measurements made over the oceans the geoid can be measured directly with considerable precision and spatial resolution. Figure 6.16 illustrates some examples of these measurements obtained from the Geos 3 satellite launched in 1975. Results such as these are of considerable interest in understanding aspects of the oceanic crust and upper mantle and most of the world's ocean areas have now been surveyed in this way by the Geos 3 and Seasat satellites.

Interpretation of gravity
With the results from satellite and gravimeter observations gravity is a better-known quantity over the Earth's surface than nearly all other geophysical quantities. Yet the interpretation of gravity is fraught with difficulty because a given gravity anomaly on the surface can be modelled by an infinite number of different density distributions. The geophysical problem is to separate the plausible from the improbable models. This inherent ambiguity can be illustrated by the simple case of gravity on a sphere of radius R and mass m. If all the mass m is concentrated at the centre of the sphere, gravity will be \( Gm/R^2 \). Likewise, if the mass is distributed in an infinitely thin layer just below the surface of the sphere gravity will also be \( Gm/R^2 \) and any intermediate radially symmetric density distribution between these two extremes will give the same value. In consequence, gravity observations on their own are not always very useful and are best interpreted together with other geophysical measurements and geological considerations. Examples of this will occur in later chapters.

Gravity anomalies
Despite this negative aspect gravity observations have provided some very useful results. One is that the variations in gravity are less than would be expected if they were due to topography alone. An example of this is illustrated in Figure 6.17, in which the result of gravity measurements at sea is plotted over a mid-ocean ridge in the form of gravity anomalies. A second example is given by the gravity anomalies over the European Alps. The observed gravity is first reduced to the geoid, the free-air correction, and then corrected for the density of the rock between the two surfaces, the Bouguer correction. A third example given by gravity across the continental margins of south-east Australia the corrections for the attraction of the topography and ocean water only increase the anomaly. All three examples indicate that the fully corrected gravity anomaly is a function of elevation. Over high elevated terrain the anomaly is markedly negative while over the oceans it becomes strongly positive, increasing in magnitude with

![Diagram of gravity anomalies](image_url)

6.16: Above, Geos 3 altimeter profile in the south-west Pacific and the ocean bathymetry. The geoid shows a pronounced negative anomaly over the South Hebridies trench. Left, Geos 3 altimeter profiles over the Derwent-Hunter seamount in the Tasman Sea. The geoid has pronounced positive anomalies over the centre of the seamount and negative anomalies over the flanks. This suggests that the load of the seamount is supported regionally rather than locally.
increasing depth. This suggests that there is much more to gravity than just topography. In particular, it suggests that regions elevated above the geoid are associated with a mass deficit somewhere in the crust below the mountain while the ocean basins are associated with a mass excess in their crust. Density anomalies which in both cases tend to compensate for the surface load.

The seismic evidence indicates that the Earth's crust varies considerably in thickness, being thin under the oceans compared with its thickness under the continents. The seismic evidence also points to an increase in density across the crust-mantle boundary, from about 2800–2900 kg/m³ to 3200–3300 kg/m³. Thus one explanation of the gravity anomalies is in terms of a crust of variable thickness; under highly elevated regions the crust is thicker and tends to compensate for the extra surface load while under the oceans the thinner crust results in the dense mantle material being closer to the surface. The crust generally behaves as a rigid layer capable of supporting stress differences but at some depth below the crust the mantle begins to behave more as a plastic material, deforming when subject to stress differences. Evidence for this response comes from laboratory measurements, from observations of post-glacial uplift and from the observation that the Earth's oblateness is very similar to what it would be if the Earth as a whole responded as a fluid to the centrifugal force. Furthermore the plate tectonic hypothesis (see Chapter 10) requires that within the upper mantle there is a region that deforms when subject to stress over a geological time scale so as to permit the lithospheric plates to move with respect to the deeper mantle.

Isostasy

Together, the gravity observations, the seismic evidence and the rheological response of the upper mantle support the models of isostasy proposed more than a hundred years ago. These were offered as an explanation of geodetic observations in India which suggested that the attraction of the Himalayas was less than it would be if this mountain range was simply on top of a radially homogeneous crust and mantle. The isostatic model assumes that the crust can be represented by blocks floating on a fluid-like and denser mantle, with each block moving vertically and independently of its neighbour. Figure 6.18(a) illustrates the model proposed by G.B. Airy (1801–92) in 1851 to explain the Indian observations. All blocks are of equal density but different heights, and they float, the longest blocks extending deepest into the mantle. At a certain depth, equal to or greater than the thickest crust, the pressure is everywhere constant and below it the mantle is in a state of hydrostatic equilibrium. The gravity anomaly according to this model consists of two contributions. One is from the visible parts of the blocks, the topography, and the second is from the 'roots', the parts of the blocks that have displaced the more dense mantle material. The two contributions are of opposite sign, and the resulting gravity anomaly is considerably smaller than if it were due to the topography alone (see Figure 6.17). At the same time as Airy expounded his model of isostatic compensation, J. H. Pratt proposed the same isostatic principle, of constant pressure at some constant depth below the surface, but with a model in which the crust is assumed to be of a variable density but with a constant depth such that under elevated areas the density is less than under low-lying areas (Figure 6.18(b)). Airy's model was based on the assumption that the topographic load stresses the Earth's crust beyond its strength-bearing capacity so that failure occurs by normal faulting, and that the crust below the load is depressed until the isostatic condition is attained. Pratt, on the other hand, assumed that the crust could be represented by blocks floating on a fluid-like and denser mantle. Figure 6.18(c) illustrates the model proposed by G.B. Airy in 1851 to explain the Indian observations.
hand, assumed that the mountains were the result of a thermal expansion of the crust as a result of a heat source in the crust itself or in the mantle below.

While both models are based on the apparently unrealistic assumption that the mantle behaves as a fluid and the equally unrealistic assumption that the response of one block is independent of neighbouring blocks, both describe surprisingly well the gravity observed over regions of variable terrain and indicate that the principle of isostasy is obeyed over much of the world.

Global anomalies
One advantage of the universality and simplicity of the isostasy hypothesis is that it is relatively simple to correct gravity measurements to take account of topography elevations and their roots. The gravity so corrected is the isostatic anomaly. If the traditional isostatic models were complete then these anomalies would be everywhere zero. Yet this is not so. Classical isostasy is not sufficient to explain the global gravity anomalies and the only satisfactory explanation is that they are due to density anomalies below the crust and not directly associated with the topography. This is perhaps the most important result that has come from satellite geodesy to date. From it two quite contradictory conclusions can be drawn about the Earth's mantle. One is that the mantle is sufficiently rigid for it to be able to support density anomalies elastically so that the present gravity field reflects conditions in the Earth at some time in its remote geological past. The concomitant stress differences in the mantle are of the order of 500 bars and the general consensus of Earth scientists is that this is excessive, that mantle materials will flow when subject to such stresses at the temperatures and pressures characteristic of the mantle. The most convincing evidence that flow would occur is seen in the postglacial rebound phenomenon in which the removal of late Pleistocene ice loads has resulted in a slow rebound of the originally depressed crust and in a flow of the mantle in response to the changes in stress which do not exceed a few tens of bars.

The alternative interpretation is that the gravity anomalies are associated with mantle convection and that the density anomalies are a consequence of temperature at a given depth not being everywhere the same. This interpretation is now widely accepted by geophysicists and is reinforced by the rather remarkable correlation that exists between the gravity anomalies and the surface expression of plate tectonics, the plate boundaries. The major subduction zones around the Pacific Ocean are all associated with rather broad gravity anomalies. The collision zones of the African and Indian plates with the Eurasian plate are also associated with positive but milder gravity anomalies as are parts of the ocean ridges, particularly Greenland and the Azores in the north Atlantic. The negative anomalies lie mostly over ocean basins and over old continental shields. These gravity anomalies are not fully understood but they do indicate that there is a relation between them and mantle observations. The question now is not so much whether convection occurs, but concerns such matters as the scale of the cells, the depth to which convection occurs, what drives it and the periodicity of the motion. The answers to these questions are still largely speculative but the gravity anomalies may make a useful contribution towards reaching an understanding.